Essays on Financial Regulation in Zambia



THOMAS QUIRK

ST CROSS COLLEGE

UNIVERSITY OF OXFORD

Submitted in partial fulfilment of the requirements for the Degree of Doctor of Philosophy in International Development

HILLARY 2023

ESSAYS ON FINANCIAL REGULATION IN ZAMBIA

THOMAS QUIRK, ST CROSS COLLEGE

DPHIL IN INTERNATIONAL DEVELOPMENT

HILARY 2023

ABSTRACT

A long literature examines how credit markets play an important role in supporting economic development in lower income countries. Yet, policymakers can have limited context-relevant research outlining how specific financial regulations might support credit market development in their economies.

This thesis contributes to the body of context-relevant research by examining the effects of two financial regulations implemented in Zambia in the early 2010s. The first is an absolute minimum bank capital regulation, implemented by the Bank of Zambia in 2012. This policy required almost all banks to raise significant sums of equity and was imposed in addition to – and we argue, due to the inadequacies of – the standard Basel bank capital rules. The second is a set of lending interest rate restrictions – more commonly known as interest rate caps – imposed on the bank and non-bank/microfinance sectors, a policy becoming increasingly popular in lower-income countries. Each policy had its justifications, but both were implemented to encourage the financial sector to support lending and local development.

Chapter 3 uses a detailed bank panel dataset to track how the imposition of the absolute capital rule: (i) forced banks to raise external equity and (ii) resulted in banks increasing their asset bases and capitalisation rates, but (iii) did not clearly lead to increased lending, at least in the short term; instead, banks resorted to holding additional treasury securities.

Chapter 5 examines the interest rate cap regulation and how it was associated with significant loan retrenchment in the microfinance market. There is also evidence that once the caps were imposed, lenders began to charge higher loan fees in lieu of interest. Further, such effects appear to have been plastic, not reverting when the caps were lifted.

Chapter 7 examines how the interest rate caps may have influenced the cost of bank financial intermediation, as measured by bank margins. We do not find evidence that the cap had an effect, possibly due to the interest rate caps being too high to bind the banking sector strongly.

DECLARATION OF AUTHORSHIP

AUTHOR: THOMAS QUIRK, 690536, ST CROSS COLLEGE

TITLE: ESSAYS ON FINANCIAL REGULATION IN ZAMBIA

SUPERVISOR/ADVISER: CHRISTOPHER ADAM

WORD COUNT: APPROXIMATELY 98,000, INCLUDING TABLES, FOOTNOTES, AND GRAPHICS

Please tick to confirm the following:

I have read and understood the University's disciplinary regulations concerning conduct in examinations and, in particular, the regulations on plagiarism (*The University Student Handbook* Section 7.7)	☑
I have read and understood the Education Committee's information and guidance on academic good practice and plagiarism	☑
The thesis I am submitting is entirely my own work except where otherwise indicated.	☑
It has not been submitted, either partially or in full, either for this Honour School or qualification or for another Honour School or qualification of this University (except where the Special Regulations for the subject permit this), or for a qualification at any other institution.	☑
I have clearly indicated the presence of all material I have quoted from other sources, including any diagrams, charts, tables or graphs.	☑
I have clearly indicated the presence of all paraphrased material with appropriate references.	☑
I have acknowledged appropriately any assistance I have received in addition to that provided by my supervisor.	☑
I have not copied from the work of any other candidate.	☑
I have not used the services of any agency providing specimen, model or ghost-written work, or unauthorised artificial intelligence in the preparation of this thesis/dissertation/extended essay/assignment/project/other submitted work. See also section 2.4 of Statute XI on University Discipline under which members of the University are prohibited from providing material of this nature for candidates in examinations at this University or elsewhere. Sections 3 and 4 of Proctors' Regulations 1 of 2003 dealing with unauthorised use of AI, also apply.	☑
I agree to retain an electronic copy of this work until the publication of my final examination result, except where submission in hand-written format is permitted.	☑
I agree to make any such electronic copy available to the examiners should it be necessary to confirm my word count or to check for plagiarism.	☑

Thomas Quirk **17.04.2023**

ACKNOWLEDGEMENTS

First and foremost, I'd like to thank my supervisor, Chris Adam, for his support and guidance throughout the thesis. He engaged in detail with the material and helped me navigate my path to submission. I'd also like to thank those who generously read my work and offered comments along the way, especially Doug Gollin, but also Margaryta Klymak (during transfer), Muhammad Meki (during confirmation), and Marijn Bolhuis. Gabriela Smarrelli and Diala Al Masri were crucial sounding boards throughout. I am thankful for the many useful comments I received over the years from participants in Cheryl Doss and Doug Gollin's QEH Quant Seminar.

This thesis was incepted during conversations with Bank of Zambia (BOZ) colleagues during my time as an ODI fellow in Zambia. It also relies heavily on data from BOZ. I'd like to thank Chilufya Kwenda, Raphael Kasonde, Richard Chirwa, and Joseph Munyoro for helpful discussions, assisting with raw data extraction, and answering my annoying questions about data from 15 years ago.

I am also grateful to the many friends in Oxford who helped make my time there enjoyable, particularly my cohort at QEH and Indrani Sigamany.

Last but certainly not least, I want to thank my family and friends for their support, including Dr Nico, Chris, Emma, Ben, Miljan, and the Lokurs. My Mum and Dad's support was unequivocal and vital throughout, but especially in proofreading swathes of material of variable quality. Finally, a special thanks to my amazing wife, Radhika, for providing not only comments on my material, but also encouragement and joy during the more and less enjoyable moments of the DPhil.

Contents

Chapter 1 - Introduction **8**

Chapter 2 – Background and a brief history of Zambia's financial system and credit market **14**

Characteristics of Zambia's Financial System 30

Chapter 3 - When Basel-inspired capital regulations are insufficient: Evidence from Zambia's use of absolute minimum bank capital requirements **38**

Introduction and Summary 38

Background 43

Related literature 53

Theoretical framework: Bank lending and deposit taking under an exogenous capital shock 57

Tracing the effects of Zambia's absolute capital regime 67

Data and Empirical Discussion 69

Data Sample 69

Results 73

Policy Implementation 73

Capitalisation Rates 81

Bank Asset Reactions 87

Robustness Tests 94

Heterogeneous Bank Reactions and Aggregate Market Structure Trends 94

Discussion and Conclusion 106

Chapter 4 - Introduction to Interest Rate Restrictions and Summary Results **110**

Background context on interest rate restrictions and their recent resurgence 110

Section contributions and summary results 113

Relevant Literature 116

Chapter 5 – Interest Rate Regulations – a case study of Zambia's non-bank/microfinance sector 120

Institutional Context and Zambia's Interest Rate Restrictions 121

The IRRs and Zambia's Non-bank sector 125

A framework model 129

Data Description 136

Empirical Methodology – Piecewise Distributed Lag Approach 137

Empirical Challenges 140

Baseline Results - Lending Growth 142

Alternative Approach – Difference-in-difference and Lending Growth 152

Results 158

The loan income shifting hypothesis and plastic policy effects 171

Summary and Discussion - IRRs and Zambia's Microfinance sector 183

Chapter 6 – An Exploratory Cross-country Analysis of MFI Interest Rate Regulations using the MiX 185

Introduction 185

Data 186

Results 192

Discussion and Conclusions 199

Chapter 7 – Zambia's banking system and IRRs: Bank Margins and Fees 202

Introduction and Background 202

Relevant Literature (bank margin determinants) 204

Theoretical Framing (IRRs and bank margins) 205
Dataset and Margin Measurement 209
Zambia Context and Margin Decompositions 211
Empirical Approach 220
Results 230
Background Results - Bank Margin Determinants in Zambia 230
Interest Rate Caps and Bank Margins 236
Results – Banks and Interest Income Switching 243
Discussion and Conclusion 247
Section Conclusion 248
Chapter 8 – Conclusion and Discussion 252

Chapter 1 - Introduction

Financial regulators in low-income countries face an unenviable task. While setting appropriate financial policy in *any* context is challenging and often requires complicated trade-offs (for e.g., between financial stability and economic output), setting policy in a less-developed financial market setting can be particularly difficult.

Lower-income or less financially developed markets often look quite different from middle- or higher-income markets. For one, they are considerably smaller and need to grow and develop; for example, the average share of domestic credit to the private sector in low-income markets was still less than 15% of GDP in 2018, whereas in more developed middle-income markets it was over 100%[1]. They are also challenging policy-setting environments in themselves: such markets can be illiquid and incomplete, market power and asymmetric information may be more likely, and institutions designed to deal with such simple market imperfections, such as credit bureaus and collateral/credit registries, may be weak or non-existent.

These differences in market and developmental characteristics mean that financial policy in less-developed financial markets may need to be context-specific and should: (i) consider objectives of supporting growth in the market itself, as well as the typical regulatory objectives followed by high-income countries, and (ii) account for higher levels of market imperfection. Take bank capital policy, for example. In addition to financial stability and output trade-offs, capital policy may need to account for the Government's specific market development objectives (like improving credit flows and financial access, attracting capital, or expanding the financial system's deposit base). Or consider consumer financial protection policies, such as interest rate restrictions. These need to balance allowing growth in nascent, higher-risk markets, with the typical objective of constraining potential usury.

These additional objectives and different market contexts can also result in policymakers setting policies that might not align with what is considered "best practice" in developed markets. Yet, we

[1] There is a long literature arguing that financial development is a critical part of the development process (see Levine (2005), Papaioannou (2007), and Beck (2012) for summaries of the core literature). While some research questions whether credit access and financial inclusion can drastically improve poverty alleviation (Banerjee, Karlan and Zinman, 2015), in this thesis we are also concerned with the broader and more structural role that financial deepening might play in the development process (see Beck (2016) for a discussion of this financial inclusion versus financial deepening distinction).

should not assume a priori that such policies are not effective or appropriate; different contexts, objectives, and politics can call for different policy solutions.

This thesis analyses two examples of financial policy solutions implemented in Zambia during the 2010s. Both regulations had a common goal of improving credit market access and market development (in addition to other objectives), but both sat outside the typical "international best practice" regulator toolkit[3].

The first policy, analysed in Chapter 3, is an absolute bank capital regime implemented by the Bank of Zambia (BOZ) in 2012. This new policy required banks to hold minimum equity of USD 100m if they were a foreign subsidiary, or USD 20m if they were domestically owned. The policy, which was a deviation from the typical Basel bank regulatory ratios, required most banks to undergo major recapitalisations[4].

BOZ implemented this policy because the Basel rules were no longer effective in achieving its multiple objectives. For one, BOZ felt the banking system was too risky, yet banks were already overcapitalised per the Basel-calibrated rules in place at the time[5]. BOZ also believed the banking sector was too small and needed to extend more lending to better support the economy and development. These two competing objectives created a dilemma for BOZ. Normally, BOZ would have been able to increase the Basel ratios to reduce the riskiness of the system (as other countries in Sub-Saharan Africa (SSA) had done). The problem was that banks could have simply reduced their loan books or risk-weighted assets to abide by any increased Basel ratio, which would have been detrimental to BOZ's developmental objectives. BOZ thus introduced an absolute capital regime to resolve this dilemma and hopefully achieve both objectives.

Although absolute minimum capital rules have previously been used in low-income contexts, there is minimal research on their effects and whether they should be seen as a useful tool for bank capital regulation, especially to achieve developmental objectives (like BOZ's). There is an empirical

[3] Both regulations have received inadequate attention in the literature, especially in a Sub-Saharan African (SSA) context.
[4] Over the past two decades, the Basel Accords – the international "gold standard" of bank regulation – have become the norm for banking capital regulation in both high-income and low-income settings. The Basel accords cover many aspects of banking regulation and have many different versions, but most relevant to this study is the Basel II capitalisation ratio, a version of which BOZ was implementing at the time it implemented the absolute capital rule, which required that banks risk weighted asset ratios be above 10%. Risk weighted asset ratios are the ratio of a bank's regulatory capital (or available capital) to its risk-weighted assets.
[5] High bank capitalisation rates in Zambia reflect the fact that banks hold many liquid non-loan securities in their portfolios and bank equity is seldom traded.

literature on bank capital, for example studies examining bank bailouts (see Li (2013); Norden, Udell, and Wang (2020); Puddu and Wälchli (2015), among others), exogenous capital shocks (see Peek and Rosengren (1997); Mora and Logan (2010); and Rice and Rose (2016)), and changes in regulatory ratio rules (see Aiyar, Calomiris and Wieladek (2014); Fang et al. (2020)). However, there is limited research on absolute forced recapitalisation and there is thus a need for empirical analysis of such regimes. Chapter 3 analyses the effects of BOZ's absolute bank capital regime and whether the policy achieved its objectives.

The Chapter's analysis uses confidential bank panel data to evaluate the regulation along a few key dimensions: (i) implementation (did banks raise capital as a result of the new rule, and more than they might have otherwise?), (ii) recapitalisation (did banks' capitalisation rates improve with the imposition of the capital regime, and did such improvements in rates persist?), (iii) asset growth (did banks actually expand their assets in response to the recapitalisations, and if so, what kind of assets did they expand into?), and (iv) heterogenous bank reactions (did some bank types react more strongly than others, and what could that mean for the market?). The Zambian context provides an especially strong setting for evaluating the implementation and effects of an absolute capital regime as the policy was neither expected by the banks and other market participants, nor introduced in response to contemporaneous economic conditions.

We find that, overall, the absolute capital policy was a partial success; banks did improve their capitalisation, but there was no evidence of increased lending. More specifically, we find: (i) strong evidence that banks that were constrained by the regulatory rule grew their regulatory capital more quickly than when the rule did not constrain them, ceteris paribus, (ii) that banks' capitalisation rates increased as they raised capital levels (and that the majority of these increases persisted), (iii) that banks grew their assets following the recapitalisations, but that much of this growth took the form of banks extending their holdings of Treasury securities, rather than growing their loan portfolios (as BOZ had perhaps hoped for), and (iv) that smaller banks (note that many domestic banks are smaller) grew their assets less quickly post recapitalisation, partly because they appear less likely to have expanded their debt financing post recapitalisation, potentially shifting market share away from domestic banks. This chapter makes a significant contribution to the literature, being the first to quantitatively evaluate a forced capital recapitalisation in a SSA, despite their not infrequent use.

The second policy, covered in Chapters 4-7, is the implementation of interest rate restrictions (IRRs) in Zambia from 2013-2016. An interest rate restriction, or interest rate cap, places a maximum limit on lenders' loan interest rates. In Zambia, these restrictions were placed on both the bank and formal non-bank lending markets. They were primarily a political response to perceptions that interest rates were too high and that the foreign-owned financial system did not support (and had never supported) Zambian development.

Broadly speaking, interest rate restrictions have long been a popular financial regulation, with governments or regulators typically imposing them with the aim of either protecting consumers from usury or lowering the cost of credit and improving credit access. There is a risk, however, that they may also result in loan retrenchment and higher use of non-interest fees.

In recent years, there has been some empirical progress made in investigating the effects of IRRs (including in middle- and lower-income countries), mainly in the form of quantitative country case studies examining effects on bank lending (see, Cubillos-Rocha et al. (2021); Romero (2021); Aiba et al. (2020); Heng et al. (2021); Roa et al. (2021))[6]. While this progress has been important given the recent resurgence in Microfinance Institution (MFI) IRR use across middle- and lower-income countries (as we document in Chapter 6), there remains a significant research gap in the SSA-specific context where there have been very few (if any) quantitative studies. Additionally, many of the previous quantitative analyses have modelled IRR regimes as a binary policy when we know (from the Zambian dataset we built for this thesis) that bindingness can vary substantially depending on the interest rate distribution within the lending market.

In Chapter 5, we examine Zambia's IRR regime and provide the first in-depth quantitative analysis of the implementation of an IRR on non-banks and MFIs in a SSA context. We constructed a dataset covering interest rate data and detailed financial statements for the non-banking sector, which helps us more accurately frame and measure how institutions were differentially affected by the cap with respect to their lending volume, something not typically done in the literature. To measure the effect of the cap on lending we use a combination of piecewise distributed lag models and event study methodologies. We also use detailed data on non-bank lending fees and commissions to help us track

[6] Many of these were published after the chapters in this thesis had been drafted.

and analyse how the IRR might have influenced lenders' pricing and fee practices, again providing a valuable contribution to the literature.

In summary, we find that Zambia's IRR likely had a material negative impact on some non-banks' lending volumes, specifically those lending at rates well above the IRR. For non-banks lending below the caps, we find little evidence of an effect. The IRRs also appear to have had a non-trivial impact on the pricing structure within the market, encouraging a strong and persistent shift away from interest income to fee and commission income. Worryingly, this change in pricing structure towards fees appears to have been plastic; it did not seem to revert when the IRRs were removed.

Almost all literature analysing the effects of IRRs (Chapter 5 included) takes the form of country case studies rather than cross-country analysis. This lack of cross-country analysis of IRRs and MFIs is justified: standardising cap regimes across countries is difficult and time-consuming, and consistent cross-country MFI data has not previously been publicly available.

In Chapter 6, however, we attempt a cross-country analysis of MFI IRRs (the first we are aware of in the literature) by constructing a new dataset. The dataset first tracks and standardises IRR regimes across countries, which we then match to the recently published cross-county Microfinance Information Exchange (MiX) dataset. We use simple descriptive statistics, finding that MFIs operating under an IRR regime grow their lending slower on average (both on the extensive and intensive margins), results consistent with our Zambia case study in Chapter 5.

As Zambia's IRR applied to both non-banks and banks, Chapter 7 examines the effect of the cap on banks, specifically on the cost of financial intermediation (measured by banks' margins). It also examines banks' use of fees and commissions in their pricing structures. As with Chapter 3, we rely on a newly constructed and detailed bank panel dataset covering all commercial banks in Zambia's credit market. We use an autoregressive distributed lag model (ARDL) to measure the cap's effect on margins and pricing structures. We do not find convincing relationships between the imposition of Zambia's IRR and either the cost of bank intermediation or banks' pricing/loan income strategies. We suspect this is because banks were relatively less bound by the cap regime, somewhat supporting the findings in Chapter 5 that less exposed lenders saw smaller effects.

Chapter 8 discusses whether the absolute capital rule and the IRRs analysed in this thesis were suitable for helping grow credit in a market economy and whether these types of policy solutions were the best choice (or were correctly calibrated) to achieve BOZ's goals, at least in the short term. We argue that neither policy directly achieved its goals of growing the credit market, perhaps suggesting that these policies were either not the best choice or were miscalibrated for the Zambian context[7]. We suspect that structural policies, like those aimed at improving market infrastructure (i.e., improving credit bureaus, collateral registries, and court processing), may have been better suited to achieve their goals. In this Chapter, we also examine possible unintended consequences of IRRs and absolute capital policies as well as the limitations of this thesis more broadly.

Finally, a note on the data used in this thesis. Much of the analysis, Chapter 6 aside, relies on newly constructed financial institution panel datasets built from raw, unconsolidated, confidential supervisory returns. This raw data was provided to me for this thesis after working at the Bank of Zambia as an Overseas Development Institute (ODI) Fellow.

The rest of the thesis is as follows: Chapter 2 provides background context on financial regulation and the credit market in Zambia. It helps frame the quantitative policy analyses to follow, placing them within an appropriate historical context. Chapter 3 is a standalone chapter analysing Zambia's absolute bank capital policy. Chapter 4 provides background on interest rate cap policies in general and summarises the findings and contributions from the subsequent quantitative chapters on IRRs. There are three semi-standalone quantitative chapters on interest rate caps: Chapter 5 analyses how Zambia's interest rate cap policy impacted its non-bank market; Chapter 6 is a preliminary exploration of MFI interest rate cap policies across lower-income countries, and Chapter 7 examines how Zambia's interest rate caps may have impacted the cost of financial intermediation. Finally, Chapter 8 briefly discusses the thesis' findings in a broader context.

[7] Note that the absolute capital regulation was successful in terms of improving bank capitalisation, a clear goal of the policy.

Chapter 2 – Background and a brief history of Zambia's financial system and credit market

The financial regulations we analyse in this thesis are strongly linked to Zambia's historical economic development challenges. As we will briefly explore in this chapter, several long-standing themes underpin financial regulation in Zambia and have helped shape the more recent policies analysed in this thesis. Most important is the deep-rooted tension between the majority foreign-owned banking system, which has long financed the larger, more formal, and economically productive sectors of the economy, and the Government's desire for a financial system that also provides broad and affordable access for smaller, more local, and rural companies and enterprises.

The quantitative chapters in this thesis will discuss the specific context around Zambia's absolute capital rules and interest rate restrictions (the two financial regulations we analyse in this thesis). Yet it is also important to understand how these regulatory decisions made in the 2010s, and the political and policy contexts that enabled them, stem from Zambia's economic development experience post-independence (in 1964) and liberalisation (in the 1990s)[8].

This Chapter has four objectives that weave through the discussion below: (i) to discuss the links between Zambia's recent financial policies and the history of its credit market and banking system, (ii) to trace Zambia's banking regulations through time, (iii) to describe Zambia's financial markets during the late 2000s and 2010s, during which the policy measures examined in later chapters were implemented, and (iv) provide background information on the Zambian financial system. All four objectives feed into and help contextualise the later quantitative chapters.

1964 – 1980s: Colonial extraction, Independence, and Bank Finance

Although implemented in the late 2010s, Zambia's recent absolute bank capital and IRR policies reflect several decades of dissatisfaction with how the country's financial sector has supported its development. The origin of the dissatisfaction has a long history and can be linked back to the

[8] In some informal sense, our empirical strategies will at times rely upon these deeper historical drivers of the policy. Along with the stable macroeconomic environment during which they were implemented, these deeper drivers of the policy help argue that their imposition were reasonably exogenous to other market developments.

extractive nature of colonial finance, structurally high-interest rates post-liberalisation, and the challenges of building a deep credit market in a low-income context.

Before independence, credit markets in what is present-day Zambia were dominated by large commercial banks servicing colonial investments, operations, and resource repatriation. Such a setup was not uncommon for the region, with colonial authorities encouraging and supporting foreign banks to set up operations. Many such banks were headquartered in South Africa and retained formal ties with London (Ailola, 1988)[9]. The banks' lending reflected this heritage – as many argue they still do today – with disbursements being made almost exclusively to expatriate civil servants or newly arrived British and South African businessmen, mines, or farmers[10]. In 1962, five banks operated 109 branches across the Federation of Rhodesia and Nyasaland (Crick, 1965). The three that operated in Zambia at independence were Barclays (est. 1918), Standard Chartered (est. 1906) and Grindlays (est. 1956). Bank regulation in these early colonial times was initially linked to a patchwork of English common law and statutes and at least partially relied on administrative policies within banks (Ailola, 1988). Formal within-territory rules were finally established with The Bank of Rhodesia and Nyasaland Act in 1956, which established a central bank responsible for monetary policy and regulation in what is present-day Zambia (for example, stipulating cash reserve requirements and basic monitoring powers)[11]. For almost all Zambians, the financial sector was exclusionary.

In 1964, Zambia won its independence, and Kenneth Kaunda became President, pursuing an agenda of modernisation. At the time, Zambia was riding the copper boom of the 1960s and had the fourth-highest GDP per capita on the African continent. However, a lack of colonial investment meant the new Government was beginning with essentially no physical or institutional infrastructure (Whitworth, 2015)[12]. Mining contributed approximately 50% of economic activity and 90% of exports. After independence, the Bank of Rhodesia and Nyasaland became the Bank of Zambia.

[9] Many of these banks were formally incorporated in the United Kingdom or had common directors between the parents and their African satellites.
[10] A racist system of discrimination called "job reservation" also meant that only Europeans could enter into skilled and well-paid jobs (Ailola, 1988), so other income-based avenues towards building towards financial access were also structurally constrained.
[11] Ailola (1988) provides a detailed analysis of regulation during the period, including discussion of the economic context in Zambia pre-Independence.
[12] The country had less than 1000 high school educated graduates and less than 100 university graduates after being a British protectorate for nearly 65 years (Martin, 1972, as cited in Whitworth, 2015)

Yet, the economic inequalities that existed between expats and Zambians at independence continued in the years after, as did the unequal access to finance and the banking system. Resident expatriates and foreign firms remained dominant in the financial sector and continued to receive the lion's share of credit disbursements. The fact that large bankable projects were still perceived to only be possible for external parties was clearly problematic. In the President's eyes, poor financial access for Zambian businesses was an issue:

> *"The banks, the insurance companies, the building societies, the hire-purchase companies and the other commercial financial institutions have not been willing to assist the Zambian businessman."*
>
> *Kaunda, Kenneth D. Zambia's Economic Revolution: The Mulungushi Declaration (1968) cited in Minogue and Molly (1974) p.115.*

The Government soon began a set of nationalist economic reforms to address these lingering inequalities. The Mulungushi Declaration of 1968 and the subsequent Matero Reforms were wide-ranging, nationalising key industries and aiming to abandon Zambia's reliance on overseas industry (Gulhati, 1991). There was also a strong move to reorient Zambia's economy towards state control[13]. The financial system was well within the intended scope of these post-independence reforms. One component of the reforms was to impose interest rate limits (deposits and lending) and foreign exchange controls as a means to constrain bank lending to expatriates and tilt credit disbursements towards the Government's economic agenda, the strategy being to limit borrowing costs for Zambians as well as indirectly ration expatriate borrowing with administrative restrictions and volume limits on foreign exchange flows (Harvey, 1972; Brownbridge, 1996). Another component of the reforms was for the Government to establish the Zambia National Commercial Bank (ZNCB, now ZANACO) to encourage Zambian and rural financial access, operating with both social and business principles. ZNCB was also to be the bank of the Government and parastatals, covering more than 75 per cent of the economy at the time (Ailola, 1998).

[13] Note that at the time state control was a norm internationally, and not just amongst communist states – large swathes of coal mining in the UK, for example, was owned by the state.

Despite the reforms, progress towards broader credit access was slow and frustration at the small flow of credit to Zambians, as well as criticism of the foreign banks' exclusionary lending and hiring practices, remained:

> *"they [the expatriate commercial banks] have been excessively conservative in their staff policies, in opening new branches in the rural areas and in their credit policies."*
>
> *Kaunda, Kenneth David. Take Up the Challenge. Zambia: Zambia Information Services, 1970, p64*

In 1970, the Government gave up on an administrative approach to regulating the banks and moved to nationalise the financial system (Harvey, 1973). As part of these reforms, many non-bank institutions operating in Zambia were nationalised and amalgamated to form parastatals. Yet, despite some willingness within the banking sector, the Government failed to agree with the foreign commercial banks on any acquisitions. So complete nationalisation of the financial sector was never achieved, and the foreign commercial banks remained independent (Brownbridge, 1996).

In lieu of attaining direct ownership, the Government reverted took a less direct approach and introduced the Banking Act of 1971 as a tool for controlling foreign-owned banks. The Banking Act introduced restrictions whereby foreign banks had to incorporate domestically and have half their board comprised of directors that were Zambian citizens or Zambian residents. Furthermore, tighter capital requirements were imposed on foreign banks, and their liquidity ratio calculations were not allowed to include balances and other assets held with their parent companies or other banks abroad (Ailola 1988). In addition, the Minister of Finance held onto the right to issue and withdraw bank licenses. Yet these new rules only gave BOZ and the Government negative powers: they still did not directly force banks to lend more than they would have otherwise to Zambians or ex-pats (Harvey, 1973).

Over the subsequent decades, the Government struggled to achieve not only its goals for the financial sector but also broad-based economic development. Between the mid-1970s and mid-1980s, per capita GDP in Zambia declined by approximately 25 per cent (IMF 1998). Although Zambia had one of the highest per-capita incomes at independence, export revenues remained undiversified and

reliant on copper. The Government crowded out private consumption and investment with what ended up being inefficient or uneconomic investments (World Bank 2002). When international copper prices declined from 1973 onwards, budget deficits grew, and the Government either borrowed or cut down on capital outlays to balance the budget (Gulhati, 1991 and Whitworth, 2014).

Banking regulation during this period restricted private financial market development. Policy was geared towards controlling inflation, the balance of payments, and financing government borrowing from banks, rather than encouraging market-based intermediation (Ailola 1998). Real interest rates were low and often negative, rationing private credit. Instead of using the market, the Government supported credit through state-owned lenders[14]. Until 1985, the banking system was characterised by high liquidity levels (partly a result of banks holding liquid assets in the "pipeline" for foreign exchange transactions and partly a result of high liquidity requirements) and high levels of government borrowing (Ailola, 1998). In 1977 and 1978, for example, Government borrowing from the commercial banking sector was 10 per cent of GDP (Gulhati, 1991).

Note that even in 1985 during the early liberalisation attempts, the Government was still concerned that the banking system was not providing broad credit access for Zambians. The creation of the Credit Guarantees Scheme in 1985, is likely evidence of this[15]. The scheme's introduction reflected the fact that the Government and the polity remained unhappy with the lack of broad-based credit access, a perception that persisted until the 2010s when the most recent caps were introduced.

Overall, the conditions for banking in Zambia between independence in 1964 and the country's switch to multi-party democracy in 1991 were challenging. Many of the structural underpinnings required for a successful credit market – efficient legal processes, collateral registries, credit bureaus, and strong prudential supervision – were not established. Coupled with the fact that many Zambian borrowers had low incomes, wealth, and collateral post-colonisation, the conditions for the formal

[14] Since independence, the Government set up several other bank and non-bank financial institutions, many with the objectives of expanding lending to small businesses and small farmers, including the Development Bank of Zambia, the National Savings and Credit Bank, the Zambia Agricultural Development Bank, and Lima Bank. At various other times, traditionally non-lending Zambian Government institutions, including INDECO and the Bank of Zambia, made disbursements to entrepreneurs and farmers; Zambian-owned deposit taking financial institutions and pension funds were also persuaded to concentrate their lending on Zambians (Ailola 1998, p5). The Government also set up other organisations supporting lending, including Industries Development Organisation, Village Industries Service, and the Small Enterprise Promotion Unit (Maimbo and Mavrotas (2003)

[15] The Scheme was created as part of the Bank of Zambia Act of 1985, one of the pieces of the legislation underpinning the broader reforms. It was intended to increase financial access by providing protection to financial institutions lending to small medium enterprises.

banking system to provide broad financial access at non-subsidised market prices were also not developed.

1980s - 2000: Attempted Liberalisation, and Liberalisation

By the mid-1980s, the accumulation of debt and a protracted economic decline over the previous decades created the context for market reform and financial liberalisation. Starting in the mid-1980s, international agencies lobbied and then persuaded President Kaunda to switch to a more liberal economic regime (Bates and Collier, 1995), and following the October 1983 election, Zambia entering into three IMF stand-by agreements between 1983 and 1986. These agreements targeted reduced government expenditure, reduced current account deficits, and devaluations in the Kwacha. The key reforms were implemented in September-October 1985 when the Government began to auction a portion of the foreign exchange market, loosening its grip on interest rates.

These initial reforms, however, were short-lived. Amid increasing inflation and a lack of political support[16], Kaunda and the Government abandoned currency auctions by 1987 and reintroduced interest rate controls. Despite these failed reforms, Kaunda still went on to win the 1988 election, although it would be his last. In October 1991, after further years of increased government borrowing and high inflation fuelling additional discontent, Zambia underwent its first democratic transition since independence and Frederick Chiluba and the Movement for Multi-Party Democracy (MMD) took power[17].

The years following the 1991 election saw change in Zambia's economic model as the new MMD Government implemented wide-ranging economic reforms[18]. The initial reforms involved removing maze subsidies and import restrictions, abolishing export licenses, allowing a faster crawl for the exchange rate, deficit reduction, and financial market reforms, including removing barriers to starting a financial institution and removing interest rate ceilings (McPherson, 1995). But later reforms also involved privatising much of the Government's commercial asset portfolio; by 1997

[16] Many Zambians associated the new foreign exchange auctions (and the reforms more broadly) with the contemporaneous decline in the value of the kwacha and increases in perceptions of inflation. Politically, this created resentment within the President's core constituencies (Bates and Collier, 1995), and, with an election due in 1988, continuing with reforms that were not popular was too risky for Kaunda.

[17] This made Zambia the first Anglophone country in Africa to undergo a peaceful and smooth transition from a one-party state to a multiparty democracy.

[18] Mcpherson (1995) p10 discusses some of the reasons why liberalisation was possible in 1992 (unlike 1987), the primary factor being a rejection of Kaunda and UNIP, and a desire for an "end to the economic stagnation and economic mismanagement associated with one-party rule".

the majority of state-own entities had been sold, and the "crown jewel" in the portfolio, Zambia Consolidated Copper Mines (ZCCM), was partly privatised in 2000.

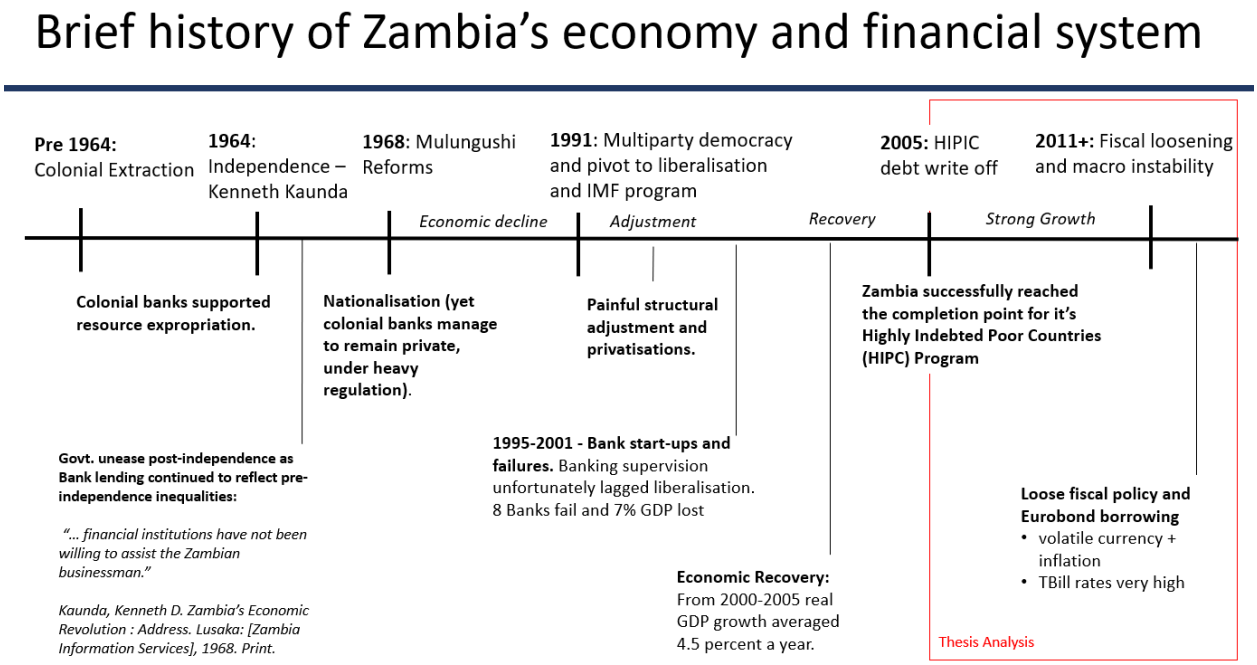
Unfortunately, the initial financial reforms in 1992 and 1993 had sequencing issues, resulting in severely negative economic outcomes. By liberalising capital flows and the financial sector before the Government had achieved fiscal balance and curtailed macroeconomic instability, Zambia experienced both a bout of extremely high inflation and high fiscal costs associated with liberalisation. Liberalising capital in an uncertain environment facilitated a reduction in demand for Kwacha, and the flight away from domestic assets, without an offset in the supply of money, resulted in high inflation and lower seigniorage revenue (Adam, 1995). Inflation, which had reduced from an annualised 400% in early 1992 to less than 50% in August, increased again to over 500% as foreign exchange was liberalised in October. Only when the Government implemented a cash budget that mechanically limited seigniorage to zero and grounded inflation expectations did inflation reduce.

The reform adjustment process was painful for Zambia and its economy. Growth was negative or less than 1% for much of the 1990s. Zambia saw de-urbanisation and deindustrialisation as previously protected government-run industries failed. The rural share of the population increased from 60% to 65% over the decade. Manufacturing value added fell from 30% of GDP to 10%[19]. The privatisations and liquidations aimed to recognise *"the need both to stem the fiscal haemorrhage from loss-making parastatals and to attract investment to enable potentially viable companies to survive"* (Whitworth, 2014, p.976); over 250 state-owned companies were liquidated, leased, or privatised between 1992 and 1999 (IMF, 2000). Yet these failures, privatisations, and the broader reform programme came with costs. McCulloch, Baulch, and Cherel-Robson (2001) find a dramatic increase in poverty and inequality in urban areas between 1991 and 1996 due to stabilisation, the removal of maize meal subsidies, and job losses resulting from trade liberalisation and the privatisation programme.

The period during and following liberalisation provided opportunities and challenges for the banking sector. In 1991, banking license regulations were relaxed both in terms of attaining a license as well as the absolute capital requirements required to open a bank. Around nine new local banks were established between 1991-1995. As Brownbridge (1996) p.11 notes, *"the primary motivation for the*

[19] Data from World Bank WDI.

establishment of most of these banks was the large profits to be earned from foreign exchange dealing, and from investing borrowed money in treasury bills, rather than from conventional lending activities". Indeed, at the end of 1993, local banks held almost double the amount of Treasury assets as loan assets. Despite this, because of high levels of previous repression between 1993 and 1995, bank credit to the private sector as a per cent of GDP still managed to increase from 4.7% to 8.4% (de Luna Martinez (2006), p20).



1990s – 2001: Liberalisation Banking Regulation and Bank Failures

In 1994 and as part of a broader set of financial market reforms, Zambia introduced the Banking and Financial Services Act (BFSA) to improve banking supervision[20]. Before liberalisation, the risks to the Government and the financial system of bank failures were limited because most banks operating in Zambia were subsidiaries of larger banking groups that were subject to their home country's supervision. This lower risk was reflected in BOZ's less hands-on approach to supervision (Maimbo 2000, p9). However, the new act clearly defined powers for BOZ to license, regulate, and supervise banks, including having the authority to define and enforce capital adequacy ratios and other prudential guidelines (Brownbridge 1996, p18). As a result, data returns that were previously

[20] See Chipimo (2014) p155-156 for a full list of financial market reforms.

infrequent, insufficiently detailed, and had no legal backing, were improved and the size of the supervision team within BOZ increased significantly.

Unfortunately, the time delay between financial liberalisation and the enforcement of the BFSA had cultivated several undercapitalised, undiversified, and illiquid local banks, planting the seeds for future bank failures. Initially, local banks competed intensively for a share of the deposits market (Simpasa, 2010). However, because the public perceived their risk as being higher than the international banks, they had to pay a funding premium. In addition, the limited lending that the local banks did do was at the higher-risk end of the market, due to the prevailing high interest rates and loan size limits (Brownbridge 1996, p12). As inflation began to fall in 1994 and the fiscal situation improved, Treasury Bill returns declined. In addition, increased competition in the exchange rate market reduced banks' transaction fees. Thus, key revenue lines for several smaller banks were being reduced, placing pressure on bank margins. Instead of letting their margins reduce, some banks increased risky lending to compensate (Maimbo and Mavrotas, 2003, p11).

Between 1995 and 2001, nine banks failed. The first three failed in 1995, including Meridian BIAO Bank, the fourth largest in the market. The loss of confidence associated with these initial failures resulted in a "flight to quality" as depositors shifted their savings from local banks to foreign banks that were perceived as less risky (Maimbo 2000, p24). This, in turn, resulted in severe liquidity problems for some banks. The next round of bank failures began in October 1997 with Prudence Bank; First Merchant Bank was last to fail in February 1998[21].

De Luna Martinez (2006) estimates that the total losses to depositors and taxpayers as a result of the failures amounted to 7% of GDP. The consequences were significant for the market: the large pre-colonial international banks took significant market share in the deposits market, growing from 35% in 1994 to over 50% in 1999 (Maimbo and Mavrotas 2000, p12), several banks put a hold on lending growth (de Luna Martinez , 2006), and financial access may have reduced as the larger formal banks had relatively higher account setup requirements and transaction costs. This shift towards foreign banks (perceived to be more exclusionary) may also have contributed to the negative narrative concerning how well the financial system was contributing to development.

[21] 1996 was an election year. Maimbo (2002), p243 argues that there was unlikely the political appetite for additional bank failures. BOZ, hence "continued futile efforts to find alternative remedies to the problems afflicting the small banks", rather than letting them fail. Otherwise, the failures may have occurred earlier.

2000 – 2010: Towards HIPC and Financial Sector Growth

Until 1996 the impact of Zambia's structural adjustment programme was adverse. The IMF estimates that per capita GDP declined by 22 percent over the first four years of the programme (IMF 1998)[22]. However, Zambia's economy did eventually rebound. From 2000-2005 real GDP growth averaged 4.5 per cent a year. Inflation, while still high, was trending downwards. The mining sector was recovering, mainly due to efficiency improvements linked to privatisation and a recovery in copper prices. Coupled with improved growth, hydrology, and expenditure constraint, the Government's fiscal performance improved, and interest rates continued to trend down (IMF 2004, IMF 2006).

In 2005, the year our datasets begin, Zambia successfully reached the completion point for its Highly Indebted Poor Countries (HIPC) Program. More than $5 billion of Zambia's external debt – around 70% of GDP – was written off, freeing up fiscal space and putting downward pressure on interest rates. Achieving the HIPC had been a long-term goal for the Government under the President, Levy Mwanawasa and was evidence that Zambia's macroeconomic management had significantly improved. HIPC completion was not, however, without political discontent: public sector salary freezes associated with expenditure constraints resulted in civil servants striking and contributed to the Government missing earlier HIPC completion points (Larmer, 2005).

The 2006 election was fought between Levy Mwanawasa, once again pursuing a political platform stressing his record on economic management and strong stance against corruption, and Michael Sata from the Patriotic Front (PF), who used a populist discourse stressing the negative impacts of liberalisation and reform, especially on the urban poor (Larmer and Frasier, 2007). Although Mwanawasa went on to win, Sata touched on a growing unrest that persists today: that the benefits of Zambia's growth and its liberalised economic system have not been equitably shared. This unrest and its emergence in the political rhetoric is important and highlights some of the political forces underpinning Sata's decision in 2013 to impose interest rate caps, one of the policies evaluated in this thesis.

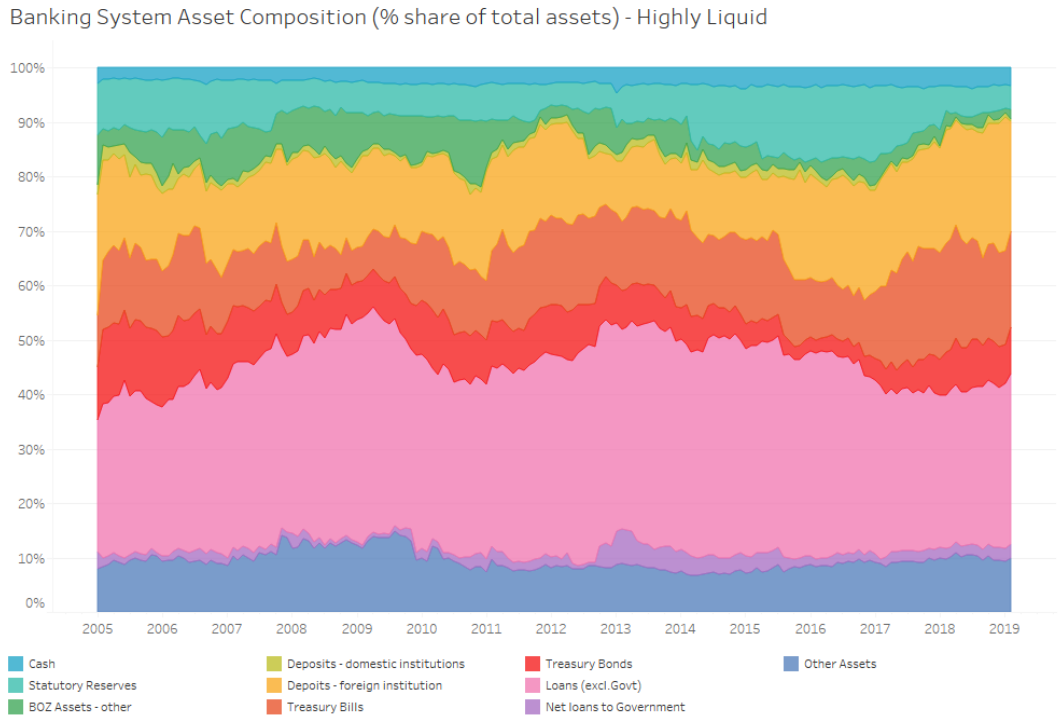
During the latter half of the 2000s, Zambia's credit markets rode the wave of improved economic growth and reduced macroeconomic volatility. The HIPC debt write-offs helped place downward

[22] Some of these losses may have been avoidable given better programme design and sequencing (IMF 1998), although The World Bank (2004) p.8 suggests, somewhat controversially, that these losses were the "price to be paid" for achieving relative price stability.

pressure on Treasury security yields, edging lending rates lower, albeit from very high nominal levels. Bank margins reduced until the Global Financial Crisis (Figure 39 in Chapter 7), and the 2006 Banking and Financial Services (Microfinance) Regulations created a much better enabling environment for smaller non-bank lenders. Indeed, the number of non-bank lenders increased from around 10 in 2005 to 35 in 2010.

Like most countries, Zambia was affected by the Global Financial Crisis (GFC). However, a lack of international financial integration meant that Zambia's banking system avoided direct financial exposures to the liquidity crisis faced elsewhere. Still, the eventual slowdown in the international copper price and a shift in capital away from emerging markets facilitated a depreciation in the Kwacha and flowthrough increases in inflation and Treasury yields. As a result, banks retrenched their lending and shifted their portfolio towards Treasury Assets (aggregate bank holdings of Treasury assets increased from 15% in 2008 to 25% in 2010 (Figure 1) and gross loans fell by 20%).

Figure 1 – Bank Asset Share Composition (% of Total Assets)



Note: Figure shows the composition of aggrgeate bank assets over the sample peroid. Each asset share type calculated as a percent of total bank assets (which is givewn on the y axis).

Source: BOZ Prudential Returns and Author's Calculations

While banks could pass on much of the rate increases into loan lending rates (as discussed in detail in Chapter 7), the GFC also resulted in banks booking significant loan losses. The commodities-driven slowdown meant that corporate and household balance sheets deteriorated, and non-performing loans (NPL) spiked, increasing from 7.2% in 2008 to 13.8% in September 2009. These elevated NPL levels persisted for several years and were one factor that led BOZ to impose absolute capital requirements in 2012, the topic of the next Chapter of this thesis[23].

2011 – 2014: The Patriotic Front, Capital Injections, and Interest Rate Restrictions

After losing in 2006, Sata and the PF went on to win the 2011 election. Soon after being elected, Sata announced his priority of cutting the cost of credit for Zambians (Piskadlo 2018)[24]. In January 2013, a blanket 9% margin above the Bank of Zambia reference rate was applied to all bank lending. Non-bank lenders also faced interest rate caps, some of which were potentially quite harsh. The Government's primary justification for implementing interest rate caps was to lower what they saw as high lending rates that were limiting financial access (IMF 2012 and BOZ 2012). Given people's perceptions that the financial sector had historically not provided broad credit access for Zambians, this policy likely resonated politically (although not with banks) and aligned well with the PF's message. The policy would have been made especially palatable because of Sata's ideological leanings, the lack of any direct fiscal implication of the policy, continued displeasure at liberalisation, and Zambia's and Sata's history with a state-controlled economy[25].

Sata presided over a mixed economic record. Buoyed by a copper price recovery post the global financial crisis and bumper harvest year, his first two years in office saw strong growth at 5.6% and 7.6% respectively. Inflation remained low at around 6.5%. During this time the Bank of Zambia also successfully implemented a new interest-rate-targeting monetary policy framework, aiming to create a more flexible monetary policy, improve the interbank market, and enhance bank competition (IMF, 2012). Although growth remained robust and inflation subdued in 2013 and 2014, the Government was beginning to loosen its fiscal stance. In 2011, the fiscal deficit was 1.8% of GDP; By 2013 and

[23] The other main objective was to and improve access to loanable funds (Gondwe (2013) pg. 3.
[24] Interestingly, Sata ended up not pushing strongly for across-the-board economic reforms when elected. As Cheeseman 2010, p.66 describes, "Recognising the widespread sympathy for Mwanawasa, and the prevailing public desire that his economic achievements not be undermined, Sata sought to…change his public image, stressing his respect for the dead president and toning down his combative manner." Interest rate controls were one exception.
[25] Sata was a Minister in Kenneth Kaunda's Government.

2014 it sat at 6.7% and 5.5% respectively, and much of these deficits were financed by non-concessional external borrowing.

On non-concessional external borrowing, it is worth noting that in September 2012 Zambia regained access to the international debt markets, an important milestone, issuing a 10-year USD 750m Eurobond at 5.625%. While this did provide an opportunity for investment, unfortunately, as with before HIPC, sovereign debt became one of the core negative narratives underpinning Zambia's economic performance in the years to come, leading up to its default in 2020 and ongoing IMF and debtor negotiations. Yet, for the majority of our dataset, and especially around the implementation of the Absolute Capital Regime and the IRR, the financial system was growing and reasonably stable.

2014 – 2020: Fiscal Dominance and a Path Towards Default.

In 2014, Sata passed away, triggering a by-election. Edgar Lungu, a minister from Sata's PF cabinet, won the by-election as well as the subsequent general election in 2016. Over his tenure, Lungu presided over both a deterioration in Zambia's fiscal position and several periods of significant vulnerability for Zambia's economy. The large fiscal deficits that began in Sata's last few years in power continued and increased under Lungu. By 2016, total public sector and publicly guaranteed debt was recorded at 60.5% of GDP, up from 21% in 2011 (World Bank, 2017)[26]. At the same time external headwinds – a weakening copper price, a depreciating currency, drought, and power shortages – exacerbated the vulnerabilities brought about by a loose and quickly deteriorating fiscal position. From the last quarter of 2015 and for much of 2016, Zambia was near crisis (World Bank 2016).

The first sign of problems and Zambia's debt vulnerabilities surfaced in late 2015 in the form of an inflation and foreign exchange crisis, occurring only a couple of months after a new $1.25m Eurobond was issued in July, Zambia's third since HPIC (the second, a USD 1 billion Eurobond, was issued in April 2014). Several structural factors lead to the crisis, including poor hydrology and power shortages, a depreciating currency amidst an increasingly dollarised debt stock, an uncontrolled budget deficit, and soft commodity prices (especially copper). But the effects were significant: from July to October 2015, the Kwacha appreciated (depreciated, as an indirect quote)

[26] Borrowing to invest in quality investments with high social returns can be a good financing strategy. However, there is evidence that this has not been the case in Zambia (World Bank, 2017)

by nearly 40%. The FX weakening was paired with a substantial inflation shock given high FX pass-through in Zambia: in October, November, and December, monthly inflation rates were 6.2%, 5.0%, and 2.1% respectively. Zambia's Eurobond-US T-bill spread widened by a factor of two as investors doubted Zambia's ability to service dollar-denominated debt. BOZ responded by tightening monetary policy and increasing reserve requirements. The policy rate, interbank lending rate, bank lending rates and Kwacha-denominated securities rose substantially. Caps on lending rates were lifted.

The banking system was placed under significant strain during the FX/inflation crisis. Depositors sought to shift into foreign-denominated accounts and dollarisation increased significantly. Nominal rates rose, and liquidity was extremely tight – the interbank rate rose well above 25%, as did Treasury yields (which had already been trending up since 2014 as the copper price slid). Many banks that were able to shift short-term security positions into offshore accounts did so (Figure 1). Bank lending weakened. BOZ continued to support the Kwacha and domestic foreign exchange reserves fell precipitously (Figure 3).

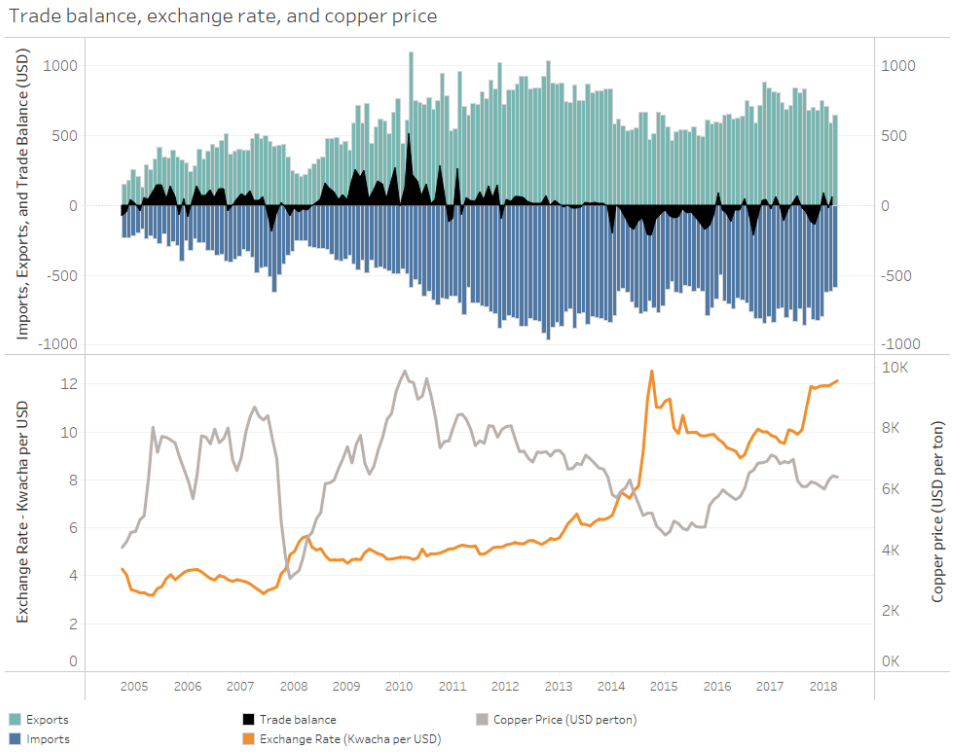
Improving copper prices in 2017 gave Lungu's regime more time. Yet the Government's loose fiscal stance continued, albeit without Eurodollar access. Domestic borrowing increased, as did non-concessional bilateral borrowing and arrears, with the Government was likely crowding out private lending and investment (IMF 2019 and Figure 3).

Summary

A few important points emerge from Zambia's financial sector history that are relevant to the policies analysed in this thesis. First, the perceptions that the financial system had not supported Zambian development are long-dated, and that the policy and political environment post-liberalisation and post-stabilisation (early to mid-2000s) still reflected this view. This context, perhaps coupled with a history of strong state intervention in the economy, underpinned and allowed for the imposition of the capital rule and the IRR regime. This is relevant for our analytical work because the more convinced we are that the timing of the imposition of these rules was politically "idiosyncratic", and not linked to the contemporaneous performance and processes in the economy, the more comfortable we can be with our empirical identification assumptions.

Second, the macroeconomic and financial sector conditions for much of our sample (2005-2019) were benign. Zambia's financial sector grew steadily from 2005 to 2015 (by a multiple of three), and the inflation and FX crisis in 2015 subsided reasonably quickly as the commodities cycle rebounded in late 2016. Our sample also stops in late 2018/early 2019, before the worst of the debt crisis. The periods around the implementation of the two financial regulations (in 2012 and 2013) were also reasonably stable and we argue were not imposed in response to acute market conditions. This is important for our empirical identification assumptions, as imposing a bank recapitalisation rule or an interest rate cap in response to adverse market conditions would likely result in the policies implementations being endogenous to the economic environment, complicating analysis and risking bias. Given the environment and market conditions when the policies were introduced, we might be less concerned about endogeneity and bias than otherwise.

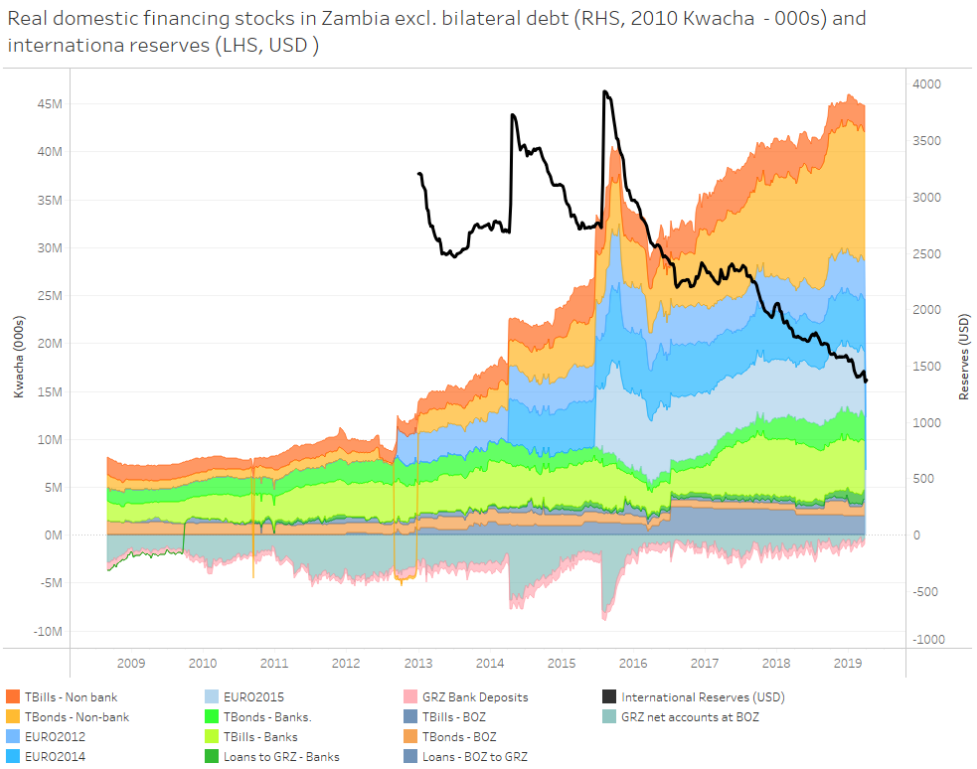
Figure 2 – Trade, Exchange Rates and the Copper Cycle



Note: The top graphic shows Zambia's monthly gross exports (positive) and imports (negative) and the subsequent net trade balance (black).

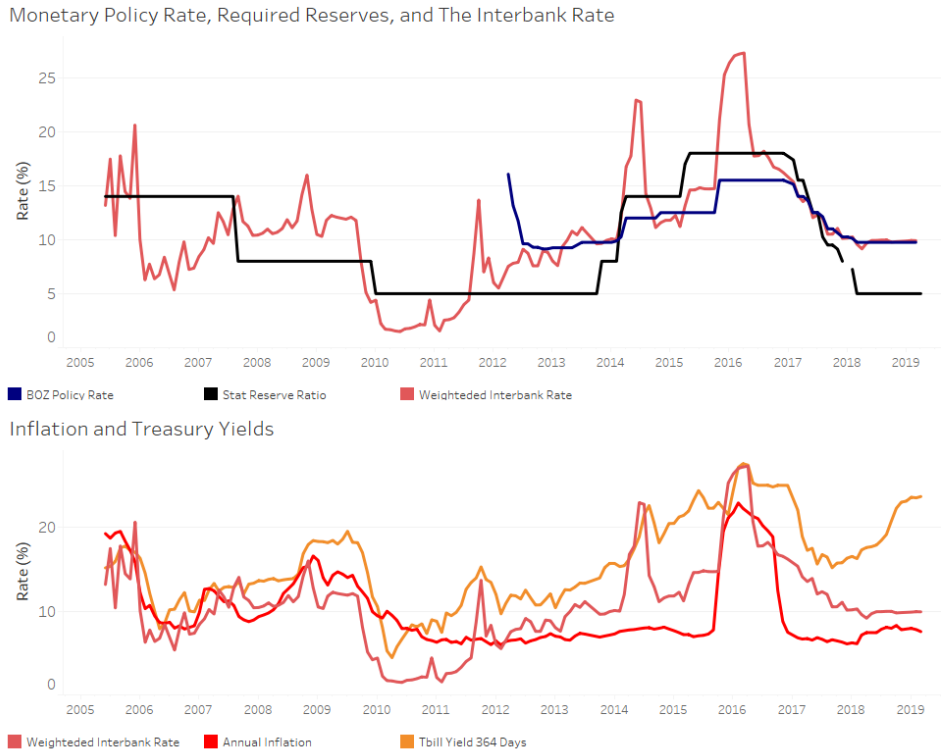
Source: BOZ Fortnightly Time Series.

Figure 3 – Accumulating Debt and Reducing Reserves



Source: BOZ Fortnightly Time Series, Capital IQ and Author's calculations

Figure 4– Monetary Policy and Interest Rates



Source: BOZ Fortnightly Time Series

Characteristics of Zambia's Financial System

Previously we examined the regulatory and economic history of Zambia's economy and financial system. Here, we provide a more static and descriptive overview of the types of lenders, loans, and borrowers in Zambia's credit market during our sample period that we analyse in later chapters. We begin by discussing the credit market as a whole, before specifically discussing the banking and microfinance sectors independently. Annex 1 provides some basic economic statistics for Zambia over the past 20 years.

Zambia's Credit Markets

By the time BOZ implemented the absolute capital rule in 2012, Zambia's formal financial lending sector comprised 18 banks and 35 non-bank lenders. The largest sector in the credit market, by value, is the banking sector, which holds 80% of all outstanding loans. By number, banks hold less than a third of the market, although clearly with a higher average loan value (see Table 1). Compared to

high-income countries, Zambia's credit market is quite small relative to the size of its economy (domestic debt to the private sector sat at 16% of GDP in 2016), though it is broadly consistent with its regional peers (Figure 5). Although Zambia's credit market has grown strongly in the decades since liberalisation, there is still substantial room for future growth. Credit or account access to formal institutions sat at 38.2% of the working age population in 2015 (Finscope, 2015).

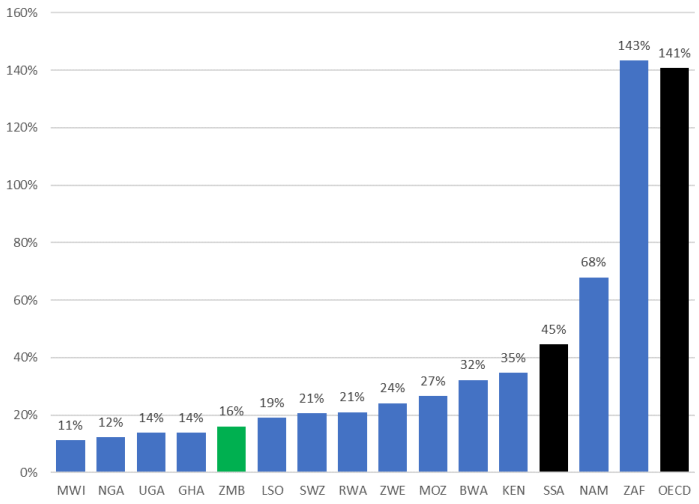
Table 1 – Lender types in Zambia's credit market (2018)

Institution Type	No. of Institutions (2018)	Amount of Credit (USD, m)	Number of loans	Average loan size (excl. mobile)	Loss Allowances (%)
Commercial Banks	**18**	**2,430 m**	**280,000**	**8,700**	**9.50%**
Non-Bank Financial Institutions	**48**	**632 m**	**472,000**	**1,300**	**13.50%**
Building Societies	3	76 m	30,000	2500	17.10%
Microfinance Institutions, incl. dep taking	35	457 m	430,000	1100	8.50%
Savings and Credit Institutions	1				
Leasing Finance Institutions	4	85m	12,000 excl. mobile	5700	32.80%
Development Finance Institutions	1	18 m mobile	1,000,000 mobile		
Financial Businesses	4				

Note: The table shows composition of Zambia's formal credit market by lender type. Mobile lending is excluded from some calculations for clarity (mobile lending can be very high volume but with low average size, skewing perceptions of the market). Earlier loan number data is unfortunately not available.

Source: BOZ CMMP.

Figure 5 – Domestic debt to the private sector as % GDP (2016)



Source: World Bank WDI.

In terms of borrowers, households and large corporate borrowers receive the most credit by volume (approximal USD 1 billion each), followed by large agricultural borrowers/commercial farms (USD 370 million). MSMEs and small subsistence farmers received significantly less (USD 200 million and 27 million, respectively). Most household lending is in the form of payday lending or account-

collateralised lending, typically disbursed to civil servants and those formally employed. MFIs provide a significant share of credit by number/volume (more than 50%), although less than a quarter by value.

Average loan sizes and the total number of loans vary significantly by lender and borrower type. For example, the loan size for a personal loan across all lender types in Zambia is relatively low; the mean is USD 960 (with high variation), while the median is much lower. Table 2, below, section shows some publicly available MFI lender data for which average loan sizes are much smaller. Predictably, loan sizes are larger in the market's more formal segments: large business and commercial agricultural lending average USD 0.18m and USD 0.25m per loan respectively.

As with many lower-income countries, non-performing loans (NPLs) have historically been high in Zambia. Over the sample period, NPLs in Zambia average over 10%, significantly higher than higher-income countries but approximately in line with regional comparators. Using the Credit Market Monitoring Programme (CMMP) dataset, we can also see significant variation across borrower and lender types. At the end of 2018, household lending, which by value is typically collateralised or income-linked, had a relatively low NPL (6.3%) compared to large businesses (13.2%), MSMEs and small agricultural borrowers (19.5%) and large agricultural/commercial farmers (24.7%). These high NPLs tell some of the story behind why we see quite high loan interest rate levels in Zambia[27].

Most credit in Zambia is lent in urban settings, approximately 90% by value and 75% by volume. The more urban regions – Lusaka and the Copperbelt – correspondingly receive the most credit. In recent years small loan value mobile lending has driven much higher credit market access across the country, although still only in small average loan sizes with short durations (Finscope, 2020).

[27] Note that NPLs are the stock of non-performing loans and are not 1:1 linked to default rates that lenders use in loan pricing – but they're clearly related.

Table 2 – Zambia credit market statistics by borrower type (2018)

Borrower Type	Number of Loans	Total Loan Value (USD, m)	Average Value (USD)	NPLs (%)
Business- Large (>100 emp)	5,524	1,030	186,524	13.26%
Business- MSME (<100emp)*	96,653	205	2,120	19.54%
Agriculture- Large (>50 ha)	1,453	373	256,475	24.68%
Agriculture- Small and Emergent (<50 ha)	7,192	27	3,779	19.48%
Households and Individuals*	1,259,766	1,204	956	6.33%
Government	1,510	185	122,461	0.08%
Other end users	324	39	120,120	29.11%
Total	**1,372,422**	**3,063**	**2,232**	**11.8%**

Note: Table contains data showing different outstanding loan numbers and values for borrowers in Zambia's credit market. Abbreviations are as follows: "emp" stands for number of employees; "ha" for hectares; and NPLs for outstanding non-performing loans. The borrower types are mutually exclusive. * indicates that the borrower type includes low average value mobile lending.

Source: BOZ CMMP

Banking Sector

In 2012, Zambia's banking sector comprised 18 banks, 4,500 employees, with 350 branches and 800 ATMs. Most bank infrastructure resides in either Lusaka (175 branches) or Copperbelt (100 branches). From 2015 onwards, growth in physical retail infrastructure slowed or stopped, reflecting a switch towards internet and telephone banking. Over the data sample, there was an average of approximately 200 bank accounts per 1000 adults in Zambia, indicating that broad-based financial access via the banking system remains low. The most frequent reason for this low cover, on the demand side, is that incomes are simply too low to justify the account fees or meet minimum account requirements (Finscope 2009, 2015).

A small number of international banks hold most outstanding loans and service most of the depositor base, in addition to ZANACO, the former and now partially state-owned bank[28]. The top three banks lent 74% of total outstanding loans in 2011 and held 67% of total deposits (Simpasa, 2013). These large banks lend across multiple products and borrower types, although typically concentrate on personal lending (salary-backed, payroll-linked, or mortgage-backed) or large corporate lending. Personal lending is typically provided to formally employed borrowers from the Government or corporates. These larger banks typically source diversified financing, attracting both short-term demand deposits and long-term deposits.

[28] In 2012, approximately 10-12 banks had more than USD 100m in total assets, but only four, Stanbic, Barclays, Standard Chartered Bank, and ZANACO had more than USD 1 billion (and none had more than USD 1.2 billion).

Some of the smaller banks, on the other hand, rely heavily on lending to corporates or the Government and use longer-term, higher-yielding term deposits, not necessarily running large retail networks.

Government securities make up a large share of banks' assets, with 80% of banks holding between 10% and 40% of their portfolios in government securities (excluding any direct ministry loans), which is a very high proportion. With reserve requirements and funds held overseas, lending by the banking sector makes up less than 50% of bank assets on average. As we will discuss in detail in Chapter 3, this low lending and high liquidity has historically led to very high risk-weighted asset ratios. Bank financing in Zambia comprises short-term demand deposits (more than 50%), term and savings deposits 25%, 10% equity, and the remaining being foreign loans and other liabilities. This is quite low leverage compared to other banking systems, contributing to the high capitalisation rates[29].

Since the banking system stabilised post the mid-1990s bankruptcies, competition in the market has been improving. In 2006 there were 13 commercial banks in Zambia, but by 2016 there were 19. Simpasa (2013) finds that foreign entrants improved the competitive environment in Zambia, at least up until 2011, and that banks have typically operated in a context of monopolistic competition. Our calculations show that bank concentration, an albeit imperfect measure of competition, has been improving.

As with other countries with a history of inflation, Zambia's banking system is highly dollarised. Approximately 40% of lending by value (in the dataset) was made in USD[30]. However, there is variation across loan types: 80-90% of lending to the mining sector from 2005-2019 was in USD, manufacturing and agricultural lending averaged approximately 60%, while only 15% of personal lending was in dollars. For our analysis, this dollarisation means that almost half of all lending was subject to foreign exchange valuation effects, which we discuss in Chapter 3 and Annex 5. Dollarisation is also present in the deposit market, with approximately 50% of demand deposits (on average) being held in foreign exchange accounts over the period.

[29] The median and aggregate risk-weighted asset ratio in Zambia over the sample ranged between 20%-27% over the sample.
[30] The median bank lends approximately 40% in USD too.

Lending interest rates in Zambia are typically floating, which means that banks pass on most interest rate risk to borrowers directly; borrowers thus face the brunt of interest rate or inflation risks. Most loans – at least sixty per cent – have interest rates that adjust within three months. This trend towards floating rates has been increasing over the sample (Figure 43 in Annex 2). Loans do typically have a longer contractual maturity than their interest rate terms, with approximately 50% of outstanding loans having more than a year until maturity at any point on average over the sample (Figure 44, also in Annex 2). For our regressions these "mechanical" constraints mean that it may take time for banks to adjust their portfolios after a policy change or a shock, and so we need to think about dynamics in our quantitative analysis carefully [31].

Non-Bank Sector

In 2012 there were 35 lenders in Zambia's non-bank credit market, up from only 11 in 2005 (it was not possible for for-profit MFIs to get a license before the 2006 Banking and Financial Services (Microfinance) Regulations were enacted). While several different types of lenders make up the non-bank market in Zambia, most new lenders entering the sample were consumer-facing Microfinance Institutions (MFIs). As with the banking sector, the non-bank sector was impacted by the GFC, the IRR in 2013, and the inflation crisis in 2015/2016. However, there has been consistently strong aggregate growth over the period, a reasonable proportion of which has been driven by a single lender, Bayport, which operates in the consumer salary-collateralised market. The non-bank market is reasonably segmented, with a considerable spread in interest rates charged across lenders, as shown in Figure 16 in Chapter 5.

Many non-bank lenders in Zambia complement commercial banks by lending to underserved rural households, MSMEs and small agricultural enterprises who would not be serviced by commercial banks (APF, 2014); however, some operate in payday or salary-collateralised lending markets, mortgages, and corporate lending and do compete directly with bank product lines. Although non-banks often have more rural networks, most of their branches and lending are still located along the 'line of rail' (i.e., from the Copperbelt Province in the north to Livingstone in the Southern Province via Lusaka).

[31] Note that this maturity and contract data has not previously been compiled and published in Zambia and so Annex 2 provides the figures.

What are the characteristics of MFIs in Zambia specifically? Table 3 provides an overview of book and loan sizes of a sample of MFIs in the market, using self-published data from the Association of Microfinance Institutions of Zambia, data form APF (2014), and other publicly available data[32]. Although not representative and geared towards MSME and small agricultural lending, Table 3 shows that average loan sizes are typically smaller than the market aggregate borrower data in Table 1 might suggest. The types of products MFIs provide are reasonably standard for non-bank lending in the region, including small business loans, trader loans, group loans, basic savings products, and payday/salaried loans to formal or for government-employed individuals. There can be variation in the types, sizes and pricing of products offered. Some MFIs are or were non-profits; however, there are also many for-profit lenders.

Other non-bank institutions include leasing companies (which have seen slow and stagnant growth over the period) and building societies, both state-owned, such as Zambia National Building Society, and privately held, such as Pan African Building Society[33]. Some of these non-MFI non-banks lend to larger companies than the MFIs, however, there can still be reasonable cross-over for small business and personal lending.

In 2015, only 2% of adults in Zambia borrowed from formal non-bank lenders. Lending in Zambia is challenging for many reasons, including low levels of formal education, informal income, low-income levels, and lack of documentation (Finscope 2009; Finscope 2015)[34]. Small loan sizes also create more difficult lending economics, and high average interest rates reflect this as well as the other challenges (see Figure 17 in Chapter 5 for the deciles of interest rates in the market over time). In Chapter 5, we will see how this high interest rate environment in the non-bank market meant that many lenders were affected by the Government's 2013 IRRs.

[32] Unfortunately, the detailed and disaggregated non-bank prudential data used in this thesis is confidential at the disaggregated lender-level. Otherwise, disaggregated data is not available unless via direct lender surveys.
[33] We do not include The Development Bank of Zambia, technically a non-bank lender, or Zambia National Savings Bank in our non-bank analysis. The former does large development infrastructure lending and the later behaves more like a deposit-taking banks than our other non-bank lenders.
[34] In 2009, the Finscope Survey found that only 2.6% of financially excluded adults had access to have the necessary documentation should lenders enforce legal documentation requirements. Most adults have a national identifier card; however, far fewer have access to proof of their stated residential address, often a requirement for loans.

Table 3 – Non-representative sample of non-bank lenders (2013)

	Microfinance Institutions							
	Vision Fund Zambia	Cetzam	FINCA Zambia	Pulse	Micro Bankers Trust	Agora	Micro-Loan Foundation	Bayport
Outstanding Portfolio (USD)	2.3	4.1	11.5	12.4	1.6	1.7	0.5	100
Number of Loans	8700	9800	39000	16400	16900	11200	3600	70000
Average Loan Size (USD)	264	418	295	756	95	152	139	1429
Average Annual Percentage Rate (Dec 2012)	121%	105%	87%	90%	88%	76%	74%	33%

Note: Average Annual Percentage Rate refers to the lending interest rate from BOZ's Cost of Borrowing Survey. Table shows simple statistics for a non-representative sample of MFIs in Zambia's credit market.

Source: BOZ interest rate survey and APF (2014).

Chapter 3 - When Basel-inspired capital regulations are insufficient: Evidence from Zambia's use of absolute minimum bank capital requirements

Introduction and Summary

Are Basel-inspired bank capital rules always enough for regulators operating in less developed financial systems? Over the past two decades, the Basel Accords have become the norm for banking capital regulation in both high-income and low-income settings[35]. But in some contexts, the Basel capitalisation rules (for example, the risk weighted asset ratio) may not be enough. In less developed financial systems (LDFSs), banks often hold very high levels of capital, frequently well above the minimum Basel-inspired capital ratio rules. Yet, at the same time, regulators often worry that banks are *still* not capitalised enough, given the riskier contexts that such regulators know that their banks operate within[36,37].

To adequately reduce bank and financial sector risk, one option available to regulators could be to increase capital ratio rules significantly above the Basel-inspired rule calibrations.[38] The problem, however, is that LDFS regulators face several downsides to directly increasing their capital ratio rules. Bank regulators in LDFSs face a more complex set of objectives as compared to their high-income counterparts: in addition to the financial stability/economic growth trade-off that Basel was designed to address, LDFS regulators also have developmental goals of increasing credit market

[35] See Jones and Zeitz (2019) and Jones (2020) for a discussion of Basel implementation across low-income countries.
[36] The risk weighted asset ratio (also known as the capital adequacy ratio or capital-to-risk weighted assets ratio) is the ratio of a bank's regulatory capital (or available capital) to its risk-weighted assets. Regulatory capital is made up of both Tier 1 capital (which captures core capital, including paid-up common shares and retained earnings, etc, and is the capital available to absorb losses without the bank ceasing trading) and Tier 2 capital (which includes revaluation reserves and some equity-like debt instruments, and is the capital available in the event that the bank is wound up). Risk-weighted assets are calculated as the weighted sum of bank assets, with the weightings being determined by the riskiness of the asset. A US Treasury bill or cash both have a zero risk-weightings, while non-secured household loans would have a 100% weighting. The exact calculation of risk-weightings depends on which Basel Accord the government is implementing and whether the bank is using its own credit risk modelling approach.
[37] It is perhaps not surprising that Basel-inspired rules calibrated for large and deep financial systems, such as the USA or France, are limited in their relevance.
[38] The average aggregate risk-weighted asset ratio (RWAR) for African countries in 2011 (many of which have LDFSs), for example, was 18.6%, already well above the original Basel RWAR benchmark of 8% and the mean minimum risk-based regulatory capital ratio in Africa of 9.7% (Source: IMF IFS).

depth, improving financial access, attracting capital, and expanding the financial system's deposit base. Ramping up the risk weighted asset ratio (RWAR) or leverage ratio requirement to reduce their banks' risks might risk banks deleveraging, putting their financial market development objectives at risk.[39]

To help solve this "over capitalised but still too risky" problem, LDFS regulators have occasionally implemented large increases in absolute minimum capital requirements.[40] Such policies require banks to raise additional capital up to some absolute threshold, but do not directly constrain banks' debt financing and relative capital ratios.

Despite the use of absolute minimum capital rules in LDFSs, there is very limited research on their potential impacts and whether they are a useful tool for bank capital regulation. There are limited detailed quantitative evaluations of whether absolute capital rules can be used to improve bank capitalisation rates (and not risking deleveraging) while supporting loan growth and market development without creating serious risks. As LDFS regulators have used and continue to use absolute minimum capital requirements, the lack of research on them (while not surprising, given the general lack of research examining the effects of the Basel-inspired capital rules in LDFSs), leaves policymakers with little quantitative support for their policymaking.

In this paper, we address this gap by empirically evaluating one episode of an LDFS regulator implementing a significant increase in its absolute minimum capital rule (in addition to the Basel RWAR rule) to improve bank capitalisation rates and financial market development. We investigate Zambia's absolute minimum capital regime, announced in January 2012, which required banks to hold a minimum capital value on their balance sheets of approximately USD 100 million for foreign banks (up from USD 3.5 million) and USD 20 million for domestic banks (also up from USD 3.5 million), independent of the size or riskiness of a bank's assets.[41]

[39] Deleveraging is when a bank tries to reduce its total financial leverage. In this paper we typically use it in relation to RWAR – in this case, deleveraging could include reducing bank assets (for example, selling or not rolling over bank loans and using the process to reduce debt financing (including deposits)) or reducing the risk-weighting of its asset portfolio (for example, by selling loans and buying Treasury Bills).

[40] Note that absolute minimum capital rules are common in Africa. What is less common is using large increases in such rules as tool for driving system-wide recapitalisations. Two examples of large increases in absolute minimum capital requirements are Ghana in 2017 (increasing from approx. USD 27 million to USD 90 million) and Nigeria in 2004 (increasing from approx. USD 5 million to USD 61 million).

[41] Almost all banks in Zambia held more capital than the minimum absolute rule at the end of 2011 (when it was at USD 3.5 million), but only one held more than the new January 2012 limit when it was introduced.

We evaluate the regulation along a few key dimensions: implementation (did banks raise capital as a result of the new rule, and more than they might have otherwise?), recapitalisation (did banks' capitalisation rates improve with the imposition of the capital regime, and did such improvements in rates persist?), asset growth (did banks actually expand their assets in response to the recapitalisations, and if so, what kind of assets did they expand into?) and heterogenous bank reactions (did some bank types react more strongly than others, and what could that mean for the market?). The Zambian context provides an especially strong setting for evaluating the implementation and effects of an absolute capital regime, as the policy was neither expected by the banks or other market participants, nor was it introduced in response to any particular contemporaneous economic conditions.

Although the Bank of Zambia's (BOZ, the financial regulator in this case) objectives for the policy were to strengthen the banking system and improve access to loanable funds, there is not a strong theoretical literature or directly relevant empirical literature to support this link. To help us begin to evaluate this link, we build a simple theoretical framework to guide our empirical analysis and establish our hypothesis.

To empirically evaluate the effects of the absolute capital regime, we construct a new bank panel dataset from the Bank of Zambia's quarterly prudential data return covering the banking system in Zambia from 2005-2019 (including approx. 17 banks and 40 quarters of data on average). As mentioned, we evaluate the policy along several dimensions, each of which has its own specification; however, most estimates rely on some form of distributed lag model with bank fixed effects (exploiting within-bank variation) and aim to connect bank recapitalisations (that were directly linked to the change in absolute minimum capital requirements) to changes in other bank outcomes, both contemporaneously and over several quarters[42]. While there are valid critiques in the bank capital literature arguing that changes in bank capital are generally endogenous to changes in lending and other bank outcomes (Peydró (2010)), we argue that changes in banks' capital made to meet the new BOZ absolute capital limit were not so. This is because the change in regulation was unanticipated and most banks recapitalised during a reasonably stable macroeconomic and lending

[42] Our robustness tests also allow for unobserved common bank shocks.

environment, and we thus treat banks' policy-induced recapitalisations as an exogenous capital shock[43].

On examining whether the regime was implemented successfully, we find strong evidence that banks that were constrained by the regulatory rule grew their regulatory capital more quickly than when the rule did not constrain them, ceteris paribus. In other words, BOZ's enforcement of the regime was successful in making capital deficient banks raise capital, ceteris paribus. This result must be seen as contingent on Zambia's good or fair-weather conditions at the time, as banks may struggle to raise capital in other contexts.

One of BOZ's key objectives of the absolute minimum capital policy was to increase bank capitalisation rates (to strengthen the banking system) and reduce financial risk. Perhaps unsurprisingly, we found that banks' capitalisation rates increased as they raised capital levels. But we also found that much of this increase in capitalisation rates persisted, at least in the five or so quarters post recapitalisation. In other words, banks do not unwind the balance sheet effects of forced equity injections, at least in the short term. This is an important result because it shows that absolute minimum capital rules can, at least in this context for Zambia, improve bank capitalisation rates. It is also not a result that must necessarily occur – under an absolute minimum capital rule, banks are free to take on additional leverage and risk until they are constrained by the Basel-inspired RWAR rule.

While we find that banks grew their assets post the recapitalisations, we also find that much of this growth took the form of banks extending their holdings of Treasury securities, rather than growing their loan portfolios (as BOZ had perhaps hoped for). Our results hence suggest that forced recapitalisation may not be a direct tool for driving additional bank lending, at least in the short term. They also show that regulators need to be aware of how their bank capital policy might impact the domestic treasury market and the banking system's role in financing domestic government debt.

[43] In terms of our assumption of an exogenous shock to bank capital, our approach is most similar to Peek and Rosengreen (1997) and Mora and Logan (2010), although our source of exogenous capital variation is a regulatory requirement rather than using cross-country loan exposures. In terms of estimating the effects of a capital regulation change or a forced equity injection, our work sits closest to Fang et al. (2020), Brei, Gambacorta, and von Peter (2013), and Homar (2016).

Finally, our results indicate that imposing a large increase in absolute minimum capital requirements may have differential effects on banks, depending on their domestication or size, which may in turn have consequences on the structure of the banking market. We find evidence that smaller banks and domestic banks (note that many domestic banks are also smaller) grew their assets less quickly post recapitalisation. This result was especially evident in the case of small banks. Part of this is because small banks and domestic banks were less likely to expand their debt financing post recapitalisation. By construction, the recapitalisation policy required international banks to raise significantly more capital than domestic banks (USD 100 million for international vs USD 20 million for domestic), potentially shifting market share away from domestic banks. But these heterogenous reactions suggest that there could be additional effects, like moving the market towards larger, international banks. In a policy sense, the results indicate that we should be aware that banks of different types might react differently to forced capital injections, having implications for the future ownership and development path of the banking sector.

This paper makes several contributions to literature. Our main results provide one of the few quantitative evaluations of a large increase in minimum absolute capital requirements on lending in an LDFS setting. In addition, we highlight an example of a regulator successfully recapitalising its banking system in a way that did not clearly lead to significant bank deleveraging, which would have been a major risk had it relied on adjusting upward its Basel-inspired RWAR. In this context, our findings provide a piece of evidence on an under-researched capital policy tool that may be of interest to other LDFS policymakers facing a similar regulatory context to Zambia in 2012 (i.e., perceiving their banks as being too risky despite their banks being over the Basel-inspired RWAR rule). We also highlight some of the potential risks of implementing such a policy.

The paper also makes two methodological contributions. First, it is the first paper to our knowledge to use detailed sector-bank non-performing loan (NPL) data to model sector demand shocks, thus helping to address Peydró (2010)'s critique of bank capital regressions[44]. Second, it is one of the first papers to use foreign exchange and write-off adjusted loan data in a low-income country setting, a vital adjustment given Zambia's highly dollarised banking system.

[44] Constructing a bank-specific time-varying proxy for loan demand follows from Fang et al. (2020) and Aiyar et al. (2014), although our use of NPL data is novel.

Finally, our paper contributes to the bank capital literature more broadly. It provides a rare example and datapoint of banks being forced to significantly recapitalise with private capital during "fair-weather" conditions[45]. The majority of the bank capital literature finds a positive relationship between various measures of bank capital shocks and lending[46]. The fact that we do not find a similar relationship could be a result of the fact that the economics of bank capital are fundamentally different in a fair-weather context (in a crisis, equity injections may affect market sentiment and liquidity constraints and have broader economic spill-overs that might flow back to lending profitability); that Zambia's LDFS context creates more inertia in its banks' reactions; or that our dataset was simply not adequately powered to pick up any effect or too short-term.

We start the paper by examining the challenges of Basel-inspired capital rules in less developed financial systems.

Background

Regulatory convergence and the challenges of Basel in less developed financial systems

Since the introduction of the first Basel accord in 1988, an overwhelming majority of countries around the world have now adopted some iteration of the Basel regime[47]. The Basel Committee on Bank Standards (BCBS), composed now of 28 typically high-income jurisdictions, sets banking regulations for internationally active banks within their jurisdictions. Low-income countries (who are not mandated to follow these rules) have slowly been implementing some of the Basel rules as well, sometimes supported by recommendations from international organisations or a desire for their own regulatory regimes to be perceived globally as being sufficiently rigorous and aligning with "international best practice"[48].

[45] The size of the recapitalisation for Zambian banks was quite significant compared to many other examples. The median Zambian bank recapitalisation for domestic and foreign banks respectively were 11% and 22% of total bank assets, while the bailouts were 3.8% in the Homar (2016) sample and in 0.6% for Gambacorta (2013).

[46] For research finding such a positive equity-lending link, see Peek and Rosengren (1997) and Mora and Logan (2010) for evidence form on exogenous capital shocks; for bank bailouts, see Li (2013), Norden, Udell, and Wang (2020), Puddu and Wälchli (2015), Berger et al. (2019), Chu, Zang, and Zhao (2019); and finally, for bank capitalisation levels, see Maurin and Toivanen (2012) and Berrospide and Edge (2010).

[47] In BIS (2021), the Bank of International Settlements estimated that out of a sample of 100 non-BCBS jurisdictions, all 100 were implementing some iteration of the Basel risk-based capital regime.

[48] International financial institutions broadly supported the roll-out of Basel regulations and "helped create a culture of receptivity to international standards and 'best practices' within regulatory authorities" (Jones 2020, pg 356). In Zambia, the IMF typically supported implementation, for example in IMF (2017) pg. 41 – "Basel II and III should be implemented, tailored to Zambian circumstances, including introducing prudential liquidity, ratios in line with Basel III'.

There are good reasons for regulators in LDFS to align their local policies with the Basel accords. Implementing Basel can help to align and facilitate home-host bank supervision, signal to international investors a commitment to following a consistent policy path, and ensure regulatory competitiveness for attracting capital[49,50].

Yet, despite the upsides, there are also problems with such regulatory convergence. First, there is the well-known problem of *implementing* increasingly complex regulations in an already capacity constrained environment (capacity constrained for both banks and regulators)[51,52]. But there are also more fundamental problems related to how well the *regulations themselves are designed and calibrated for LDFSs*. This paper examines one example of where an LDFS regulator – the Bank of Zambia – deviates from and expands on the core Basel capital toolkit because the fundamental design and calibration of the accords did not adequately provide for its own context. Specifically, we examine the Bank of Zambia's use (and implementation) of a fair-weather absolute capital regime to support lending and recapitalisation, whereby banks were forced to recapitalise and inject equity, despite sitting well above Zambia's Basel-inspired capital rules of the time.

Before we examine Zambia's absolute capital policy specifically, we first examine the broader context behind why the Basel capital rules might not be as relevant or sufficient for LDFSs as for the BCBS members[53].

[49] Jones, Beck and Knaack (2018) and Jones and Zeitz (2019) discuss the drivers of regulatory convergence.

[50]"Proportional Implementation", the idea that national authorities can adapt regulatory requirements for non-internationally active banks operating in their jurisdictions can also, in theory, help mitigate the negative effects of the more recent and more complex Basel Accords. Magaldi de Sousa (2015) p.g 5 is unconvinced proportional implementation will necessarily work in LDFSs because supervisors may lack the capacity to appropriately apply it.

[51] The Basel rules have become increasingly complex over time, with the regulations progressing from 30 pages in 1988 for Basel I, to 347 pages for Basel II, and over 616 pages in 2010 for Basel III (Haldane, 2012).

[52]Research discussing how challenging the more recent Basel reforms may be for less developed financial systems to implement includes Jones (2020); Jones, Beck, and Knaack (2018); Kasekende, Bagyenda, and Brownbridge (2015); Fuchs, Losse-Mueller, and Witte (2013); Gottschalk (2016), and Gottschalk and Griffith-Jones, (2006). Ward (2002) is relevant for analyses of earlier versions of Basel. One example of this complexity is the potential use of internal model-based approaches for assessing risk and weighting assets, and how this might exacerbate information asymmetries between banks and supervisors (Jones, 2020).

[53] Note that there are examples of other non-capital Basel rules that are also questionable in an LDFSs setting. The Net Stable Funding Ratio (NSFR), which aims to reduce the likelihood that disruptions to a bank's regular sources of funding will erode its liquidity position by "requiring banks to maintain a stable funding profile in relation to the composition of their assets and off-balance sheet activities" pg1 (BIS, 2014), is one example. Thankfully, proportional implementation has typically been applied to the NSFR and few lower-income countries have moved towards implementing it (Hohl et al. (2018)).

The Problem: Basel Capital Ratios, Rule Convergence, and High-Capital High-Risk Banking Systems

The Basel Accords' capital rules aim to constrain banks from taking on too much risk by setting limits on ratios of their risky assets to their capital. The Accords' capital rules originally comprised of a risk-weighted asset ratio set at around 8% but, more recently, leverage ratios and other capital buffers have been added to the rule set[54]. As mentioned previously, only the members of the BCBS are mandated to implement the specific limits set by Basel, but in practice the Basel capital rules have become the norm and a benchmark worldwide, including for LDFSs[55].

What are the problems with LDFS regulators being strongly influenced by the capital Basel norms? We see two. First, there may be an issue of rule calibration. Because of higher baseline levels of risk in LDFSs (and their underlying economies), as compared to BCBS members, both banks and regulators have a preference for higher capital levels, meaning that Basel-inspired capital rule calibrations may be less relevant as a tool for managing financial system risk. To begin with, banks in LDFSs often *choose* to hold much higher levels of capital than those in BCBS countries because their underlying asset values are more volatile (Brownbridge et al., 2015).[56] For example, the average aggregate RWAR for African countries (many of which have LDFSs) was 18.6% in 2011, already well above the original Basel RWAR benchmark of 8%[57]. In addition, regulators are also likely to prefer higher levels of capital ratios than the baseline calibrations provided in Basel, and potentially higher still than what their banks currently hold. This is because regulators may know that: (i) the

[54] In 1998, the BCBS set the minimum risk-weighted asset ratio (RWAR) for its members at 8%. The RWAR ratio, which is calculated as the ratio of risk-weighted assets to regulatory capital, was and remains one of the core instruments for capital regulation in the Basel accords. "Risk-weighted assets" were originally calculated as the sum of a bank's assets, with each asset being weighted by its relative credit risk level (e.g., unsecured lending is weighted 100% and cash is weighted 0%). In each iteration since Basel I, there have been changes in the definitions of capital as well as how risk-weightings are calculated. Additional capital rules, such as the leverage ratio set at 3% (tier 1 capital/total assets) and the Tier 1 RWAR ratio, have also been added. Additional capital conservation buffers and counter-cyclical buffers were required as a part of Basel III.

[55] In 2011, African countries completing the World Bank's Bank Regulation and Supervision Survey recoded a mean minimum risk-based regulatory capital ratio of 9.7%, not far from the Basel minimum. The minimum ratio was 8% (37.5% of African countries), the maximum, 15% (6.25% of countries). In 2016, several African countries began to implement Basel II and Basel III and their additional risk requirements; the mean minimum risk-based regulatory capital ratio increased to 10.6%.

[56] Berger (1995 pp. 395) provides a useful distinction between "market capital requirements" and "regulatory capital requirements" that helps frame the discussion. A "market capital requirement" is the "capital ratio that maximizes the value of the bank in the absence of regulatory capital requirements" and which "the bank would tend towards without regulatory capital requirements". The regulatory capital requirement is typically the statutory limit. In LDFSs, banks have typically sat well above the "regulatory capital requirements" and have instead likely been driven by "market capital requirements". In LDFSs, the "market capital requirements" are likely high because of a lack of deposit insurance, information asymmetries between banks and bondholders, long and expensive bankruptcy processes, and higher capital buffers.

[57] Even if banks in LDFSs hold large capital buffers, we have not found strong evidence that their capital and risk decisions are driven by Basel-calibrated RWAR rules.

lack of a liquid equity market can restrict an appropriate market valuation of bank equity, and there could be a wedge between the true value of the bank's equity and the recorded book value (Brock and Suarez, 2000); (ii) their financial systems are riskier and less diversified than the BCBS jurisdictions on which Basel was calibrated; (iii) local government securities are zero-weighted and there is often no deposit insurance (so strong capital positions are essential for keeping depositors engaged in the financial system); and (iv) in many LDFSs, data frameworks supporting transparency, reporting, accounting, and audit are underdeveloped, so there can be variations in the quality of capital (Fuchs, Lossee-Muller, and Witte, 2013)). Taken together, the higher capital held by banks and the higher capital preferences of regulators means that rules based on the Basel calibrations may be less relevant as a tool for managing financial system risk in LDFSs[58]. The later iterations of Basel may have reduced the gap between Basel's risk calibration and the higher risk in LDFSs (for example, with capital add-ons to the baseline 8% RWAR in Basel III), but LDFSs have, due to reasons mentioned above, been slow to implement these new rules.

The second problem of Basel-benchmarked capital rules for LDFSs is one of design. The Basel toolkit may not always be enough to meet the additional regulatory objectives often pursued by regulators in LDFSs. In addition to the financial stability and economic growth trade-off that the Basel tools were developed to address, LDFS regulators often need to consider additional developmental goals, like credit market development (especially to higher-risk sectors like medium-small businesses), improving financial access and inclusion, attracting capital, and expanding the financial system's deposit base.[59] Having such multiple objectives can complicate regulators' decision-making, even for a simple rule like the Basel I risk-weighted capital asset ratio. Consider a regulator looking to improve bank capitalisation and financial stability while also aiming to improve

[58] The potential for policy miscalibration is all the more worrisome given the normative policy anchoring created by the Basel regime, whereby there are strong reputational and competitive forces that increase the stakes for low-income country regulators of deviating significantly from Basel, even if Basel calibrations were insufficient and such a deviation were appropriate.

[59] The centrality of Basel has at least in part crowded-out a first-principles analysis of capital regulation in LDCs, leaving such regulators with an incomplete toolkit for achieving their usually more complicated financial system objectives. While the Basel standards are designed to give regulators scope in adapting the regulations for local implementation (proportionality), as highlighted by Jones (2020), there is limited evidence that even the fundamental Basel Core Principles are even an appropriate starting point for designing and implementing banking regulation in LDFSs. Note, also, that for LDC policymakers there is limited empirical evidence even linking capital adequacy rules to financial crisis, or other banking market outcomes, using developing-country data. Chiuri et al. (2001) is one exception and provides some evidence for the trade-off existing in emerging market economies. Intuitively, we should be sceptical of how well the developed financial system literature and modelling can be applied to LDFSs. For starters, banks operating in LDFSs cannot easily and cheaply access capital markets. There is also limited securitisation post origination, and businesses also do not typically issue commercial paper, instead relying on bank lending or retained earnings. Moreover, fundamental market institutions and structures, such as a well-functioning credit reference bureau or a long and deep treasury market, do not always exist.

financial depth in the country. Increasing the regulatory RWAR may improve bank capitalisation rates, but it may also risk banks deleveraging to achieve the ratio (rather than raising new capital). While deleveraging is not necessarily a bad outcome for developed financial systems that are size-satiated, it could clearly work against an objective of improving financial depth in a LDFS setting.

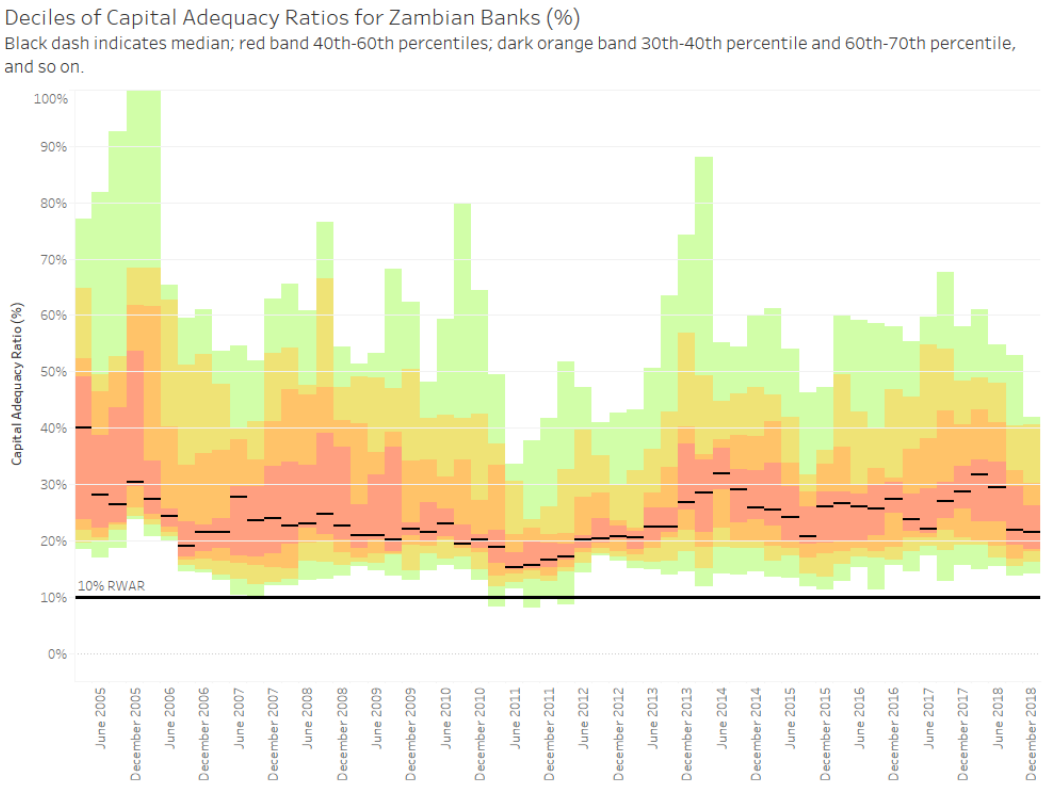
Context: Zambia, Recapitalisations, and expanding the Basel Toolkit.

At the beginning of 2012, BOZ found itself negotiating this very tension: wanting to increase bank capitalisation (in a banking system in which most banks already sat well above the Basel-inspired RWAR rule), while also supporting financial market growth.

In the wake of the 2008 financial crisis, several of Zambia's banks recorded large loan provisions and losses, especially in the construction, mining, and manufacturing sectors. Although the Zambian economy recovered quicky post crisis as commodity prices re-strengthened over 2009-2011, the banking system still faced major losses from 2008, with the aggregate non-performing loan (NPL) ratio increasing from 6% in early 2008 to nearly 12% in 2011. As losses from the crisis were slowly booked or written off by banks, capitalisation eroded from what had historically been very high rates: median (average) bank RWARs reduced from approximately 24% (29%) in June 2007 to 15% (20%) in June 2011. And herein lies the problem of tool calibration: with a median RWAR of 15% in June 2011, all but one bank remained well above BOZ's Basel-inspired RWAR rule of 10%, so there was little direct regulatory pressure to recapitalise. Clearly, the Basel-calibrated settings would not be enough to realign capitalisation levels in Zambia's banking system to what BOZ perceived to be an appropriate level of risk[60]. Figure 6 provides percentiles of banks' RWAR ratios in Zambia.

[60] The evidence that BOZ perceived capitalisation to be too low can be seen in one of BOZ's recorded objectives of the recapitalisation policy: to "*Make the banking system strong enough to withstand shocks and instability in the economy should they occur*" Gondwe (2013) p3. Also note that there was likely some urgency in the decision, as in late 2011 the global economy began to deteriorate once again. In August and September 2011 commodity prices – including the copper price, a key determinant of Zambia's export earnings – decreased significantly (approx. 30%). Although commodities did stabilise in 2012, BOZ must have been deeply concerned about the potential for further commodities deterioration. Finally, it is worth highlighting that Zambia does not have an explicit deposit guarantee scheme, which means that strong bank capital positions are essential in keeping savers and depositors engaged with the banking system and stemming off bank runs.

Figure 6 – Deciles of Zambian Banks' Risk Weighted Asset Ratios Over Time



Note: The horizontal black line provides the median capital adequacy ratio for Zambia's banks in a given quarter. The colour bands provide decile ranges for banks' capital adequacy ratios. The red band in the middle, for example, shows the 40th-60th percentile range; the dark orange, the 30th-40th and 60th-70th percentiles, etc. The long black horizontal line at 10% shows the BOZ's regulatory limit for a bank's capital asset ratio.

Source: BOZ Prudential Returns and Author's calculations.

What was stopping BOZ from simply increasing its RWA requirements up from 10%? As discussed previously, it would have been in tension with Zambia's other financial market objectives and may have run the risk that banks might deleverage to achieve the new RWAR rule, rather than by raising capital[61]. We also know from Gondwe (2013) that BOZ was concerned with whether bank financing was supporting investment and growth as it should; increasing RWAR was unlikely to allay this concern. There might also be an argument that BOZ was not keen to increase its RWAR too far from either the Basel norm or its regional comparator countries. Zambia's minimum RWAR already sat at 10%, 2% higher than the 8% Basel norm and approximately in line with the regional average. If BOZ increased the RWAR too far above its country comparators, Zambia's banking system might

[61] Myers and Majluf (1984) suggest that information asymmetries on prospective projects between firm managers and new investors may make equity raising less attractive to firms (and managers prefer not to rely on new equity finance). In Zambia, banks are either held privately or are capitalised directly from their parents abroad. There is no deep market of bank equity securities. As a result, this exact form of Myers and Majluf (1984) asymmetric information might be less problematic. Instead, the asymmetric information may exist more between managers (in country) and headquarters in another country, which may have a similar effect of reducing the attractiveness of recapitalising a bank by injecting equity.

have seemed like a less attractive place for international banks to place their bank capital (even if an absolute capital rule requires an equity injection, at least it wasn't telling banks to constrain their operations like a higher RWAR might). An increase in Zambia's RWAR rule too far above the Basel norm might also have reputational consequences for BOZ policymakers; as Jones (2020) argues, there can be reputation incentives for policymakers to converge on Basel standards more broadly.

All of these factors highlight that the Basel capital tools alone might not have been enough for Zambia and BOZ, explaining why the regulator looked to expand on the Basel toolkit.

Absolute and Fair-weather capital regulations as a policy solution

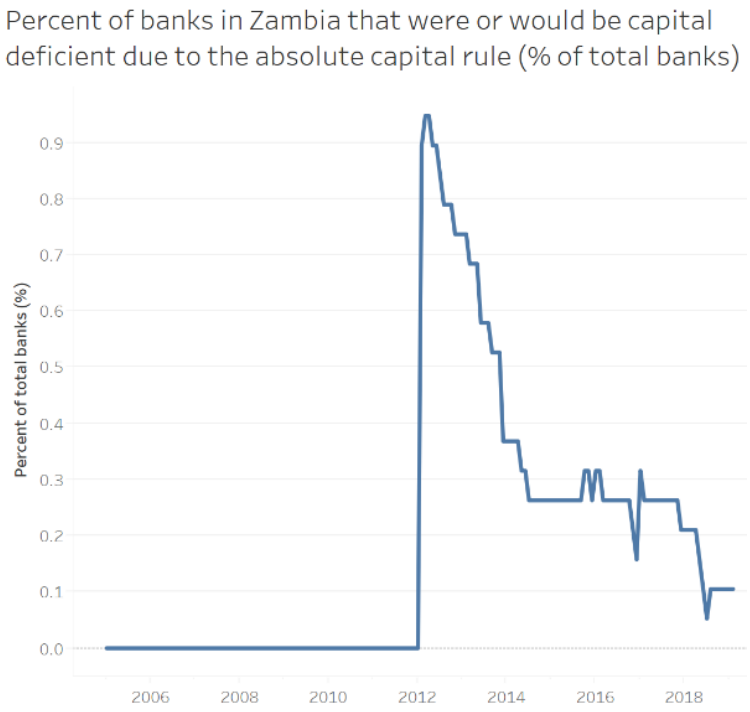
In January 2012, BOZ took action to address its perceived low bank capitalisation levels and market development concerns. It announced that banks in Zambia would need to meet a new absolute capital minimum by the end of the year, in addition to the Basel-inspired 10% RWAR. The size of the new absolute minimum would be determined by bank ownership type, with majority foreign owned banks needing to hold 520 million Kwacha (approx. USD 100 million) in regulatory capital and domestic banks needing to hold 104 million Kwacha (approx. USD 20 million). Assuming all banks managed to recapitalise to the new absolute limit, the policy would increase aggregate regulatory capital by USD 490 million (from USD 530 million to USD 1.02 billion), a 93% increase worth 2% of GDP[62]. The required recapitalisations were also significant at the individual bank level. At the time of the announcement, 18 out of 19 banks in Zambia did not have sufficient regulatory capital to meet the new limit (as shown in Figure 7), and many were deficient by a large margin. The median required growth rate in regulatory equity was 132% (or, as a proportion of total bank assets, 20 percentage points). This constituted a significant capital raising for most banks and for the system in general. While there are other cases of countries using absolute minimum capital requirements as a binding tool for bank capital policy (although such rules are sometimes framed as changes to banking license requirements), Zambia's policy required an especially large and quick recapitalisation[63].

[62] An important caveat on these numbers: the capital requirement is kwacha denominated, and so has reduced over time, especially given Zambia's high inflation/foreign-exchange crisis in late 2015 and then the fiscal-linked increase in inflation beginning in 2019. By January 2019, at the end of our data sample, the real value would have reduced by approximately 35-40%.

[63] The closest recent example of such a large recapitalisation is for Ghana in 2017, where the minimum capital requirement was increased from $25.9m to $86.4m. Most other examples of increases in minimum absolute capital requirements in the region are smaller; for example, in 2007 Kenya's minimum core capital requirement increased from approximately USD 4 million to USD 12 million by 2012 (Gudmundsson, et al. 2013), which is more typical for the region. See Ifeanyi Attama and Nfor Yuni (2021) for a summary of several short recapitalisation case studies in Africa.

An important point to note about the policy is that it applied in addition to (not instead of) Zambia's 10% RWAR rule. At any point in time, banks needed to satisfy both rules; whichever rule required the highest level of regulatory capital defined the bank's capital requirement. Additionally, banks were initially allowed one year to recapitalise, but at the end of 2012 this was then extended by a further year to December 2013[64]. As a result, the timing of banks' recapitalisations is potentially endogenous, something we will discuss in detail later.

Figure 7 – Banks Bound by the Capital Rule



Note: Y-axis in decimals and calculates share of banks that would/were not meeting the new absolute rule, i.e., 0.9 would mean 90% of banks in Zambia did not, at that point, hold sufficient regulatory capital to be capitalized above the new absolute limit.

Source: BOZ Prudential Data.

The official BOZ justification for the increases in absolute requirements was to improve financial stability and support financial market development. But implicit in the policy was a wager: while BOZ saw increased capital as a means to *"Make the banking system strong enough to withstand shocks and instability in the economy"* and *"Make banks contribute more to the growth of the economy; and improve access to loanable funds"* (both Gondwe (2013) pg. 3), there was clearly a

[64] Initially, BOZ gave banks one year to raise any capital shortfall. But in late 2012, this period was extended for another year at least partly due to significant lobbying by the banks, especially the smaller international banks facing the new ~USD 100 million limit. As reported in Lusaka, "*Market analysts have said the move has caused panic among banks in Zambia with a few international banks having written to the central bank indicating their difficulties in complying with the December deadline to increase their capital*" Lusaka Times (2012).

risk that banks might not see the Zambian market as sufficiently profitable and give up their licenses. Indeed, the IMF was concerned: *"there could be a withdrawal of some foreign banks that either serve small, niche markets or have no corporate interest in expanding/merging, possibly resulting in a decline in competition and the scope of financial services"* IMF (2012) pg. 19. There was also model risk; requiring banks to raise more capital would not necessarily result in them lending more, as we will explore further in the theoretical framework.

What likely made the policy less of a wager for BOZ was the "fair-weather" context of the recapitalisation. Typically, regulators would force banks to recapitalise and raise equity during financing and banking crises. But in the case of Zambia's fair-weather recapitalisation, the economy and banking system were doing well. Indeed, from 2008 to 2012 Zambia had grown steadily, and was projected to continue to do so (IMF 2012). In fair-weather conditions, the Zambian banking market was likely an attractive place for an investment. BOZ likely believed that international banks may have wanted to protect any option value they had for charter/franchise rents in the Zambian market in the future and so judged the bank-withdrawal risk as low. The high returns on domestic Zambian government debt were also likely a factor that made bank withdrawal less likely – even if banks were not interested in expanding into relationship-based lending (which was BOZ's objective), the portfolio value of having access to Zambian domestic debt may have been high. Finally, given that other countries in the region were not requiring such large bank recapitalisations at the time, it was more likely that banks' international parents might be able to finance the injections.

There were potentially some other strategies behind BOZ's choice to force equity recapitalisations that merit mention. One strategy may have been that the recapitalisation policy would attract additional capital from overseas (and not an insignificant sum at that). Domestic capital markets were extremely limited in Zambia, and this was potentially a strategy to tap outside sources of capital (or reduce bank dividends) at little cost to the country and without any capital crowd-out (assuming, of course, that banks did not leave)[65]. Politically, the capitalisations likely also made sense for BOZ. Given the long history of the financial sector not providing broad access for Zambians (the perception was that banks only wanted to serve large corporates), this was likely a policy that would have resonated more broadly with the public (although not the banks), as well as aligning with the

[65] Capital crowd-out, the idea that investments into bank capital would come at the cost of other domestic investment in Zambia would have been more of a risk if more capital was to be raised internally.

newly elected President, Michael Sata, and his goal of pushing for more available and lower costs of credit (Piskadlo, 2018).

There may also have been other target outcomes of the recapitalisation policy in addition to those BOZ announced officially. One such outcome could have been to increase the role of the banking system in financing Government domestic debt (or increasing its liquidity). As credit infrastructure in Zambia was weak (the credit bureau and collateral registry were not effective in 2012), it was perhaps not clear that banks could easily and profitably have expanded their lending post recapitalisation as BOZ's developmental objectives might have had in mind. In such a case, banks might simply park their additional capital in Government securities, providing additional financing for the government (and shifting a higher share of the banking system to government assets).

Whether or not it was part of BOZ's objectives, the capital policy may also have impacted the structure of the banking market itself. The recapitalisation policy might, for example, shift market share towards internationally owned companies, putting pressure on domestically owned banks. Because of the much larger capital requirement for international banks, the recapitalisation would naturally shift equity market share towards international banks. Whether and how different types of banks would then leverage the equity and expand into new asset types is an empirical question that we will need to examine. Skewing the market towards international banks could be a worrying consequence if BOZ was concerned about keeping banking system rents within country (especially if depositors perceived that the newly recapitalised foreign banks were now much safer). Perhaps BOZ felt that the additional quantum of foreign capital was worth the possibility of reducing the potential role that domestic banks played in the country's future banking markets, or that international banks or a smaller set of larger consolidated banks were actually better for long-term market growth in Zambia. Although Zambia's financial market context in 2012 was different, BOZ must have been aware of Nigeria's recapitalisation policy in 2004 and the consequences the policy had on the number of banks in that market[66].

The purpose of this paper is to evaluate the outcomes of a capital rule that falls outside the Basel toolkit, through the examination of Zambia's absolute capital regulation in 2012. Thus far, we have

[66] After the 2004 increase in minimum absolute capital requirements, the number of banks in Nigeria reduced from 89 to 25, with 14 failing to meet the new limit and losing their license, and all but four taking part in some type of merger (Cook, 2011))

discussed the context for the policy (the limitations of a Basel-based approach and how it pushed BOZ towards an absolute rule) as well as BOZ's objectives for the policy and other potential outcomes. Yet we still need to build some empirical or theoretical priors to help frame our evaluation, which we will do by examining the related literature on bank capital and building a simple theoretical framework to fit the Zambian context.

Related literature

This paper aims to investigate the potential impacts of Zambia's absolute capital regulation on banking market outcomes. Unfortunately, absolute bank capital limits or forced market recapitalisations during fair-weather conditions have not seen much attention in the banking literature to date, despite the fact that forced recapitalisations are not uncommon in LDFSs, especially in Sub-Saharan Africa. The limited existing literature typically has an aggregate approach and examines bank profitability or bank competition, rather than examining a broader set of bank outcomes or reactions[67]. This paper hence adds to a thin literature quantitatively examining banking capital regulations in LDFSs, and the limitations of the Basel accords in such a setting.

Given the limited directly relevant literature, the purpose of this literature section is twofold: (i) to highlight other areas of the banking capital literature that sit close to Zambia's recapitalisation experience, even if the contexts are not directly comparable; and (ii) to help us build some hypotheses for what the impacts of the absolute capital policy might be. While in the previous sections we outlined the BOZ's and the IMF's beliefs about the potential impacts of the policy, we do not know what our priors should be based on relevant empirical and theoretical research on bank capital.

The first relevant literature concentrates on historical government bank bailouts, and how such equity injections impact bank outcomes, especially bank lending. This literature is relevant to this paper because, despite the clear difference in contexts between a bailout and a fair-weather recapitalisation, there is a mechanical similarity between Zambia's capital regime and a forced bailout. In a mechanical balance sheet sense, bank bailouts (or forced injections of capital into banks by the government) have similarities with Zambia's regime, in which banks were being forced to recapitalize themselves under a form of an absolute capital rule. In both cases, banks undergo a

[67] Cook (2011) and Sandow et al. (2021) are two examples for Nigeria and Ghana respectively, but both concentrate on bank performance, efficiency, or competition, rather than tracing bank reactions.

positive absolute capital shock, which may impact their lending and capital structure decisions. The cause and context of the change in capital are clearly different between the bailout and Zambia's regime: one derives from a change in regulatory limit under "fair-weather" conditions, the other is a forced injection, usually during an economic crisis. But there is a strong link.

The bailout literature typically examines bailout episodes stemming from Japan's housing crisis in the 1990s or programs responding to the Global Financial Crisis in the US and in Europe to investigate how bank bailouts supported bank lending growth and bank risk taking. Looking across countries and bailout episodes, there is evidence that bailout equity injections supported growth in bank lending or the easing of lending conditions (Li (2013); Norden, Udell, and Wang (2020); Puddu and Wälchli (2015); Berger et al. (2019); Chu, Zang, and Zhao (2019)). But the evidence remains mixed, with some research only finding significant lending effects when the bailout was large (Gianetti and Simonov (2012); Homar (2016)), or when it reaches some certain thresholds (Brei, Gambacorta, and von Peter (2013)), and sometimes not at all (Nakashima (2016); Duchin and Sosyura (2014); Bassett and Demiralp (2014)). This bailout literature is relevant because it shows that it is plausible that, economic context obviously aside, after large equity injections (like Zambia's forced recapitalisations) banks may look to extend or grow their lending books or assets. If we were to treat the contextual differences as being second order to the role of injections of capital[69], such studies may support BOZ's hypothesis that the capital injections will support additional lending.

Another body of relevant literature, which is conceptually similar to the bailout research, explores the effects of non-bailout but arguably exogenous shocks to bank capital on bank lending. This research also finds a generally positive relationship between bank capital shocks and bank lending. The approach is to use bank exposures across international jurisdictions (or unexpected shocks) to isolate loan supply and other balance sheet responses from loan demand factors, allowing for a plausibly clean identification[70]. Peek and Rosengren (1997) examine the lending impacts of capital

[69] This is a big caveat. The market effects of large bailout programmes, for example in easing liquidity constraints in interbank funding, or supporting banks to restart securitisation, might also create mechanisms supporting positive changes in bank lending. On the other hand, during a crisis loan demand is likely to be low and the valuation of loan collateral may also have been reduced, creating a context of weak loan growth. These factors unlikely existed in Zambia's recapitalisation context (Zambia's capital regime was introduced during a period of good economic growth, and although banks had lower capital levels than in the past, they were still well above international comparators).

[70] The challenge for this capital shock research, and research on bank capital in general, is that shocks to bank capital will often be correlated with conditions in the underlying economy, loan demand, and the financial performance of borrowers. This then creates an identification problem in that the capital shocks (impacting loan supply) may covary with firm/household shocks (impacting loan demand) and disentangling the loan supply effect from the loan demand effect may

shocks to Japanese parents of subsidiaries in the US; Mora and Logan (2010) examine UK bank lending after a bank's overseas subsidiaries suffers losses; and Rice and Rose (2016) use US community bank exposures to the Fannie Mae and Freddie Mac receiverships during the Financial Crisis. All find a positive relationship between capital shocks (the shocks are typically negative) and bank lending. As with the bailout literature (but using a different and often negative source of capital variation), this research again provides tentative evidence that there could be a positive relationship between Zambia's recapitalisations and bank lending, assuming that bank responses are symmetric regarding capital gains and capital losses.

Other studies examine how the level of bank capitalisation influences bank lending, but without seeking clean exogenous identification. Bernanke and Lowe (1991) examine the 1990s recession in the USA and find evidence that poorly capitalised banks expanded lending less rapidly than those that were better capitalised (using variation both across US states and across banks in New Jersey). Berrospide and Edge (2010) find similar results using a longer panel of banks[71]. This is an important piece of context for Zambia's recapitalisation policy because, if we believe that Zambia's banks were indeed undercapitalised in early 2012, there might be reason to believe that the recapitalisation would lead to improved lending[72].

Maurin and Toivanen (2012) and Berrospide and Edge (2010) evaluate how banks' deviations from an estimate of their internal target capital ratio might impact lending (i.e., whether banks hold a capital ratio surplus or deficit compared to an estimate of their target ratio[73]). Both find positive but modest effects of bank capital ratios on loan growth, but other factors, such as loan demand, appear to drive the majority of bank lending growth. In relation to Zambia's regulatory regime, such a policy would have constituted a major negative shock to banks' target capital ratio (hence a positive shock to the deviation), likely implying a small positive lending effect.

be difficult. Clearly, this would also likely be a problem for the bank recapitalisation literature if the reason for the recapitalisation is not independent of loan demand (or that selection for which banks were to be recapitalised were correlated with bank-specific loan demand).

[71] Berrospide and Edge (2010) find that a "1-percentage-point increase in the capital ratio leads to a long-run increase in annualised BHC loan growth that is only between 0.7 and 1.2 per- centage points." pg 31. Bernanke and Lown (1991) find that a 1-percentage-point increase in the equity capital ratio resulted in a 2-percentage-point increase in loan growth.

[72] Zambia's growth at the time of the recapitalisation was strong and the level of capitalisation in Zambia was unlikely as weak as that in New Jersey – so the context difference between the US and New Jersey vs. Zambia is significant and questions the externally validity in this case.

[73] A bank's optimal capital ratio is estimated as a function of the regulatory ratio requirement.

The final relevant empirical literature examines how regulatory changes in relative capital ratio rules impact bank lending and other banking system outcomes. In the wake of an increase in required regulatory ratios, there is evidence that banks reduced loan volumes (Aiyar, Calomiris and Wieladek (2014); Fang et al. (2020)), but also that banks may tend to adjust the composition of their assets (substituting into lower risk-weighted assets) rather than reducing volumes (Francis and Osbourne (2012)), and that there might be heterogenous responses in lending types (Bridges et al. (2014)) or by bank characteristics (Fang et al. (2018)). These papers have typically used micro-prudential regulatory variation for identification because of the lack of policy variation in capital ratio rules both within and across countries[74]. Regulatory changes in relative capital ratio rules – such as changes to banks required RWARs – are relevant for Zambia's absolute recapitalisation experience because of the direct mechanical links between the two capital regimes in the specific case when a bank raises capital as a means to improving its relative capital ratio (indeed they are mechanically the same in this case, as opposed to when a bank improves its risk-weighted capital ratio by deleveraging[75]).

Taken together, this research provides some support for BOZ's implicit hypothesis that increased bank capital levels might support improved bank lending, albeit heavily caveated given the clear differences in financial system and recapitalisation contexts[76]. The lack of literature on the impacts of increased minimum capital requirements and forced recapitalisations during fair-weather conditions also clearly highlights that this paper should be able to contribute to the empirical bank capital literature.

In terms of theory, it is common in the bank-capital / bank-bailout literature to use the irrelevancy propositions set out in Miller and Modigliani (1958) (henceforth, referred to as MM) as a motivating theoretical baseline. Indeed, much of the current debate concerning optimal bank capital ratio policy in developed countries concerns financial frictions, discussions of optimal debt, and whether or to

[74] This is because the BCBS set the RWA ratio at 8% and the rule has broadly stayed there; most countries and banks implement similar versions of the rule. But, both the Bank of England and the European Banking Authority have imposed some form of bank-specific capital ratios which have been used for research. The more recent changes in regulations due to Basel III and Basel IV will likely create more variation in the future.

[75] Granted, improving a capitalisation ratio by increasing equity is an endogenous choice for banks constrained by a newly changed relative capital rule, and many banks in these studies do recapitalise by deleveraging (which is not an option for banks under an absolute capital rule).

[76] Keep in mind significant differences in economic context may bring into question the external validity of these studies for a Zambian context, for example, the lack of a traded bank equity market, the limited scope of the lending and deposit markets, etc.

what extent MM irrelevance will hold for banks (for a small sample of a much larger literature and debate on optimal bank capital regulations, see Aiyar, Calomiris, Wieladek (2014); Miles et al. (2012); Admati, DeMarzo, Hellwig, Pfleiderer (2013); Admati (2016); Thakor (2014); Admati (2015); Berger (1995)[77]).

In Zambia, MM is likely far less relevant as markets are incomplete and asset returns are potentially subject to size/scale limitations[78]; most bank equity and non-deposit debt in Zambia's banking system is not traded. It is questionable, as a result, to assume that an MM interpretation is the best framework to work from, or that we can easily model a large equity injection in fair-weather conditions as simply pushing banks off their static optimal leverage ratio (defined here and throughout the paper as regulatory capital / total assets), from which they can then revert (a more plausible suggestion for a small bank operating in a large, deep, and developed financial market, but not in Zambia). It is unlikely that Zambia's banks, having raised equity, could all drastically increase their callable deposit financing or other debt financing.

To help frame our empirical work for the characteristics of the Zambian market and to allow us to rely less on either MM as a theoretical benchmark or on empirical estimates that may not be externally valid for Zambia's context, we build a simple partial equilibrium framework to guide our evaluation in the next section.

Theoretical framework: Bank lending and deposit taking under an exogenous capital shock

Below we outline a simple framework to help us think through how an unexpected and forced equity injection might impact bank decision-making in Zambia. The goal is to frame Zambia's equity

[77] Firms (and banks) almost certainly do have an optimal financing mix because of debt tax shields (Miller and Modigliani (1963) and Miller (1977)), costs of financial distress (Bradley et al. 1984), pecking order financing and information asymmetries (Myers (1984) and Myers and Majluf (1984)), to name but a few reasons. Specifically for banks, other factors may pull them away from capital structure indifference, for example, deposit insurance and implicit safety nets (Berger 1995). If we believed that Zambian banks had optimal capital ratios and that such ratios were independent of the absolute level of capital held by the bank, then a forced recapitalisation would not impact the calculation of that optimal ratio, and we might expect that BOZ's policy would result in banks increasing their debt financing and asset holdings after a fair-weather recapitalisation. As discussed later, we are not convinced this is a likely scenario.

[78] A strict application of MM for Zambia's banking market – a questionable approach given that MM's assumptions do not fit well with the Zambian context – implies that a forced recapitalisation would not lead banks to make further balance sheet adjustments after a forced recapitalisation; bank assets would increase 1:1 with the forced equity investment, but bank debt would remain unchanged. Under MM, the higher (and more expensive) equity share of the bank's financing should be offset by a reduction in its cost of equity and debt. The bank's beta is assumed constant (so that it would receive a fair return for any additional equity-financed investments) and the forced investment would not change its weighted average cost of capital (because reductions in debt costs would offset any higher equity costs), hence no further adjustment.

capital injection in a simple context that fits several of the stylised facts of banking in Zambia and the characteristics of the absolute equity rule (and the subsequent recapitalisations). Specifically, it tries to reflect the limited and imperfect loan and deposit markets in Zambia, an equity shock that was unexpected and arguably exogenous, the high frictions involved for non-institutional depositors to access the treasury market, and limited non-bank options for firm finance.

The framework sits especially close to Peek and Rosengren (1995), but also Courakis (1984), the simple model in Khwaja and Mian (2008), and Freixas and Rochet (2008)'s summary model of Klein (1971) and Monti (1972). Compared to these other models, our framework emphasises (crudely) the potential impacts of an equity shock to a bank's balance sheet[80]. The framework is static and there is no interbank market or government asset. We do not model banks' intra asset allocations – for example the decision to lend to firms or the government in the form of Treasury securities – and we ignore information asymmetries or imperfections as well as issues relating to bank asset liability management.

Agent Types and Markets

There are three agent types in this framework: banks (i), firms (j), and depositors. The baseline model assumes that banks are monopoly lenders in their own specialised market segment and do not compete on price. Banks are assumed to operate in an imperfectly competitive deposit market (i.e., the quantities of deposits can be affected by the interest rate offered by the bank). We assume that there is only one bank-firm relationship and that both banks and firms can witness a firm's productivity level before they negotiate a loan rate. In Annex 3 and Table 4 (at the end of the framework discussion), we examine some different market structure assumptions (for example, that banks lend in a competitive loan market).

Firms' Optimisation Problem

Each bank (i) lends to a single firm (j). Each firm is assumed to have a return on investment, $\bar{r}_{ij}$, and there are no alternatives to bank financing for funding investments. The firm is trying to maximise

[80] The framework is not trying to model how and why banks and shareholders endogenously raise capital and how they make long-term financing and investment decisions. Instead, it is trying to isolate the short-term tactical reactions of bank managers to an exogenous requirement to hold more capital. Along this line, the framework assumes that: (i) bank capital is predetermined; and (ii) that the recapitalisation can be modelled as an unexpected exogenous capital shock. The model is clearly unrealistic, but given that Zambia's absolute capital rule was unanticipated, was not in response to a crisis or bank failures, and that new equity is very expensive to issue in Zambia, the exogenous equity assumption is perhaps more innocuous than it seems.

its profit, π_{ij}, and requires borrowing from bank i in the form of loans, L_{ij}. We assume that firms face diminishing returns to scale in their investments, hence revenue per unit of borrowing reduces by $-\alpha.{L_{ij}}^2$, whereby α is a parameter that determines the scale of decreasing returns to firm investment (this assumption is similar to Khwaja and Mian (2008)).

Firms maximise their profits in the period according to:

$$\max_{L_{ij}} \quad \pi_{ij} = (\bar{r}_{ij}).L_{ij} - r_{L_{ij}}.L_{ij} - \alpha.{L_{ij}}^2 \tag{1}$$

The first order conditions for firms are:

$$\frac{d\pi_{ij}}{dL_{ij}} = \bar{r}_{ij} \; - \; r_{L_{ij}} - \alpha.2L_{ij} = 0$$

Each firm has an inverse demand for loans such that:

$$r_{L_{ij}} = \bar{r}_{ij} - \alpha.2L_{ij}$$

And loan demand curve:

$$L_{ij} = \frac{1}{\alpha.2}.\bar{r}_{ij} - \frac{1}{\alpha.2}.r_{L_{ij}}$$

In this setup, firms are assumed to make zero abnormal profit after financing costs because of monopoly loan supply. No arbitrage is allowed across loan markets. Given the underdeveloped equity markets in Zambia, a heavy loan requirement for firm financing is not a terrible assumption. Future versions of the framework might expand on borrowers' equity options as an alternative to bank financing.

Banks' Optimisation Problem

Banks are assumed to have a monopoly in their specific loan market. The interest rate they offer to their firm is based on each firm's observable productivity, $\bar{r}_{ij}$. The bank's margin is hence a function of the firm's return on investment. There is no credit rationing or other informational imperfections.

Banks are assumed to maximise:

$$\max_{L_{ij},D_i} \quad \pi_{ij} = L_{ij}\left(r_{L_{ij}}(L_{ij})\right) - (\phi_{i1} - \phi_{i2}\bar{r}_{ij}) - r_{D_i}D_i - C(D_i, L_{ij}) \tag{2}$$

π_i is bank i's profit in the period. Loan losses associated with lending volumes are dealt with crudely: ϕ_{i1} is a constant that captures the (constant) portion of any loan that banks would expect to result in loan losses if they held a diversified loan portfolio. ϕ_{i2} is a simple constant that reduces the size of this expected actuarial loss given the observed productivity of the firm (i.e., more productive firms result in lower loan losses per unit of lending for the bank). The crude approach is necessary because: (i) we are not trying to directly model loan risk and portfolio diversification; and (ii) in future econometric analysis we want firm productivity to influence the profitability of bank lending in addition to the direct effect via banks receiving a higher interest rate from their loans to more productive firms. We do not consider corner cases and assume that $\phi_{i1} - \phi_{i2}\bar{r}_{ij} > 0$, i.e., an additional dollar of lending from the bank to each firm will always be associated with a positive actuarial loan loss. r_{D_i} is the rate the bank pays to receive a volume of deposits, D_i, from the deposit market. Because we are not modelling bank risk, $r_{D_i}$is not a function of the bank's leverage (clearly a simplifying assumption). $C(D_i, L_{ij})$ represents the cost of making loans and collecting deposits, which we assume is linear and separable, i.e., $C(D_i, L_{ij}) = \tau(D_i + L_{ij})$.

Banks cannot force firms to accept a given interest rate, but they do observe each firm's inverse demand function:

$$r_{L_{ij}} = \bar{r}_{ij} - \alpha.2L_{ij}$$

We assume the bank operates in an imperfect deposit market characterised by an increasing linear supply of deposits, consistent with Peek and Rosengren (1995).

$$D_i = d_0 + d_1\left(r_{D_i} - \overline{r_D}\right) \tag{3}$$

Banks can attract deposits by offering above the mean market rate, $\overline{r_D}$. The idea is that if a bank offers a deposit rate greater than $\overline{r_D}$, it can attract elastic depositors from other banks (both inside and outside the banking region), but also from large financial institutional investors (such as the National Pension Scheme Authority (NAPSA)) for which commercial banks deposits are close substitutes to other short term liquid assets, such as Treasury securities. Another framing might be to think of $\overline{r_D}$ as large depositors' opportunity cost of bank deposits. d_0 is best thought of as an inelastic demand for liquid deposits and transactional deposits, or deposit customers who would face very high costs of changing their deposits to another bank. This is a clearly a simplified setup given

that the market rate for deposits may clearly change in the case of large equity injections across multiple banks and that there is no modelling of consumers or consumers' savings decisions.

Solving for r_{D_i} from the deposit market demand function, we can then substitute the rearranged deposit demand equation into the bank's profit maximisation problem:

$$r_{D_i} = \frac{D_i - d_0}{d_1} + \overline{r_D}$$

$$\underset{L_{ij},D_i}{max}\pi_{ij} = L_{ij}(r_{L_{ij}}(L_{ij}) - (\phi_{i1} - \phi_{i2}\bar{r}_{ij})) - \left[\frac{D_i - d_0}{d_1} + \overline{r_D}\right].D_i - \tau(D_i + L_{ij}).$$

We then use the balance sheet accounting identity to substitute out D_i from the bank profit function:

$$D_{it} = L_{ijt} - E_{it}$$

And now banks solve:

$$\underset{L_{ij},D_i}{max}\pi_{ij} = L_{ij}(r_{L_{ij}} - (\phi_{i1} - \phi_{i2}\bar{r}_{ij})) - \left[\frac{L_{ijt} - \overline{E_\iota} - d_0}{d_1} + \overline{r_D}\right].(L_{ij} - E_i) - \tau(D_i + L_{ij})$$

$\overline{E_\iota}$ is the beginning of period equity and is assumed exogenous.

The bank makes a loan to the firm, assuming the firm wishes to borrow. The bank can adjust the size of each loan (or the associated interest rate, taking into account each firm's demand curve). The bank's FOC will be:

$$\frac{\partial \pi_i}{\partial L_{ij}} = \bar{r}_{ij} - 4\alpha.L_{ij} - \phi_{i1} + \phi_{i2}\bar{r}_{ij} + \frac{-2L_{ij0} + 2\overline{E_\iota} + d_0}{d_1} - \overline{r_D} - 2\tau = 0$$

Solving for L_{ij} we find the equilibrium loan market outcome:

$$L_{ij0} * = \frac{\bar{r}_{ij} - \phi_{i1} + \phi_{i2}\bar{r}_{ij} + \frac{2\overline{E_\iota} + d_0}{d_1} - \overline{r_{D0}} - 2\tau}{2\left(2\alpha + \frac{1}{d_1}\right)} \quad (4)$$

Given an exogenous shock to bank equity, equilibrium lending should change such that:

$$\frac{\partial L_{ij}}{\partial E_i} = \frac{\frac{2}{d_1}}{2\left(2\alpha + \frac{1}{d_1}\right)} = \frac{1}{\alpha d_1 + 1} > 0 \quad (5)$$

Which is greater than 0 but less than 1 for most reasonable values of αd_1.

Solving now for the bank's deposit market outcome, where we again substitute in the balance sheet identity, but this time for L_{ij}:

$$\max_{L_{ij},D_i} \pi_{ij} = L_{ij}(r_{L_{ij}} - (\phi_{i1} - \phi_{i2}\bar{r}_{ij})) - \left[\frac{D_i - d_0}{d_1} + \overline{r_D}\right].D_i - \tau(D_i + L_{ij})$$

The bank's FOC will be:

$$\frac{d\pi_i}{dD_{ij}} = \bar{r}_{ij} - 4\alpha D_i - \alpha 2E_i - \phi_{i1} + \phi_{i2}\bar{r}_{ij} + \frac{-2D_i + d_0}{d_1} - \overline{r_D} - 2\tau = 0$$

Solving for deposit market equilibrium:

$$D_i * = \frac{\bar{r}_{ij} - \alpha E_i - \phi_{i1} + \phi_{i2}\bar{r}_{ij} + \frac{d_0}{d_1} - \overline{r_D} - 2\tau}{2\left(2\alpha + \frac{1}{d_1}\right)} \tag{6}$$

And for an exogenous equity shock, the change in equilibrium deposits:

$$\frac{\partial D_{ij}}{\partial E_i} = \frac{-\alpha}{2\left(2\alpha + \frac{1}{d_1}\right)} = \frac{-\alpha d_1}{2(\alpha d_1 + 1)} < 0 \tag{7}$$

After the equity shock, it appears that the bank partially offsets the increase in the size of its balance sheet. This makes intuitive sense because the equity shock pushes the bank's marginal cost curve further down its downward-sloping marginal revenue curve.

To help clarify the intuitions that loans will increase (but less than 1:1 with the equity increase) and that deposits will decrease, Figure 8 provides a visual aid showing a demand and supply representation of the two markets, the balance sheet constraint, and an exogenous equity shock[81]. Figure 8 shows that, given an exogenous equity shock, ΔE_i, the bank will respond by increasing its lending by $L_{ij}^{**} - L_{ij}^{*}$ and reducing its deposits by $D_{ij}^{**} - D_{ij}^{*}$. How much banks increase lending and reduce deposit-taking for a given equity shock is a function of the slope of the marginal revenue of loans curve (which is itself a function of the competition in the loan market and the marginal productivity of firm investment, captured by α) and the marginal cost of deposits (i.e., d_1),

[81] The marginal cost and marginal revenue curves for the loan and deposit market in Figure 8 are derived from differentiating the bank revenue and cost functions with respect to bank lending and bank deposit-taking.

respectively. The intuition is that, after banks are forced to inject equity, the marginal cost of lending reduces and the marginal revenue that can be gained by raising an additional dollar of deposits will simultaneously shift downwards. The marginal cost of lending shifts downward because a larger portion of the bank's lending is now funded by "free" equity. The marginal revenue of deposit-taking shifts downwards because each additional loan taken by the firm is less productive. Because our simple framework does not model risk or, say, allow for the marginal cost of deposits to be a function of equity (a MM-type intuition), the "free" increase in equity financing (from the bank managers perspective) partly crowds out the bank's need for deposit financing (at least when the marginal cost of deposit financing is upward sloping). The bank's assets increase less than the quantity of the equity injection because loans are diminishing in return. Deposits are reduced because they are expensive. Taken together, this means that the bank's capitalisation rate (equity to total assets) will also increase or improve.

Figure 8 – Simple framework

Simple banking market framework with monopoly loan pricing, reducing returns to borrowing, an incomplete deposit market, and an exogenous equity shock

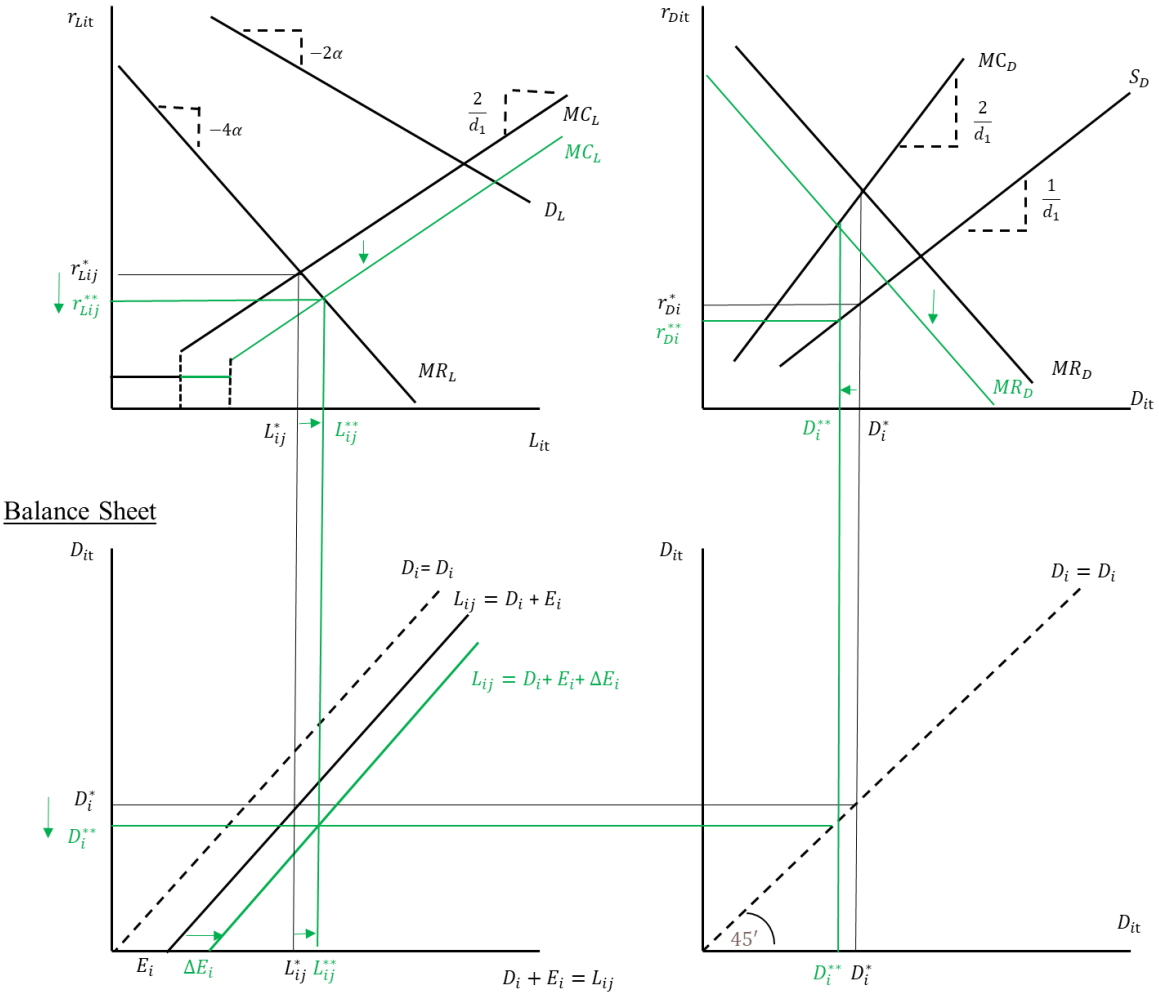
Note: The graphic shows an exogenous positive equity shock for the bank, and how it impacts loan and deposit outcomes. The shock expands equity on the balance sheet (see column 1, row 2, in the figure below, and the green rightwards shift in the balance sheet identity line, equal to ΔE_i). Black curves indicate pre-shock curves or curves that are not shifted by the shock. The equity shock impacts both the loan market and the deposit market, as described below. A shock means that the bank manager can access a larger stock of "free" capital, but each additional loan that it makes brings in less revenue. r_{Lit} and r_{Dit} are the interest rates on loans and deposits, respectively. L_{ij}, D_i, E_i represent loans, deposits, and equity. α and d_1 are parameters that determine the slope of the loan marginal revenue curve and the slope of the deposit cost curve. These parameters determine the final change in loans and deposits associated with an equity shock (via the comparative statics terms in the text) and are defined in the text. MC_L and MR_L, and MC_D and MR_D are the marginal cost and benefit of lending and taking deposits, respectively. A * represents the initial equilibrium and ** the post capital injection equilibrium.

Loan Market

The equity increase reduces the amount of costly deposits needed for the same level of loans, shifting down the marginal cost of loans curve. But each new loan is yielding less revenue, so the pass-though of new equity to new loans is less than 1:1.

Deposit Market

Each additional deposit is associated with a reduction in marginal revenue (because in the loan market the MR of loans is downward-sloping). The bank hence partly substitutes out of expensive deposit financing in the presence of an equity shock.



The balance sheet is directly connected to the loan market because banks can only invest in loans. The green lines show the shift associated with the equity shock, with loans increasing less than the equity shock (and deposit financing reducing).

This is a 45-degree line connecting the deposit market to the balance sheet (and then through to the loan market).

BOZ may have been concerned if the equity injections meant that banks would have been less willing to accept deposits, as predicted in Figure 8. Such a reduction would have clearly worked against their objectives of improving financial access and inclusion.

In practice, BOZ may have been less worried about reducing deposits because banks do have access to low-risk assets that do not exhibit downward-sloping returns (in addition to making loans, as in this simple model) and so are less likely to stop accepting deposits. These assets might be Treasury securities or term deposits at other institutions (noting that the interbank market is quite illiquid). In the simple framing above, we might imagine that the bank might choose to invest in a Treasury security if it was forced to increase its equity and it had reached the point where the marginal return from lending fell below the Treasury bond yield (for much of the period deposit rates were lower than Treasury returns). If banks want to strategically hold market share in the deposit market or there are frictions in reducing deposits, we might expect these alternative assets to matter even more.

The framework above also makes certain market structure assumptions that may more-or-less hold in practice (for example, monopoly pricing in lending and decreasing returns to firm borrowing). In Annex 3, we re-examine some of these assumptions to help us understand other potential outcomes within this framework, specifically, when the bank loan market is competitive, when firms do not face a reducing return to investment, and, as an extreme example, when banks wish to follow a constant leverage ratio (which we assume is exogenously determined). Table 4, below, shows the comparative statics for loan and deposit equilibrium under each of these assumptions[82].

[82] For example, if we assume that firm investment is not diminishing in volume, banks will increase lending 1:1 with the equity investment but will leave deposit financing unchanged. The bank makes a profit on each additional dollar of lending associated with the new "free" equity, so it lends more, but the slope of the deposit cost curve and the loan market curve have not changed and so the volume of lending financed by deposits will also not change. In the case that banks wish to hold a specific leverage ratio, banks will take on additional financing and acquire additional loans to ensure the ratio holds. This is clearly an unrealistic scenario given the size limitations on both asset returns and financing costs in Zambia, but it does provide a corner case and is more relevant in a world where banks are facing binding regulatory RWARs. If we assume that the loan market is competitive, then we expect banks to still increase lending and reduce deposits.

Table 4 – Baseline Framework Extensions

Market Structure	Adjustment - Bank Balance Sheet Component	
	Assets	Deposits
<u>Baseline – Monopoly Loan Market</u> • Downward-sloping loan demand • Monopoly loan pricing • Imperfectly competitive deposit market	Positive (+) $\frac{\partial L_{ij}}{\partial E_i} = \frac{1}{\alpha d_1 + 1} > 0$ Likely less than 1 for reasonable values of α and d_1	Negative (-) $\frac{\partial D_{ij}}{\partial E_i} = \frac{-\alpha d_1}{4\alpha d_1 + 2} < 0$
<u>Alternative (i) – Competitive Loan Market</u> • Downward-sloping loan demand • Competitive loan pricing • Imperfectly competitive deposit market	Positive (+) $\frac{\partial L_{ij}}{\partial E_i} = \frac{1}{\alpha d_1 + 1} > 0$ Likely less than 1 for reasonable values of α and d_1	Negative (-) $\frac{\partial D_{ij}}{\partial E_i} = \frac{1}{\alpha d_1 + 1} - 1 < 0$
<u>Alternative (ii) – Constant returns to loan investment</u> • Perfectly elastic loan demand • Monopoly loan pricing • Imperfectly competitive deposit market	Positive (+) $\frac{\partial L_{ij}}{\partial E_i} = 1$ 1:1 with the equity injection	No Change (0) $\frac{\partial D_{ij}}{\partial E_i} = 0$
<u>Alternative (iii) – Constant Leverage Ratio</u> • Constant leverage ratio constraint, c • Downward-sloping loan demand • Monopoly loan pricing • Imperfectly competitive deposit market	Positive (+) $\frac{\partial L_{ij}}{\partial E_i} = c$ Could be a large multiple increase	Positive (+) $\frac{\partial D_{ij}}{\partial E_i} = c - 1$ Could be a large multiple increase

Note: This table provides loan and deposit outcomes due to an exogenous equity shock using the simple theoretical framework from the text, but under different market structure assumptions (these are provided in each row and defined in the leftmost column). These alternative market structures are described in some detail in Annex 3. Notation is defined in the text.

Taken as a whole, the simple framework (including the alternative scenarios and the comparative statics) suggests that, unless banks do have an exogenously determined and static optimal capital ratio, at least in the short-term, forcing an equity injection onto banks may be less likely to result in large loan increases than BOZ may have hoped. If the Zambian loan market was extremely profitable, or the equity injection resulted in positive system-wide externalities, loan growth might be more likely. Furthermore, given that in practice banks have access to low-risk loan alternatives,

which we know offered high real returns at the time, it is perhaps more likely that the equity financing would end up in low-risk Treasury assets, and that banks wouldn't reduce deposits to limit expansions in their size[83]. Richer models – for example where equity has system-wide externalities or reduces bank financing costs – may make banks more likely to extend lending post equity injection, which we also accept as a possibility.

Tracing the effects of Zambia's absolute capital regime

Now that we've set up the relevant literature and framework, we can start examining empirical data and testing our (and BOZ's) hypotheses. But first, we need to pin down what specific questions we will try to answer with the data, and why they are important and for whom.

It is unlikely that the conditions facing the country in 2012 (a banking system that was capitalised above the Basel RWARs yet was not capitalised enough to allay the regulator's financial stability concerns) were or are unique to Zambia. These background tensions underpinning Zambia's absolute capital requirement and the limits of Basel's capital regulations in high-equity, high-risk financial systems are familiar across Sub-Saharan Africa. Banks do typically hold higher levels of capital (Beck and Cull, 2013) and regulators, may, for the same reasons mentioned earlier, have a higher preference for bank capital. We also know that other countries in the region have and do implement similar rules (although not often as large, and potentially for other reasons too).

For policymakers, however, there is limited guidance on the potential effects of implementing these alternative policies. The empirical and theoretical bank capital literatures we discussed clearly do have some role to play in helping policymakers think through their bank capital policy, but there is little research that is either directly relevant or contextually appropriate. This paper aims to provide one datapoint on the short-term effects that absolute capital requirements might have on banks, which could be a useful tool for bank regulators in similar contexts.

The structure of the empirical analysis hence follows from what we consider would be the most relevant and useful to policymakers, while also being consistent with testing BOZ's objectives (improving capitalisation rates and bank lending) and contributing to the literature on bank capital. The approach examines several partial and short-term outcomes of the capitalisation policy,

[83] From 2012 onwards, Zambia's Treasury Bill rates were some of the highest real yields in the world, sitting easily in the top decile of countries submitting data to the IMF for the International Financial Statistics.

accepting that in many cases these outcomes are interdependent, and we can only capture transitory effects[84].

We will evaluate the recapitalisation policy along the following four dimensions:

1. Policy Implementation. Did banks actually recapitalise in response to BOZ's rule? If so, within what timeframe and in what way? How bound were banks by the new rule? Was the implementation of the absolute capital rule associated with banks growing their capital bases more quickly than previously, ceteris paribus? For policymakers interested in Zambia's experience, they would want to know whether banks were actually able to raise the requested capital (contingent, of course, on Zambia's economic context at the time and the costs of non-compliance)[85].

2. Capitalisation Rate. As banks raised capital in response to the higher absolute capital rule, did bank capitalisation rates improve, and, if so, did these improved bank capitalisation rates persist? As banks were not constrained by the absolute rule to hold a certain capitalisation rate, they could have, in theory at least, increased their deposits or borrowing in line with their recapitalisations (to keep their relative capitalisation rates unchanged). Our hypothesis is that, given the large size of the recapitalisations, the limited size of Zambia's market, and frictions and imperfections in debt financing markets, the recapitalisation policy would have strengthened capital ratios, at least in the short term. For policymakers looking to improve bank capitalisation rates with an absolute capital rule but who also would like not to significantly increase RWAR rules above the Basel norms, this is clearly a relevant outcome to analyse.

3. Bank Asset Reactions. Is there evidence that banks extend their assets and lending after being forced to recapitalise, at least in the short term? Is there evidence that banks expanded into one type of asset rather than another? We will test BOZ's hypothesis that a forced recapitalisation under fair-weather conditions will improve bank lending (within a six-quarter horizon). In terms of the Basel framework, we will be testing whether, in this context, it is possible to recapitalise banks without risking banks retrenching their lending (a risk regulators would face if they were to simply increase banks required RWARs). We will also evaluate whether, assuming banks grow their assets bases,

[84] A stronger capitalisation rate, for example, is quite likely going to be associated with one (or many) of reduced debt financing, reduced lending or risky assets, or increased capital.
[85] Given the fair-weather context of the new capital rule and the costs of non-compliance, our hypothesis is that most banks did raise capital.

they expand into lending or liquid short-term securities (for example, Treasury securities, which may also have important consequences for the banking system's role in financing domestic government debt). How banks adjust their asset portfolios post recapitalisation is an important consideration for policymakers and will help us understand the limits of this recapitalisation policy as a tool for directly supporting financial market development (or at least not constraining it).

4. Unintended Consequences, heterogeneous bank reactions, and distributional effects. We know that smaller and less internationally connected banks with less deposit and lending infrastructure may have found it harder to access debt financing (and, indeed, increase lending) post recapitalisation. Is there evidence that this resulted in domestic or smaller banks growing less than larger or international banks? Is there evidence of smaller banks merging or dropping out of the market?[86] Regulators in LDFSs will have goals for how they want their banking markets to develop. Having a better understanding of how forced recapitalisations may impact different kinds of banks is clearly relevant (for example, if regulators have long term objectives of retaining domestic ownership of the banking system).

Data and Empirical Discussion

Data Sample

To investigate the impact of Zambia's capital regime, we collected quarterly financial and regulatory data from 2005 to 2019 for each of the 19 commercial banks operating in Zambia[87]. Over the sample period, six banks entered the market, one failed, and two banks merged. We removed the bank that failed because it had been financially distressed since well before the capital regime change. To adjust for the merger, created a weighted average of the two banks pre-merger[88]. For newly registered banks, we remove the first two quarters of data from when the bank entered the market. This is because financial data for new banks can be quite abnormal during their initial capitalisation period and as they begin to expand their lending capabilities, neither of which is representative of a bank operating in business-as-usual conditions. We provide results using both unadjusted and winsorised data at 5% to ensure the results are not being driven by outliers, although the direction and

[86] This paper stops short of undertaking an aggregate analysis of changes in market power or market competition post recapitalisation, or aggregate cross-country comparisons, but this would be another avenue of research.
[87] The quarterly dataset is constructed from raw monthly prudential returns collected by BOZ and consolidated by the author. We do not use monthly data because of balance sheet recognition and timing issues.
[88]The merger was in 2016, far enough after both banks' recapitalisation experiences.

significance of results are not very sensitive to this. Winsorising likely helps us better sample our desired bank population for growth-rate regressions because we want our results to primarily reflect banks operating in normal conditions, and not be driven by volatility in banks' balance sheets that is due to the fact they are either new to the market or are going through financial distress.

Empirical Approach and Discussion

To trace the effects of Zambia's absolute capital regime we use fixed effects panel data models, relying primarily on within-bank variation and argue for the plausible exogeneity of the forced bank recapitalisations. We use a slightly different methodology, specification, and identification approach for each of the four dimensions over which we wish to evaluate the policy, each of which is explained in their respective results section. The aim in each approach, however, is to exploit the arguably exogenous imposition of the absolute capital rule (and the exogenous within-bank variation in capital growth that it created).

The Zambian context provides an especially strong setting for evaluating the implementation and effects of an absolute capital regime, as the policy was neither expected by the banks or other market participants, nor was it introduced in response to any particular contemporaneous economic conditions (the losses that pushed banks' capitalisation rates lower than BOZ's risk threshold and inspired the policy were a result of economic conditions from two or three years prior). The challenge in much of the empirical bank capital literature is that changes in bank capital are inherently endogenous to bank asset/loan demand and the underlying economic context and collateral valuations, meaning that any estimate linking changes in bank capital to changes in banking market outcomes will be biased[89]. The endogeneity in this case is driven by similar factors impacting both the demand for bank loans and debt financing and the evolution of bank equity. If, however, we are prepared to examine only changes in capital linked to recapitalisations under the rule and accept that the imposition of the rule was unexpected and broadly exogenous to these factors, such endogeneity should be limited[90].

[89] This is the basis of Peydró (2010)'s critique of Berrospide and Edge (2010).
[90] This assumption or approach is most similar to Peek and Rosengreen (1997) and Mora and Logan (2010), although, in our case, the exogenous capital changes are limited to an unexpected regulatory change, rather than by exploiting cross-border banking relationships.

Also, that there is not strong evidence that banks "selected" into being recapitalised based on their characteristics provides another reason why this Zambian context provides a strong setting for analysis. A legitimate concern with analysing the effects of this policy is that banks with certain levels of unmeasured time-varying characteristics – riskier or weaker banks, for example – might have a higher capital deficiency under the policy, thus producing biased estimates of their responses[91]. For Zambia's capital regime, the capital rule was applied across the market, with the only selection criteria being a bank's absolute capital level. So, while there may be significant differences between large banks and small banks in Zambia that will correlate with the required size of bank recapitalisation, banks neither directly self-selected which group they would be in, nor does bank size necessarily correlate with bank health. For example, there is no significant linear relationship between banks' riskiness/distress, as measured by banks' RWARs, and a banks' capital shortfalls immediately after the recapitalisation was announced (the coefficient on banks RWAR in a simple bivariate cross-sectional regression explaining the capital deficit as the policy was announced as a function of the RWAR is not statistically significant). Further, even if there was a form of selection bias for Zambian banks, it is not clear that the direction of bias would necessarily be in support of our null hypotheses; if weaker banks were over-represented in the recapitalisation shocks, it is possible that asset/lending growth is actually weaker for them relative to other banks (because of bank infrastructure constraints on expansion and debt financing access, for example).

Additionally, detailed and commercially sensitive prudential data for Zambia's banks has been constructed (for both before and after the recapitalisations), allowing us better control for bank covariates as well as creating better demand controls[92]. First, the better data allows our identifying assumptions to be made conditional on several bank-specific time-varying characteristics, making what will clearly be strong identifying assumptions more defendable. Second, it allows us to create new bank-level demand control variables that are time-varying and sector-specific, similar to the approach used by Aiyar et al. (2014) and Fang et al. (2020). This new variable allows us to better control for the fact that banks in our dataset can have quite different loan sector concentrations, and each sector can potentially be exposed to its own shocks. For low-income countries, such detailed

[91] This is often a problem in the bank bailout literature as banks that need bailouts are often those in severe financial distress, and accounting for the bank bailout selection mechanisms is a challenge for much of this literature.
[92] The constructed dataset contains the usual balance sheet, income statement, and regulatory capital calculations, as well as detailed sector-level NPL data, and foreign exchange asset and liability denominations.

controls are rare and our approach of using bank-sector non-performing loans (NPL) accounts is new. Annex 4 provides a detailed description of how we created this variable, its limitations, the novelty in using NPL data to do so, and the methodological contribution to the literature[93]. The detailed dataset allows us to make important adjustments to loan and deposit data to account for currency exposures and revaluations as well as loss recognitions and write-offs. Such adjustments can be vital for highly dollarised financial sectors with less diversified banks (especially when the exchange rate is volatile) but such data is seldom available for research to make such adjustments in LDFSs and this is one of the first papers to do so. Annex 4 shows how these modifications were made and derives a simple foreign exchange adjustment given Zambia's data and shows our approach to making such adjustments (which sit most closely with Bridges et al. (2014)).

One weakness of our empirical setup is that the timing of banks recapitalisations was an endogenous bank choice: within the two-year implementation timeframe, banks chose when to inject the capital. The problem is that banks could have held off recapitalising in an attempt to either finance some of the capital shortfall with retained earnings (which are a function of profit) or wait for better market conditions to raise capital. Both of these decisions may be structurally related to the bank outcomes we are measuring, and, if true, could bias our estimates[94]. Though we are concerned about this possible endogeneity, several factors do lessen the likeness that it will have unacceptably biased out results. First is that the economic conditions over the two years were reasonably stable, minimising the chances that large swings in banking sector conditions drove bank recapitalisations. It is also worth keeping in mind that 37% of the recapitalisations also occurred within the original timeframe (1-year), and that the recapitalisations were very large (indeed a large multiple of a normal year's annual retained earnings). This further limits the likeliness that our results are fully biased, as does the fact that we attempt to directly control for demand conditions at the bank-sector level. So, while we want to be clear that this timing endogeneity is a problem, we hope that it is only second-order.

As with many papers in the bank capital literature, we will attempt to account for banks' unobservable time invariant characteristics by including bank-fixed effects in our regressions. For

[93] Unfortunately, bank-sector demand controls still will not help control for loan/borrower-specific shocks, the key reason why many more recent papers look to use borrower-fixed effects and corporate bank-borrower-matched loan level data with multiple lender relationships, as introduced by Khwaja and Mian (2008).

[94] For example, take the case of our lending growth outcome. A bank recapitalising with retained earnings will likely have accumulated such profits during a time that banking conditions and loan demand were strong. But this means our "exogenous" capitalisation variable and our loan supply outcome are both likely correlated with loan demand conditions.

cross-bank time-varying shocks (like commodity price shocks), we include a set of macroeconomic controls and their lags and use time-fixed effects and two-way fixed effects as a robustness check[95]. Time-series cross-sectional data often also contain both contemporaneous correlation across units and unit level heteroskedasticity (Bailey and Katz (2011)), and our Zambian bank panel is likely no exception. To account for such non-spherical error structures in our regressions, we will follow Homar (2016) and Bahaj and Malherbe (2017) from the banking capital literature and use Driscoll and Kraay (1998) standard errors that allow for arbitrary error correlations across banks and time. Specifically, we use Hoechle (2007)'s procedure for implementing Driscoll and Kraay (1998) that allows for unbalanced panels. In our robustness tests, we will also use standard errors clustered at the bank level (allowing for autocorrelation within banks), as with Fang et al. (2020), as well as Panel-Corrected Standard Errors (PCSE) as recommended in Beck and Katz (1995) and Beck (2001), and supported by Monte Carlo analysis in Moundigbaye, Rea, and Reed (2017) for a dataset of our dimensions.

Results

Policy Implementation

We begin by examining the implementation of the capital rule. Specifically, we examine how many banks were constrained by BOZ's new absolute capital rule, whether they actually recapitalised in response to the rule, how large any recapitalisations were, and within what timeframe banks recapitalised.

The first main finding is that ninety-five per cent of banks – all but one – were required to raise at least some capital because of the absolute capital rule. The fact that the regulation bound such a high proportion of banks was shown previously in Figure 7 (which shows the per cent of banks that held regulatory capital under Zambia's absolute capital rule when it was announced). Most banks did

[95] Many papers in the bank lending panel literature employ both bank and time fixed effects in their regressions (relevant papers in the literature include Brei et al. (2013), Bridges et al. (2014), Fang et al. (2020), among others). The idea is that time fixed effects should absorb common time effects across banks, absorbing any economy-wide firm productivity shocks or common cross-bank shocks. There are two issues: one is that it is not clear how one can actually interpret estimates from TWFE estimations (Kropono and Kubinec, 2020); the other is that, in a staggered, continuous, dynamic treatment setting with no clear never-treated group, such as these recapitalisations, we may be inviting bias (see Goodman-Bacon (2021), Sun and Abraham (2021), Callaway and Saint'Anna (2020), and others). For now, our approach is to include TWFE regressions in the sensitivity analysis, but to let the main baseline recapitalisation results be driven by only within-bank variation, attempting to control for unobserved common trends with relevant macroeconomic controls. Typically, our results are not sensitive to including time fixed effects and using TWFE.

recapitalise, and by July 2014, half a year before the extended deadline, 75% of banks had managed to raise capital to hold regulatory capital above their respective limits (ZMW 504 million for foreign banks and ZMW 104 million for domestic banks), as per Figure 47 in Annex 11. The next five years saw the remaining banks slowly make progress or drop out of the market [96,97].

Not only did most banks have to recapitalise, but the size of the recapitalisations was large too (both for the banking system in general, and for banks individually). If we compare Zambia's aggregate regulatory tier 1 capital growth to a sample of its regional peers (Figure 9), there is a clear jump in growth in 2012 and 2013 resulting from BOZ implementing the new absolute limit – for both 2012 and 2013, growth in aggregate real tier 1 capital sat at above 50%[98]. Bank capital in LDFSs can be volatile and noisy (as risks are typically less diversified and provision recognition practices can be less accurate and the quality of capital less certain (Kasekende et al, 2015 and Caprio and Honohan, 1999)). But even with this noise, we can see the consequences of Zambia's new regulation.

Figure 10 is a stacked bar chart and shows the total kwacha value and timing of aggregate recapitalisations undertaken by banks that were regulatory deficient (the columns are a stack of individual bank Kwacha recapitalisations). The figure shows that majority of the value of the recapitalisations were completed before the December 31, 2013, deadline, although there was clearly also a tail.

Capital decomposition analysis (unpublished) shows that much of the capital growth in Zambia took the form of new share equity and non-disbursed earnings (and coupled with this, reduced dividend rates). There was, however, large transfers across capital accounts for banks that had high levels of retained earnings and needed to convert this into paid-up share equity.

[96]BOZ likely did not wish to withdraw a bank's license if there was a reasonable chance that the bank would reach the rule.
[97] Much of the capital growth in Zambia took the form of new share equity and non-disbursed earnings (and coupled with this, reduced dividend rates). There was, however, large transfers across capital accounts for banks that had high levels of retained earnings and who needed to convert this into paid-up share equity.
[98] Total regulatory capital, the directly relevant metric, was not available in the IMF International Financial Statistics. Tier 1 capital captures only core equity assets of a bank and consists primarily of common stock and disclosed reserves (or retained earnings) but may also include non-redeemable non-cumulative preferred stock. It misses some forms of regulatory capital, for example revaluation reserves or hybrid capital instruments.

Figure 9 – Regional Comparisons of Real Capital Growth

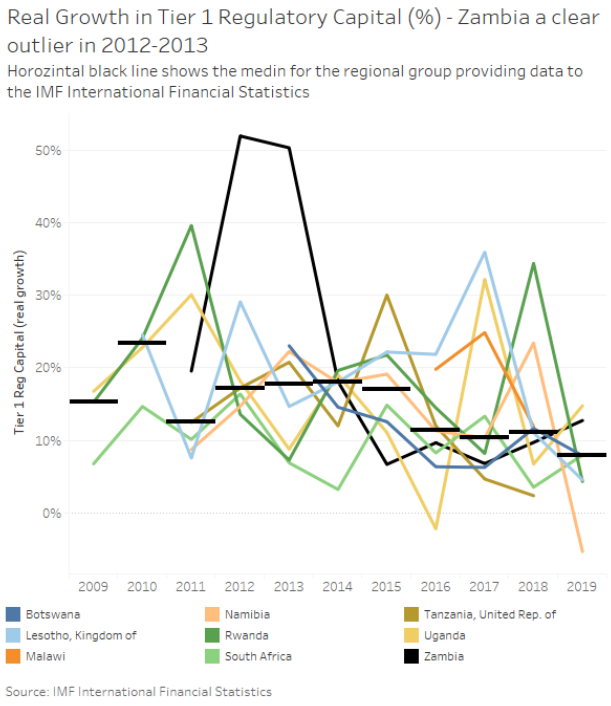
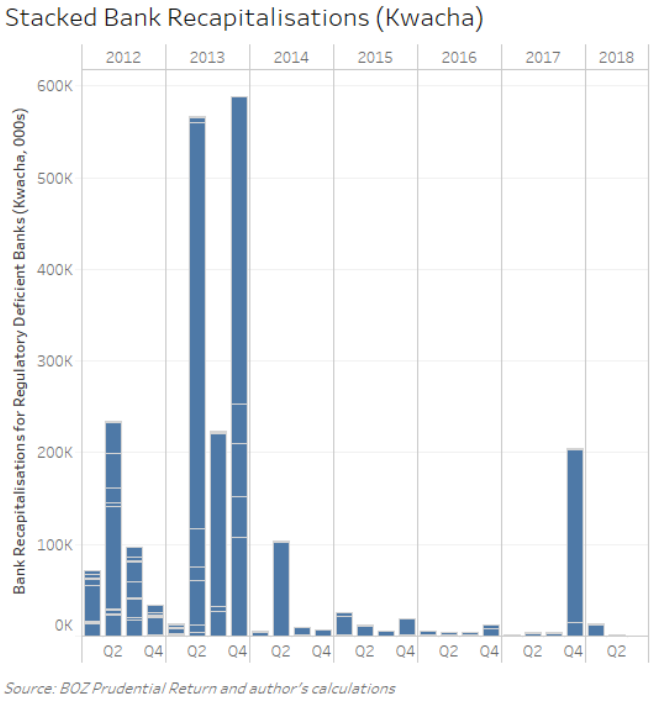


Figure 10 – Total Kwacha Value and Timing of Bank Recapitalisations per quarter (Stacked, Kwacha)



At the bank level, there is strong evidence that banks were required to recapitalize, and sometimes significantly. After the announcement of the new rule, the upper quartile of banks by absolute regulatory capital were holding approximately 70% of the capital that they would need to be compliant with the new rule; the lower quartile, only 20%. This shows that most banks had to undergo considerable recapitalisations.

We see this in the data when we break down the forced recapitalisations as a percentage of total bank assets (Figure 11)[99]. Denominating the equity shock as a percent of total assets helps us think about the recapitalisation in terms of a bank's size, normalising across banks sizes and their required recapitalisations. After some banks raised equity in the second quarter of 2012 (typically the larger,

[99] Specifically, we calculate bank recapitalisations as:

$$Recapitalisation_t^i = \frac{\Delta def_t^i}{TA_t^i}$$

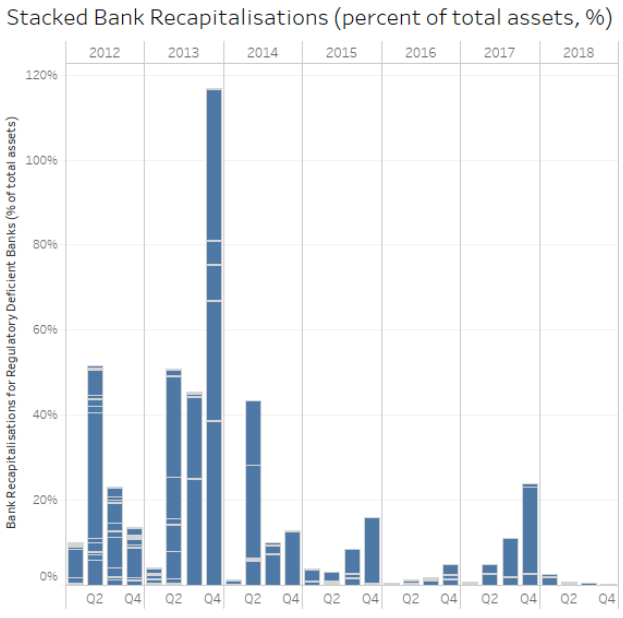
$$def_t^i = \begin{cases} C_t^i - C_t^{i,Reg}, & C_t^i - C_t^{i,Reg} < 0 \\ 0 \quad , & C_t^i - C_t^{i,Reg} \geq 0 \end{cases}$$

Where $\Delta \mathrm{def}_{t-k}^{i} / \mathrm{TA}_{t-k}^{i}$ measures the growth in regulatory equity a bank underwent in a period (as a percent of total assets) contingent on the equity growth being associated with the bank being in regulatory deficit. $\mathrm{C}_{t}^{i,\mathrm{Reg}}$ is the regulatory capital requirement at time t, C_{t}^{i} is the bank's equity at time t, TA_{t}^{i} is the bank's total assets.

more diversified banks), many of the remainder recapitalised at the end of 2013, or in mid-2014. Figure 12 provides a histogram of bank recapitalisations for capital deficient banks, measured as a percent of total bank assets as per the above definition. Note that some banks recapitalised with infrequent but large transactions, while others took a more gradual approach. This assists in explaining the high number of small recapitalisations (which are driven primarily by banks accumulating retained earnings while they were capital deficient and often waiting until they sourced capital for a large transaction to meet the target).

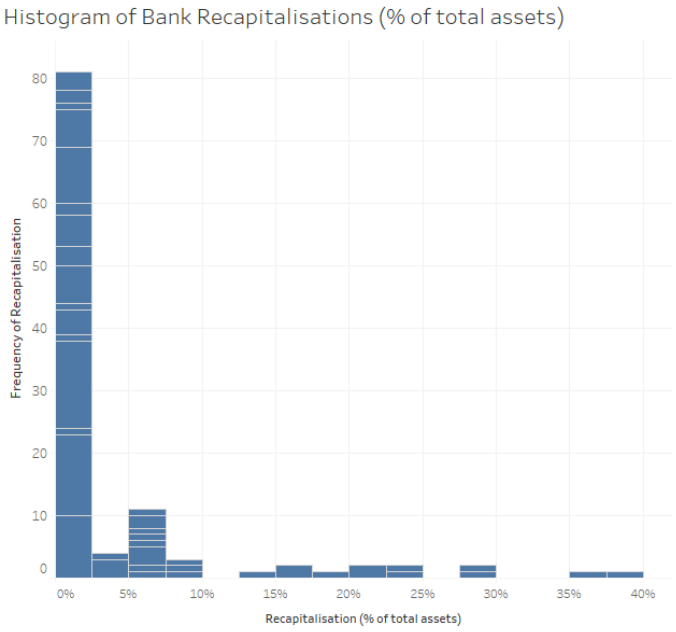
Table 5 shows the cumulative total recapitalisations by size and by bank count, essentially summing over the quarterly data by bank. Table 6 provides the summary statistics on the recapitalisations. What is clear is that some banks had to undertake significant recapitalisations.

Figure 11 – Total Value and Timing of Bank Recapitalisations per quarter (Stacked, % of Total Assets)



Source: BOZ Prudential Return and author's calculations

Figure 12 – Histogram of Bank Recapitalisations per quarter (measured as a % of Total Bank Assets)



Source: BOZ Prudential Return and Author's calcualtions

Table 5 – Total Cumulative Bank Recapitalisations by size (% of total assets)

Total Cumulative Bank Recapitalisations by size (% of total assets)

Cumulative Recapitalisation	0-10%	10%-20%	20%-30%	30%-40%	40%-50%	50%-60%	60%-70%	70%-80%
Number of Banks	7	1	2	1	3	1	0	2

Note: The table provides Total Cumulative Bank Recapitalisations by size bin (% of total assets). For example, a bank that recapitalised twice in two different quarters, by 20% of total assets each time, would record a cumulative recapitalization of 40% in this table. Recapitalisation calculated as $Recapitalisation_t^i = \frac{\Delta def_t^i}{TA_t^i}$ and $def_t^i = \begin{cases} C_t^i - C_t^{i,Reg}, & C_t^i - C_t^{i,Reg} < 0 \\ 0 \quad , & C_t^i - C_t^{i,Reg} \geq 0 \end{cases}$ where $\Delta \text{def}_{t-k}^i / \text{TA}_{t-k}^i$ measures the growth in regulatory equity a bank underwent in a period (as a percent of total assets) contingent on the equity growth being associated with the bank being in regulatory deficit. $C_t^{i,Reg}$ is the regulatory capital requirement at time t, C_t^i is the bank's equity at time t, and TA_t^i is the bank's total assets.

Source: BOZ Prudential Returns.

Table 6 – Descriptive Statistics for Bank Recapitalisations (% of total assets)

	N	Mean	S.D.	Min	25%	Median	75%	Max
Bank Recapitalisations (% of total assets)	111	0.042	0.076	0.0004	0.005	0.011	0.028	0.388

Note: Table shows the summary statistics for banks recapitalisation as a percentage of total assets, aggregated by quarter. For example, a bank that recapitalised twice to reach the new capital rule, in two different quarters, by 20% of its total assets each time, would have two entries of 20% in this table. Recapitalisation calculated as $Recapitalisation_t^i = \frac{\Delta def_t^i}{TA_t^i}$ and $def_t^i = \begin{cases} C_t^i - C_t^{i,Reg}, & C_t^i - C_t^{i,Reg} < 0 \\ 0 \quad , & C_t^i - C_t^{i,Reg} \geq 0 \end{cases}$ where $\Delta \text{def}_{t-k}^i / \text{TA}_{t-k}^i$ measures the growth in regulatory equity a bank underwent in a period (as a percent of total assets) contingent on the equity growth being associated with the bank being in regulatory deficit. $C_t^{i,Reg}$ is the regulatory capital requirement at time t, C_t^i is the bank's equity at time t, and TA_t^i is the bank's total assets.

Source: BOZ Prudential Returns

What these data and figures do not tell us with clarity is whether the bank recapitalisations were any more or less than what we would have expected to see had the rule not come into place, ceteris paribus. Banks naturally accumulate capital, and oftentimes more so when economic conditions are favourable. It is possible that banks might have raised capital at the same rate in lieu of the policy, given the background conditions in the banking market and the economy (in which case the recapitalisation absolute regulation regime may actually have had little effect on bank behaviour). To investigate whether capital growth rates were higher, ceteris paribus, we estimate a simple bank panel regression, taking into account other macroeconomic and bank specific drivers of capital growth that are common in the literature:

$$\Delta capital_t^{ij} = \alpha^{ij} + \beta_1 . def_ind_t^i + \sum_{p=1}^{P} \boldsymbol{\beta}_P{}' . \boldsymbol{X}_{t-p}^i + \sum_{L=1}^{L} \boldsymbol{\beta}_L{}' . \boldsymbol{M}_{t-p} + \varepsilon_{it} \quad (8)$$

$$def_ind_t^i = \begin{cases} 1 & , C_t^i - C_t^{i,Reg} < 0 \\ 0 & , C_t^i - C_t^{i,Reg} \geq 0 \end{cases}$$

The approach is crude because it ignores the dynamics of banks' recapitalisations, simply attempting to pick up average differences in capital growth recorded by banks that were constrained during a quarter, compared to when they were not constrained. We define $\mathrm{def_ind}_\mathrm{t}^\mathrm{i}$ as an indicator dummy variable which equals one if the bank was constrained by the capital rule in a given time period and zero otherwise (def stands for "in capital deficit" and equals one if the bank's regulatory capital level, $\mathrm{C}_\mathrm{t}^\mathrm{i}$, sits below its regulatory threshold, $\mathrm{C}_\mathrm{t}^\mathrm{i,Reg}$). $\Delta\mathrm{capital}_\mathrm{t}^\mathrm{ij}$ is the growth rate in bank capital for bank i and capital type j from quarter t-1 to t. α^{ij} captures unobserved time-invariant bank characteristics that can be correlated with $def_ind_\mathrm{t}^\mathrm{i}$. In this case, it might capture the fact that some banks naturally raise more capital (or retain more earnings) than others. We control for bank-specific characteristics by including $\mathbf{X}_{\mathrm{t,t-p}}^\mathrm{i}$, a vector of bank-specific control variables that are broadly consistent with the bank capital lending literature. This vector includes z_{it}, a bank-specific measure of loan demand, discussed in Annex 4, $logA_{t-p}^i$, the log of a bank's total assets, ROA_{t-p}^i, the ratio of the bank's net income to its total assets, LIQ_{t-p}^i, the ratio of a banks liquid securities to its total assets, RWA_{t-p}^i the bank's risk-weighted asset ratio, and $Prov_{t-p}^i$, provisions made by the bank in the quarter, as a percentage of its total loans. We control for broad macroeconomic variables by including $\boldsymbol{M}_{t-p}$, a vector of macroeconomic variables. The set includes the inflation rate, π_{t-p}, the Treasury Bill rate, $TBill_{t-p}$, Reserve Requirements, RR_{t-p}, and the growth in the exchange rate, fx_{t-p}. Dynamics are captured by the lag structures, P, L. In most specifications we use four-to-six lags for the recapitalisation variable, and then two lags for the macroeconomic and bank-specific variables, but the results are not sensitive to adjusting covariate dynamics. Our choice of lag structure is consistent with the literature (for example, Francis and Osbourne (2012)), and our results are robust to using alternatives. Flow variables are accumulated over the three-month period. Where appropriate, we lag variables to avoid simultaneity bias driven by balance sheet identities. A table of descriptive statistics is available in Annex 10.

Table 7 reports the results for equation (8). We find that banks that were constrained by the absolute capital regulation in a quarter experienced a capital growth rate that was approximately 5.9-10.5%. higher on average per quarter than when they were not constrained (see columns 3 and 4 in Table 7, our preferred results). Using time fixed effects or two-way fixed effects as a sense check, the higher average growth rate drops only slightly. The estimates obviously smooth out the fact that some banks recapitalised quickly while for others it was a more drawn-out experience. But given the mean (median) capital growth rate per quarter in the dataset was 2.99% (1.37%), the estimate is economically significant. Note that the non-winsorised capital growth data can be noisy. Winsorised data produce smaller but still statistically significant estimates. The full TWFE estimate including covariates (the final column of Table 7) fails to find a significant relationship at conventional levels, although this is not the case for the winsorised results.

Table 7 – Capital Growth and Capital Deficient Banks

Notes: This table presents estimates regressing banks' regulatory capital growth rates (the dependent variable, in percent) against a dummy variable indicating whether or not the bank was deemed capital deficient under the absolute regulatory limit in a given quarter. The regression are approximately based on equation (8) in the text. Each column represents a regression using a different estimation approach (OLS = Ordinary Least Squares, Within = Bank Fixed Effects, Time = Time Fixed Effects, TWFE = Two-way Fixed Effects). Data is quarterly and was constructed from the Bank of Zambia's monthly Prudential Return. Bank-Specific controls include the following variables (for two lags, including t-1 and t-2): log of total bank assets, return on assets, liquid assets as a percent of total assets, RWAR ratio, a sector-specific loan measure, provisions for the quarter as a percent of beginning total assets. Macro variables include the following (for two lags, including t-1 and t-2): the statutory reserve ratio, quarterly growth in the exchange rate, the average inflation rate over the quarter, and the average Treasury Bill Yield (365 day) over the quarter. Driscoll Kraay or Robust Errors were used and are provided in parentheses. Data not winsorised. *** p < 0.01, ** p < 0.05, * p < 0.1.

Dependent Variable	Total Regulatory Capital Growth (Percent)							
	OLS	OLS	Within	Within	Time	Time	TWFE	TWFE
Capital Deficit Dummy	0.083***	0.056**	0.105***	0.059**	0.055***	0.044**	0.079**	0.047
	[0.022]	[0.022]	[0.025]	[0.023]	[0.019]	[0.021]	[0.035]	[0.041]
Observations	774	738	774	738	774	738	774	738
R-Squared	0.024	0.147	0.024	0.149	0.1	0.208	0.1	0.205
Bank-Specific Controls	X	√	X	√	X	√	X	√
Bank Fixed Effect	X	X	√	√	X	X	√	√
Time Fixed Effects	X	X	X	X	√	√	√	√
Macro Controls	X	√	X	√	X	X	X	X

Table 8 presents results for equation (8) but using the disaggregated components of capital (share capital, retained earnings, and tier 2 capital[100]). The idea here is to get a feel for the key themes underpinning how banks recapitalised across capital accounts. We estimate a much higher average growth rate on the share capital growth coefficient (101.5%, column 2), a capital type that we know

[100] Tier 2 capital captures several legitimate components of bank capital that are typically counted as regulatory capital but still do not quite offer the protection of, say share capital, or other examples of tier 1 capital. Tier 2 capital here captures eligible preferred shares, eligible subordinated term debt, eligible general provisions, and revaluation reserves, among others.

many banks were forced to recapitalize with because BOZ required that 80% of the recapitalisation be in the form of paid-up share capital[101]. The same cannot be said for growth (or reductions) in retained earnings, which will reflect profitability and write-offs. It is also consistent with the fact that banks did not frequently inject share capital during other periods in our sample, and instead grew by accumulating retained earnings. The negative but insignificant coefficients on retained earnings growth rate coefficients (columns 4-6) are consistent with the fact that several banks reallocated capital between their surplus capital and share capital accounts, and a lower growth rate coefficient is consistent with this. Such balance sheet reshuffling helped banks that had accumulated retained earnings to not have to raise additional share capital (originally BOZ required 80% of the recapitalisations to be share capital), but it would not have impacted their absolute level of regulatory capital (as both share capital and retained earnings/retained surplus capital accounts count as regulatory capital).

Table 8 – Components of Capital Growth and Capital Deficient Banks

Notes: This table presents estimates regressing compoenents of banks' regulatory capital growth rates (the dependent variables, including Paid-up Share Capital, Retained Earnings, and Tier 2 Capital, in percent) against a dummy variable indicating whether or not the bank was deemed capital deficient under the absolute regulatory limit in a given quarter. The regressions are approximately based on equation (8) in the text. Each column represents a regression using a different estimation approach (OLS = Ordinary Least Squares, Within = Bank Fixed Effects, Time = Time Fixed Effects, TWFE = Two-way Fixed Effects). Data is quarterly and was constructed from the Bank of Zambia's monthly Prudential Return. Bank-Specific controls include the following variables (for two lags, including t-1 and t-2): log of total bank assets, return on assets, liquid assets as a percent of total assets, RWAR ratio, a sector-specific loan measure, provisions for the quarter as a percent of beginning total assets. Macro variables include the following (for two lags, including t-1 and t-2): the statutory reserve ratio, quarterly growth in the exchange rate, the average inflation rate over the quarter, and the average Treasury Bill Yield (365 day) over the quarter. Driscoll Kraay/Robust Standard Errors were used and are provided in parentheses. Data not winsorised. *** p < 0.01, ** p < 0.05, * p < 0.1.

Dependent Variable (non-Windsorised)	Regulatory Capital Component Growth (Percent)								
	Paid-up Share Capital Growth			Retained Earnings			Tier 2		
	OLS	Within	TWFE	OLS	Within	TWFE	OLS	Within	TWFE
Capital Deficient Dummy	1.133**	1.015**	0.606*	-1.179	-2.007	-2.899	-2.267	-1.925	0.023
	[0.496]	[0.414]	[0.337]	[0.811]	[1.620]	[2.118]	[1.548]	[1.246]	[0.726]
Observations	746	712	712	746	712	712	495	481	481
R-Squared	0.07	0.09	0.23	0.00	0.03	0.10	0.01	0.07	0.18
Bank-Specific Controls	X	√	√	X	√	√	X	√	√
Bank Fixed Effect	X	√	√	X	√	√	X	√	√
Time Fixed Effects	X	X	√	X	X	√	X	X	√
Macro Controls	X	√	X	X	√	X	X	√	X

Taken together, these results support the idea that regulatory capital deficient banks underwent recapitalisations, and potentially more than they would have otherwise. At the least, there is a broadly statistically significant association between capital growth rates and banks being subjected to the policy. Taken at face value, the results also support the idea that BOZ was successful in forcing

[101] Note: These series are noisy as some banks shift capital across their capital accounts. But the results are robust, albeit of a much smaller magnitude, if we use winsorised data.

banks to raise capital. Whether this increased capital also leads to higher capitalisation rates is at this point unclear.

As we can see from this initial analysis, banks that were bound by the BOZ's recapitalisation rule did end up increasing their capital levels, and likely more than they would have otherwise, ceteris paribus. However, we see that banks are likely to rearrange their capital accounts to minimise new capital issuance; regulators need to be careful about how they design recapitalisation policy in the future to account for this.

Capitalisation Rates

One of BOZ's primary goals with the absolute capital regulation was to improve the soundness of the banking system (Gondwe 2013)[102]. In this section we test for an association between policy-linked bank recapitalisation levels and reductions in banks' financial risks, as measured by their capitalisation rate (leverage ratio and RWA ratio). As banks were not constrained by the absolute rule to hold a certain capitalisation rate (leverage ratio and RWAR), they could have, in theory at least, increased their deposits, borrowing, or asset riskiness in line with their recapitalisation levels, leaving their relative capitalisation rates unchanged[103]. They may also have unwound the capitalisation impact of raising capital over time as they expanded their holdings or risky assets, essentially implying that the impact of the policy on capitalisation rates was only short term.

We evaluate this link between the absolute capital rule and bank capitalisation in two parts. First, following on from the previous section, we examine whether there is evidence of a positive association between banks being bound by the absolute capital rule and changes in RWA or leverage ratios, holding constant other factors that might drive changes in capitalisation rates. Second, we examine how bank capitalisation rates responded to policy-induced recapitalisation episodes, both contemporaneously and then throughout the following 6 quarters. This allows us to test whether banks retained (or slowly unwound) the effects of their capital raisings on their capitalisation rate.

As to the first part, if the imposition of the policy was associated with improvements in bank capitalisation rates, we would expect the percentage point change in RWARs to increase more for a

[102] Gondwe (2013) pg.3 stated one goal was to, "Make the banking system strong enough to withstand shocks and instability in the economy".
[103] Note that the minimum Basel-based CAR rule of 10% still applied and acted as a regulatory floor.

bank when it was deemed capital deficient under the rule than otherwise. Hence, we examine whether there is evidence of a positive association between banks being bound by the absolute capital rule and changes in RWA or leverage ratios:

$$\Delta RWAR_t^i = \alpha_i^j + \beta_1 . def_ind_t^i + \sum_{p=1}^{P} \boldsymbol{\beta}_P{}' . \boldsymbol{X}_{t-p}^i + \sum_{L=1}^{L} \boldsymbol{\beta}_L{}' . \boldsymbol{M}_{t-p} + \varepsilon_{it} \quad (9)$$

$$def_ind_t^i = \begin{cases} 1 & , C_t^i - C_t^{i,Reg} < 0 \\ 0 & , C_t^i - C_t^{i,Reg} \geq 0 \end{cases}$$

Where $\Delta RWAR_t^i$ is the percentage point change for bank i for quarter t-1 to t, and $def_ind_t^i$, $\mathbf{X}_{t-p}^i$, and $\boldsymbol{M}_{t-p}$ are as explained previously.[104]

In principle, our policy treatment in equation (9), and in many of the remaining sections of this paper, is similar to a set of papers in the bank bailout literature that essentially investigate the effect of a shock to banks' capitalisation levels on banking market outcomes. These papers typically do not have bank-specific measures of bailout size and so use a treatment dummy variable to capture the recapitalisation, independent of its bank-specific size (there are many examples, but two are Li (2013) and Puddu and Wälchli (2013)). The one benefit the bailout papers have compared to this one is that they have many non-bailout banks that can be used to create a control group. Although we do have some banks and deposit-taking non-banks that were not directly bound by the regime, a simple DiD approach is not feasible given our sample. Another point to note is that we are not trying to uncover a general estimate of bank equity/capitalisation or equity/bank lending elasticity, as with other papers in the literature (e.g., Berrospide and Edge (2010)). Instead, we are only interested in the specific context in which the capital raising is due to an absolute regulatory requirement – we make no claim that the estimates are relevant outside such a context.

We find evidence that banks that were capital deficient did increase their capitalisation rates, in addition to their capital levels. On average, banks that were constrained by the capital rule increased their capitalisation rate by approximately 5.38 percentage points, compared to when they were not constrained, ceteris paribus (see column 4 from Table 9, which shows the estimates from equation (9) and constitutes our preferred estimate). Given that mean (median) capitalisation rate in the sample

[104] We also used $\Delta Leverage_t^i$ as the dependent variable as another measure of bank risk, where $Leverage_t^i$ is defined as Total Bank Equity/Total Bank Assets. The results are broadly consistent.

was 30.5% (22.6%), this is an economically meaningful average per quarter change in capitalisation rates.

Note that a bank's capitalisation rate at a point in time is a composite measure reflecting many aspects of a bank's business: it reflects a bank's decisions on both the volume and characteristics of its investments (including lending), how it finances these investments, dividend policy and whether it sets capital structure targets, risk preferences, as well as the valuation/performance of such assets and liabilities (which may be the result of market prices and forces, including, for example, the demand for and performance of loans). As a result, there are many factors that can drive a change in a bank's capitalisation rate, and bank capitalisation rates can be volatile. Nevertheless, through this noise we still estimate a level difference in changes in banks' capitalisation rates associated with the policy. The result is robust to using the leverage ratio as an alternative measure of the capitalisation, although this is not unsurprising given their similarities.

Table 9 – Changes in Risk Weighted Asset Ratios and Capital Deficient Banks

Notes: This table presents estimates regressing changes in banks' RWARs (the dependent variable, in percentage points) against a dummy variable indicating whether or not the bank was deemed capital deficient under the absolute regulatory limit in a given quarter. The regressions are approximately based on equation (9) in the text. Each column represents a regression using a different estimation approach (OLS = Ordinary Least Squares, Within = Bank Fixed Effects, Time = Time Fixed Effects, TWFE = Two-way Fixed Effects). Data is quarterly and was constructed from the Bank of Zambia's monthly Prudential Return. Bank-Specific controls include the following variables (for two lags, including t-1 and t-2): log of total bank assets, return on assets, liquid assets as a percent of total assets, RWAR ratio, a sector-specific loan measure, provisions for the quarter as a percent of beginning total assets. Macro variables include the following (for two lags, including t-1 and t-2): the statutory reserve ratio, quarterly growth in the exchange rate, the average inflation rate over the quarter, and the average Treasury Bill Yield (365 day) over the quarter. Driscoll Kraay or Robust Errors were used and are provided in parentheses. Data is not winsorised. *** p < 0.01, ** p < 0.05, * p < 0.1.

Dependent Variable	Change in Risk-weighted Asset Ratio (Percentage Points)					
	OLS	OLS	Within	Within	TWFE	TWFE
Capital Deficit Dummy	0.0325*	0.0509***	0.0715***	0.0538***	0.0873**	0.0562**
	[0.0167]	[0.0179]	[0.0249]	[0.0180]	[0.0310]	[0.0277]
Observations	746	712	746	712	746	712
R-Squared	0.04	0.05	0.04	0.06	0.12	0.12
Bank-Specific Controls	X	√	X	√	X	√
Bank Fixed Effect	X	X	√	√	√	√
Time Fixed Effects	X	X	X	X	√	√
Macro Controls	X	√	X	√	X	X

While equation (9) and Table 9 tell us that there is a level or percentage point shift in changes in capitalisation rates associated with a bank being capital deficient under the absolute capital regime, what remains to be seen is whether, once a bank recapitalises in association with the absolute regulatory rule, it continues to retain and operate at a higher capitalisation rate.

This thus takes us to the second part of our analysis. It may take banks time to readjust their lending, investing, and financing post recapitalising, and so we might not expect to see banks immediately

return to their post-recapitalisation target level of capitalisation. To investigate whether banks "retain" the recapitalisation effect of the capital injection, we estimate the following:

$$\Delta RWAR_t^i = \alpha_i + \sum_{k=0}^{K} \beta_k \frac{\Delta def_{t-k}^i}{TA_{t-k}^i} + \sum_{p=1}^{P} \boldsymbol{\beta}_P{}' . \boldsymbol{X}_{t-p}^i + \sum_{l=1}^{L} \boldsymbol{\beta}_L{}' . \boldsymbol{M}_{t-l} + u_{it} \qquad (10)$$

$$def_t^i = \begin{cases} C_t^i - C_t^{i,Reg}, & C_t^i - C_t^{i,Reg} < 0 \\ 0 \quad , & C_t^i - C_t^{i,Reg} \geq 0 \end{cases}$$

Where $\Delta \mathrm{RWAR}_{\mathrm{t}}^{\mathrm{i}}$ is the percentage point change for bank i for quarter t-1 to t, and $\mathbf{X}_{\mathrm{t-p}}^{\mathrm{i}}$,and $\boldsymbol{M}_{t-p}$ are as explained previously. def_t^i captures a bank's regulatory capital deficiency at time t, and so $\Delta \mathrm{def}_{\mathrm{t-k}}^{\mathrm{i}}/\mathrm{TA}_{t-k}^{i}$ measures the growth in regulatory equity a bank underwent in a period (as a percent of total assets) contingent on the equity growth being associated with the bank being in regulatory deficit. This is the same as the data used in Figure 11. β_{k} hence estimates the association between an equity injection in a given period (but only those injections that are closing a regulatory gap) and the percentage point change in the bank's capitalisation rate for the period t-k. For example, if k=0, then $\beta_{\mathrm{k=0}}$would be an estimate of the instantaneous effect of the equity injection on capitalisation. The idea is that, if there is persistence in the effect of the policy-induced recapitalisations, we should find that the summation of the β_{k} should be positive (i.e., their estimated total effect is greater than zero) and jointly statistically significant, even as we examine longer sets of k (as we examine longer lags there is additional time for banks to unwind the effects). Compared to equation (9), equation (10) associates larger recapitalisations in a specific period with changes in capitalisation rates in that period as well as for the subsequent periods. Table 10 shows that there is not strong evidence that banks aggressively unwound the capitalisation level effect of the regulatory-induced recapitalisation. Each column of Table 10 estimates equation (10) for increasing lengths of deficit recapitalisation lags (i.e., k), with each row providing estimates for the effect for a given deficit recapitalisation lag (k). The results are robust to using winsorised data or a TWFE setup. The "Total Estimated Effect" is calculated at the bottom of the table, as well as the hypothesis that it is significantly different form zero (i.e., $\sum_{\mathrm{k=0}}^{\mathrm{K}} \beta_{\mathrm{k}} = 0$).

We estimate that a 1 percentage point policy-induced equity shock (as a percent of total assets) was associated with an instantaneous increase in RWAR of approximately 1.6 percentage points, with

the total estimated effect remaining at around 1 percentage points as we estimate equation (10) for longer sets of k. For example, using a six-quarter lag (i.e., k=6, columns 11 and 12 in Table 10), our estimate of the total effect of a 10 percent recapitalisation remains statistically significant at 10.04 percentage points. The consistently positive and significant long-term effect (the lack of negative estimated coefficients for the shock lags) is evidence that banks did not readjust their balance sheets to retain their previous capitalisation rates after recapitalising, at least within a six-quarter timeframe. Given the significant frictions and imperfections in Zambia's financial markets and the longer-term timeframes underpinning banks' capital structure strategies (which might slow bank reactions), the possibility clearly remains that banks revert to their previous levels of capitalisation over long horizons. Indeed, there is a slow reduction in the estimated total effects as we estimate over longer periods. When we use winsorised data, which will arguably mute the strongest recapitalisation experiences, we see a smaller but still statistically significant effect and one that drops more quickly. The results are robust to a TWFE approach (Annex 6).

In any case, at least in the short term, these estimates support the idea that BOZ's absolute recapitalisation policy during fair-weather conditions did reduce bank riskiness as measured by relative capitalisation rates, one of BOZ's key goals for the policy. In terms of the challenges of using RWAR rules directly to improve capitalisation, the absolute capital policy also appears to have allowed BOZ to improve bank capitalisation rates without needing to increase regulatory capital asset ratios and risk banks deleveraging.

From an MM framing, the results indicate that banks do not appear to have adjusted back to pre-recapitalisation capital targets, at least in the short run. This could be because banks changed their targets after the recapitalisation, frictions/imperfections in that market made it uneconomic to make further capital structure changes, or banks make long-term capital targets that they will adjust to over several years.

Table 10 – Changes in Risk Weighted Assets and Bank Recapitalisations

Notes: This table presents estimates regressing banks' risk weighted asset ratios (the dependent variable, in percentage points) against a recapitalisation variable and its various lags. The recapitalisation variable is the ratio of equity raised by a bank in deficit under the absolute minimum capital rule in a quarter divided by the bank's total assets at the beginning of that quarter. The "Cumulative Effect" row calculates the sum of all the estimate lag coefficients included in that regression. The "P-value (H0: Cumulative Effect=0)" column is testing the hypothesis that the linear combination of the regulatory deficit variable and its lags are significantly different to zero. The regressions are approximately based on equation (10) in the text. Each column represents a regression using a different estimation approach (OLS = Ordinary Least Squares, Within = Bank Fixed Effects, Time = Time Fixed Effects, TWFE = Two-way Fixed Effects) and a different lag length (for example, (0,4) indicates that lags from 0 to 4 are included). Data is quarterly and was constructed from the Bank of Zambia's monthly Prudential Return. Bank-Specific controls include the following variables (for two lags, including t-1 and t-2): log of total bank assets, return on assets, liquid assets as a percent of total assets, RWAR ratio, a sector-specific loan measure, provisions for the quarter as a percent of beginning total assets. Macro variables include the following (for two lags, including t-1 and t-2): the statutory reserve ratio, quarterly growth in the exchange rate, the average inflation rate over the quarter, and the average Treasury Bill Yield (365 day) over the quarter. Driscoll Kraay or Robust Errors were used and are provided in parentheses. Data are not winsorised. *** p < 0.01, ** p < 0.05, * p < 0.1.

Dependent Variable	Change in Risk-Weighted Asset Ratio (Percentage Points)											
Lags (Beg, End)	(0,0)		(0,2)		(0,3)		(0,4)		(0,5)		(0,6)	
Regression Type	OLS	Within	OLS	Within	OLS	Within	OLS	Within	OLS	Within	OLS	Within
Recapitalisation Variable	1.7144***	1.6854***	1.6408***	1.7025***	1.6308***	1.6847***	1.6532***	1.6028***	1.5969***	1.6463***	1.5978***	1.6462***
	[0.3022]	[0.2967]	[0.3150]	[0.3122]	[0.3207]	[0.3194]	[0.3294]	[0.3230]	[0.3190]	[0.3254]	[0.3208]	[0.3273]
Recapitalisation Variable (t-1)			-0.2258	-0.2277	-0.235	-0.2425	-0.25	-0.259	-0.2645	-0.2555	-0.2689	-0.2542
			[0.2415]	[0.2547]	[0.2372]	[0.2482]	[0.2394]	[0.2337]	[0.2376]	[0.2425]	[0.2404]	[0.2473]
Recapitalisation Variable (t-2)			-0.0355	-0.0418	-0.0212	-0.0338	-0.0483	-0.0453	-0.0509	-0.044	-0.0541	-0.0537
			[0.1026]	[0.0905]	[0.0989]	[0.0905]	[0.0989]	[0.1012]	[0.0977]	[0.0966]	[0.0979]	[0.0986]
Recapitalisation Variable (t-3)					-0.1733**	-0.1470*	-0.1557**	-0.1958**	-0.2012**	-0.1726**	-0.2043***	-0.1706**
					[0.0849]	[0.0784]	[0.0728]	[0.0799]	[0.0785]	[0.0715]	[0.0750]	[0.0690]
Recapitalisation Variable (t-4)							-0.0416	-0.0378	-0.0491	-0.0366	-0.0525	-0.0429
							[0.0572]	[0.0516]	[0.0536]	[0.0568]	[0.0557]	[0.0577]
Recapitalisation Variable (t-5)									0.0352	-0.0035	0.0425	0.0081
									[0.0356]	[0.0596]	[0.0468]	[0.0632]
Recapitalisation Variable (t-6)											-0.0712	-0.0869
											[0.1109]	[0.0903]
Cumulative Effect	**1.714**	**1.685**	**1.380**	**1.433**	**1.201**	**1.261**	**1.158**	**1.065**	**1.066**	**1.134**	**0.989**	**1.046**
P-value (H0: Cumulative Effect=0)	**0.000**	**0.000**	**0.000**	**0.000**	**0.001**	**0.000**	**0.001**	**0.002**	**0.002**	**0.003**	**0.004**	**0.006**
Observations	746	712	714	712	698	696	680	682	666	664	650	648
R-Squared	0.174	0.236	0.206	0.24	0.215	0.25	0.276	0.252	0.258	0.281	0.26	0.285
Bank-Specific Controls	X	√	X	√	X	√	X	√	X	√	X	√
Bank Fixed Effect	X	√	X	√	X	√	X	√	X	√	X	√
Time Fixed Effects	X	X	X	X	X	X	X	X	X	X	X	X
Macro Controls	X	√	X	√	X	√	X	√	X	√	X	√

Bank Asset Reactions

One of BOZ's key objectives with its recapitalisation policy was to improve lending in the economy[105]. Yet, as we show in the simple framework previously, inducing large lending responses with forced capital injections could be a challenge; unless the capital reduced market frictions or creates externalities, it is not likely to improve the economics or profitability of bank lending opportunities, thus limiting the likelihood that a capital injection would result in a large, leveraged increase in bank lending (as BOZ perhaps had hoped)[106]. Banks also had access to high-yielding Treasury securities which would have offered an investment alternative to lending, meaning that banks may have invested into Treasury securities what could not be economically allocated to loans or other securities.

In this section we investigate whether banks expanded their assets post-recapitalisation, and, if so, what type of assets they typically extended into[107]. We will also investigate what assets (if any) banks expanded into, testing the hypothesis that the recapitalisation policy may have supported government financing.

First, the empirical setup. The goal is to evaluate how bank assets (in total and of different kinds) reacted to the recapitalisation policy. As with equation (10) in the previous section on capitalisation rate, we want to allow banks to react with a lag, and we are primarily interested in the total impact of the policy (for a given equity injection). As such, we estimate a simple distributed lag model:

$$\Delta y_t^i = \alpha^i + \sum_{k=0}^{K} \beta_k \frac{\Delta def_{t-k}^i}{TA_{t-k}^i} + \sum_{p=1}^{P} \boldsymbol{\beta}_P{}' . \boldsymbol{X}_{t-p}^i + \sum_{L=1}^{L} \boldsymbol{\beta}_L{}' . \boldsymbol{M}_{t-l} + u_{it} \qquad (11)$$

$$def_t^i = \begin{cases} C_t^i - C_t^{i,Reg}, & C_t^i - C_t^{i,Reg} < 0 \\ 0, & C_t^i - C_t^{i,Reg} \geq 0 \end{cases}$$

[105] For example, to "…improve access to loanable funds" p3 (Gondwe, 2013) and "a solid capital base allows banks to increase their risk taking and improves the amount of lending in the economy" p1 (Dr Ngandu, the Deputy Governor, in Lusaka Times (2012).

[106] Table 4 and Figure 8 form the Theoretical Framework provides sone intuition for this.

[107] Stickly speaking, it remains an empirical question whether banks will increase their assets in response to the equity injection. Our simple theoretical framework and the fact that banks can invest in non-loan securities suggests that they would likely expand their balance sheets, at least in the short term. But, if Zambia's debt financing markets operated with little friction and there were no good options for investment, we might expect otherwise (i.e., that banks might keep their assets base constant, and simply reduce debt financing).

Where Δy_t^i is the growth in a specific asset class for bank i for quarter t-1 to t. Each asset class is estimated separately, including the following: total assets, loans, Treasury assets (including bills and bonds of all tenures), and other assets (comprised primarily of assets held in foreign banks and statutory reserves, but also including other bank assets and infrastructure, which are typically a small component). $\mathbf{X}_{t-p}^i$ and $\boldsymbol{M}_{t-p}$ are time-varying bank-specific and macroeconomic controls, respectively, and as explained previously. def_t^i captures a banks regulatory capital deficiency at time t, and so $\Delta \mathrm{def}_{t-k}^i / TA_{t-k}$ measures the percentage point change in total assets that resulted from a policy-induced equity raising in a given period[108].

In terms of identification, our preferred approach is to use within-bank variation and link deviations of a bank's asset growth (from its average) to variation in the same bank's capital growth (from its average), conditioned on the capital growth being a result of closing a regulatory deficit gap (which we argue is exogenous). Assuming BOZ's hypothesis that capital raising would increase lending was true, we would expect that banks would record larger asset/loan growth during/immediately after a recapitalisation episode compared to when the same bank was not recapitalising (holding all else equal). In this approach, banks that were not capital deficient under the regulatory regime will be zero-weighted in their contribution to the recapitalisation coefficient (Kropono and Kubinec (2020)).

The key fixed effects assumption, applied to this context, is hence:

$$E(u_{it} | \Delta def^{\boldsymbol{i}}, \boldsymbol{X^i}, \boldsymbol{M}, \alpha^i) = 0$$

This implies that, conditional on bank-specific factors (both time-varying and time-invariant) and macro covariates, bank recapitalisations due to the new capital regime from any period should not be related to the error in time t. As discussed in the empirical discussion in the data section, we believe this is a reasonably defendable assumption for examining banks' lending growth, especially given the non-bailout and non-selective nature of the recapitalisation.

As discussed in Peydró (2010), changes in bank capital are typically endogenous to loan demand and the economic context, meaning that any estimate linking changes in bank capital to changes in

[108] In some ways, our policy treatment is similar to some papers in the bank bailout literature, which essentially investigate the effect of a shock to banks' capitalisation levels on banking market outcomes. These papers typically do not have bank-specific measures of bailout size and so use a treatment dummy variable to capture the recapitalisation, independent of its bank-specific size (there are many examples, but two are Li (2013) and Puddu and Wälchli (2013)).

banking market outcomes will be biased. The endogeneity is driven by similar factors impacting both the demand for loans and the evolution of bank equity.

Our primary approach to this endogeneity in this paper is to treat Zambia's regime-forced bank recapitalisations as an exogenous source of capital variation. The idea is that, while changes in capital are generally endogenous, changes in banks' capital to meet the new BOZ absolute capital limit were not. This approach is most similar to Peek and Rosengreen (1997) and Mora and Logan (2010), although, in our case, the exogenous capital changes are limited to a shock to a subset of banks within a given period, rather than by exploiting cross-border banking relationships.

We choose to use $\Delta \text{def}^{\text{i}}_{\text{t}-\text{k}}/TA_{t-k}$ as our main exogenous policy variable because it assists to normalise, at least slightly, across banks with different business models[109]. It is also not uncommon to use banks leverage as a shock/policy variable in the capitalisation and bank lending literature (rather than, say, policy induced equity growth, $\Delta \text{def}^{\text{i}}_{\text{t}-\text{k}}/E_{t-k}$ or as a percent of RWARs $\Delta \text{def}^{\text{i}}_{\text{t}-\text{k}}/RWA_{t-k}$). Annex 7 provides the results when using policy-induced equity growth and policy induced changes to banks RWAR (rather than the shock as a percent of total assets). In any case, the direction and statistical significance of this paper's results are quite robust to using different forms of the exogenous equity shock or using winsorised data.

In our analysis, we find evidence that banks did expand their asset bases post injecting equity, but that little of this expansion appears to have been into additional lending. Instead, we find that banks increased their holdings of Treasury assets. We did not find statistically significant coefficient estimates to support the claim that banks increased their holdings of other asset types, although this is quite possibly a data depth issue[110].

Columns 1-3 of Table 11 show the results of equation (11) using banks' total asset growth as the dependent variable. Each "Recapitalisation Variable (t-k)" row provides the coefficient for the equity shock variable of the k-th lag. The seventh row shows the total estimated effect ($\sum_{\text{k}=0}^{\text{K}} \beta_{\text{k}}$) and the p-value testing for the hypothesis that the cumulative estimate effect is different from zero. We find

[109] We don't necessarily want to pick up differences across banks portfolio composition when measuring the shock, which would happen more if we used lagged RWAR in the denominator rather than total assets. Using equity growth is a legitimate option and we provide these additional results in Annex 7.

[110] The other asset class is a catch-all for all non-loan, non-treasury assets. The data series are quite noisy because banks often fully enter and fully exit positions in these asset classes, often with only a short holding period. As such, we do not have enough bank data to meaningfully partition the effect of the equity injection from what are noisy growth series. Future work might look to use an aggregate approach for this case.

that in the quarter of the equity injection, banks increase their asset holdings, and they do so approximately in line with the size of the equity injection (the coefficients in row 1 of columns 1-3 span from 0.99 to 1.03, very close to 1). This is perhaps unsurprising given that we know that it may take time for banks to leverage any additional equity they injected (or, on the contrary, reduce their debt financing - changing domestic deposit holdings would take time). However, considering that the total estimated effect of the recapitalisation after five months was 0.823 (see Table 11, specifically the "within" column under the Total Assets heading (column two) and the row titled "Cumulative Effect"), there is evidence that banks likely did not further expand their assets in addition to this initial 1:1 equity for assets growth. We also cannot reject the null hypothesis that the cumulative effect is different to 1. These results sit reasonably well with the simple theoretical framework; in the short term at least, there is no strong evidence of banks leveraging and drastically increasing their asset holdings (or trying to return to their previous leverage ratios).

Given this initial evidence that banks did expand their assets, is there evidence that banks expanded into additional lending as BOZ had hoped? Or did banks expand into other asset classes? Using equation (11), we estimate whether the equity shock is associated with increased growth rates for different asset classes, specifically, total loan growth (columns 4-6 in Table 11), treasury asset holding growth (columns 7-9), and other asset growth (columns 10-12). We do not find evidence that banks expanded their lending, at least within the first five quarters post injection. The coefficient on the contemporaneous loan growth effect is not statistically different from zero, and neither is the estimated total effect over the next five quarters (row 7 column 6). Using other versions of the equity shock variable does not change this outcome (see Table 33-Table 34 in Annex 7). This result is noteworthy because it suggests, contrary to BOZ's objective, there is not evidence that forced absolute recapitalisations are likely to spur large increases in bank lending, at least within a year-and-a-half time period. Instead, we find strong evidence that banks increased their holdings of treasury securities post equity injection. Including the quarter after the bank recapitalised, for each 1 percentage point growth in a bank's equity/total asset ratio that would have resulted from the forced recapitalisation, the bank's growth rate in treasury assets increased by over 3 percentage points (Table 11, column 8, estimate rows 1 and 2). Over the full five quarters post recapitalisation we estimate the total effect to be 4.3 percentage points. This is evidence that, at least in the short term (five quarters post capitalisation quarter) and holding all else constant, the recapitalisation likely had

an impact on the banking system's capacity to support the Government's domestic debt issuance and borrowing. We cannot tell from this data and setup whether this higher treasury growth actually crowded-out the banking sector's financial intermediation in an absolute sense, but these treasury growth estimates are prima facie evidence that the equity injections may have had some impact on domestic debt financing.

We find little some evidence that banks grew their holdings of "Other Assets" contemporaneously in wake of a forced capitalisation (row 1 column 10-12 in Table 11). Over a longer six-quarter period, however, our estimates of the total effect of the injections were not statistically distinguishable from zero. We know from examining the raw data series that some banks did appear to contemporaneously increase their holding of overseas deposits post recapitalisation, essentially recycling some of the financing associated the equity injection back to their parents as deposits (i.e., the Zambian subsidiary making a deposit at its parent). However, because of the very noisy nature of the "other assets" series for some banks in the dataset, we would likely struggle to detect such effects statistically over longer periods, even if they did exist (the "Other Assets" series is noisy because banks often take large short-term positions with their parents that can then be quickly reversed). Results using banks Off Balance Sheet liabilities (i.e., off balance sheet guarantees, indemnities, and letters of credit – all a form of banks expanding their role in financing or supporting firms in Zambia) suffer from a similar problem, and show no effect. Future research might examine aggregate bank assets in Zambia's banking market in relation to the forced recapitalisations to better manage the effect of volatile bank-level data series or employ bank-level case studies.

The results in Table 11 also show some of the weaknesses in our data and empirical strategy. The negative and statistically significant Treasury Growth coefficients and the positive Other Asset coefficients in row 4 of the coefficient results (columns 7-12) are likely reflecting a tightening of monetary policy by BOZ in March 2014 (whereby a 6 percentage point increase in required reserve holdings would have forced some banks to switch from Treasury assets to holding reserve assets), soon after several banks recapitalised (Table 42 in Annex 11 tracks monetary policy in Zambia over the data sample). Such effects should have been picked up in our macroeconomic controls or in the time fixed effects and highlight the limitations of our assumed exogeneity of some bank recapitalisations.

How do these asset response results sit with the bank capital literature? Framed within the bank bailout literature, our results of a weak-at-best lending reaction is most consistent with Nakashima (2016) who finds little evidence that Japanese banks expanded their lending because of government equity injections in response to a string of bank failures in the late 1990s. This is in contrast to other research from the US and Europe which finds stronger links between equity bailouts and loan growth (Li (2013), Puddu and Wälchli (2015), Brei, Gambacorta, and von Peter (2013), and Homar (2016)). Importantly, however, there is a substantial qualitative difference in both economic and financial context between a "bailout" equity injection and a "fair-weather" recapitalisation. In a crisis, equity injections may affect market sentiment differently and have broader economic spillovers that might flow back and substantially improve the attractiveness to banks of lending. It's not clear that such dynamics applied in Zambia's case, which might explain the difference in findings between our work and others. The literature examining the link between bank capitalisation levels and lending growth, for example Bernanke and Lown (1991) and Berrospide and Edge (2010), suggest that better capitalised banks lend more. While we do not find a negative loan growth coefficient in our results, our results are obviously not aligned. The difference may again simply be a result of contextual differences (both in term of the banking sector's performance at the time and structural differences in the banking and economic systems). Another explanation could be that threshold effects exist, whereby improvements in bank capitalisation only led to stronger lending growth in the case where banks are badly undercapitalised or under stress, which was not as clearly the case for the majority of Zambian banks.

Table 11 – Bank Asset Growth and Bank Recapitalisations.

Notes: This table presents estimates from regressing the growth rates for different types of bank assets (the dependent variable, in percent) against a recapitalisation variable and its various lags. The recapitalisation variable is the ratio of equity raised by a bank in deficit under the absolute minimum capital rule in a quarter divided by the bank's total assets at the beginning of that quarter. The "Cumulative Effect" row calculates the sum of all the estimate lag coefficients included in that regression. The "P-value (H0: Cumulative Effect=0)" column is testing the hypothesis that the linear combination of the regulatory deficit variable and its lags are significantly different to zero. The regressions are approximately based on equation (11) in the text. Each column represents a regression using a different estimation approach (OLS = Ordinary Least Squares, Within = Bank Fixed Effects, Time = Time Fixed Effects, TWFE = Two-way Fixed Effects) and a different asset type (the set includes the following: Total Assets, Bank Loans, Treasury Assets, and Other Assets). Data is quarterly and was constructed from the Bank of Zambia's monthly Prudential Return. Bank-Specific controls include the following variables (for two lags, including t-1 and t-2): log of total bank assets, return on assets, liquid assets as a percent of total assets, RWAR ratio, a sector-specific loan measure, provisions for the quarter as a percent of beginning total assets. Macro variables include the following (for two lags, including t-1 and t-2): the statutory reserve ratio, quarterly growth in the exchange rate, the average inflation rate over the quarter, and the average Treasury Bill Yield (365 day) over the quarter. Driscoll Kraay or Robust Errors were used and are provided in parentheses. Data are not winsorised. *** p < 0.01, ** p < 0.05, * p < 0.1.

Dependent Variable	Percentage Growth in Asset Type (Percent)											
	Total Assets			Loans			Treasury Assets			Other Assets		
Regression Type	OLS	Within	TWFE	OLS	Within	TWFE	OLS	Within	TWFE	OLS	Within	TWFE
Recapitalisation Variable	0.9916***	1.0435***	1.0324***	-0.289	-0.289	-0.4006	2.1020**	2.0654**	2.2842**	1.4975***	1.6350***	1.5295***
	[0.1270]	[0.0933]	[0.0995]	[0.2139]	[0.2139]	[0.2492]	[0.9847]	[1.0055]	[1.1271]	[0.3752]	[0.3602]	[0.3264]
Recapitalisation Variable (t-1)	-0.0946	-0.0806	-0.1377	-0.4535	-0.4535	-0.3719	1.2816	1.4385	1.311	-0.3788	-0.3243	-0.3014
	[0.1279]	[0.1672]	[0.1738]	[0.4545]	[0.4545]	[0.4957]	[0.9135]	[0.9771]	[1.1658]	[0.4042]	[0.4473]	[0.4647]
Recapitalisation Variable (t-2)	0.1545	0.1249	0.1313	0.5009	0.5009	0.3635	-0.0022	0.2469	0.3943*	0.3333	0.0851	0.0721
	[0.1602]	[0.1412]	[0.1277]	[0.3902]	[0.3902]	[0.3490]	[0.1976]	[0.2196]	[0.2204]	[0.3489]	[0.4065]	[0.4372]
Recapitalisation Variable (t-3)	-0.0348	-0.0564	-0.0642	0.1459	0.1459	0.0732	-1.1036***	-0.9428***	-0.8326***	0.136	0.1114	0.2336
	[0.1121]	[0.0903]	[0.0931]	[0.2019]	[0.2019]	[0.2201]	[0.2193]	[0.3324]	[0.2585]	[0.1880]	[0.1507]	[0.1850]
Recapitalisation Variable (t-4)	0.0367	0.0216	0.0839	0.6144**	0.6144**	0.4635	0.3607	0.6989*	0.7207*	-0.2149	-0.5812*	-0.5235
	[0.0987]	[0.0633]	[0.0567]	[0.3041]	[0.3041]	[0.3122]	[0.5712]	[0.3987]	[0.4010]	[0.1329]	[0.3264]	[0.3456]
Recapitalisation Variable (t-5)	-0.0513	-0.2302	-0.2256	-0.0586	-0.0586	0.0072	0.7876*	0.7828	0.729	-0.5689**	-0.9778**	-0.8617**
	[0.0685]	[0.1549]	[0.1633]	[0.1625]	[0.1625]	[0.1274]	[0.4640]	[0.4853]	[0.5679]	[0.2662]	[0.3911]	[0.3976]
Cumulative Effect	**1.002**	**0.823**	**0.82**	**0.46**	**0.46**	**0.135**	**3.426**	**4.29**	**4.607**	**0.804**	**-0.052**	**0.149**
P-value (H0: Cumulative Effect=0)	**0.006**	**0.028**	**0.030**	**0.516**	**0.516**	**0.853**	**0.011**	**0.001**	**0.007**	**0.321**	**0.959**	**0.881**
Observations	666	664	664	664	664	664	645	643	643	666	664	664
R-Squared	0.06	0.27	0.35	0.09	0.09	0.16	0.08	0.19	0.25	0.03	0.17	0.26
Bank-Specific Controls	X	√	√	X	√	√	X	√	√	X	√	√
Bank Fixed Effect	X	√	√	X	√	√	X	√	√	X	√	√
Time Fixed Effects	X	X	√	X	X	√	X	X	√	X	X	√
Macro Controls	X	√	X	X	√	X	X	√	X	X	√	X

Robustness Tests

Annex 10 provides some simple robustness tests for our core results. We take three simple approaches: for investigating bank outliers, we rerun all our main baseline specifications but drop one bank at a time (Table 39); for covariate sensitivity, we run each baseline regression removing each covariate in order from each regression (Table 40); and finally, to investigate our choice of lag structure in the distributed lag models, we run each model over various differing lag structures (Table 41). Our results are not overly sensitive to the inclusion or exclusion of specific banks or variables or our choice of lag structure.

Throughout the paper we typically provided results using both unwinsorised and winsorised data (winsorised at 1%, i.e., for the top 1% and bottom 1% of the distribution). Given the noisiness in some of the growth rate series, we were concerned that outliers might drive specific results. While the coefficients' magnitudes typically sit lower when using winsorised data, the statistical significance of estimates typically holds.

Heterogeneous Bank Reactions and Aggregate Market Structure Trends

One component of BOZ's absolute capital policy was that it required differentiated absolute minimum capital standards depending on whether a bank was majority foreign or domestic owned: international banks were required to meet a minimum capital requirement of Kwacha 520 million, approximately USD 100 million, while Zambian owned banks needed to meet a minimum of Kwacha 104 million or USD 20 million.

By the nature of its construction, such a differentiated policy would impact the market structure in Zambia's banking system, shifting equity market share towards larger international banks (assuming none decided to leave the market, which none did). In aggregate, international banks were required to raise USD 414 million and domestic banks approximately USD 80 million, which BOZ would have ex-ante calculated as a likely outcome[111].

[111] In a policy sense, there are arguments both for and against increasing foreign bank ownership. While foreign ownership might mean banking sector rents are remunerated to another country and that good borrowers might be cherry-picked (Detragiache, Gupta, and Tressel, 2008), there is also evidence that foreign banks operating in Eastern Africa can be better than domestic banks at procuring loans (Zins and Weill, 2018) and can have lower interest rate spreads and be more efficient (Beck et al. 2014). Claessens and van Horen (2012) discuss the debate in much more detail. It's possible that for

What BOZ might not have known was whether different bank types would react differently to the capital injections, potentially furthering the expansion in the market of some bank types over others. The foreign/domestic ownership distinction is an obvious grouping. Still, another logical candidate might be the size of banks, where larger banks may be more able to leverage their balance sheets post recapitalisation because they: (i) already have more capacity and network infrastructure to attract deposits; (ii) may seem additionally less risky to depositors after the equity injection; and (iii) may have better access to overseas funding. Annex 8 provides details on the number of banks in each category.

In this section, we first examine the possibility that the forced equity injections may have had differential impacts on banks. In addition to the direct differences in recapitalisation sizes required for different bank types, differential impacts may have added to a shift in market share from one type of bank to another. Our approach is piecemeal and partial, but we aim to get a high-level feel for whether heterogeneity in bank responses might have been a feature of the post-recapitalisation banking market. We begin by examining whether banks of different types responded differently to the shock – strong evidence of heterogenous asset growth reactions might hint at unintended market share consequences. We depart from our previous empirical specification and instead use a setup similar to Fang et al. (2020) that allows us to more clearly test for differences in asset growth slope parameters based on bank type[112]. To examine the possibility that banks of different types could better leverage the new equity in the deposit market, we test for heterogeneity in banks' deposit reactions. We then move on to examine some high-level aggregate market structure data and its evolution post recapitalisation. We stop short, however, of directly linking bank specific recapitalisations to changes in market shares (or another potentially relevant outcome, competition).

To examine whether banks of different types responded differently to the shock, we want to test for heterogeneous responses to banks' forced recapitalisations. Evidence of heterogenous responses might hint at unforeseen consequences of the policy for the structure of the banking market. Our

Zambia in 2012 this debate was only second order compared to the reductions in financial sector risk associated with the recapitalisations.

[112] Examining parameter heterogeneity across bank types within the previous distributed lag model becomes cumbersome because of the multiple lag parameters.

previous empirical setup, a distributed lag model, does not easily allow for such tests. Instead, we borrow an approach from Fang et al. (2020) and estimate:

$$\Delta y_{t,t+r}^{ij} = \alpha^{ij} + (\beta_1 + \beta_{type}.T^{type}).\frac{\Delta def_t^i}{TA_t^i} + \boldsymbol{\beta}'\boldsymbol{X}_t^i + \delta_t + \varepsilon_{it} \quad (12)$$

where $\Delta y_{t,t+r}^{ij}$ captures the growth rate for bank i for asset j, beginning at time t and ending at time t+r. $T^{type}$is a dummy indicator variable for the bank groups across which we want to test for parameter heterogeneity (we test for parameter equality for the equity shock variable between the two groups for each model (i.e. H0: $\beta_1 = \beta_{type}$). In this case β_1 represents the parameter for one bank coded as zero in β_{type}. The two group classifications we use for this analysis are: (i) foreign and domestic banks, based on recapitalisation requirements; and (ii) and bank size, with groups determined by the median of banks' recorded total assets before the recapitalisation (Annex 8 provides more detail on these categorisations, as well as the limited sample size underpinning this analysis). Bank-specific control variables, $\boldsymbol{\beta}'\mathbf{X}_t^i$, are measured at the beginning of the growth period to avoid spurious relationships (which would exist given that equity is directly related to the balance sheet ratios used as controls via the accounting identity). As with Fang et al. (2020), we use TWFE, so δ_t, is a time dummy. Note that equation 12 implicitly imposes, ad-hoc, a dynamic structure on the data, based on the choice growth period length (i.e., t+r).

In summary, we find tentative evidence that large banks and small banks reacted heterogeneously to forced recapitalisations. Between foreign and domestic banks, the evidence for heterogeneous reactions was not convincing, at least in the short term. In a policy sense, the results indicate that we should be aware that banks of different types might be able to react differently to forced capital injections, while acknowledging that we may be too poorly powered to pick up any effect. Table 12 provides the estimates from equation (12) for each asset-type dependent variable (total assets, loans, etc) and each group classification (i.e., large banks versus small banks and foreign vs domestic banks). Each column under each asset type captures an estimate for a different growth period, moving from t=0 and r=1 (the instantaneous effect) to t=0 and r=6. The rows with estimates provide the values for β_1 and β_{type} from equation (12) for the two sets of bank groups. The "Test of equality (p-value)" row provides results from testing the hypothesis H0: $\beta_1 = \beta_{type}$, with a low p-value

indicating that there is evidence that we can reject the null hypothesis of parameter equality[113]. If we can reject the null, there is evidence of heterogeneous reactions.

The evidence for heterogenous responses between large vs. small banks appears consistent. Large banks appear to have had much stronger asset growth reactions across most asset types and most periods. For example, in the period up to the fourth quarter after the equity injection, we estimate that large banks grew their total assets by 1.92 percentage points for every 1 percentage point increase in equity/total asset ratio due to the forced recapitalisation (column 3, row 2 of the coefficient estimates in Table 12). We estimate that smaller banks grew only 0.72 percentage points. The difference between the two is statistically significant with a p-value of 0.020. We also found a significant difference in Treasury security growth response, with large banks growing their holdings much more (see the "Treasury Securities" column block and the positive and statistically significant coefficients, as well as frequent rejection of H0: $\beta_1 = \beta_{\text{type}}$).

There is less consistent evidence that foreign banks and local banks responded heterogeneously. The second block of rows in Table 12 under the "ownership" row header provides the estimates and parameter equality tests. For example, six quarters after the recapitalisation, the growth rates for foreign banks' total assets and treasury assets were higher for foreign banks than domestic banks, although only significantly (and inconsistently so) for the Treasury assets (foreign bank growth effects were 0.91 percentage points and 2.49 percentage points for every 1 percentage point increase in equity/total asset ratio due to the forced recapitalisation, respectively, compared to 0.16 percentage points and -2.38 percentage points for domestic banks). Annex 9 provides estimates using winsorised data where the evidence is slightly more supportive of heterogeneous responses, perhaps because the data in the more volatile asset allocating banks are dampened. Using fixed effects with macro controls yields similar results (Table 37), as does estimating equation (12) without controls (Table 38).

In all of this work, however, the noisiness in the small/domestic bank parameters suggest that we should be cautious and only view the results as speculative, especially given the small sample sizes highlighted in Annex 8. While there is significant overlap between the size grouping and the ownership grouping, four banks out of 17 did change grouping, suggesting that the difference in

[113] Bolded p-values are cosmetic and simply indicate significance at the 10% level.

results we find between the two groupings is plausible. This is a reminder that we are underpowered and using a small sample.

This possibility of heterogeneous responses carries over to how banks' debt financing changed after their forced equity recapitalisations. Intuitively, examining banks' debt financing (and especially deposit-taking) post recapitalisation helps us better understand whether and how banks might have financed the asset growth we recorded in Table 12 (in addition to the forced equity injections). We also examine any differences in debt financing responses across bank types.

Perhaps unsurprisingly, given the heterogeneity we found in bank asset growth, we find that larger banks appear to have expanded their debt financing more than smaller banks. Given that most banks in Zambia rely on deposits for their debt financing, we also test specifically for differences in deposit growth reactions (during the sample period, only two banks in the dataset had median deposits of less than 80% of total debt financing, with the majority being much higher). Table 13 provides the estimates using equation (12) but considers banks' total debt financing growth (which includes loan and wholesale parent financing, etc.) and total deposit growth, rather than asset class growth. The table can be read in very much the same manner as Table 12. For example, in the year following a bank equity recapitalisation of 1 percentage point of total assets, large banks grew their deposit financing by an additional 1.57 percentage points (the fourth column within the second column group titled "Total Deposits" and in the second row titled Large under the "Large" grouping). This compares to an insignificant estimate of 0.23 percentage points for small banks. The difference between the two is significant (i.e., we can reject H0: $\beta_{\text{Large}} = \beta_{\text{Small}}$). A similar theme carries across different growth period lengths, debt financing types (total and deposit), and the two groups. It would have been interesting to have been able to investigate the more granular story of whether the source of this estimated heterogeneity in debt financing was driven by zero-sum deposit shifting between banks and bank types, or whether this deposit came from other margins (e.g., new deposits or increases in deposits from existing depositors).

Across both the asset class growth estimates and the debt financing estimates, these differences between groups suggest that the forced recapitalisation may have been associated with larger banks expanding, potentially taking a more significant role in the Zambian banking system. Of course, we have not shown here that this growth comes at the cost of smaller or domestic banks and our analysis

is in no way comprehensive or definitive (and suffers from the econometric and sample shortcomings previously mentioned). But, at the least, these results indicate that this type of differential reaction may be a feature of such policies and warrants discussion when policymakers are considering absolute capital requirements.

Table 12 – Heterogeneous Bank Reactions to the Recapitalisation Across Asset Types

Notes: The regressions in this table are based on equation (12) in the text and the approach is loosely borrowed from Fang et al. (2021). This table presents estimates from regressing the growth rates for different types of bank assets over different periods (the dependent variable, in percent) against a recapitalisation variable. The recapitalisation variable is the ratio of equity raised by a bank in deficit under the absolute minimum capital rule in a quarter divided by the bank's total assets at the beginning of that quarter. Columns show different dependent variables (and are asset-type and growth period-specific). The main rows containing coefficients provide the beta on the recapitalisaion variable for the different bank types. The "Test of equality (p-value)" row provides the p-value from testing the hypothesis H0:B1=B(Type), which tests whether the asset growth betas for the different types of banks are equal. Bolded p-values are significant at 10%. Each column represents a regression using a different asset type (the set includes the following: Total Assets, Bank Loans, Treasury Assets, and Other Assets) and a different growth period (for example, (0,4) is estimating the growth rate for the period t=0 to t=4, i.e. beginning at the start of the period in which the recapitalisation occurred, covering up until the beginning of the fourth quarter). Data is quarterly and was constructed from the Bank of Zambia's monthly Prudential Return. Bank-Specific controls include the following variables (with one lag): log of total bank assets, return on assets, liquid assets as a percent of total assets, RWAR ratio, a sector-specific loan measure, provisions for the quarter as a percent of beginning total assets. Driscoll Kraay Standard Errors were used and are provided in parentheses. Data are not winsorised. *** p < 0.01, ** p < 0.05, * p < 0.1.

	Dependent Variable: Growth Rate for Each Asset Type	**Total Assets**					**Total Loans**					**Treasury Securities**					**Other Assets**				
Group	**Growth Period (t,t+r)**	(0,1)	(0,2)	(0,3)	(0,4)	(0,6)	(0,1)	(0,2)	(0,3)	(0,4)	(0,6)	(0,1)	(0,2)	(0,3)	(0,4)	(0,6)	(0,1)	(0,2)	(0,3)	(0,4)	(0,6)
Large	Small	0.7707***	0.5584**	0.7232**	0.2929	0.0379	-0.7791**	-1.1417**	-1.5041**	-0.9435	-0.6612	1.5756	2.3876***	2.9827***	1.0892	-2.0041*	1.3289***	0.6652	0.7351*	0.2966	-1.2415*
		[0.1376]	[0.2447]	[0.3084]	[0.3593]	[0.4655]	[0.2955]	[0.5182]	[0.5786]	[0.5690]	[0.9997]	[1.1098]	[0.8083]	[1.0520]	[1.0788]	[1.1349]	[0.4123]	[0.6295]	[0.3982]	[0.4942]	[0.6986]
	Large	1.5922***	2.1850***	1.9224***	2.5595***	1.9993***	0.0434	-0.7041	-1.2146	-0.1094	-1.6254	4.6982***	6.4192***	6.5534***	5.5510***	5.2602***	1.8793***	1.6778***	0.5068	2.7651**	2.2565**
		[0.2026]	[0.4490]	[0.6814]	[0.7991]	[0.6114]	[0.6426]	[1.0779]	[1.3286]	[1.1006]	[1.9845]	[0.8196]	[0.8409]	[1.1996]	[1.5478]	[1.7970]	[0.3628]	[0.5096]	[0.7471]	[1.1110]	[1.0303]
	Test of equality (p -value)	**0.009**	**0.003**	**0.121**	**0.020**	**0.027**	0.204	0.691	0.787	0.376	0.547	**0.011**	**0.000**	**0.059**	**0.034**	**0.003**	0.252	0.299	0.807	0.109	**0.023**
	Obs.	729	713	697	681	649	729	713	678	646	598	704	689	673	658	628	729	713	697	681	649
	R2	0.26	0.32	0.41	0.47	0.54	0.17	0.23	0.30	0.28	0.28	0.19	0.28	0.26	0.27	0.3	0.21	0.25	0.28	0.32	0.36
	Bank-Specific Controls	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√
	Bank Fixed Effect	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√
	Time Fixed Effects	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√
	Macro Controls	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Ownership	Domestic	0.8062***	0.7175**	0.8470**	0.567	0.1581	-0.8666***	-1.1259*	-1.6380**	-1.0553	-0.8402	2.2658*	2.2851**	2.9440**	1.3905	-2.382	1.1459**	1.0415	1.0304**	0.8473***	-1.1765
		[0.1266]	[0.3106]	[0.3679]	[0.4177]	[0.5302]	[0.2705]	[0.6182]	[0.6442]	[0.6809]	[1.0747]	[1.3289]	[0.8606]	[1.1834]	[1.3828]	[1.5724]	[0.5591]	[0.7193]	[0.4181]	[0.3048]	[0.7246]
	Foreign	1.1565***	1.1613**	1.1568***	1.0398	0.9162	-0.1685	-0.9303	-1.1005	-0.2786	-0.9029	2.1743	4.7540***	4.9975***	3.0478**	2.4983	1.9645***	0.5356	0.0722	0.6516	0.5703
		[0.2638]	[0.5562]	[0.3301]	[0.7113]	[0.6273]	[0.4776]	[0.8733]	[1.0369]	[0.8273]	[1.6213]	[1.7999]	[1.1467]	[0.9092]	[1.2651]	[1.9042]	[0.3109]	[0.7825]	[0.5786]	[0.8554]	[0.9772]
	Test of equality (p -value)	0.178	0.461	0.500	0.566	0.314	0.224	0.858	0.600	0.407	0.967	0.967	**0.073**	0.163	0.360	**0.034**	0.153	0.623	0.121	0.847	0.209
	Obs.	729	713	697	681	649	729	713	678	646	598	704	689	673	658	628	729	713	697	681	649
	R2	0.25	0.32	0.40	0.46	0.54	0.17	0.23	0.30	0.28	0.28	0.18	0.27	0.26	0.27	0.3	0.22	0.25	0.28	0.32	0.35
	Bank-Specific Controls	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√
	Bank Fixed Effect	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√
	Time Fixed Effects	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√
	Macro Controls	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X

Table 13 – Heterogeneous Bank Debt Financing Reactions to the Recapitalisation

Notes: The regressions in this table are based on equation (12) in the text and the approach is loosely borrowed from Fang et al. (2021). This table presents estimates from regressing the growth rates for total bank financing growth and total bank deposit growth over different periods (the dependent variable, in percent) against a recapitalisation variable. The recapitalisation variable is the ratio of equity raised by a bank in deficit under the absolute minimum capital rule in a quarter divided by the bank's total assets at the beginning of that quarter. Columns show different dependent variables (and are asset-type and growth period-specific). The main rows containing coefficients provide the beta on the recapitalization variable for the different bank types. The "Test of equality (p-value)" row provides the p-value from testing the hypothesis H0:B1=B(Type), which tests whether the debt financing growth betas for the different types of banks are equal. Bolded p-values are significant at 10%. Each column represents a regression using one of the two debt financing variables and a different growth period (for example, (0,4) is estimating the growth rate for the period t=0 to t=4, i.e. beginning at the start of the period in which the recapitalisation occurred, covering up until the beginning of the fourth quarter). Data is quarterly and was constructed from the Bank of Zambia's monthly Prudential Return. Bank-Specific controls include the following variables (with one lag): log of total bank assets, return on assets, liquid assets as a percent of total assets, RWAR ratio, a sector-specific loan measure, provisions for the quarter as a percent of beginning total assets. Driscoll Kraay Standard Errors were used and are provided in parentheses. Data are not winsorised. *** p < 0.01, ** p < 0.05, * p < 0.1.

	Dependent Variable	**Total Debt Financing Growth**					**Total Deposit Financing Growth**				
Group	Growth (t,t+r)	(0,1)	(0,2)	(0,3)	(0,4)	(0,6)	(0,1)	(0,2)	(0,3)	(0,4)	(0,6)
Large	Small	-0.1287	-0.1829	0.3265	0.2299	-0.1138	-0.1811	-0.1325	0.1991	0.0303	-0.4754
		[0.1686]	[0.2901]	[0.4040]	[0.4324]	[0.4877]	[0.1345]	[0.2936]	[0.3327]	[0.4157]	[0.5698]
	Large	0.5865***	1.4299***	1.2660***	1.9015***	1.9719***	0.9416***	1.3679***	1.5926***	2.1936***	2.5148***
		[0.1241]	[0.2908]	[0.3615]	[0.5068]	[0.7031]	[0.1795]	[0.3263]	[0.3297]	[0.5551]	[0.6409]
	Test of equality (p -value)	**0.002**	**0.000**	**0.005**	**0.004**	**0.005**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**
	Obs.	729	713	681	649	601	729	713	681	649	601
	R2	0.34	0.42	0.42	0.38	0.44	0.36	0.42	0.42	0.39	0.45
	Bank-Specific Controls	√	√	√	√	√	√	√	√	√	√
	Bank Fixed Effect	√	√	√	√	√	√	√	√	√	√
	Time Fixed Effects	√	√	√	√	√	√	√	√	√	√
	Macro Controls	X	X	X	X	X	X	X	X	X	X
Ownership	Domestic	-0.0984	-0.0608	0.3471	0.3462	-0.2385	-0.1793	-0.2249	0.1051	-0.0419	-0.9258***
		[0.1664]	[0.3756]	[0.4691]	[0.5414]	[0.5350]	[0.1560]	[0.2924]	[0.2252]	[0.4378]	[0.3284]
	Foreign	0.208	0.480	0.8067***	0.9416*	1.2422**	0.4313	0.8595***	1.1373***	1.3547***	1.9126***
		[0.2701]	[0.5098]	[0.2154]	[0.5494]	[0.5830]	[0.4112]	[0.2873]	[0.2867]	[0.4041]	[0.4955]
	Test of equality (p -value)	0.242	0.397	0.281	0.452	**0.031**	0.151	**0.002**	**0.000**	**0.023**	**0.000**
	Obs.	729	713	681	649	601	729	713	681	649	601
	R2	0.34	0.41	0.42	0.38	0.44	0.35	0.42	0.42	0.39	0.45
	Bank-Specific Controls	√	√	√	√	√	√	√	√	√	√
	Bank Fixed Effect	√	√	√	√	√	√	√	√	√	√
	Time Fixed Effects	√	√	√	√	√	√	√	√	√	√
	Macro Controls	X	X	X	X	X	X	X	X	X	X

Market Structure Trends

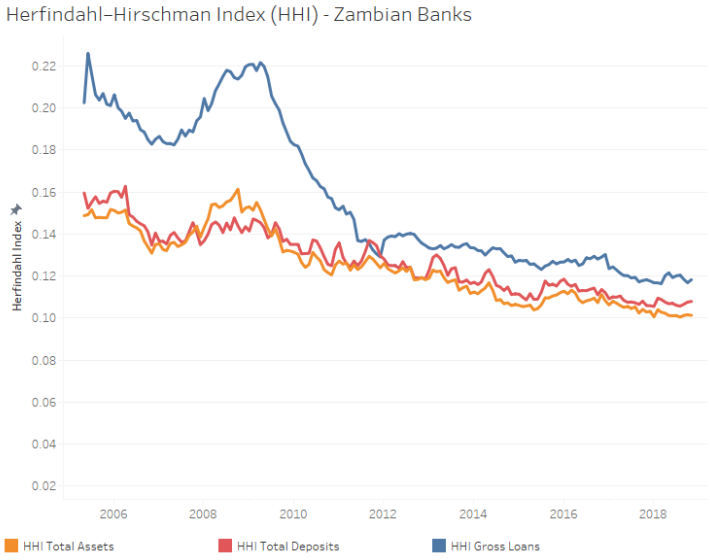
While this heterogeneity analysis indicates that larger banks potentially adjusted more to the IRR regime per unit of recapitalisation, our analysis can only comment on relatively short-term and partial reactions. To understand how, in a narrative sense, Zambia's banking market structure evolved after the recapitalisation, we now examine evidence for changes in the market more broadly.

Interestingly, at a casual glance the recapitalisations do not appear to have had a decisive impact on Zambia's banking market structure. Measures of market concentration for Zambia's banking sector have been slowly trending downwards since the GFC, and there are no clear changes in the aggregate series around the recapitalisations. For example, analysis using the Herfindahl–Hirschman index[114]

[114] We calculate the Herfindahl–Hirschman index for each market, j, as, $HH_t^j = \sum_{i=1}^{N} \left(\frac{Asset_{it}^j}{\sum_{i=1}^{N} Asset_{it}^j} \right)^2$. The HHI is a measure of the size of firms in relation to the industry/market they operate in and is an indicator of the amount of competition among firms.

for Zambia's banks' total assets, total deposits, and total loans, show the index trending downward, somewhat undeterred, towards increasing levels of competition and disbursed market shares (Figure 13). Given the weightings in the Herfindahl–Hirschman index (HHI) and the large number of factors that can drive market shares, it is hard to know exactly what lies behind this trend. But at a high level the recapitalisations do not appear to have had a sharp or immediate impact on market structure as measured by the HHI.

Figure 13 – Herfindahl-Hirschman Index for Zambian Banks



Note: The figure provides monthly Herfindahl–Hirschman index calculations for several banking markets in Zambia.

Examining the market shares of different bank size groups provides a slightly more granular picture, with medium sized banks appearing to grow, not by taking market share from smaller banks, but at the expense of the largest incumbents.

We first break the banks into four groups (the largest nine banks in groups of three, sorted by decreasing size, and then one group of the smaller, remaining banks) and track the groups' market share changes over time (Figure 14). From 2012, there is a slow but steady trend for the smallest of the large banks to gain total bank asset market share.

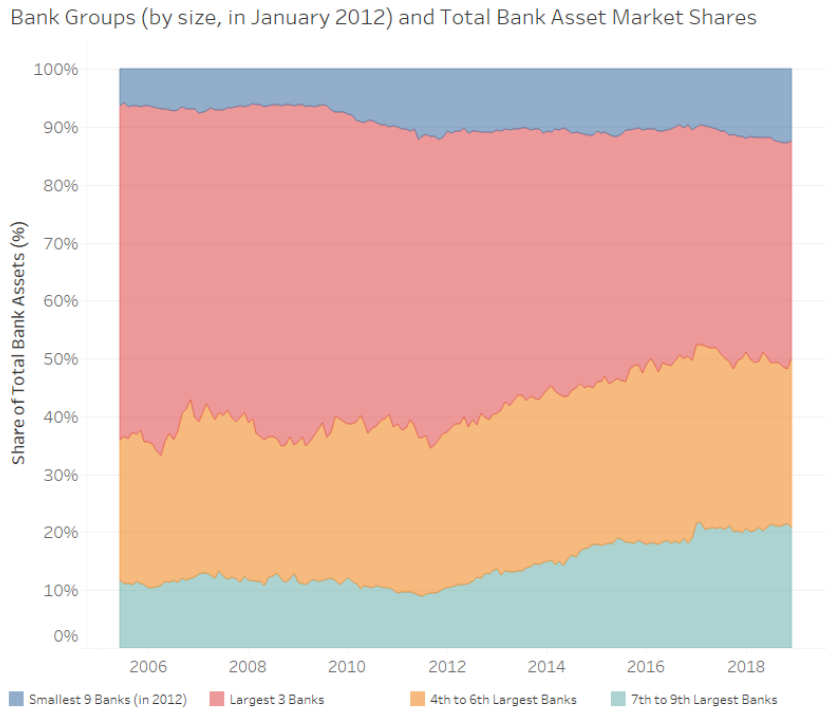
A generous interpretation might be that the recapitalisations helped encourage some middle-sized banks to increase their foothold in the Zambian market and improve competition with the largest incumbents. It was these medium sized-internal banks, after all, who underwent the largest absolute

recapitalisations (as the largest three banks were essentially almost capitalised). But clearly further research would be needed to argue such a claim, as these banks may have increased their Zambia market share in any case, independent of the recapitalisations. We find a similar story to Figure 14 when we analyse Zambia's loan market and deposit market shares – large incumbents appear to be losing market share to smaller (but still reasonably sized) retail banks.

Note that Zambia's banking markets grew steadily in real terms from 2010 onwards. We can therefore know that at least some of the medium-sized banks' growth will have been into new lending, deposits, and securities, rather than cannibalising existing loan demand, and deposit and treasury security supply.

Taken together this analysis might allay some of our previous policy concerns that larger bank recapitalisations in Zambia were to the detriment of smaller lenders, at least over longer horizons. It is also not inconsistent with some of the previous heterogeneity results, in that there may be a smaller group of medium-sized "large" banks that drove much of the regulatory variation and growth, and perhaps at a greater rate than the smaller banks. In saying this, our quantitative results cover only a short time horizon and we do not have enough data and power here to properly understand these dynamics.

Figure 14 – Total Bank Asset Market Shares by Bank Size Groups



Note: Figure shows bank group Total Bank Asset market shares over time. Banks are grouped by Total Assets share in January 2012 into four groups: largest three, 4th-6th, 7th-9th, and the smallest 9 banks.

Market Structure and differences between Nigeria and Zambia's Recapitalizations

Nigeria also underwent its own absolute capital banking reforms in 2004, with Charles Soludo, Nigeria's Central Bank Governor, implementing a radical set of reforms. Yet there is an interesting and somewhat stark difference between market structure outcomes following the reforms: in Zambia there were no direct bank mergers or failures associated with recapitalisations, whereas in Nigeria, the number of banks fell from 89 to 25, amidst a series of mergers and bank failures. Standard market structure measures also highlight this difference: Zambia's HHI changed very little (reducing slightly) post the banks' recapitalisations, while in Nigeria, CBN (2013) estimates that HHI increased from 0.0486 in 2003 to 0.0808 in 2006 (after the reforms were complete).

There are good reasons for these two experiences being different, even if the recapitalisations initially appear similar. For policymakers, understanding the reasons and context around these two experiences (and why they might link to different outcomes) is clearly important.

The clearest reason is due to the policies being different in form and magnitude. In Zambia, the recapitalisations were two tiered: banks who wanted to have more than 51% ownership overseas were required to raise regulatory capital to approximately USD 100m, while for majority domestic owned banks, the constraint was only USD 20m. This meant that smaller banks could still raise capital, albeit painfully, and do so without consolidating. In Nigeria, there was a single recapitalisation requirement of approximately USD 190 for all banks (Cook, 2011). This meant that small banks (of which there were many, most banks in Nigeria had capital less than USD 10m (Soludo, 2004)), had to either raise capital, merge, or lose their license. Because few banks in Nigeria had foreign parents (Claessens and van Horen,2012), accessing foreign equity capital was also more challenging[115].

These differences in policy arose because the rationales for each recapitalisation were fundamentally different. In Zambia, BOZ aimed to improve bank capitalisation rates and increase bank lending directly because of the recapitalisation (as discussed in detail earlier in this chapter). In Nigeria, the

[115] In saying this, Nigeria's stock market is more liquid and vibrant than Zambia's, so in a comparative sense some Nigerian banks did have this option for sourcing additional capital; yet the market did not have the capacity to provide capital to all banks (Ogowewo and Uche, 2006).

recapitalisation was designed specifically to reform the banking market's structure, with the idea that bank consolidation would improve market efficiency, capitalisation, and market functioning, and remove uncompetitive banks (Soludo, 2004). Zambia's market was already quite consolidated before the recapitalisations; the five largest banks in Zambia held nearly 74% of banking sector assets, and even more so for deposits. In Nigeria, the tail of small banks in the banking market was much larger. The 22 largest banks held only 67% of the market share in 2004, indicating far higher levels of asset dispersion. Nigeria's policy also had a strategic goal: to better position its banks to compete with international banks in the future (Soludo, 2004).

The enforcement of the regulation in each country was also quite different, and this almost certainly contributed to the different market structure outcomes we see in the data. Zambia's enforcement of non-complying banks was relatively lenient, with undercapitalisation resulting in a stronger form of BOZ oversight and some loss of operating independence. But undercapitalisation in January 2014 did not result in an immediate license withdrawal or restrictions on deposit-taking. Several Zambian banks did take longer than the specified time to recapitalise, but they did not cease operating and were not liquidated. It is also not clear that BOZ would have had the political and popular support to withdraw licenses for banks that were not clearly failing (and were merely operating under the limit). Nigeria's enforcement, on the other hand, was far harsher, following a policy of "zero tolerance in the regulatory framework" (see p.5 in Soludo (2004)). Indeed, in January 2016, directly after the 18-month recapitalisation limit was up, CBN announced the removal of 11 directors of unsuccessfully recapitalised banks, leading to a sale (and consolidation) of those banks' assets (Ogowewo and Uche, 2006). These different enforcement approaches again reflect the difference in policy objectives and the assumed economic mechanisms via which the policies were supposed to operate. Zambia was not aiming for market consolidation, so was likely more willing to tolerate some non-compliance before removing banking licences.

In summary, the differences in market structure outcomes between Zambia and Nigeria do make some sense when we consider the form, objectives, and enforcement of the two policies, as well as the background market context in which they took place. If Zambia had implemented only the one higher capitalisation requirement that was strictly enforced, there very likely would have been more pressure for consolidation, especially among the 9 small banks in Zambia that together made up only 12% of total bank assets in January 2012.

Discussion and Conclusion

Taken as a whole, our findings suggest that Zambia's absolute capital policy was broadly successful at improving bank capitalisation, even if there wasn't strong evidence the policy improved bank lending in the short term. Importantly, the results provide an example of a regulator successfully recapitalising its banking system in a way that did not lead to significant bank deleveraging. This would have been a significant risk had it relied on adjusting its Basel-based RWAR upward. Our findings also provide a piece of evidence on an under-researched capital policy tool that may be of interest to LDFS policymakers facing a similar regulatory context to Zambia in 2012 (i.e., perceiving their banks as being too risky despite their banks sitting above the Basel-inspired RWAR rule). In such circumstances, absolute minimum capital regulations may provide an additional tool for managing the trade-offs between financial sector and development objectives, rather than relying only on Basel-inspired RWARs. Our results suggest, however, that even large capital injections were not strongly associated with increased bank lending on average, although our estimates are clearly only short-term and unable to capture how the capital may have influenced banks' lending in the longer term. This is an important caveat because banks' strategies and investments that might be needed to enable large lending increases (such as the branch network), may be a constraint in the short term.

More broadly, bank regulators in LDFSs should be encouraged by international standard setters to think not just in terms of "proportional implementation" of Basel, but also in terms of implementing other policies that are designed for their context. Such policies, appropriately calibrated and quite possibly implemented in conjunction with Basel, might help LDFSs better match their multiple stability and development objectives. At the least, there could be more debate and research on policy options from the LDFSs' perspective.

Our results do not suggest that absolute minimum capital rules are a policy tool for all contexts and at all times (indeed, quite the opposite). There is a clear external validity caveat associated with our findings: Zambia's recapitalisation experience and the associated outcomes must be seen as being contingent on the fair-weather economic conditions facing Zambia at the time, both in terms of economic performance and the relatively strong position of the Government's balance sheet. Such conditions likely kept the value of bank charters and investment willingness high, making it easier

for banks to raise capital. In addition, strong domestic Treasury returns may have made international banks more willing to leave assets in-country. There was also some leniency in BOZ's treatment of non-complying banks after the 2014 deadline (for example, not automatically withdrawing licenses after the 2014 deadline if a bank was not compliant), possibly reducing the likelihood of bank mergers or bank exits. All of these factors, as well as other differences in banking-sector context, may help explain why there were limited bank closures or mergers in Zambia, especially compared to other counties that have had recapitalisation policies but have seen different results (in Nigeria, for example, the number of banks fell from 89 in 2004 before its absolute capital policy to 25 afterwards (Ifeanyi Attama and Nfor Yuni (2021)).

In addition to these issues regarding external validity, there are also some dimensions of caution for policymakers looking to implement minimum absolute capital requirements. Our findings suggest that LDFS regulators need to be aware of how large changes in absolute minimum capital policies might also impact their domestic Treasury market and potentially support government financing; we found a stronger link between the recapitalisation policy and growth in Treasury securities than between the policy and lending growth. The heterogenous bank reactions we record in the final section also suggest that this policy was associated with differential growth across bank types, resulting in the big banks getting bigger and small banks, relatively smaller. LDFS countries implementing large changes in minimum absolute capital rules should thus be aware that their policies may have consequences for the structure of their banking markets. Note that our results rely on bank-level estimation and so cannot directly infer implications on the market in aggregate; for example, the impact the regulation had on, say, bank competition. Nevertheless, this is a clear avenue for future work.

Since the imposition of the capital limits and banks' recapitalisation, Zambia's economy has experienced several negative macroeconomic episodes: first a currency and inflation crisis in 2015/2016, then a sovereign debt crisis from 2017 that is still ongoing, and now COVID-19[116]. While we do not have rigorous quantitative evidence for this, we suspect it is likely that the recapitalisation policy has played at least some role in providing additional buffers for the banking system and has helped ensure that banks could weather recent economic crises. Since the recapitalisation policy,

[116] These challenges are reflected in Zambia's headline growth figures: from 2005-2013 economic growth averaged 7.6%; from 2014 to 2019, it declined to an average of 3.3%; in 2020, growth was negative.

only one bank, Intermarket Bank, has failed, despite the increased pressure on the banking system.[117]. Future research might build a richer counterfactual narrative for Zambia's banks, attempting to piece together how they may have fared without the capital injections.

[117] Aggregate non-performing loans (as a percent of total loans) in Zambia increased from 6% in 2012 to over 11% in 2019.

Interest rate restrictions in lower-income countries: Lessons from Zambia and the Microfinance Information Exchange database (MiX)

The objective of the next four chapters is to investigate the effects of interest rate restrictions (IRRs) on financial market outcomes, particularly in Zambia but also other lower-income countries. It consists of three empirical chapters, each addressing a different research question on IRRs. In Chapter 5, we examine how Zambia's IRR, implemented in 2013, impacted non-bank lending. Chapter 6 constructs a cross-country dataset of Microfinance Institutions (MFI) IRRs and tests whether these regimes were associated with poorer lending and self-sufficiency outcomes for MFIs. Chapter 7 assesses how Zambia's IRR may have affected the cost of financial intermediation for banks in Zambia. The conclusion provides a policy discussion. Prior to delving into the analytical chapters, Chapter 4 presents the context and background for interest rate policies, as well as summarizing the results and contributions of the section.

Chapter 4 - Introduction to Interest Rate Restrictions and Summary Results

Background context on interest rate restrictions and their recent resurgence

Interest rate restrictions – or usury laws – are neither new nor uncommon. Since antiquity, sovereigns and religious authorities have been limiting the legal interest rate in one form or another. The first record of interest caps or restrictions dates back to Hammurabi, an ancient Babylonian king, who, in around 1800 BCE, set the maximum interest rate in his kingdom for lending grain at 33% per annum; for lending silver, at 20% (Sylla and Homer, 1991). The Romans were also not partial to excessive interest, setting their first maximum interest rate at one ounce of silver per pound of silver lent per annum, or approximately 8.3%. In more recent history, Adam Smith and Jeremy Bentham famously argued – Smith for, Bentham against – the necessity of usury laws (Jadlow, 1977), and since independence, many American states have imposed some form of usury law on consumer lending (Peterson, 2008). In the European Union, 21 of 28 countries had some form of restriction on lending rates in 2017 (Faherty et al. 2017).

Why are interest rate restrictions so ubiquitous? Governments have typically justified interest rate caps or restrictions based on one of two arguments: the first is to protect consumers from usury – the act of charging exorbitant, unfair, or predatory interest rates; the second is to lower the cost of credit and improve credit access. Historically, caps were mostly used by governments to limit usury. The more recent trend is for higher-income countries to impose usury limits at high rates to protect consumers while lower-income countries more frequently use rate caps as a structural policy tool to lower the price of credit more broadly (Ferrari et al., 2018). It is the latter that we are predominantly concerned with in this thesis[118].

[118] The different objectives of interest rate policy result in divergent regulatory designs for IRRs. Ferrari et al. (2018) provide a useful taxonomy of interest rate caps, highlighting five key design characteristics: (i) the scope of the caps – i.e. whether the caps apply to one small market segment (e.g. payday loans), to a specific product or provider type, or to the whole market; (ii) the number of caps in the market; (iii) whether the cap is an absolute cap or is linked to some benchmark interest rate, and which rate it might be linked to (central bank rate or market average rate); (iv) how the cap is calculated – often as either a fixed spread or a multiple of a benchmark rate; (v) and whether the cap includes fees.

Many lower- and lower-middle income countries in SSA – our primary interest – also have a history of employing IRRs. In the 1950s-1970s, many governments in SSA ran tightly controlled interest rate regimes, as did many other governments around the world. These IRR policies were often one component of more extensive financially repressive regimes[119]. From the late 1980s until the early 2010s, however, there was a broad trend towards financial sector liberalisation in many lower-income countries. Notably, this often included a policy shift towards market-based loan rate determination. Over this period, many LICs and MICs de-regulated lending interest rates, while at the same time loosening their capital accounts, auctioning treasury bills, and exposing their exchange rates to market pressures (William and Mahar, 1998; Galbis, 1993). Across East Africa, for example, many countries liberalised their interest rates during the broader push for economic reforms in the 1980s and 1990s (for example, Malawi in 1987, Zambia in 1992, Tanzania in 1994, Uganda in 1993, and Rwanda in 1995). But, again, this was a global shift rather than uniquely an African phenomenon and was often supported by international institutions' recommendations (for example, the reforms in Jamaica in 1981, Mexico in 1989, Haiti in 1995, and Samoa in 1998).

In the last decade, however, this trend towards interest rate liberalisation has begun to reverse, with several countries having re-introduced lending IRRs. The data clearly shows a gradual unwinding of the shift toward interest rate liberalisation and market-determined rates[120]. Ferrari et al. (2018), for example, found that at least 30 emerging or developing market economies have introduced some form of IRR since 2011, and Calice et al. (2020) found that more than 50% of high-income countries had some form of IRR in 2020. Indeed, our own newly constructed dataset on IRRs for this thesis confirms this trend for MFIs around the world – we find a substantial increase in the proportion of countries regulating lending rates in their microfinance sector between 2002 and 2018.

Why are we seeing this shift back towards interest rate repression in LICs? One explanation of the reversal towards IRRs is that it is a symptom of deeper structural challenges and tensions that have emerged from economic reforms in the 1980s-1990s. A decade or more after financial liberalisation, nominal lending rates in many of the de-liberalised countries remain extremely high despite being market-determined. This often reflects the fact that deeper structural market reforms which drive the

[119] IRRs were not unique to countries in SSA before the 1980s – many governments around the world were running such regimes at the time. William and Mahar (1998) provide a survey of financial sector liberalisation across a sample of both high-income and low-income countries.
[120] The IRRs are often imposed even if some components of the liberalisation reforms, such as Treasury auctions, persist.

market to provide intrinsically lower lending rates and solve information imperfections have not been achieved, and that macroeconomic conditions remain volatile (Feyen and Huertas, 2020)[121]. The recent resurgence in political appetite for lower lending rates – from citizens and politicians alike – is thus understandable. Interest rate caps and restrictions are the easiest and most visible means for policymakers to effect change.

The problem with this reversion towards IRRs – a problem this section aims to contribute understanding – is that not enough is known about how IRRs may impact financial market outcomes, especially in LICs and SSA. While the short-term political benefits of lower headline lending rates are clear, the consequences for the health and breadth of the credit markets are unfortunately less so. One issue facing policymakers is that there is not a clear theoretical prior that outlines the impacts of lower headline lending rates, and especially so in lower-income settings where market imperfections are likely more pervasive. For example, a key concern international organisations have with interest rate caps, especially in developing markets, is that they may inadvertently reduce loan supply and harm credit access (see Ferrari et al. (2018), Maimbo and Gallegos (2014)). Indeed, a simple static analysis of a price restriction in a competitive credit market might predict that lending volumes should reduce following an interest rate cap being imposed. However, this result can be easily contested under alternative assumptions of lender pricing power, predatory pricing, moral hazard, or when lenders can avoid regulations by changing their business models to target non-interest loan-related income streams (such as fees). In such a theoretically contingent and ambiguous setting, empirical analysis is essential in supporting policymaking.

Fortunately, in recent years there has been some empirical progress made in investigating the effects of IRRs, mainly in the form of quantitative country case studies examining effects on bank or MFI lending (see, Cubillos-Rocha et al. (2018); Romero (2021); Aiba et al. (2020); Heng et al. (2021); Roa et al. (2021)). Yet there remains a significant research gap in the SSA-specific context where there have been very few quantitative studies and much quantitative analysis has been blunt, modelling IRR regimes as a binary treatment effect. Moreover, for countries without any domestic

[121] In a credit market context where information exchanges are less efficient (because of a lack of a well-functioning collateral registry or credit bureau), bankruptcy proceedings remain backlogged, and average loan sizes remain small (loan and collateral cost economics are hence poor), it is perhaps unsurprising that credit markets will continue to clear at high nominal lending rates.

market-specific case studies, there is also a lack of cross-country quantitative evidence that might provide at least some indication of the effects of interest rate restrictions.

Given the increasing number of LICs implementing IRRs and the fact that the theoretical underpinnings of IRRs are context-dependent, there is a clear need for market-specific empirical research to help guide policymakers' decisions. In addition, there is a vast literature arguing that financial development is a critical part of a country's economic development process (Levine (2004), Papaioannou (2007), and Beck (2012) summarise the core literature), and policymakers need to impose IRRs with an awareness of how such policies might (or might not) contribute to such broader financial development. This section of the thesis looks to contribute to the academic literature on IRRs, strengthening the evidence base that policymakers can use to support future decision making (or at least those in Zambia or in other countries with similar financial contexts).

Section contributions and summary results

This section makes three contributions towards the literature regarding IRRs' potential impacts on financial market outcomes. The first contribution, made in Chapter 5, is to provide the first in-depth quantitative analysis of the implementation of an IRR on non-banks/MFIs in a SSA context. We do this by examining Zambia's interest rate cap regime, which was implemented in 2013 and then removed in 2015[122]. Zambia's IRR is an excellent candidate for empirical analysis because, as we later argue, the imposition of the policy was likely exogenous to contemporaneous macroeconomic conditions. We are also fortunate to have had access to confidential, raw financial returns for the entire financial system, which we have used to construct a detailed panel database. Such data is rare in a SSA context. Our first case study examines the impact of the cap on Zambia's microfinance sector, employing a panel dataset containing every formal non-bank lender in Zambia. The analysis uses both distributed lag and event study models. Because our newly constructed dataset covers both interest rate data and detailed financial statements, we can more accurately frame and measure how different institutions were differentially affected by the cap, something not typically done in the literature. Detailed data on lending fees also helps us track and analyze how the caps influence lenders' pricing and fee practices, again providing a valuable contribution to the literature.

[122] Zambia's IRR regime applied across both the microfinance and banking sectors, albeit differentially. The cap on the banking sector was initially set at 18% and then loosened to 28%. The caps of MFIs and other non-banks were set at 42% and 30% respectively, and then loosened to 64% and 46%. The regime was implemented in 2013 and removed in late 2015.

The section's second contribution, in Chapter 6, is to provide a broader view of MFI interest rate cap regimes around the world, offering descriptive cross-country analysis of IRRs and MFI outcomes. To start, we use secondary sources to build a new longitudinal dataset tracking microfinance IRRs across countries. We then link this to a recently released World Bank dataset on MFIs – the Microfinance Information Exchange (MiX). We use the dataset to provide some non-causal descriptive statistics examining lenders bound by IRRs, for example, whether lenders that were bound by IRRs grew their lending more slowly than those that were not. We analyze the association between IRRs and MFI outcomes across lending growth (by value and number) and operational self-sufficiency. Data coverage and annual data frequency do place caveats and limits on the creditability of the analysis.

Our final case study in Chapter 7 moves away from an analysis of microfinance lenders. Instead, it examines the impact of Zambia's IRR on its banking sector, specifically concentrating on whether the cap might have influenced banks' costs of financial intermediation and pricing/income structures[123]. As with Chapter 5, we rely on a newly constructed and detailed bank panel dataset covering all commercial banks in Zambia's credit market. Given the persistence in banks' intermediation costs, we employ an Autoregressive Distributed Lag Model (ARDL) approach.

To results. The overarching finding of this section is that policymakers should seriously consider the potential downsides of implementing binding IRRs, supporting the recent findings from the empirical literature[124]. We find evidence that (i) imposing caps risks both adverse short-term effects on MFI loan growth, as well as increasing the opacity of pricing practices within the market, and (ii) that an IRR may have heterogenous effects on lenders depending on how bound they were by the cap (with more bound lenders seeing more substantial effects). However, there is also nuance in these findings: we do not find that the caps significantly affected MFIs or banks that price below or near an IRR.

Some of our stronger results derive from our Zambia MFI case study in Chapter 5. The case study shows that Zambia's IRR likely had a material impact on non-banks, especially those previously lending well above the IRR. Specifically, we estimate that a lender pricing 35% above the IRR

[123] Financial intermediation is the process by which a financial system pools and transfers excess funds from deposits to borrowers. The efficiency of financial intermediation is important because it is a driver of the depth, breadth, and accessibility of the banking system.

[124] Note: we make no claim that under certain conditions, such as when an IRR is imposed at a high interest rate or when lenders wield pricing power, that IRRs will not have positive effects on lender outcomes.

ceiling (approximately the median-priced lender in the market) reduced growth in their net outstanding loan book by a cumulative 72% in response to the imposition of the cap compared to what we would have expected such a lender to do otherwise. Compared to lenders unaffected by the cap (i.e., those pricing below the IRR ceiling), using an event study design we estimate that the median lender reduced its net outstanding loan book by 53% in the four quarters after the cap was imposed. The IRRs also appear to have had a non-trivial impact on the pricing structure within the market, encouraging a strong and persistent shift away from interest income to fee and commission income. We estimate that, on average, the caps were associated with a cumulative 20 percentage point increase in the fee share of lenders' income after five quarters, unlikely a positive development for market pricing transparency. Our analytical results align with the anecdotal and qualitative evidence that the caps were a painful shock for the market, as well as the fact that in the quarter after the cap was imposed four lenders handed in their lending licenses to BOZ.

Our cross-country analysis of IRRs on MFIs in Chapter 6 broadly supports our Zambian findings. We find a negative association between an IRR regime being in place and MFI loan growth rates. For example, an MFI bound by an IRR saw its loan portfolio growth rate reduce by 13.8 percentage points compared to when it was not bound. Likewise lending grow rates by the number of outstanding loans also reduced (by 10.9 percentage points). Further research might be able to unpick a more causal effect of IRRs in a cross-country setting.

Zambia's banking system appears to have been less affected by the IRRs. Although we do not test for changes in lending volumes, we do not find convincing relationships between the imposition of Zambia's IRR and either the cost of bank intermediation or banks' loan pricing/income strategies. We suspect this is because the IRR placed on the banking system was less severe for most banks than for the MFI market, although the absence of evidence is not evidence of absence. We also did not find evidence that banks increased their fees in the wake of the cap, again perhaps suggesting that banks were not strongly bound.

Our data, analysis, and findings do infer some policy recommendations. The final part of this section briefly explores these. Our primary suggestion is for policymakers to take a diagnostic and granular approach to setting IRRs in less-developed financial systems. Such an approach might try to marry both high-level market competition analysis with a granular approach to understanding the interest

rate distribution and business models within the credit market (and better predict which market segments will not be viable under a given cap).

The rest of this chapter discusses the literature that is relevant to the three core chapters of this section, as well as a more explicit demonstration of where the results of this thesis fit within the IRR literature more broadly.

Relevant Literature

The three chapters in this section draw from, and relate to, a common literature on IRRs. Each chapter will necessarily have specific contextual and methodological references; however, some research is relevant to them all. Such research, and our contributions to it, are discussed here.

Our analysis and findings nest within three related works of literature examining IRRs. The core contributions of the section lie closest to an empirical literature examining the impact of interest rate caps on financial market outcomes in lower-income settings. In recent years, there have been several strong contributions to this literature, many of which find evidence that stronger or newly imposed IRR policies have a repressive effect on market access (see Cubillos-Rocha et al. (2021); Romero (2021); Aiba et al. (2020); Heng et al. (2021); and Roa et al. (2021)[125]). Several non-quantitative case studies or policy papers also find similar results (Safavian and Zia (2018); Alper et al. (2019); Central Bank of Kenya (2018); Heng et al. (2015); and Ferrari et al. (2019)). However, not all empirical work supports IRRs having repressive effects. By examining the imposition and removal in 2009 and 2011 of Bangladesh's 13% cap on business lending, Miyauchi (2019) found that credit

[125] Cubillos-Rocha et al. (2021) and Romero (2021) examine the loosening of Columbia's microfinance rate from 20% to 33% in 2007, with both finding evidence that the loosening policy encouraged and facilitated financial market access. Aiba et al. (2020) use credit bureau data covering the universe of regulated loan accounts in Cambodia to investigate the effects of an 18% cap imposed in 2017 and find that both deposit and non-deposit taking MFIs adjusted to the interest rate cap policy by increasing average loan sizes and decreasing the number of newly disbursed loans, while they kept increasing the number of borrowers per loan account (a finding corroborated by Heng et al. (2021)). In 2014 Bolivia introduced one of the lowest IRRs for a low-income country: a cap of 11.5% for micro and SME credits by MFIs, likely a highly binding policy given the high cost of lending. Roa et al. (2021) examine Bolivia's policy in a two-way fixed effects setup and find strong evidence of credit rationing associated with the cap, with the number and value of loans reducing, and the average value of loans increasing. Cuesta and Sepũlveda (2021) and Maderia (2019) examine a large decrease in Chile's interest rate regulation for small consumer loans between 2013 and 2015. Cuesta and Sepũlveda (2021) use individual-level administrative data in a difference-in-difference set up (with treatment groups aggregated in bins according to loan size and borrower risk and exploit policy variation across loan size). They find that policy change improved the borrower risk pool, but both the number of loans and value of lending were deterred by lower caps. Maderia (2019) uses a representative household survey and a RDD design to compare consumers with risk-adjusted interest rates slightly above and slightly below the legal interest rate ceiling. He finds that being above the interest rate cap reduces the probability of credit access.

supply increased, mostly due to expansion on the extensive margin (i.e, the number of loans disbursed)[126].

Chapter 5 adds to this literature by providing an additional datapoint/case study supporting the idea that IRRs may be repressive, as well as providing some of the first quantitative analyses in SSA of an interest rate cap regime[127]. The section also provides a methodological contribution to the literature by using lending rate data to allow for lenders' likely heterogeneous and potentially non-linear exposures to the IRR – one of the first to do so.

While the empirical literature tends to support the idea that interest rate caps repress lending and market access, there is less clear-cut guidance from theory. Early research on interest rate regulations framed restrictions within a static, competitive, and informationally frictionless loan market, typically arguing that price controls would result in credit rationing (via loan quantities or non-interest loan terms) and reduced access (see Ostas (1976), Vandenbrink (1980), and Chicago Federal Reserve (2001)). In the financial development literature, a broadly similar theme emerges in McKinnon (1973) and Shaw (1973) with their arguments for financial system liberalisation and the critical role liberalisation plays in supporting investment and development in LICs. Moreover, much of the policy recommendations from the Bretton Woods institutions arguing for liberalisation in the 1990s seem to draw from this theoretical vein (Demetriades and Luintel (2001)). More recently, Miller (2013) uses a perfectly competitive loan market framework to argue against the use of interest rate caps in lower-income settings.

A competitive loan model benchmark only tells part of the story, however. Richer models of banking or lending can justify the use of IRRs, at least in some contexts. The simplest context is if the loan market resembled a non-competitive market structure; under such an assumption, interest rate restrictions might actually enhance financial access and development. Examples are given in Courakis (1984) or Demetriades and Luintel (2001), who assume non-competitive loan markets and show that IRRs could increase lending volumes and improve financial development[128]. Non-

[126] Miyauchi's analysis was for bank lending only, but conceptually the imperfect market context underpinning his results may also apply for microfinance.

[127] There is a clear gap in the empirical literature for SSA countries. Our data and empirical approach is clearly not as comprehensive and convincing as some higher-income countries which use loan-level data (Cuesta and Sepulveda (2019) and Galenianos and Gavazza (2022)) – but it should still make a useful contribution given the scarcity of such data in the region.

[128] Clearly this result requires caveats: The level of the cap in relation to the counterfactual non-monopoly market equilibrium will determine whether lending is increased by the cap. Repressing the interest rate far below the counterfactual

competitive pricing may also arise from geographically segmented, remote, or rural lending (Cubillos et al., 2018). A non-competitive market setting may matter additionally if we allow for a model of relationship lending: Miyauchi (2017) shows how IRRs might not only increase loan supply in the period in which they were implemented – a static result deriving from the imperfect competition in the banking sector – but also lending in later periods, i.e., IRRs might support future lending that would not have been financed otherwise. If interest rates can have sorting or incentive effects, the analysis of IRRs is also more complicated than in a simple demand and supply setting (Stiglitz and Weiss, 1981). High interest rates may induce moral hazard; for example, Coco and Meza (2009) show that a just binding usury law may create welfare gains.

To help frame our empirical analysis we build a simple and stylised segmented market lending model that allows for differing lender pricing and cost structures (rather than assuming a representative lender). While the model provides few deep theoretical insights, it does highlight the importance of considering heterogeneity in lenders' exposure to the cap and the potential for heterogeneous reactions.

Although not usually considered in the theoretical modelling, IRRs may also push lenders to apply pressure to non-interest margins to offset the effects on interest income from an IRR (Ferrari et al. (2018)). The most common non-interest margins are usually additional loan fees and commissions. We make an empirical contribution to the literature by examining this mechanism and showing that this offset needs to be taken seriously; in our Zambia MFI sample, we document a strong increase in the use of unregulated fees and commissions associated with the cap. Such a shift in margins may work to dampen effects associated with the caps if they are not also restrained.

Finally, this section relates to a broader policy literature examining interest restrictions. As the use of IRRs has become more prevalent in the past decade, research documenting the different types of IRRs and their specific forms has been growing (see Maimbo and Gallegos (2014); Asian Development Bank (2016); Ferrari et al. (2018); Calice et al. (2020a); Calice et al. (2020b)). These papers provide a taxonomy of interest rate regulations and highlight the significant variation in interest regulations across countries, offering a cross-sectional snapshot at a point in time. The

equilibrium would likely repress lending (this is what we find in our simple model). This point shows that governments need to be well aware of the cost structures underpinning lenders in their markets (and these will differ across market segments), and imposing a cap comes at the risk of mispricing, meaning the IRR policy could do more harm than good.
.

chapter on IRRs and MFIs across countries contributes to this literature by providing a longitudinal and standardised database of MFI interest rate caps – albeit based on secondary sources – for over 90 countries. By linking this database to the World Bank's MiX database, we also contribute to this policy literature by providing some initial cross-country analysis of interest rate caps.

Chapter 5 – Interest Rate Regulations – a case study of Zambia's non-bank/microfinance sector

> *"…the interest rate capping set on financial institutions in 2013 choked individual companies' revenues" - Kelvin Milambo, previous CEO of CETZAM (a non-bank lender), cited in Manic (2015)*

> *"Zambians, especially women and the youth have bright ideas and business proposals but they cannot implement them because of high interest rates on money available in the banking and financial services sector," – Alexander Chikwanda- Zambian Minister of Finance (2011–2016), Sep 30, 2011, cited in Mfula (2011).*

In early 2013, the Zambian Government imposed IRRs on all formal lenders operating in the credit market. The IRRs were applied differentially across various types of lenders, but some lenders in the non-banking sector were particularly bound. In this chapter we provide a quantitative case study of how Zambia's microfinance/non-bank sector responded to the implementation of IRRs. The chapter sits squarely within the empirical literature evaluating the effects of IRRs, as discussed in the previous chapter. The chapter begins by exploring the context surrounding Zambia's implementation of IRRs and examining how the IRR likely impacted non-bank lenders heterogeneously. We then build a very simple non-bank lending framework to help build variation in lender loan rates into our analysis and quantitative approach. The core of the paper provides two pieces of quantitative analysis. The first examines how the introduction of IRRs may have impacted lending volumes in Zambia's microfinance sector, using both distributed lag and event study/difference-in-difference approaches; the second analyses how the IRRs may have encouraged lenders to shift loan income from interest income into charging fees and other commissions.

Institutional Context and Zambia's Interest Rate Restrictions

The economic and political context regarding interest rates in Zambia and the imposition of Interest Rate Restrictions.

In January 2013, after less than a year in office, a newly elected Government under Michael Sata imposed interest rate caps on Zambia's credit market. Prior to these caps, no IRRs had been in place in the microfinance market since the market liberalising reforms in the 1990s. The caps were broad (applying across the whole market), benchmarked to the Bank of Zambia Policy Rate, and differentiated by the credit provider. Banks' lending was limited at 9% above the policy rate (18.25% when implemented), while microfinance institutions and non-bank financial institutions were capped at 2.302 and 1.644 times the bank rate, respectively (i.e., an absolute rate of 42% and 30%)[129,130]. The caps applied to new lending only and did not include any new regulations on fees[131].

The Government's stated objectives in implementing the caps were to lower the cost of credit, enhance credit access, and mitigate risks of over-indebtedness (FSDA and Davel, 2013). But the IRRs also likely reflected both (i) a deep and historical political perception that, even nearly 50 years after independence, Zambia's financial system was still inadequately serving the country's development and its citizens, and (ii) the perception that Zambia's market liberalisation reforms from the 1990s had inadequately reduced lending costs and that some lenders were simply charging too much.

As discussed in detail in Chapter 2, throughout Zambia's history there has been a long-held perception among both politicians and the polity that Zambia's financial system was still inadequately serving the country's development. These perceptions can be linked back to both the extractive nature of colonial finance as well as the challenges the new Government faced post-

[129] Non-bank financial institutions include the following institution types: Leasing Finance Institutions; Building Societies; the Development Bank of Zambia; and the National Savings and Credit Bank.
[130] BOZ Circulars CB 25/2012 and NB 08/2012 introduced the caps. Compared to other IRRs from around the world, Zambia's regime appears not to have been abnormal. The IRR on the banking sector (9% above the BOZ rate) was more generous than Kenya's 2016 cap (of only 4% above the Bank of Kenya rate) or Bolivia's current cap (11.5% absolute), yet still more restrictive than a regime like Uruguay's (50% absolute). It was also a relatively simple rule to implement; unlike some regimes, Zambia's rule applied as a blanket rule across all loan types and categories (Vietnam, as a counter example, has several ceiling rates for different priority sectors, and South Africa's applied caps differentially based on credit type). Likewise, Zambia's 42% microfinance cap appears less harsh than those for many other countries (India, Bangladesh, and Cambodia, for example set caps at approximately 26%, 27%, and 18% respectively). Even within Africa, Zambia's cap sits higher than the 27% IRRs set for WEAMU countries. In saying this, and especially for MFIs, the distribution of interest rates within a market may be as important as the absolute cap level.
[131] Fees are regulated in Zambia via STATUTORY INSTRUMENT No. 179 of 1995, but this instrument carves out significant room for lenders to apply non-interest loan fees, which they did.

independence in increasing financial access via foreign banks, a struggle some would argue is ongoing[132]. Although regressive financial policies and a flailing economy during the late 1970s-1980s did constrain intermediation, we believe that these perceptions created a context in which strong administrative rules were seen as a justified solution to banks' lack of lending. This historical context underpinned the imposition of the IRRs in 2013.

If long-held grievances against financial access lay the groundwork for IRRs, the more recent flashpoint was the political dissatisfaction at high-interest rates in Zambia post the 1990s market reforms and financial liberalisation.

While for much of the 1990s Zambia struggled to adjust to this newly liberalised economic regime, by 2001 the macroeconomic shockwaves and volatility emanating from the liberalisations had begun to dampen and the run of banking failures in the late 1990s had ceased. Inflation (although still high in 2005 at 18%) was reducing, and Zambia recorded real average growth of over 7% per annum between 2005 and 2012. The financial sector was also growing, with supportive reforms for the banking and non-banking sectors enacted in 1995 and 2006 respectively[133]. In addition, the HIPC debt write-offs in 2005 lessened the Government's debt burden freeing up fiscal space and putting downward pressure on interest rates.

Throughout this period leading up to the imposition of IRRs in 2013, bank lending rates and spreads were slowly trending downwards from their post-reform peaks. Yet, there remained an issue: lending rates remained elevated in an absolute sense, especially when compared to levels seen in the artificially depressed pre-reform era[134]. This was particularly the case for smaller, riskier, and less formal borrowers, who would have naturally struggled to access finance at low rates in any case and was likely due to poor credit market infrastructure, such as a functioning bureau, land and collateral registry, and an efficient debt litigation system (de Luna Martínez, 2006). The quote at the beginning of this chapter by Alexander Chikwanda, Zambian Minister of Finance from 2011–2016, reflects this perspective and the dissatisfaction leading to IRRs.

[132] Specifically, see the section titled "1964 – 1980s: Colonial extraction, Independence, and Bank Finance" and the quotes form Kenneth Kaunda, both in Chapter 2.
[133] Domestic credit to the private sector by the banking system grew from 6.5% of GDP in 2005 to 11.9% in 2012. The number of non-banks grew from 14 in 2006, when the reforms were implemented, to 34 in 2012.
[134] Remember that loans were heavily rationed in the pre-reform period, but such rationing/quantity effects are harder to see and are potentially less political than higher interest rates.

Until 2011 there had been political support for Zambia's financial market reforms. However, 2011 marked a turning point for the political narrative in Zambia, recreating the policy space for IRRs and a stronger role for the state in finance. In September, the Patriotic Front (PF)'s Michael Sata won Zambia's presidential election from Rupia Banda, providing a clear political demarcation from the more pro-economic reform era of Levy Mwanawasa (Banda's predecessor) and the Movement for Multi-Party Democracy (MMD). Sata's populist discourse stressed the negative and unequal impacts of liberalisation and reform, especially on the urban poor. High lending rates and low lending volumes for small business were part of this discourse, and soon after being elected, Sata announced his priority of cutting the cost of credit for Zambians (Larmer and Frasier, 2007; Piskadlo, 2018). Whether in truth or in perception, the benefits of Zambia's post-reform growth were not clearly being distributed across all, and high lending interest rates were one focal point.

The IRRs implemented by Sata in early 2013 were hence a credible signal from his new Government that it too saw the cost of credit as an issue. The IRRs aligned with the new President's broader political narratives and his pushback against the 1990s reforms, as it constituted a reasonable turnaround from the previous decades' policies of financial sector liberalisation. The Government's historical role in controlling the economy and interest rates also likely made implementing IRRs more palatable for Sata and other politicians than otherwise might have been the case (Sata and other politicians within the PF had worked in government in the 1980s under Kaunda when controlling the financial system was seen as business as usual).

Finally, for the microfinance sector specifically, there was also concern within BOZ that MFIs were charging unjustified interest rates and that consumers needed protecting (Siwale and Kimmitt, 2017). Given that BOZ was responsible for enabling the regulations (passed in 2006) that facilitated growth in consumer-lending MFIs and some of the highest interest rate lenders, they may have felt exposed. Coupled with this was also an increasingly public discourse against high MFI rates. In June 2012, one local paper wrote: *"We should not accept a system where loan sharks masquerading as responsible micro-lenders or banks steal what little our workers and the poor have in high interest charges. The government needs to move in and set some limits on interest charges to the poor and bring it to the same level as the interest that the rich pay."* (The Post, July 2012, cited Siwale and Kimmitt 2017, pg 21).

In summary, the discussion above and our research provide us with some confidence that the imposition of the IRRs was thus a response by Sata to both the long-term historical challenges regarding credit access for Zambians and high post-reform interest rates, rather than a short-term response to acute macroeconomic or political events of the time[135]. This arguable exogeneity of the IRR's imposition is an important component of our identification strategy, as we will discuss later in this chapter.

Loosening and removal of interest rate caps

In June 2014, the bank cap rate was loosened by BOZ, increasing the cap's lending margin above the BOZ Policy Rate by 6.75%, from 9% to 15.75%. This loosening was undertaken primarily to weaken liquidity constraints in the banking sector, which had been under severe pressure since BOZ tightened monetary and liquidity conditions earlier in the year (liquidity was surprisingly tight in the months before the cap was loosened, with interbank lending rates spiking at 25%). In practical terms, this cap loosening meant that the maximum lending rate for banks increased from 21% to 28%, likely a significant shift downwards in the volume of loans likely exposed to the cap. For microfinance lenders and non-banks, loosening of the cap influenced their own cap formulas, increasing their maximum rates from 48.4% to 64.5% and from 34.5% to 46.0% respectively.

The caps were entirely removed in November 2015 as a part of BOZ's response to calm the markets and mitigate the damage to Zambia's currency and inflation crisis. Several factors had led to this crisis, including weakening copper prices, unsustainable fiscal policy, and poor hydrology leading to the Kwacha depreciating by almost 45% from the beginning of 2015 to the cap's removal, and monthly inflation hitting 6.23% in October. As interbank rates again trended towards 20% in October and monetary policy tightened to dampen inflation, BOZ moved to provide the credit market with more flexibility; removing the caps was one option to help achieve this (Homley, 2017). Hence the loosening and eventual removal of the caps were clearly endogenous to broader financial market conditions at the time, something we need to be mindful of in our empirical work and approach.

[135] In addition to the fact that interest rates in Zambia had been decreasing since the HIPC debt write-off in 2005-2006, we have not found evidence that the caps were either (i) a reaction to any other short-term lending rate hikes, or (ii) that banks or non-banks were expecting the cap or that they were re-optimizing their balance sheets pre-emptively. We suspect that elevated interest rates post-reform were likely driven by structural micro-market factors.

Note: Table shows the IRR level and calculation for different lender types over the cap regime. Bolded figures show the cap levels. See BOZ Circulars CB 25/2012 and NB 08/2012 for the specific calculations and reference rates.

Source: BOZ and Author's Calculations

Figure 15 – Zambia's Interest Rate Cap Regime

Lender Type	Imposed - Jan 2013	Loosened - July 2014	Removed - Nov. 2015
Banks	**18.25%** (9%+BOZ rate)	**28%** (1.78*9%+12%)	-
MFI/Consumer	**42%** (2.302*18.25)	**64%** (2.302*18.25)	-
Non-bank (other)	**30%** (1.644*18.25)	**46%** (1.644*18.25)	-

The IRRs and Zambia's Non-bank sector

By year-end 2012, just before the cap was implemented, Zambia's microfinance market included 34 institutions, comprised of 24 MFIs, 3 development or savings banks, and 7 building societies or leasing/financing institutions. With a reasonable macroeconomic environment and supportive reforms being implemented in 2006, the market had seen strong growth over the previous half-decade, even if it remained a relatively small part of the credit market[136].

The introduction of the cap regime was a large and unexpected shock for many non-bank lenders. Four small lenders failed or ceased business in the quarter immediately after the caps were implemented – a clear break from the growth witnessed in previous years. Another microfinance lender, Cetzam, priced its loans at 100%, by no means the highest-priced lender in the market, and yet still struggled:

> *"…your revenue-earning capacity is pushed down by more than 50 per cent and yet you are not given time for you to be able to streamline your operations…I think we were only given about two weeks for us to implement the revised interest rate capping, and that really became very difficult for institutions, not Cetzam alone, to survive in that environment" - Kelvin Milambo, Cetzam CEO, 2015 (cited in www.Manic.com)*

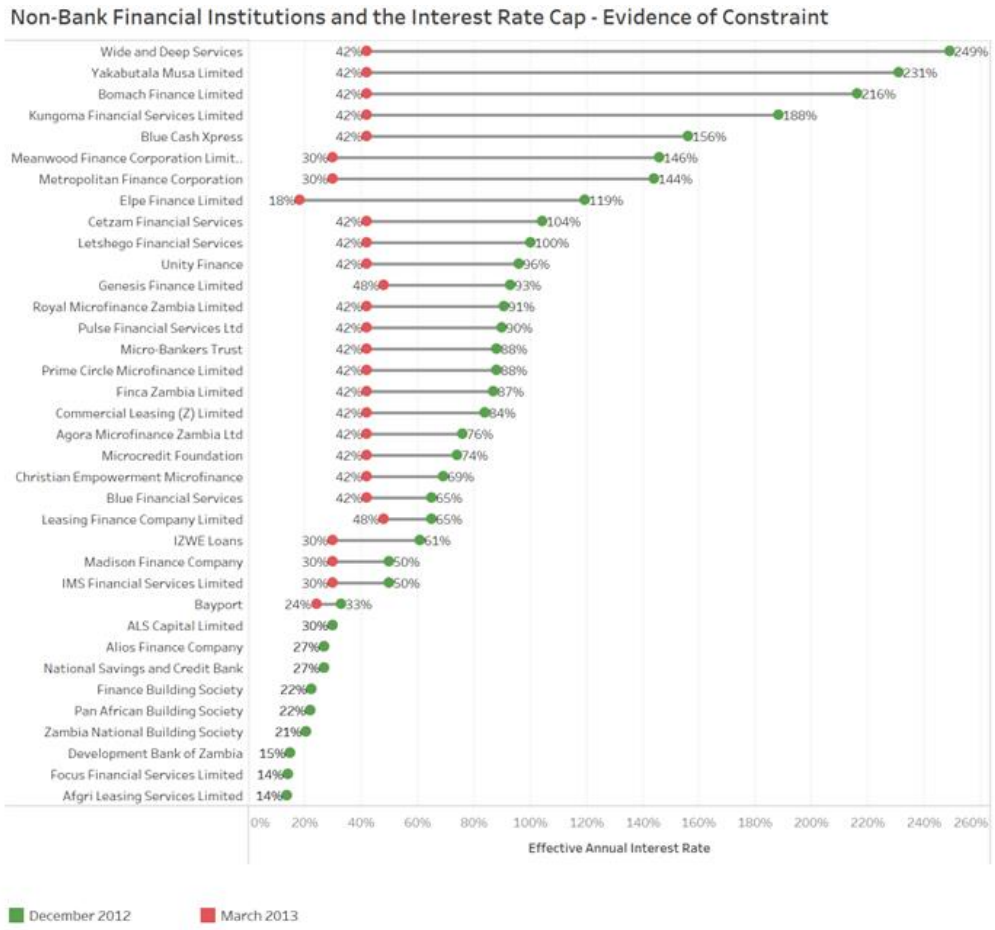
[136] In 2007, there were only 16 institutions, five of which were microfinance lenders. Note: our analysis only includes formal lenders who have to submit data to BOZ (i.e., those that had capital greater than $10,000). Informal lenders, who are not captured in official data, were not affected by the cap, or at the least BOZ did not enforce the rule.

Indeed, the caps affected many non-bank lenders. In the month before the IRRs were implemented, many lenders were pricing above the cap. There was also significant variation by how much lenders were bound: out of the 34 non-bank institutions operating in Zambia in our dataset when the cap was enforced, all but nine were typically pricing loans at a higher rate than the cap. Most of the lenders in the market hence likely faced significant exposure to the new cap regulation[137]. The most exposed lender, a small microfinance organisation, was pricing at almost 250% (nearly 200% above the cap). Other lenders, for example larger cooperative and savings institutions (non-banks), typically lent at rates well below their cap limits. Figure 16 shows non-bank lender pricing in Q4 2012 (before the cap) and Q1 2013 (after the cap was imposed).

[137] Unlike our analysis using the MiX database, the Zambia sample includes interest rate data for each institution, at least over the relevant period regarding the cap. Our interest rate data is based on a survey completed by non-banks in which they complete pricing and fees data for their "typical" customer. Non-banks will hold a portfolio of loans and many use differentiated pricing. As such, their loan portfolios will be comprised of a distribution of interest rates. We only measure the "typical" loan price, and so a lender with a "typical" lending rate. Thus, even a lender whose "typical" borrower sits below the cap may still have some proportion of loans that are directly impacted by the cap. We do not have loan level data and so cannot calculate the true exposure for each lender. However, as a broad measure of cap exposure, the average is likely not too bad as lenders usually concentrate on only a few related market segments and the variance in pricing at the institution level is limited.

Source: BOZ quarterly interest rate survey

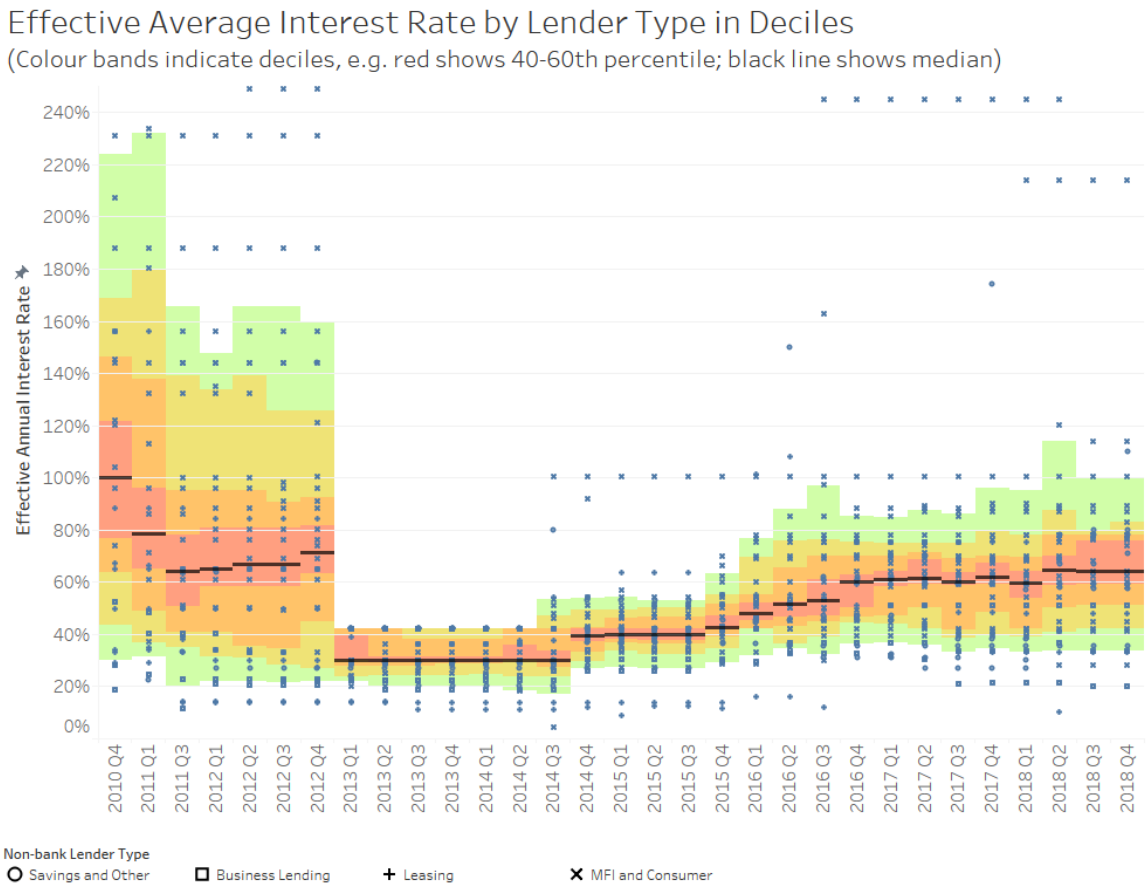
Figure 16 – Zambia Non-Bank Lending Interest Rates Before and After IRR Imposition



Note: Figure shows interest rates recorded by BOZ for non-bank lenders both before and after the caps were imposed. No red dot for a lender (i.e for a March 2013 interest rate) means that there was no change in recorded rate; for example, Pan African Building Society lent at 22% both before and after the IRR was imposed.

The IRRs resulted in a dramatic shift in pricing across the market, as shown by Figure 17 (which provides the decile spread of interest rates - each colour band represents a decile). Note the sharp drop and squeezing of interest rate deciles in Q1 2013, as well as the fact that as the regime was further adjusted, the distribution of lenders' pricing typically also shifted (for example, the loosening of the IRRs in 2014, or as the core reference rate shifted). After the cap was removed, the variance in market pricing remained subdued. This could be due to the IRRs changing the market structure, but it could also simply be because other underlying drivers of interest rates had changed.

Figure 17 – Effective Annual Interest Rate by Non-Bank Lender Type



Note: Each square, dot, or cross provides a lending interest rate for a given MFI in a given quarter. The colour banks provide decile ranges across all lenders' interest rates in a given quarter. The red band in the middle shows the 40th-60th percentile range, the dark orange the 30th-40th and 60th-70th. The horizontal black line provides the median interest rate for a given quarter.

Source: BOZ quarterly interest rate survey. Note: based on raw BOZ data

There is also qualitative evidence that suggests the cap had an impact on the market. In 2015, Siwale and Kimmitt (2017) undertook unstructured and anonymised interviews of 6 MFIs. They found evidence that the cap resulted in lenders reducing lending as a whole and/or reducing lending to higher risk market segments. One MFI commented:

> *"MFI X would not have survived the interest cap, given that 80% of our clients are at the bottom end. We survived and* [are] *able to recapitalise because our foreign shareholders are willing to bear the loss and allow the MFI* [to] *adapt and respond to the challenge." (CEO, MFI 4, July 2016, cited in Siwale and Kimmitt 2017, p.19):*

Another MFI indicated that the IRR was linked to reducing outreach and lower lending to higher risk borrowers:

> *"We are currently contemplating closing some branches and only keeping the profitable ones as well* [as] *targeting segments that are cheaper to service" (MFI 4, July 2016, cited in Siwale and Kimmitt 2017, p.21)*

In July 2013, soon after the cap was introduced, FSDA and Davel (2013) undertook anonymous interviews with MFIs in Zambia's market. Again, they found evidence that the caps were having a negative effect on the MFI market:

> *"…We incurred significant losses when the caps were first introduced as we had no choice but to stop lending…"*
>
> *"…We are now adopting a much more cautious approach to growth, particularly in outlying areas…"*
>
> *"…We have put a halt to plans to extend our branch network until government's policy direction becomes clearer… "*
>
> *"…During the time that we stopped lending, we are aware that some of our clients had to resort to informal money lenders…"*
>
> *(Anonymous MFI quotes, cited in Financial Sector Deepening Africa (FSDA) and Davel, 2013 p.27)*

Based on this large shift in interest rate distribution due to the IRRs (Figure 17), the significant variation in how severely different lenders were bound (Figure 16), and the qualitative interview evidence suggesting harsh impacts of the cap, the next part of this chapter builds a simple framework to help outline how we might expect different lenders to react to the IRRs as well as highlighting some of the econometric challenges we will face in investigating the effects of the IRRs in Zambia.

A framework model

The primary goal of this chapter is to estimate the effects of Zambia's interest rate caps on its microfinance institutions' lending. In our empirical approach we need to account for the fact that, within Zambia, IRRs may impact lenders differentially. Theory sometimes models credit markets as consisting of a single representative lender and borrower, but for an empirical analysis of interest rate caps in Zambia, such an assumption may be problematic. Take Zambia's MFI market in 2012 (see Figure 16, which shows the lending rates for MFIs operating in Zambia's market before the cap was introduced). Some lenders, such as Bayport, lend at low interest rates and are unlikely to be

bound by an IRR, while others, such as Cetzam, lend at rates well above the IRR. It is unreasonable to expect that these two lenders will react similarly to a uniform IRR, as was imposed in Zambia. In our framework, we hence need to allow for multiple lender interest rates and heterogeneity in lenders' reactions, while also accounting for how this might affect the empirical analysis of the impacts of IRRs.

We outline a simple econometric model to clarify our thinking and frame our analysis. The model is stylised and aims to model simple loan rationing to highlight the econometric challenges of investigating IRRs, rather than contributing in any theoretically interesting way. The key structure of the model is that it assumes multiple loan market "segments" in which a lender and a borrower interact. It is this "segmented" structure across markets that lets us model lender heterogeneity in a clear (albeit simple) way. We require the use of segmented markets because they help us understand how the level of the interest rate cap compares to a given lender's loan interest rate (and cost structure). This comparison helps understand whether a firm fails (a problem in our dataset), lends as if there was no cap in place, or undertakes disequilibrium loan rationing. In our econometric analysis, we will be sampling across market segments in which there may be lenders in all three categories. There is some evidence that credit markets in lower-income countries do exhibit segmentation (Ghosh, Mookherjee, and Ray, 2000).

Our framework here is similar in some respects to the simple model in Chapter 3, and so also draws from Peek and Rosengren (1995), Courakis (1984), and Khwaja and Mian (2008). The analysis of IRRs within a setting like this is not especially novel, although we have not come across a framework that links simple analytical market segments to the econometric challenges and truncation we know exists in our dataset.

Model and Market Summary

The model comprises of N segmented loan product markets, with a specific lender providing financing for a specific firm in each segment. The firm seeks finance from a single lender in each loan market segment and sells its goods in a competitive market (whose pricing is assumed exogenous). Borrowers/firms and lenders are subject to demand and supply shocks, respectively. This approach to loan market segmentation is clearly highly stylised, but it aims to reflect different classes of borrowers, different lending technologies used to service such borrowers (by different

firms), significant variations in borrower risk, and geographic segmentation. An example is that a collateralised payday loan in Lusaka is clearly within a different market to a group microfinance product in Chipata). Even within Zambia's banking sector there is some empirical evidence of industry segmentation (Simpasa and Pla, 2016). We assume there are no information asymmetries. There are two policy "regimes" we will consider. In the first, interest rates are set freely. In the second, the government imposes an IRR capping the maximum lending rate at $r_{CAP,t}$.

Borrowers/Firms

Each borrower sells goods to maximise profit. Firms have no capital and rely entirely on borrowing in each period to fund their product inputs. Each borrower sells goods in a competitive external market whereby we assume that the price of its good, P_j, is fixed and exogenous. C_j is the exogenous cost of a borrower's inputs, which it needs to finance by borrowing and paying interest of $r_{it,L}.C_j.Q_{jt}$. Costs per unit increase for firms as production increases because of diminishing marginal returns and finite market sizes[138]. In terms of decision-making variables, each period the firm chooses to produce Q_{jt} units of outputs or chooses to get financing for a given quantity from the lender at rate $r_{it,L}$, but not both. In other words, it can reject financing at a given interest rate, but this will impact its ability to produce. There are no retained earnings across periods. A borrower is exposed to two exogenous shocks that impact its demand for financing: ε_{jt} is a borrower-specific shock and ε_t is a common borrower shock[139]. Both shocks operate through the price the firm receives for its good.

Period t (assuming no cap)

$$\max_{Q_{jt}} \pi_{jt} = \left(P_j + \varepsilon_{jt} + \varepsilon_t\right).Q_{jt} \; - \; C_j.Q_{jt} - r_{it,L}.C_j.Q_{jt} - \alpha_j.Q_{jt}^2 \tag{5.1}$$

$$\frac{\partial \pi_{it}}{\partial Q_{it}} = P_j + \varepsilon_{jt} + \varepsilon_t - C_j - r_{it,L}.C_j - 2\alpha_j Q_{jt} = 0$$

$$Q_{jt} = \frac{P_j + \varepsilon_{jt} + \varepsilon_t - C_j(1 + r_{it,L})}{2\alpha_j}$$

[138] This is modelled as the $\alpha_j.Q_{jt}^2$ term in the cost function.

[139] For simplicity, we assume these shocks follow: $E\left[\varepsilon_{jt}\right] = 0, var\left(\varepsilon_{jt}\right) = \sigma_{\varepsilon 1}^2,\ E[\varepsilon_t] = 0, var(\varepsilon_t) = \sigma_{\varepsilon 2}^2, and\ cov\left(\varepsilon_{jt}, \varepsilon_t\right) = 0$.

$$r_{it,L} = \frac{P_j + \varepsilon_{jt} + \varepsilon_t - 2\alpha_j Q_{jt}}{C_j} - 1$$

Lending Organisation

Each lender operates in a competitive market segment, i.e., they do not have any pricing power[140]. They act as intermediaries, borrowing externally at $r_{it,C}$ (from an inefficient or finite exogenous capital market) and lending L_{it} to firms at $r_{it,L}$. Lenders face an upward-sloping marginal cost of financing function, which is modelled linearly as $r_{it,C}(L_{it}) = \gamma_i . L_{it}$ (our disequilibrium results explored later do not rely on this, but it helps us ensure an upward-sloping loan supply curve). The intuition is that capital markets are likely finite for many lenders and attaining more finance can be additionally costly. We leave the i subscript in γ_i to show that in principle this might differ across segments, but in all our empirical work we assume homogeneity and hence that $\gamma_i = \gamma$. Lenders have a market/segment specific variable cost of lending, c_i. Lenders are faced with a common and lender-specific marginal cost shock, δ_t and δ_{it} respectively[141].

Period t (assuming no cap)

$$\max_{L_{it}} \pi_{it} = L_{it}\left(r_{it,L} - r_{it,C}(L_{it}) - c_i - \delta_{it} - \delta_t\right) \tag{5.2}$$

$$\max_{L_{it}} \pi_{it} = L_{it}(r_{it,L} - \gamma_i . L_{it} - c_i - \delta_{it} - \delta_t)$$

$$\frac{\partial \pi_{it}}{\partial L_{it}} = r_{i,Lt} - 2\gamma_i . L_{it} - c_i - \delta_{it} - \delta_t = 0$$

$$L_{it} = \frac{r_{i,Lt} - c_i - \delta_{it} - \delta_t}{2\gamma_i}$$

$$r_{it,L} = 2\gamma_i . L_{it} + c_i + \delta_{it} + \delta_t$$

Market and Disequilibrium Outcomes

In period t, assuming no cap, each market segment clears, and we see equilibrium lending in each segment such that:

$$L_{it}^* = \frac{P_j + \varepsilon_{jt} + \varepsilon_t - C_j(c_i + \delta_{it} + \delta_t + 1)}{2(\gamma_i C_j + \alpha_j)}$$

[140] We expand on this assumption at a later point.

[141] Again, we assume these shocks follow: $E[\delta_{it}] = 0, var(\delta_{it}) = \sigma_{\varepsilon 1}^2,\ E[\delta_t] = 0, var(\delta_t) = \sigma_{\varepsilon 2}^2, and\ cov(\delta_{it}, \delta_t) = 0.$

However, if an interest rate cap, r_{CAP}, is implemented, things become more complicated. Under a cap regime, there are now three possible outcomes for each lender, each determined by where each lender's cost structure in Period t sits with respect to the imposed cap.

Outcome 1: Assuming a lender prices its loans below the cap in equilibrium – in other words the cap is not binding (i.e. that $r_{CAP,t} \geq r_{i,Lt}$) – we should see the market clear as it would have otherwise, with the market interest rate and loan volume adjusting for the demand and supply shocks as above. This lending outcome would also be the value of lending in the market in the counterfactual case where an interest rate cap was not imposed (which we shall denote with L_{it}^{*}).

Outcome 2: If we instead assume the lender is bound by the cap but the cap is not so binding that all loans are unprofitable (i.e. $r_{i,Lt} > r_{CAP,t} \geq r_{it,Zero\ profit} = c_i + \delta_{it} + \delta_t$), we would expect to observe some lending, but supply would be constrained by the cap compared to the market-determined outcome that would have occurred without the cap. $r_{it,Zero\ profit}$ is the interest rate at which the lender will drop out of the market because even a near-zero loan amount will be loss making, as it equals the marginal cost of lending for the first unit of lending. The level of lending under a binding and rationing cap is given by:

$$L_{it}^{cap} = \frac{r_{CAP,t} - c_i - \delta_{it} - \delta_t}{2\gamma_i}$$

In this case we see disequilibrium. At the capped interest rate, lenders supply less than the borrower firms demand and there is a form of rationing. A shift up or down in the cap shifts the lending outcome along the lender's marginal cost curve.

Outcome 3: Finally, if the cap is so low that it is not profitable for the lender to make a loan (i.e. $r_{it,Zero\ profit} = c_i + \delta_{it} + \delta_t > r_{CAP}$) then we would expect the firm to drop out of the market and we would not collect data on their lending.

In period t under a cap regime, we would thus expect:

$$L_{it}^{cap} = \begin{cases} \dfrac{P_j + \varepsilon_{jt} + \varepsilon_t - C_j(c_i + \delta_{it} + \delta_t + 1)}{2(\gamma_i C_j + \alpha_j)}, & r_{CAP,t} \geq r_{i,Lt} \\ \dfrac{r_{CAP,t} - (c_i + \delta_{it} + \delta_t)}{2\gamma_i}, & r_{i,Lt} > r_{CAP,t} \geq r_{it,Zero\ profit} \\ unobserved, & r_{it,Zero\ profit} > r_{CAP,t} \end{cases}$$

The effect of the cap on lending

To isolate the "effect" of the cap on lending, we compare the market outcomes in period t under the cap with what we would have expected to have observed without the cap (i.e., the counterfactual). As a first step, let us examine what the market outcome would be in period t if no cap were put in place, which we would *expect to hold for all market segments*:

$$L_{it}^{*} = \frac{P_j + \varepsilon_{jt} + \varepsilon_t - C_j(c_i + \delta_{it} + \delta_t + 1)}{2(\gamma_i C_j + \alpha_j)} \; and \; r_{it,L}^{*} = \frac{\alpha_j(c_i + \delta_{it} + \delta_t) + \gamma_i\left(\left(P_j + \varepsilon_{jt} + \varepsilon_t\right) - C_j\right)}{\alpha_j + C_j\gamma_i} \quad (5.3)$$

To determine the effect of the cap, we need to find the difference in outcomes between the actual cap-influenced outcome (L_{it}^{cap}) and the hypothetical of no cap (L_{it}^{*}):

$$\Delta L_{it}^{cap} = L_{it}^{cap} - L_{it}^{*}$$

$$= \begin{cases} \dfrac{P_j + \varepsilon_{it} + \varepsilon_t - C_j(c_i + \delta_{it} + \delta_t + 1)}{2(\gamma_i C_j + \alpha_j)} - \dfrac{P_j + \varepsilon_{it} + \varepsilon_t - C_j(c_i + \delta_{it} + \delta_t + 1)}{2(\gamma_i C_j + \alpha_j)} = 0 & r_{CAP,t} \geq r_{it,L}^{*} \\ \dfrac{r_{CAP} - (c_i + \delta_{it} + \delta_t)}{2\gamma_i} - \dfrac{P_j + \varepsilon_{jt} + \varepsilon_t - C_j(c_i + \delta_{it} + \delta_t + 1)}{2(\gamma_i C_j + \alpha_j)} < 0 & r_{it,L}^{*} > r_{CAP,t} \geq c_i + \delta_{it} + \delta_t \\ unobserved & c_i + \delta_{it} + \delta_i > r_{CAP,t} \end{cases}$$

Inserting the value of r_{it}^{*} into the above expression, rearranging, and reparametrizing yields:

$$\Delta L_{it}^{cap} = \begin{cases} 0 & r_{CAP} \geq r_{it,L}^{*} \\ \left(r_{CAP,t} - r_{it,L}^{*}\right).\left(\dfrac{1}{2\gamma_i}\right) = \beta_i\left(r_{CAP,t} - r_{it,L}^{*}\right) & r_{it,L}^{*} > r_{CAP,t} \geq c_i + \delta_{it} + \delta_t \\ unobserved & c_i + \delta_{it} + \delta_t > r_{CAP,t} \end{cases} \quad (5.4)$$

This equation sits at the core of our empirical work. It shows that we might expect non-linearities in loan outcomes if we estimate the effect of the cap on lending in a cross-section of market segments. There is clear disequilibrium when the cap is binding and potentially two different forms of truncation when either the lender is unbound or when the lender leaves the market. A lender's response type post IRR is a function of where the IRR lies relative to the lender's marginal costs, i.e., the IRR will impact lenders in each market segment differently depending on the distance between the IRR and lender's marginal costs. Different cost and production structures across firms and borrowers imply that, given a common interest rate cap, lenders providing loans at higher interest rates will reduce their lending by more than those who price only marginally above the cap.

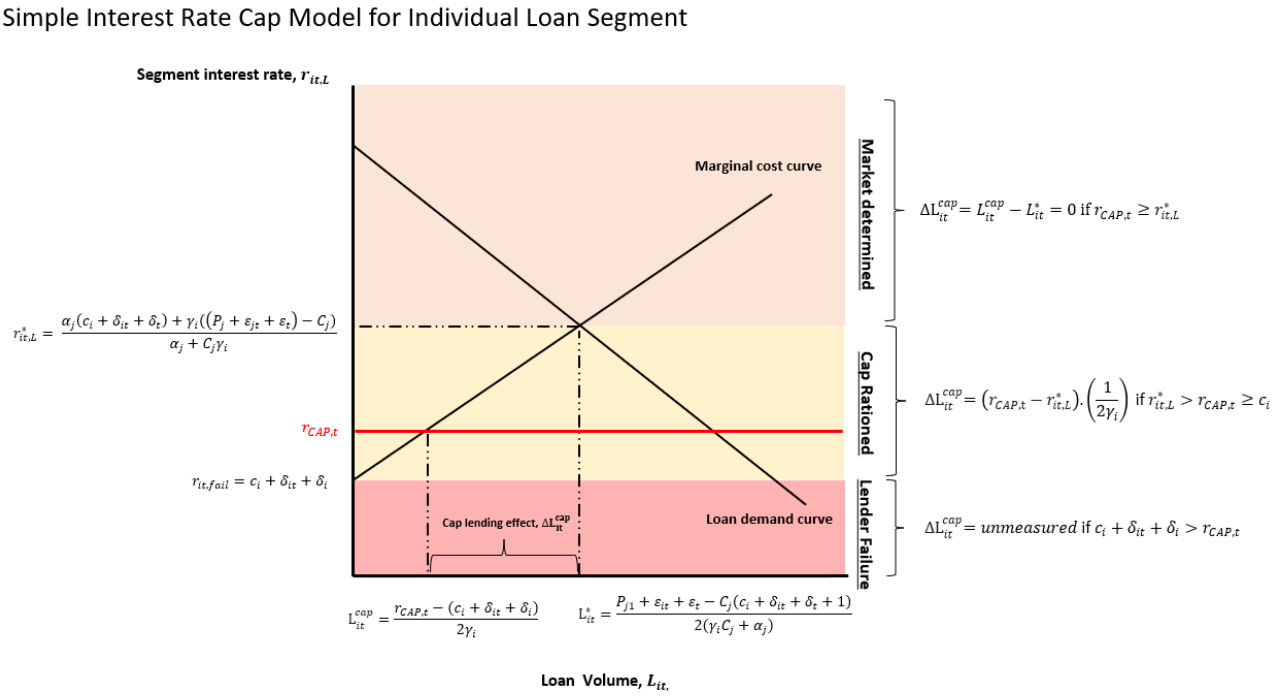
For the regulator, there are hence three possible outcomes for each market segment: (i) if the cap is placed above the market clearing interest rate, we do not expect to see any change in lending; (ii) if the cap is placed below the market clearing price but above the zero-profit rate (i.e above $c_i + \delta_{it} + \delta_i$), the lender will reduce its supply of credit at a rate of $1/2\gamma_i$ per percentage points below the counterfactual market clearing rate, and (iii) if the cap is below the marginal cost of lending, then the lender will exit the market, stop lending, and not return data.

Figure 18, below, provides a simple graphical example of one lender segment and one assumed interest rate cap. Our sample is assumed to be made up of N of these market segments, all of which will have their own assumed cost structures and their own regions of cap rationing and lender failure. For simplicity, Figure 18 assumes there are no shocks in period two.

In our baseline empirical work, we attempt to estimate β using a piecewise regression approach, which will indicate how much lenders (but only those who were both bound by the cap and did not fail), changed their supply of lending due to the IRR[142]. The heterogeneity in lender pricing is allowed for in the estimate of β (at least linearly), as lenders who were more bound by the cap because of their higher costs (i.e $r_{CAP,t} - r^*_{1t,L} < r_{CAP,t} - r^*_{2t,L}$) should reduce their lending by more $\Delta L_{1t}^{cap} = \beta(r_{CAP,t} - r^*_{1t,L}) < \Delta L_{2t}^{cap} = \beta\left(r_{CAP,t} - r^*_{2t,L}\right)$. In this setting, our primary hypothesis to test would be $H_0: \beta \neq 0$.

[142] We assume that all lenders face the same γ_i cost parameter, and differences in interest rates are only the result of differences in cost of providing finance ($c_i + \delta_{it} + \delta_i$), a limiting but necessary assumption.

Figure 18 – Individual Market Segment in Simple Two-period Interest Rate Cap Model



Note: The figure provides a simple graphical example of one lender segment operating under an interest rate cap, based on the simple theoretical framework. Our actual sample is assumed to be made up of N of these market segments or lenders, each of which will have its own assumed cost structures and hence regions of "market determined", "cap rationing" and "lender failure" outcomes. The red line indicates the interest rate cap, which could have been placed in any of the regions by the Government, and in turn determines how the lender reacts. For simplicity, the figure assumes there are no shocks to the lender or the borrower (although note how a marginal cost supply shock would increases the "lender failure" region, as $c_i + \delta_{it} + \delta_i$ represents the break-even lending interest rate). The volume of loan rationing is show by the distance between $L_{it}^{cap} - L^*_{it}$ on the x-axis and is calculated, after some rearrangement, as $(r_{CAP,t} - r^*_{it,L}).\left(\frac{1}{2\gamma_i}\right)$.

Data Description

Our analysis uses the full panel of non-bank lending institutions in Zambia from 2005-2019. We employ an original database of quarterly MFI prudential returns that we compiled for this thesis from a repository of raw/uncleaned BOZ excel returns[143]. The data covers the universe of non-banking financial institutions (NBFIs) in Zambia from 2005 to 2019 and provides balance sheet and income statement data. Before microfinance reforms in 2006, there were only 10-14 NBFIs operating in Zambia; by the time the cap was introduced in Q1 2013, however, there were over 34. There is significant variation in the number of quarters per institution in the database, but the median number of quarters across the sample institutions is 20. The panel is strongly unbalanced, a result of the high turn-over of institutions during the 15-year period, especially among the smaller lenders.

[143] The database combines the bank of Zambia's quarterly supervisory MFI return and BOZ's quarterly market pricing and cost of credit survey. Our data – and the enforcement of the cap – are limited to formal lending institutions who were large enough to be registered with BOZ (i.e. they had capital greater than approx. USD 10,000).

The loan interest rate data we use to model lender heterogeneity begins in 2010 and comes from BOZ's quarterly interest rate survey completed by microfinance lenders. The data covers lending institutions' effective annual interest rate, averaged across their loan portfolios[144]. This dataset is quite rich, especially for the sub-Saharan Africa region, although is still limited compared to the loan-level data used in some of the higher-income country case studies on IRRs.

Descriptive statistics are provided in Table 14 below.

Table 14 – Zambia Non-Bank Descriptive Statistics

Variable	Full Sample					Bound Lenders			Non-Bound Lenders		
	Obs.	Mean	S. dev.	Min	Max	Obs.	Mean	S.dev.	Obs.	Mean	S.dev.
Growth in real net loans (%)	1,323	0.03	0.28	-1.12	1.15	976	0.03	0.30	347	0.04	0.23
Change in fees and commissions share of income (%-point)	1,302	0.00	0.08	-0.47	0.58	1,011	0.00	0.08	357	-0.01	0.08
Log total lender assets	1,382	9.71	2.23	4.61	14.87	1,024	9.18	2.16	358	11.23	1.63
Leverage ratio (Total Regulatory Capital/Total Assets)	1,390	0.38	0.45	-4.01	2.81	1,031	0.46	0.44	359	0.16	0.42
Loss allowance (Recognized losses/Gross Loans)	1,382	0.14	0.18	0.00	1.00	1,024	0.15	0.20	358	0.09	0.13
Operating expense ratio (Non-interest Expense/Total Assets)	1,378	0.20	0.38	0.01	5.97	1,020	0.24	0.43	358	0.10	0.14
Age of firm	1,561	17	13	1	56	1,170	15	11	391	22	15
Tbill Yield (364)	1,561	16.0	5.2	5.2	27.6	1,170	16.0	5.2	391	15.9	5.2
Annual Inflation	1,561	9.4	4.3	6.0	22.2	1,170	9.3	4.3	391	9.7	4.4
Kwacha/USD	1,561	6.7	2.5	3.2	12.3	1,170	6.8	2.5	391	6.8	2.5

Number of Observations Per Lender by Percentile	min	5%	25%	50%	75%	95%	max
	2	2	10	20	33	56	56

Empirical Methodology – Piecewise Distributed Lag Approach

We take two empirical approaches to investigate the effects of Zambia's IRRs on lending. The first, which is covered directly below, is our baseline. It uses a bespoke distributed lag (DL) model with unit fixed effects that link closely with the simple theoretical framework from the previous section. The approach builds on the idea that bound and unbound lenders might react differently to an IRR,

[144] Unlike our analysis using the MiX database, the Zambia sample includes interest rate data for each institution, at least over the relevant period regarding the cap. Our interest rate data is based on a survey completed by non-banks in which they complete pricing and fees data for their "typical" customer. Note, however, that non-banks will hold a portfolio of loans and many use differentiated pricing. As such, their loan portfolios will be comprised of a distribution of interest rates. We only measure the "typical" loan price, and so a lender with a "typical" lending rate, so even a lender whose "typical" borrower sits below the cap may still have some proportion of loans that are directly impacted by the cap. We do not have loan level data and so cannot calculate the true exposure for each lender. However, as a broad measure of cap exposure, the average is likely not too bad as lenders usually concentrate on only a few related market segments and the variance in pricing at the institution level is limited.

and that more bound lenders might have a greater reaction. Note that the IRR reaction parameters in these DL results rely only on within-unit variation, with a lender's previous lending experiences acting as the counterfactual.

The second, alternative, approach is to use an event-study difference-in-difference design, whereby unbound lenders are used to create a counterfactual for bound lenders. This approach sits closer to the literature, and we discuss it in detail later in the chapter.

To begin, our baseline estimates rely on a piecewise DL lender fixed effects regression. Our parameters of interest are β^j and β^k for the bound and unbound lender reactions respectively. β^j is a set of slope effect parameters for the bound group of lenders and β^k is a set of binary treatment parameters for the unbound. β^j and β^k are sets of parameters because we need to allow for dynamics.

$$\Delta Y_{it} = \alpha_i + \sum_{j=d}^{J} \beta^j . I . \Delta[T . (r_{i,\, t=2012q4} - r_{cap,t=2013,q1})] + \sum_{k=d}^{K} \beta^k . (1 - I) . \Delta T + \sum_{l=1}^{L} \beta^l X_{it}^l + \sum_{m=1}^{M} \theta^m M_t^m + \varepsilon_{it} \qquad (5.8)$$

$$I = \begin{cases} 1, & r_{i,t=2012q4} - r_{cap,\, t=2013q1} \geq 0 \\ 0, & r_{i,t=2012q4} - r_{cap,\, t=2013q1} < 0 \end{cases}$$

$$T = \begin{cases} 1, & t > 2012q4 \\ 0, & t \leq 2012q4 \end{cases}$$

ΔY_{it} captures lending growth (the log difference in the real net outstanding loan portfolio); α_i is a lender fixed effect; I is a group indicator variable (coded 1 for lenders bound by the cap when it was imposed and 0 for those that were not); $. \Delta[T(r_{i,t=2012q4} - r_{cap,t=2013q1})]$ captures the change in cap bindingness due to changes in policy (where $r_{cap,t=2013q1}$ is the value of the cap and $r_{i,t=2012q4}$ is the interest rate the lenders was charging in the period before the cap was introduced). T is the IRR indicator, indicating when a cap regime was in place or not. X_{it}^l includes several lender-level covariates, capturing firms' size (log of total lender assets), leverage (leverage ratio), cumulated losses (loss allowance), business model and efficiency (operating expense ratio), and firm age.

Allowing for dynamic effects is a necessity in our modelling approach. This is because the IRRs only applied to new lending, and so there will naturally be some delay in the IRR's effects. Lenders might be able to stop issuing new loans immediately, but reducing their portfolio would require

waiting for loans to mature naturally. Lenders might also want to wait to ensure that the policy was going to remain in place before adjusting their business models. We hence expect there to be a lag in lenders' loan response to the cap. We model lag structures up to a length J and K for the policy variables and L and M covariates to capture this[145] [146]. To allow for anticipation effects, we let J and K be negative. We provide sensitivity analysis over the choice of lag structure in the main results by providing estimates from an array of regressions stemming from equation (5.8), beginning at d=-5 with J=-1, spanning to d=0 and K=5, allowing us to test for anticipation effects and sensitivity to dynamic specifications.

The intuition behind the piecewise structure in equation (5.8) is that, because there is some theoretical justification for non-bound lenders reacting differently to bound lenders, we need to at least allow for a different lending reaction and estimate two different sets of reaction parameters ($\beta^j \; and \; \beta^k$). The indicator variable I allows for this break. The "intensity" of the policy and heterogeneity across lender exposure within the bound group is captured by the term $\Delta[T.\left(r_{i,\,t=2012q4} - r_{cap,t=2013q1}\right)]$, which essentially captures the IRR policy shock. As there is not a strong theoretical reason why lenders pricing below the IRR will have a heterogenous response, we model their exposure to the cap as a binary treatment (i.e, ΔT). In our base-case, fixed effects version of the model, we are only using variation in policy bindingness for each lender across time, tracking changes in outcomes that are driven by such variation (i.e., within unit variation). In all models, we restrict lenders to react homogeneously to all non-policy variables and linearly with respect to policy exposure.

The cumulative impact of the policy – our primary hypotheses of interest - can be calculated from equation 5.8 as $\sum_{j=0}^{J} \beta^j$ and $\sum_{k=0}^{K} \beta^k$ for bound and unbound lenders respectively, with the long-term effects being tested via $H_{0A}: \sum_{j=0}^{J} \beta^j = 0$ and $H_{0B}: \sum_{k=0}^{K} \beta^k = 0$. The figures provided in the results section provide point estimates and confidence intervals for the cumulative effect of the policy up to (or from) the number of quarters shown on the x-axis.

As with Chapter 3, these estimates are not causal. The key parameters we are estimating are a set of marginal parameters based on within-MFI variation in a fixed effects model. Unbiased estimation

[145] Both dependent variable series, already in growth rates, show little persistence so we do not model them dynamically (i.e., in an ARDL setup). Although we do have some time dimension in the data, the panel is strongly unbalanced.
[146] This dynamic lags for the interest income switching hypothesis (our second hypothesis) likely required additionally long lags as this would have required more fundamental changes to their business models and internal processes.

requires $E(\varepsilon_{it}|I.\Delta[T.(r_{i,\ t=2012q4} - r_{cap,t=2013,q1}), \boldsymbol{X^i}, \boldsymbol{M}, \alpha^i) = 0$, which is to say that conditional on specific non-bank factors (both time-varying and time-invariant) and macro covariates, the level of bindingness and imposition of the IRR regime should not be related to the error in time t. In terms of the simple framework from earlier, this condition would be violated if, for example, unmeasured loan demand shocks were to be related to the imposition or severity of the cap. We can never be certain that our estimates are unbiased. However, based on the economic and political context around the time of the cap, we believe this is at least defendable.

Empirical Challenges

Drawing from the simple theoretical model, there are three empirical challenges that require discussion: incidental truncation and firm drop-out in the post IRR sample; imperfect measurement of cap bindingness, and alternative segment market structures. The discussion can be underpinned by more formal analysis.

The first empirical issue is incidental truncation and firm drop-out: several institutions in our sample failed due to the imposition of the cap (four fail in the quarter after the cap was introduced and several more in the subsequent years), and if the firm failure is not independent of our lending and income share variables, our sample may not be a representative sample of microfinance institutions in Zambia, biasing our estimates.

Using the basic model from above and assuming a linear reaction to cap bindingness, we do not believe this form of truncation will cause bias. We suspect that this form of lender failure is based primarily on the level of cap bindingness (an independent variable, based on both the level of the cap and the marginal cost of lender lending, which incorporates exogenous price shocks), and so should not bias estimates. Firm failure due to the cap is hence essentially a setting in which non-random sample selection is being determined by independent variables. So, assuming (i) that there is no measurement error or uncertainties in the selection equation, and (ii) that the effect of bindingness is linear through its range (an assumption we have made), then OLS should be unbiased and consistent (although perhaps not efficient).

The second empirical issue is imperfect measurement and attenuation bias. In reality, we cannot measure the hypothetical equilibrium interest rate (or loan volume) that a segment would have charged (or lent) in the period the cap was imposed. Instead, our dataset and analysis rely on a proxy:

the interest rate from the end of the period before (i.e., $r^{*}_{it-1,L}$). We assume that this price would have remained unchanged. This is less of a problem for the set of lenders that are either theoretically unaffected by the cap (which we can observe given that $r_{i,L2} > r_{cap}$) or lenders that drop out due to the cap ($r_{cap} < c_i + \delta_{i2} + \delta_i$), as these lenders cleanly self-identify even without being able to measure their counterfactual interest rate. However, it is a problem for the subset of lenders for whom it is still profitable to lend under the cap but not at the same level as if there were no cap (i.e. those for which $r_{i,L2} *> r_{CAP} \geq c_i + \delta_{i2} + \delta_i$). For these lenders we can only measure $r_{i,L1}$ and L_{i1} and then assume that $r_{i,L1} = r_{i,L2} *$ and that $L_{i1} = L_{i2} *$. This introduces a source of mismeasurement that will lead to attenuation bias and classical measurement error (under reasonable assumptions), which we will need to accept may be a feature in our results.

The third issue is that assuming other market structures – like lender pricing power – will result in the loan outcome function becoming *additionally* piecewise. Under monopoly lending power, for example, a cap imposed below the monopolist's interest rate but above the competitive price will increase lending, yet if it were placed below the hypothetical competitive price for the monopolist, it will induce rationing. Clearly such variation in market structures and lender reactions would bias any simple piecewise estimate of β^j. Given that lenders were not clearly making supernormal profits (see Box 1 at the end of the section), we are not overly concerned.

The final issue we discuss results from the simple assumption that there is only one borrower in each segment and that lenders make loans at only one interest rate. In reality, most lenders in the market will hold a portfolio of loans which will have a distribution of interest rates, with borrowers' rates being priced more granularly based on their risk and loan terms (although, typically, there is not much variation because lenders often concentrate on only one market segment). This means that lenders who price near the cap may have portfolios that spread or bleed across the cap-induced bound/non-bound threshold. We only measure one interest rate for a lender (the rate for their "typical" loan and customer), so there is scope for some noise in what the bindingness measure is capturing and how bound a lender's portfolio likely is. This might result in a negative lending reaction to the cap even if they were not technically "bound" by the cap based on their average lending rate. Such a lender might also be able to shift its lending across risk types within its portfolio, resulting in the opposite outcome. To help address some of these concerns, we will run sensitivity analysis over various definitions of the bound lenders.

Baseline Results - Lending Growth

We find evidence that in the quarters after the IRR was implemented, institutions bound by the cap reduced their lending; those that were more bound by the cap reduced their lending by more. Further, we find little evidence that those lenders pricing below the cap reduced their lending. In summary, our results confirm the broad consensus in the empirical IRR literature that binding caps might risk rationing credit supply.

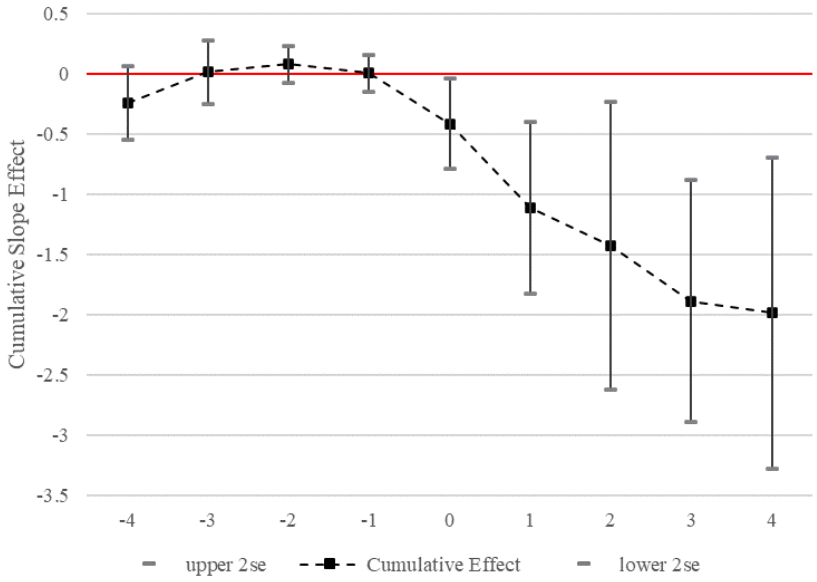
Figure 19 provides results for our baseline specifications using equation (5.8). Each black square provides a point estimate for the cumulative effect of the policy up to the given number of quarters on the x-axis (i.e., $\sum_{j=0}^{J} \beta^{j}$, where J is provided by the x-axis). The figure shows the sensitivity of the cumulative effect estimates to varying lag structures, examining for evidence of anticipation (up to four quarters before the policy), and then examining the cumulative effect of the policy up to 3-4 quarters after. Figure 20 does the same but for the non-bound set of lenders. Table 15 provides the full regression estimates for both bound and non-bound lenders.

We find that for each percentage point that an institution is bound by the IRR, over the following year its lending reduced by approximately 2 percent compared to when it was not bound, with a 95% confidence interval of approximately 0.75 percent to 3.5 percent. For a lender who was pricing its loans 35% above the cap (35% being approximately the median), we find that it reduced its lending by a cumulative 72% compared to when it was not bound by the cap, ceteris paribus (the cumulative slope parameter point estimate in Figure 19 (-2), multiplied by 0.36).

Lenders with lower pre-IRR interest rates also continue to see meaningful effects. We estimate that a lender only bound by the cap by 10 percentage points saw a reduction in lending over a one-year period with a 95% confidence interval of between 7.5% and 32.5%, with a mid-point estimate of 20% - a meaningful reduction. For lenders that were unbound by the cap, we do not find evidence that the cap drives negative loan growth (Figure 20). Indeed, there is some limited evidence that unbound lenders grew in the wake of the cap, possibly reflecting the reasonably strong underlying economic conditions at the time, not all of which will be captured in our modelling.

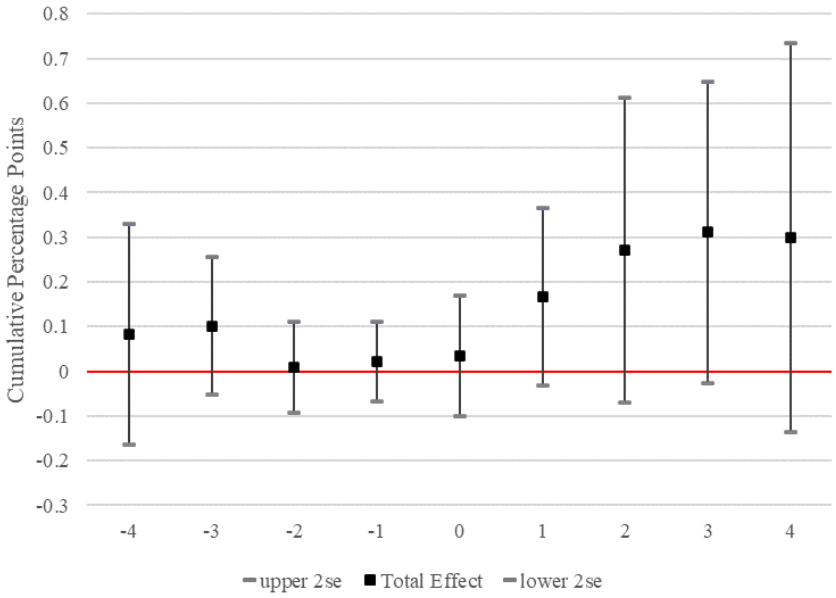
Most of the policy impulse for the bound lenders appears to occur in the first three to four quarters after the imposition of the cap, which makes sense given that the durations of many MFI loan portfolios are quite short. The cumulative effect then flattens out, which is what we would expect.

Figure 19 – Marginal Effect Piecewise Distributed Lag Models for Bound Lenders – Cap Implementation



Note: The figure provides results from the distributed lag model regressing growth in lenders' net real outstanding loan books against the change in its cap bindingness due to changes from imposition of the IRR policy. It captures estimates for bound lenders only. The figure shows cumulative marginal effect slope parameter estimates, $\sum_{j=d}^{J} \beta^{j}$, for equation 5.8, spanning over various lag structures (x-axis), whereby each x-axis period provides an estimate from its own regression. Controls include both lender-specific macroeconomic control variables, as per Table 15. The regression includes lender fixed effects. Standard errors are robust and clustered at the lender level. Confidence interval at 95%

Figure 20 – Binary Effect Piecewise Distributed Lage Models for Unbound Lenders – Cap Implementation



Note: The figure provides results from the distributed lag model regressing growth in lenders' net real outstanding loan books against the change in a binary treatment dummy for the imposition of the IRR policy. The estimates are for non-bound lenders only. The figure shows cumulative estimates, $\sum_{k=d}^{k} \beta^{k}$, for equation 5.8, spanning over various lag structures (x-axis), whereby each x-axis period provides an estimate from its own regression. Controls include both lender-specific macroeconomic control variables, as per Table 15. The regression includes lender fixed effects. Standard errors are robust and clustered at the lender level. The y-axis is given in decimals. Confidence interval at 95%.

The results are economically significant, even if our small sample does result in large confidence intervals. If the median bound lender reduces their lending by approximately 70% due to the IRR (compared to when they were not bound by the cap), policymakers need to seriously weigh the benefits of an interest rate cap against such a large reduction in loan volumes. But policymakers would also need to be concerned if the true effect was actually much less – even at the lower 95% confidence limit of our point estimate, within a year the median bound lender would see a cumulative loan volume reduction of approximately 25%. As discussed later, sensitivity and robustness analysis support these findings that many lenders struggled in the immediate aftermath of the caps being implemented. The quantitative results in this chapter also corroborate a less scientific, case-by-case, examination of the lender data in which we saw significant retrenchment and several eventual failures post IRR imposition. The exposed/bound institutions that did grow post cap also typically either (i) increased loan-linked fees and commissions, or (ii) only recently joined the market and were expanding their portfolios and market shares as part of this process.

These results add to the broad consensus in the literature that stronger IRR policies may have a repressive effect on market access (Cubillos-Rocha et al. (2018); Romero (2021); Aiba et al. (2020); Heng et al. (2021); and Roa et al. (2021)). These results add to the literature by showing that specific lenders who are/were more exposed to the IRR appear to have reduced their lending more, an important finding for better understanding how such policies have heterogenous effects. While Cubillos-Rocha et al. (2021) find that microlenders increased lending relative to banks when IRRs were reduced, we show that we might expect heterogenous responses within such a group of microlenders, and that such heterogeneity will be a function of their pricing (and market segment). Thus, we argue that countries considering IRRs should carefully consider the distribution of interest rate pricing in the market, and where the IRR will sit with respect to this distribution. The results also suggest – although do not prove – that a standard form of rationing may have been the dominant mechanism via which the Zambia's IRR operated.

Table 15 – Baseline Results – Piecewise Distributed Lage Models for Cap Implementation

Independent Variables		Anticipation			Static	Lagged Effects				
		(+3)	(+2)	(+1)	0	(-1)	(-2)	(-3)	(-4)	(-5)
Bound Lenders Slope Effects	Forward Lead (t+3)	-0.031 [0.0799]								
	Forward Lead (t+2)	0.011 [0.0503]	0.036 [0.0694]							
	Forward Lead (t+1)	0.039 [0.0811]	0.049 [0.0743]	0.006 [0.0772]						
	Static Effect (t=0)				-0.4216** [0.1899]	-0.4476** [0.1966]	-0.4519** [0.2030]	-0.4752** [0.2000]	-0.4751** [0.2224]	-0.4745** [0.2275]
	Lagged Effect (t-1)					-0.6668*** [0.2449]	-0.6862** [0.2616]	-0.7282** [0.2784]	-0.7301** [0.2764]	-0.9295*** [0.2802]
	Lagged Effect (t-2)						-0.291 [0.2778]	-0.311 [0.2771]	-0.316 [0.2783]	-0.323 [0.2899]
	Lagged Effect (t-3)							-0.378 [0.3244]	-0.372 [0.3142]	-0.391 [0.2960]
	Lagged Effect (t-4)								-0.095 [0.3346]	-0.081 [0.3412]
	Lagged Effect (t-5)									-0.153 [0.1490]
	Cumulative Effect	**0.019**	**0.085**	**0.006**	**-0.422**	**-1.114**	**-1.429**	**-1.892**	**-1.988**	**-2.351**
	Test - H0: Cumulative effect=0	**0.884**	**0.265**	**0.940**	**0.030**	**0.003**	**0.020**	**0.000**	**0.003**	**0.002**
Unbound Lenders Binary Effects	Forward Lead (t+3)	0.078 [0.0500]								
	Forward Lead (t+2)	-0.010 [0.0419]	0.000 [0.0529]							
	Forward Lead (t+1)	0.019 [0.0463]	0.009 [0.0443]	0.012 [0.0460]						
	Static Effect (t=0)				0.027 [0.0674]	0.023 [0.0673]	0.015 [0.0727]	0.024 [0.0780]	0.015 [0.0810]	0.002 [0.0839]
	Lagged Effect (t-1)					0.1271* [0.0740]	0.1419* [0.0789]	0.1516* [0.0839]	0.1492* [0.0847]	0.1749* [0.0890]
	Lagged Effect (t-2)						0.105 [0.0752]	0.085 [0.0745]	0.083 [0.0754]	0.076 [0.0786]
	Lagged Effect (t-3)							0.049 [0.0919]	0.021 [0.0917]	0.017 [0.0938]
	Lagged Effect (t-4)								0.008 [0.0970]	-0.032 [0.0999]
	Lagged Effect (t-5)									-0.026 [0.0674]
	Cumulative Effect	**0.088**	**0.009**	**0.012**	**0.027**	**0.150**	**0.262**	**0.310**	**0.275**	**0.212**
	Test - H0: Cumulative effect=0	**0.275**	**0.865**	**0.789**	**0.695**	**0.133**	**0.128**	**0.074**	**0.204**	**0.449**
Lender-specific Covariates	Log of Total Assets (t-1)	-0.0662*** [0.0185]	-0.0701*** [0.0182]	-0.0735*** [0.0197]	-0.0652*** [0.0188]	-0.0612*** [0.0173]	-0.0543*** [0.0171]	-0.0524*** [0.0189]	-0.0550*** [0.0190]	-0.0633*** [0.0197]
	Capital Asset Ratio (t-1)	-0.045 [0.0545]	-0.027 [0.0488]	-0.016 [0.0517]	0.012 [0.0577]	0.017 [0.0570]	0.034 [0.0659]	0.025 [0.0709]	0.035 [0.0784]	0.041 [0.0848]
	Loss Allowance Ratio (t-1)	0.108 [0.1409]	0.053 [0.1226]	-0.021 [0.1146]	-0.085 [0.1220]	-0.106 [0.1208]	-0.084 [0.1246]	-0.090 [0.1252]	-0.059 [0.1241]	-0.051 [0.1266]
	Operating Expense (t-1)	0.066 [0.1154]	0.025 [0.1267]	0.061 [0.0869]	0.072 [0.0853]	0.098 [0.0685]	0.099 [0.0606]	0.1191* [0.0606]	0.1169* [0.0602]	0.1276** [0.0557]
	Firm Age	0.000 [0.0022]	0.001 [0.0020]	0.002 [0.0018]	0.002 [0.0017]	0.002 [0.0018]	0.002 [0.0019]	0.003 [0.0020]	0.0038* [0.0023]	0.0051** [0.0024]
Macroeconomic Controls	Treasury Bill Rate (t-1)	0.0045* [0.0026]	0.004 [0.0028]	0.003 [0.0027]	0.003 [0.0027]	0.003 [0.0026]	0.003 [0.0025]	0.003 [0.0022]	0.003 [0.0023]	0.002 [0.0024]
	Annual Inflation (t-1)	0.000 [0.0027]	0.001 [0.0028]	0.001 [0.0027]	0.001 [0.0027]	0.001 [0.0027]	0.001 [0.0028]	0.002 [0.0027]	0.002 [0.0028]	0.002 [0.0029]
	Foregin Exchange Rate (t-1)	-0.005 [0.0086]	-0.007 [0.0086]	-0.010 [0.0077]	-0.0132* [0.0067]	-0.0150** [0.0067]	-0.0151** [0.0066]	-0.0204*** [0.0071]	-0.0234*** [0.0073]	-0.0259*** [0.0078]
	Copper Price (t-1)	0.0216* [0.0112]	0.0234** [0.0107]	0.0209** [0.0098]	0.0212** [0.0093]	0.0204** [0.0092]	0.0232** [0.0099]	0.0220** [0.0110]	0.0209* [0.0111]	0.0205* [0.0119]
	Constant	0.4754* [0.2393]	0.5120** [0.2331]	0.5649** [0.2235]	0.5005** [0.2139]	0.4814** [0.2060]	0.3808* [0.2270]	0.377 [0.2493]	0.412 [0.2638]	0.4950* [0.2853]
	F	8.819	7.884	11.242	8.903	7.919	7.127	8.422	15.721	14.497
	Observations	1096	1150	1207	1258	1258	1212	1158	1104	1048
	R-squared	0.064	0.054	0.059	0.067	0.087	0.088	0.092	0.097	0.114
	Lender Fexed Effects	√	√	√	√	√	√	√	√	√
	Time Fixed Effects	X	X	X	X	X	X	X	X	X

Note: Table 15 provides results from equation 5.8 across various treatment lengths. The dependent variable is growth in lenders' net real outstanding loan books. Each column provides results from a different regression and lag structure regarding the treatment variables (across both the bound and unbound lenders). For bound lenders the estimates are marginal effects, for unbound lenders, a binary treatment effect. The table begins by testing for policy leads (t-3) and finishes estimating a model allowing for five lags (t-5). The bolded rows provide cumulative effect parameter estimates, $\sum_{j=0}^{J}\beta^{j}$ and $\sum_{k=0}^{k}\beta^{k}$, as well as p-values from testing H0: $\sum_{j=0}^{J}\beta^{j}=0$, and are explained in detail in the text. Controls include both lender-specific macroeconomic control variables (defined in Table 14) and the regressions include lender fixed effects. Standard errors are robust and clustered at the lender level. ***p<0.01, **p<0.05, *p<0.10.

Robustness and sensitivity

To ensure that these results are not overly sensitive to modelling or data decisions, we re-run differing versions of equation (5.8) but under alternative settings/restrictions. This exercise is necessary given our small sample and relatively large treatment effects. Overall, we find that our results are acceptably stable and that they are unlikely a feature of any one specific modelling choice or data outliers. We investigate the following alternative settings/restrictions: (i) loosening the definition of a bound institution to be defined by the median institution (rather than by having direct exposure to the cap); (ii) allowing for policy variation due to the imposition, loosening, and removal of the cap, (rather than just the imposition); (iii) controlling for common trends by including time fixed effects (albeit at the loss of a clear parameter interpretation); and (iv) estimating the model removing outliers and covariates.

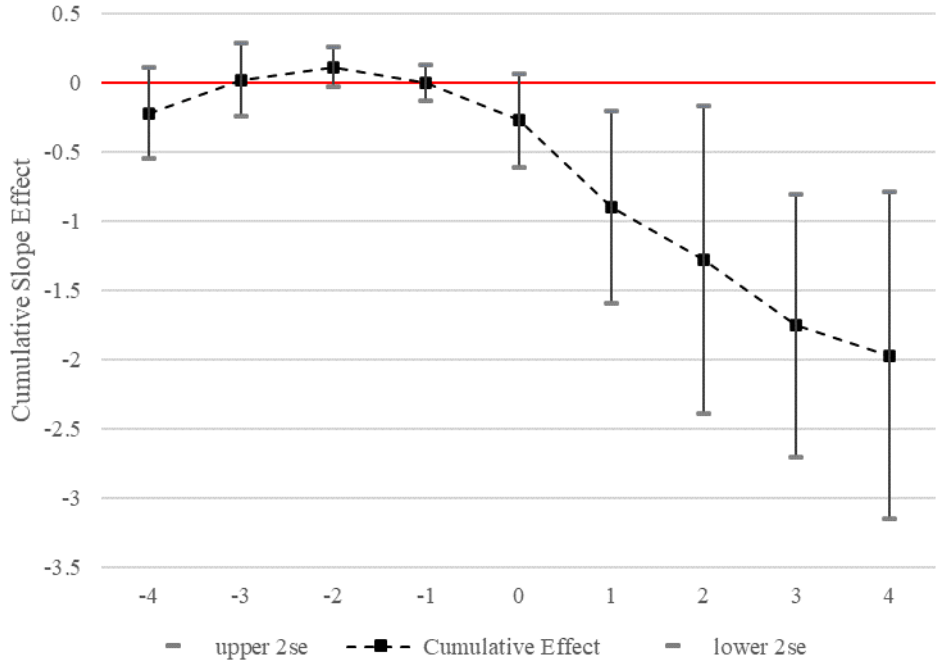
Redefining treatment - Adjusting the definition of IRR bindingness

A simplification of our theoretical framework and empirical approach is that they assume lenders make loans at only one interest rate, and it is this rate that defines whether or not that lender was bound by the cap. For lenders lending near the IRR, however, it is feasible that they held loans issued both above and below the IRR. If we want our primary cap reaction parameter to only capture loan portfolios bound by the IRR and not include loans with interest rates lower than the cap, we need a more conservative choice of interest rate boundary. Further, while we restrict lenders to react linearly to being bound, lenders pricing close to the cap might be better placed than those pricing much higher to use fees and other non-interest income sources to offset the IRR's effects. A more conservative interest rate binding definition will remove such lenders from the main "bound" sample.

Figure 21 provides results from re-running equation 5.8, but with redefining I_{it} and bound lenders by the median level of bindingness (35% bound), rather than a lender's interest rate being higher than the IRR. This approach reallocates approximately nine or so marginally bound lenders across groups and requires that:

$$I = \begin{cases} 1, & r_{i,\, t=2012q4} - r_{cap,t=2013q1} > 0.35 \\ 0, & r_{i,\, t=2012q4} - r_{capt,t=2013q1} \leq 0.35 \end{cases}$$

Figure 21 – Marginal Effect Piecewise Distributed Lag Models for Bound Lenders (Median Defined) – Cap Implementation



Note: The figure provides results from the distributed lag model regressing growth in lenders' net real outstanding loan books against the change in a binary treatment dummy for the imposition of the IRR policy. The estimates are for lenders for whom the cap bound by more than 35%. The figure shows cumulative estimates, $\sum_{k=0}^{k} \beta^k$, for equation 5.8, spanning over various lag structures (x-axis), whereby each x-axis period provides an estimate from its own regression. Controls include both lender-specific macroeconomic control variables, as per Table 15. The regressions include lender fixed effects. Standard errors are robust and clustered at the lender level. The y-axis is given in decimals. Confidence interval at 95%.

If we compare Figure 21 to Figure 19, we see that the baseline results do not appear to be sensitive to changing the group definition to the median interest rate (i.e. switching lenders from the bound group to the unbound group at the IRR bindingness margin), suggesting that the parameter estimates are primarily driven by lenders pricing substantially above the interest rate cap. The mechanics of the fixed effects estimator in a continuous treatment setting might also be contributing to this lack of sensitivity, given that lenders who are 'close' to the cap will not exhibit much policy variation and so will not be highly weighted in the distributed lag (DL) estimates. The magnitude and profile of the non-bound lenders remain consistent with Figure 20.

Analysing the imposition, loosening, and removal of the IRR

Our baseline results rely on variation in the cap policy due solely to the imposition of the IRR. This was because the IRR's imposition was the only source of policy variation that we could be comfortably certain was (i) not anticipated, and (ii) not driven by contemporaneous shifts in the

macroeconomic environment. This assumption sits at the heart of our identification strategy. Yet the simple theoretical framework does suggest that the loosening and removal of the cap in mid-2014 and late-2015 should result in a symmetric response by lenders, whereby bound lenders increasing lending after an IRR removal. Ideally, our results should remain directionally consistent if we were to adjust 5.8 and allow for all source of policy variation (as follows)[147]:

$$\Delta Y_{it} = \alpha_i + \sum_{j=d}^{J} \beta^j . I . \Delta[T . (r_{i,\, t=2012q4} - r_{cap,t})] + \sum_{k=d}^{K} \beta^k . (1 - I) . \Delta T + \sum_{l=1}^{L} \beta^l X_{it}^l + \sum_{m=1}^{M} \theta^m M_t^m + \varepsilon_{it} \qquad (5.9)$$

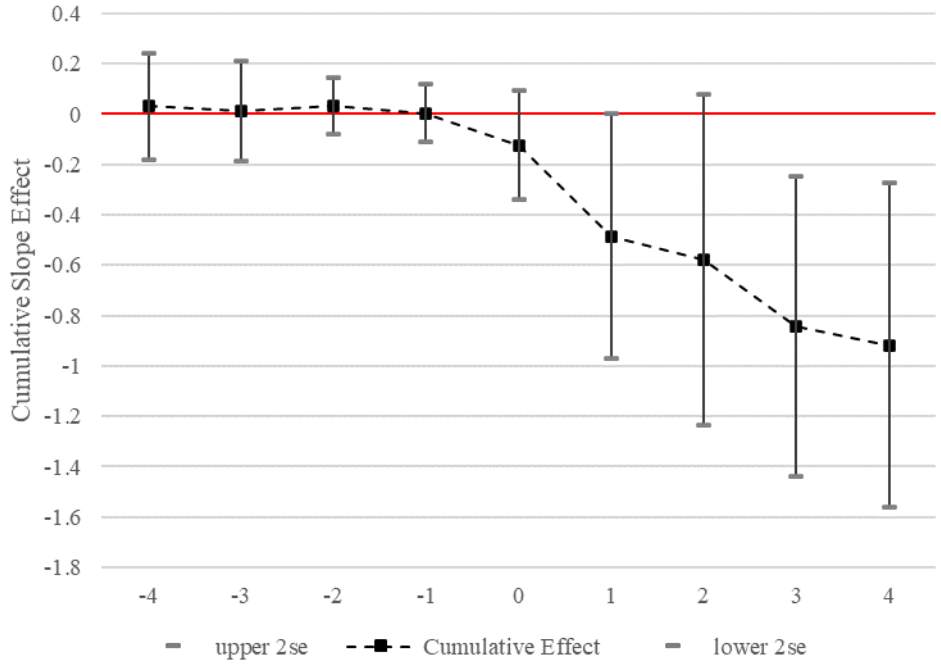
$$I = \begin{cases} 1, & r_{i,t=2012q4} - r_{cap,\, t=2013q1} \geq 0 \\ 0, & r_{i,t=2012q4} - r_{cap,\, t=2013q1} < 0 \end{cases}$$

$$T = \begin{cases} 1, & r_{cap,t} > 0 \\ 0, & r_{cap,t} \leq 0 \end{cases}$$

Figure 22 shows that the results remain consistent if we allow for all sources of cap policy variation. The estimates tell a directionally similar story, although the point estimates are lower in magnitude. This suggests that the variation in growth due to the removal of the caps might not be as clearly associated with loan growth as when the IRRs were imposed. In saying this, the loosening and removal of the cap in 2014 and 2016 respectively were likely policy decisions made endogenously to the contemporaneous macroeconomic environment, and so we need to be sceptical of whether we can reliably learn much from this variation.

[147] Of course, this symmetry relies on the assumption that lenders did not change their business models over time in response to the cap, which seems implausible.

Figure 22 – Marginal Effect Piecewise Distributed Lag Models for Bound Lenders – All Cap Variation

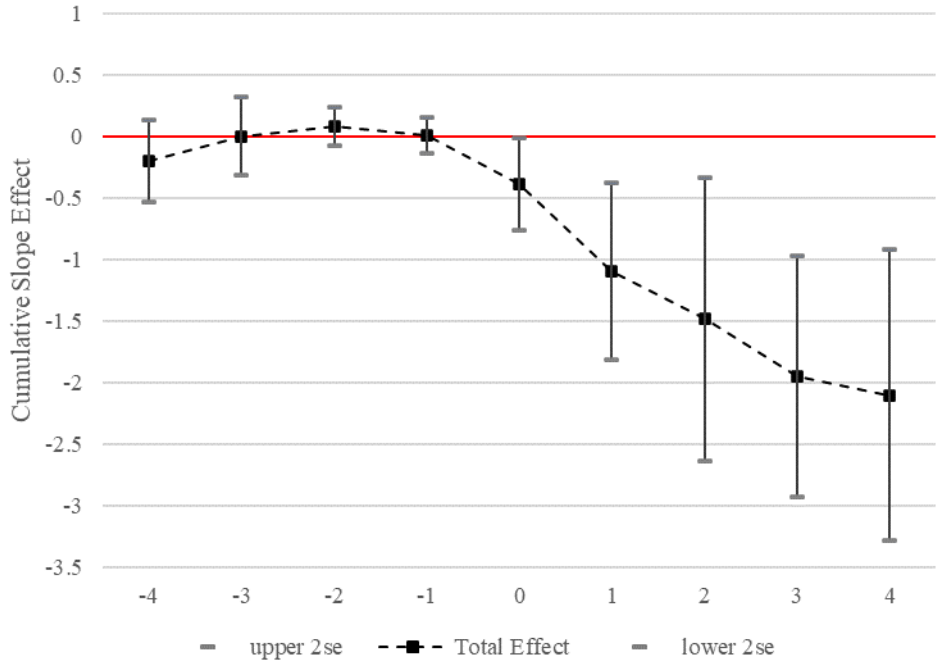


Notes: The figure provides results from the distributed lag model regressing growth in lenders' net real outstanding loan books against the change in its cap bindingness due to the imposition, loosening, and removal of the IRR policy, as well as any changes in the BOZ reference rate. It captures estimates for bound lenders only. The figure shows cumulative marginal effect slope parameter estimates, $\sum_{j=d}^{J} \beta^{j}$, for equation 5.9, spanning over various lag structures, whereby each x-axis period provides an estimate from its own regression. Controls include both lender-specific macroeconomic control variables, as per Table 15. The regressions include lender fixed effects. Standard errors are robust and clustered at the lender level. Confidence interval at 95%.

Controlling for common time trends

In our baseline results, we control for common macroeconomic factors in our analysis by including lagged values of the Treasury Bill rate, inflation rate, foreign exchange rate, and the copper price. Unfortunately, the sparsity of other quarterly macroeconomic data series in Zambia limits what we can control for and hence limits our confidence in the approach – there is a risk that we may be misattributing other common lender shocks that might correlate with policy changes in the cap. One solution to this is to include time dummy variables in 5.8, α_t, which will absorb the common component of lender shocks, essentially transforming equation 5.8 into a form of Two-Way Fixed Effects (TWFE) regression. The cost of this approach is that it is harder to intuitively interpret the coefficients or be certain what variation is driving the results (Kropro and Kubinec, 2020). Nevertheless, our baseline results are robust to modelling common trends in a TWFE setup (see Figure 23).

Figure 23 – Marginal Effect Piecewise Distributed Lage Models for Bound Lenders – Cap Implementation and Two-way Fixed Effects



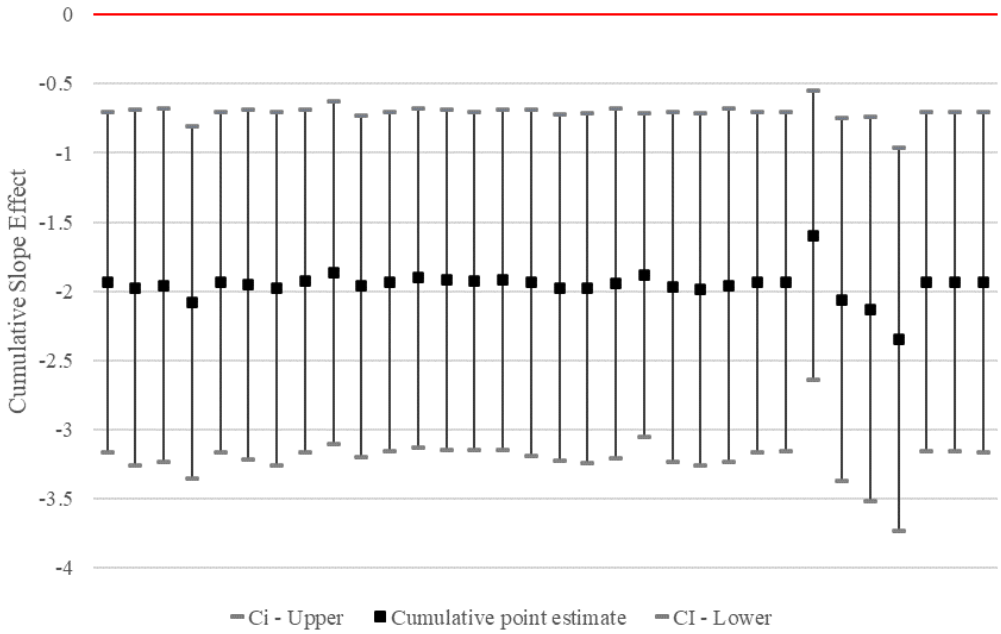
Notes: The figure provides results from the distributed lag model regressing growth in lenders' net real outstanding loan books against the change in its cap bindingness due to the imposition of the IRR policy. It captures estimates for bound lenders only, but allows for IRR policy variation due to the imposition of the cap. The figure shows cumulative marginal effect slope parameter estimates, $\sum_{j=0}^{J} \beta^{j}$, from a version equation 5.8 which includes a time dummy variable, α_t. It spans over various lag structures, whereby each x-axis period provides an estimate from its own regression with a lead/lag structure of that dimension. Controls include lender-specific control variables only (lender-specific controls can be found in Table 15). The regressions include lender fixed effects. Standard errors are robust and clustered at the lender level. Confidence interval at 95%.

Removing institutions and variables

We want to ensure that no individual lender drives the qualitative narrative in our results. Given our relatively small sample, checking for outlier sensitivity is vital. Similarly, we also want to ensure that certain groups of lenders, covariates, or regression settings are not driving the results. Overall, our results appear robust in this respect.

We begin by testing that no one lender drives the results. We run equation 5.8 approximately 30 times, in each regression removing one lender (with replacement). Ideally the point estimates should be reasonably stable. Figure 24 provides the results for the point estimate for the cumulative effect of the IRR on real net lending growth over five quarters post implementation. There are minor deviations in the point estimates, but there is little evidence that the results are being driven by a single lender.

Figure 24 – Marginal Effect Piecewise Distributed Lage Model for Bound Lenders After Five Quarters for Cap Implementation Only with Each Lender Removed (with replacement)

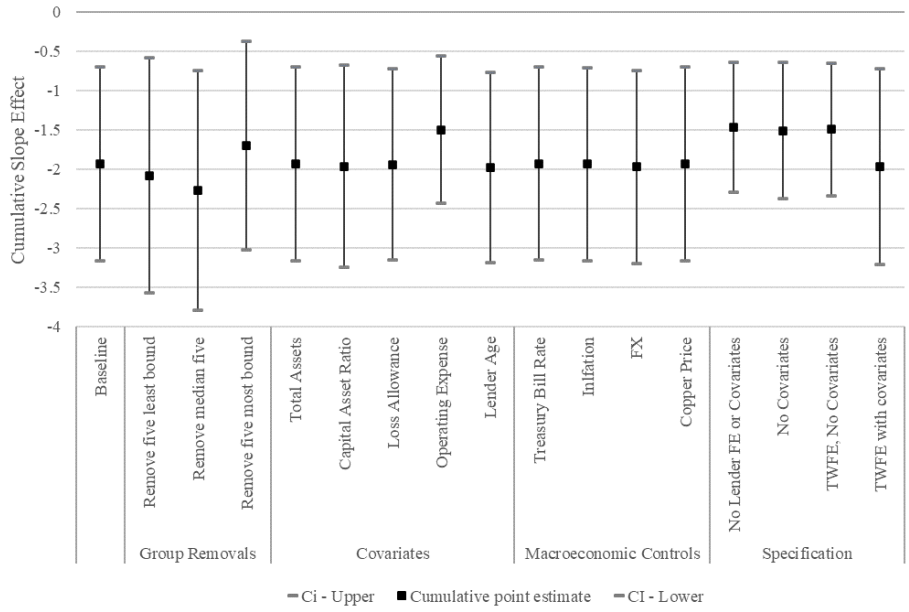


Notes: The figure provides results from the distributed lag model regressing growth in lenders' net real outstanding loan books against a change in lenders' cap bindingness due to the imposition of the cap. Each estimate is from a regression of equation 5.8 (with five policy variable lags) but with one lender removed (with replacement) as we move along the x-axis. Each estimate provides the cumulative marginal effect slope parameter, $\sum_{j=0}^{5}\beta^{j}$. Controls include lender-specific and macro control variables (specific controls can be found in Table 15). The regressions include lender fixed effects. Standard errors are robust and clustered at the lender level. Confidence interval at 95%.

Figure 25 provides the baseline results across several of the other sensitivity tests. First, we removed groups of five lenders from across the cap exposure distribution: the five least exposed, the five most "centrally" exposed, and the five most exposed. Our point estimate results are reasonably stable, especially in terms of the overarching narrative. This is a reassuring result, as it shows us that it is not just the very exposed lenders driving our estimates. Next, we run the baseline regressions excluding each covariate one at a time. There is some sensitivity to the inclusion of lenders' operating expenses in the baseline regression, but not in a sense that changes the narrative direction of the results[148]. Using other regression specifications also does not drastically change the point estimates.

[148] Operating expense is defined in this study as non-interest expenses as a proportion of net loans. It captures the non-interest cost per Kwacha of loan lent. It includes brick-and-mortar expenses, as well as loan write-off expenses. As some lenders entered financial distress and began to retrench their lending post cap, non-interest expenses may not have scaled back with the reduction in lending, driving this ratio up (and creating a correlation with the IRR and its imposition).

Figure 25 – Marginal Effect Piecewise Distributed Lage Model for Bound Lenders After Five Quarters for Cap Implementation Under Various Sensitivity Scenarios



Notes: The figure provides results from the distributed lag model regressing growth in lenders' net real outstanding loan books against a change in lenders' cap bindingness due to the imposition of the cap. Each estimate is from a regression of equation 5.8 (with five policy variable lags) but with different sensitivity analysis or specifications as we move along the x-axis. Each estimate provides the cumulative marginal effect slope parameter, $\sum_{j=0}^{5} \beta^j$. Controls include control variables (controls can be found in Table 15). TWFE regressions are run without macro controls. All regressions include lender fixed effects. Standard errors are robust and clustered at the lender level. Confidence interval at 95%.

Overall, our DL results argue that the imposition of the IRR in Zambia likely resulted in bound lenders retrenching their loan portfolios, especially for those who were lending in high interest rate segments of the market. For the median exposed lender, for example, we estimate that the loans volumes were approximately 70% lower one year after the imposition of the IRR, compared to if that lender had not been bound. The robustness and sensitivity analysis argues that, accepting the econometric and modelling assumptions we have made, the results are stable.

In the next section of this chapter, we move to examine the IRRs under a different and alternative set of econometric and modelling assumptions.

Alternative Approach – Difference-in-difference and Lending Growth

We now take a different approach to evaluating Zambia's cap regime and employ a difference-in-difference (DID) method. Our previous DL results relied only on within-lender variation in IRR bindingness and within-lender variation in loan portfolio growth to estimate how IRRs impacted

lending[149]. Here we again want to evaluate how lenders' outstanding loan portfolios changed when the IRR was imposed, but this time we want to do so against an arguably stronger counterfactual of what the lender's outcome would have been had they not faced the IRR (based on a counterfactual projection and the experience of lenders who were not bound by the cap). The primary assumption underpinning this approach will be that the lenders who were pricing below the IRR when it was implemented can be assumed to have not been affected by the cap, and hence act as an untreated control group.

There are inherent benefits to using a DID approach over the previous distributed lag (DL) modelling. Unlike the fixed effects DL model, a DID approach should help to provide more confidence that time-varying common shocks/trends in MFI loan growth are not driving our results. This is a legitimate concern in the DL modelling which relies solely on a small set of macroeconomic variables to capture time-varying shocks[150].

The DID analysis results corroborate the qualitative narrative from the DL analysis. We find that the median bound lender in the market, pricing approximately 36% above the cap, would have respectively reduced their lending by the end of the year by more than 60% compared to what they might have otherwise, with a 95% confidence interval implying a lending reduction of between 30% and 120%.

Exploring Parallel Trends

Before we begin with formal estimation, we first want to establish the baseline plausibility of the parallel trend assumption – one of the fundamental assumptions underpinning the DID analysis and our results. The validity of this assumption sits at the heart of event study and DID identification. For our IRR policy, the assumption requires that absent the imposition of the IRR, lenders bound by the IRR (treated lenders) will on average follow the same trend as lenders who were not bound by it (the control lenders)[151]. Later, we will test for a difference in pre-trends within our regression framework, but for now we want to show that such an assumption is at least plausible given our data. To do so, we compare the outcomes of our two groups over time.

[149] In other words, how did a lender's growth rate change when the IRR was implemented, compared to its own growth experience before the IRR was enforced.
[150] In saying this, a DID clearly relies on other assumptions regarding the credibility of bound lenders' hypothetical counterfactual.
[151] This is the simple case in which the IRR is modelled as a binary treatment effect.

Figure 26 plots the average log real outstanding net loan series for the lenders who were pricing below the IRR (red) and lenders pricing above (blue). The first figure covers the period over the event window for which the panel is reasonably balanced and compares trends across levels; the second provides a levels comparison over a longer time-period but suffers some composition effects[152]. Both figures show that bound and unbound lenders saw strong growth leading up to the IRR's imposition in 2013 and a reasonably clear deviation in growth trends soon after IRR was implemented, suggesting that a DID approach might be reasonable. Visually, there appears to be co-movement between the series, suggesting similar reactions by both lender types to common shocks. The convergence in growth trends post 2016 (when the IRR was removed) align with the idea that bound lenders expanded post cap-removal, while unbound lenders became relatively more caught up in the inflation and foreign exchange crisis of 2015-16.

Figure 26 – Common Lending Trends for Bound (treated) and Unbound Lenders (control)



Note: The Figure plots the average log real outstanding net loan series for the lenders who were pricing below the IRR (red) and lenders p above (blue). The first figure covers the period over the event window for which the panel is reasonably balanced and compares trends levels; the second provides a levels comparison over a longer time-period but suffers some composition effects. The key difference betwe two are that the LHS graphic uses a dual axis to show how the trends between bound and unbound lenders compared leading up to th impositions. The RHS graphic shows the level difference between the two and how they trend over a longer time horizon (but with a less sample). The red vertical lines indicate a change in IRR regime.

[152] Due to the longer time period the longer sample is not balanced and will exhibit some compositional effects, especially towards the end of the period. The first red vertical line indicates the imposition of the IRR in 2013, the subsequent two the loosening and removal.

One challenge for our analysis, highlighted by Figure 26, is the large level difference in average log loans that exist between the bound and unbound groups – a magnitude of 4-5 times on average, in Kwacha terms. Such a size difference might prove problematic for our DID approach as we need to ensure that, conditional on covariates, the mechanisms that led to this difference in loan levels are not also related to differences in growth trends. The size difference derives directly from the fact that in Zambia's credit market, smaller lenders typically lend in higher interest rate markets, and so the lenders bound by the IRR will naturally be smaller on average (increasingly so as IRR bindingness increases). So, to the extent that there are inherent scale economies in loan production or differences in fundamental growth potential in market segments, parallel trends between the two might be questionable. The existence of market power or network economies (whereby an incumbent lender might grow easily) would also be problematic. A traditional "balance table" is provided in Annex 13, and there are clearly level differences in characteristics for bound vs. unbound lenders.

Yet, there are several structural reasons to draw comfort from our groupings and to be less concerned with the size differential between bound and unbound lenders. On the supply side of lending, non-banks across the size spectrum will face similar input markets (financing, labour, etc) and so would respond similarly to shocks in such markets. Likewise, while MFI demand might be segmented, we expect positive covariance in loan demand across borrowers and across economic sectors (a copper shock would likely impact Bayport's public servant clientele as well as lenders providing to lower-income borrowers; poor hydrology might impact both smaller farms and larger producers (via electricity input costs). The simultaneous dip in loan growth in 2009 after the Financial Crisis and the contemporaneous growth from 2010 to 2013 provide some reassurance that the two groups respond similarly to their macroeconomic environment. Finally, all lenders are subject to the same set of regulatory rules (and changes to these rules), all are regulated by the same institution (BOZ), and all face the same business and lending environment (tax structure, levels of corruption, court bankruptcy delays, usefulness of the credit reference bureau, collateral registry, etc). All these provide suggestive evidence that the parallel trend assumption might be reasonable.

Identification and parameter interpretations

In this chapter, we aim to evaluate whether binding IRRs influenced non-bank lending and whether lenders more bound by the cap were incrementally affected. A standard observational binary DID analysis typically attempts to estimate an average treatment effect on the treated (ATT) – or some

related estimand. Zambia's IRR regime exposed lenders to the IRR differentially based on where in the market they were before its implementation. As a result, in our analysis we will target estimating a form of marginal effect slope parameter, in this case the Average Causal Response on the Treated, $ACRT(\mathrm{t}, z|z)$[153].

The $ACRT(\mathrm{t}, z|z)$, estimates the incremental effect of the z-th percent exposure to the cap on a lender's outcome, given they experienced a cap exposure z at time t. It is a meaningful parameter because it allows us to make statements about how changing the IRR severity for a given lender impacts its loan outcomes on the margin[154].

We can estimate the $ACRT(\mathrm{t}, z|z)$ within a standard dynamic continuous treatment DID event study design, with the approach relying on several analogous assumptions as the canonical DID setup, namely random sampling, support, and no anticipation[155]. The $ACRT(\mathrm{t}, z|z)$ however, requires a more stringent form of the parallel trends assumption; i.e., the average change in outcomes across time and across lenders (given a certain amount of exposure to the IRR) must be the same as the average change in outcomes over time for all units that actually received that exposure (Callaway et al., 2021)[156,157].

How reasonable are these assumptions? We think that the no anticipation assumption is reasonable given that (i) we found no media coverage discussing IRRs before they were implemented, and (ii) in our background contextual research we have not come across any evidence suggesting that BOZ gave warning to, or had detailed discussions with, lenders prior to the IRRs. We test for pre-trends

[153] We are primarily concerned with investigating the effects of the implementation of the IRR, rather than the loosening or removal of the cap. This is because we can be more certain that the cap's imposition was not anticipated by lenders – we can't be sure of this with the loosening and removal of the cap. This setting also means that we are only estimating a more standard, continuous treatment, non-staggered DID setup that includes a set of never-treated units, simplifying analysis (and perhaps avoiding some of the challenges posed by the new DID literature, see Roth et al. (2022) for a summary).

[154] In the past, continuous DID has typically been seen as an extension of binary TWFE and analysts have often interpreted continuous DID slope parameters as a causal effect of a change in doses. However, until recently there has been limited research concerning identification and interpretation of dose-response DID designs (Callaway et al. 2021).

[155] See Callaway et al. (2021) pg. 22 for the formal assumptions required for the $ACRT(\mathrm{t}, z|z)$.

[156] Note that this involves potential outcomes under various exposures, z, rather than only untreated potential outcomes.

[157] Intuitively, this different and stronger parallel trends assumption is required because we will implicitly be making comparisons between lenders with different IRR exposures, rather than only making comparisons between bound lenders and unbound lenders. Another option would be to assume homogenous lender responses to IRR bindingness, which would allow identification (Callaway et al. 2021). Our theoretical framework (conveniently) assumes away any heterogeneous effects across exposures/dosages because all lenders face the same linear marginal cost curve - in essence, lenders' ATT (t,z|z) functions are the same line and the slope of this line (i.e. the ACRT(z|z)) will be constant and equal). So, if we were to take the framework seriously – a big if – this would allow us to identify the ACRT(t, z|z). Our other option is to assume the strong parallel trend assumption.

equality in our results, something that should hold if we think the parallel trends assumption is reasonable.

Empirical implementation

We implement a version of an event study DID that draws directly from following Freyaldenhoven et al. (2021), a reasonably standard implementation that allows for a continuous treatment dosage. We estimate a version of:

$$Y_{i,t} = \alpha_i + \gamma_t + X'_{it-1}\psi + \sum_{m=-G}^{M} \beta_m \, . \, z_{i,t-m} + \varepsilon_{it} \qquad (5.10)$$

$$z_{i,t} = \begin{cases} T \, . \, r_{i,\, t=2012q4} - r_{cap,t=2013q1}, & r_{i,\, t=2012q4} - r_{cap,t=2013q1} > 0 \\ 0, & r_{i,\, t=2012q4} - r_{capt,t=2013q1} \leq 0 \end{cases}$$

$$T = \begin{cases} 1, & t \geq 2013q1 \\ 0, & t \leq 2012q4 \end{cases}$$

Where Y_{it} is the log of a lender's outstanding net real loans. $z_{i,t}$ is a treatment variable, set equal to how bound a bound lender was when the IRR was implemented, and zero otherwise (i.e. $z_{i,t}$ will take 0 for all lenders before the cap was implemented, and after the cap was implemented it will take $r_{i,\, t=2012q4} - r_{cap,2013,1}$ for lenders for whom $r_{i,\, t=2012q4} - r_{cap,2013,1} > 0$ and 0 for a lender for whom $r_{i,\, t=2012q4} - r_{cap,2013,1} < 0$). To help build $z_{i,t}$, we use a binary indicator, T, to indicate whether there was an IRR in force at time t. M defines the periods over which we assume the outcome can be affected by the treatment after it was implemented (M≥0), while G defines the pre-policy periods over which we assume the policy can affect the outcome (G≥0, but we typically assume it is equal to zero). We test for evidence of differences in pre-trends in the results section. The parameters we are most interested in are the $\{\beta_m\}_{m=-G}^{M}$, which we summarise as an approximate cumulative percent growth effect for a given exposure and time path, $\sum_{m=-G}^{k} \hat{\beta}_m$, for different horizons, k. We plot these summary parameters in our event study graphics (see Figure 28). Note that the actual implementation of Freyaldenhoven et al. (2021) is slightly more involved than equation 5.10 and is discussed in Annex 14

The model is fully saturated with policy variables, so we normalise on the period before the policy was implemented (k=-1).

Results

IRR implementation – is there evidence that the IRR regime was binding?

Before we examine how the imposition of the IRR affected the volume of non-bank lending between bound and unbound lenders, we first need to have confidence that the cap was actually implemented, that the imposition of the cap persisted, and the cap resulted in an exposure differential between our bound and unbound lenders. These are important because if bound non-banks continued to lend at high rates after the imposition of the cap, we might not expect to see a significant difference in a lender response between bound and unbound lenders.

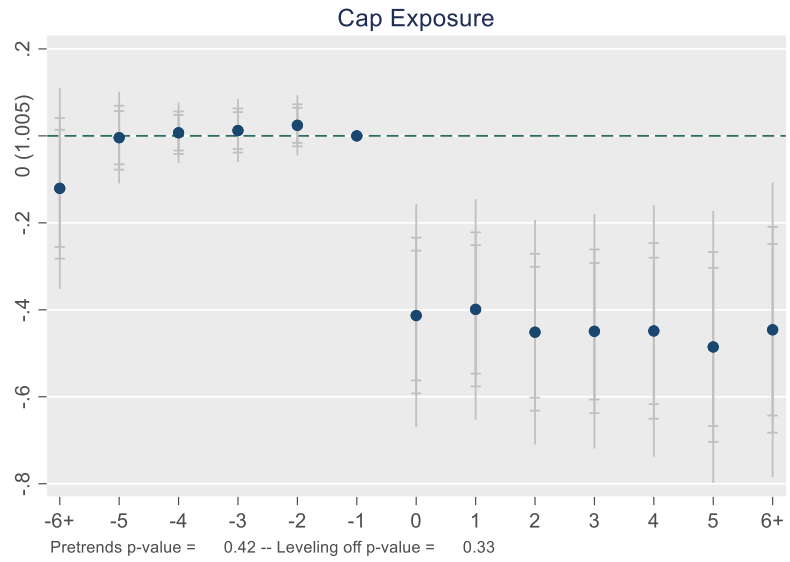
We hence track how the average level of recorded non-bank interest rates changed over time and across the imposition of the IRR by regressing a binary treatment version of equation 5.10, whereby $z_{i,t}$ takes a simpler form:

$$z_{i,t} = \begin{cases} T, & r_{i,\, t=2012q4} - r_{cap,t=2013q1} > 0 \\ 0, & r_{i,\, t=2012q4} - r_{capt,t=2013q1} \leq 0 \end{cases}$$

Figure 27 provides the results. There is strong evidence that the policy was binding on higher interest rate lenders. There is a distinct drop in the cumulative difference in recorded interest rates between bound and unbound lenders, suggesting that the cap did result in significant and lasting exposure for the bound group. Indeed, in the quarter after the cap, the average interest rate for bound lenders dropped nearly 40% more, on average, compared to unbound lenders[158]. The cumulative difference also appears to persist. This gives us confidence that the IRR policy was implemented and that bound lenders were exposed to the policy.

[158] Note that this is also a conservative estimate as some exposed lenders stopped completing the interest rate survey in the year after the interest rate cap was implemented and we are using un-adjusted BOZ data in Figure 27.

Figure 27 – Cumulative Difference in Recorded Interest Rates Between Bound and Unbound Lenders



Note: The figure provides coefficient estimates of $\sum_{m=-G}^{k} \hat{\beta}_m$ from equation 5.10 using lenders' average interest rates as the dependent variable and a binary treatment variable, for different horizons, k (x-axis). The comparison group is formed from lenders pricing below the IRR in Q4 2012. The Y-axis shows percentage point differences. Pointwise confidence intervals provided at 90% and 95% (horizontal lines grey lines), and the 95% uniform confidence band (vertical line grey line). Standard errors clustered at lender level. No covariates. The approach follows Freyaldenhoven et al (2021). See Annex 14 – Event Study Implementation, for additional details. The Pretrends p-value statistic provides a Wald test for evidence of pretends.

Lending Outcome results

In summary, we find evidence that compared to an unbound counterfactual, the imposition of a binding IRR negatively impacted loan growth for bound Zambian lenders, corroborating the results from the DL model. We again find that lenders that had more exposure to the cap reduced their lending by more than lenders with less exposure to the cap. The results also appear robust to simple parallel trends tests and other sensitivity analyses discussed below. However, and perhaps predictably, the high levels of uncertainty in our previous DL estimates also persist in these results, a reminder that we need to frame the results as directional rather than being precise point estimates.

Equation 5.10, implemented as in Annex 14, produces our baseline results. We find that for each percentage point that a lender was bound by the IRR, it reduced its lending by 1.5 percent after a year, as compared to its unbound or less bound counterfactual[159]. Figure 28 provides the coefficient plot for the estimates $\sum_{m=-G}^{k} \beta_m$ for different horizons k both with and without covariates, which

[159] This is the baseline ACRT(t,z|z) result and requires assuming homogeneity in treatment response for a given dose or the stronger cross-exposure parallel trend assumption.

provide the percent growth effect of the IRR implementation per percentage point of cap bindingness over the event window.

From an individual lender perspective, these results imply that a non-bank pricing at the lower end of the pre-IRR interest rate distribution, say at 10 percentage points above the IRR threshold, would expect to grow by 15% lower after five quarters than what it might have otherwise[160]. The median lender in the market, however, pricing approximately 35% above the cap, is estimated to have reduced their lending by the end of the year by more than 53% (1.5*0.35), with 95% confidence interval implying a lending reduction of between 30% and 120%, compared to what they might have otherwise[161]. Allowing for a linear group-specific trend difference as suggested in McKenzie (2020) does not meaningfully alter the results.

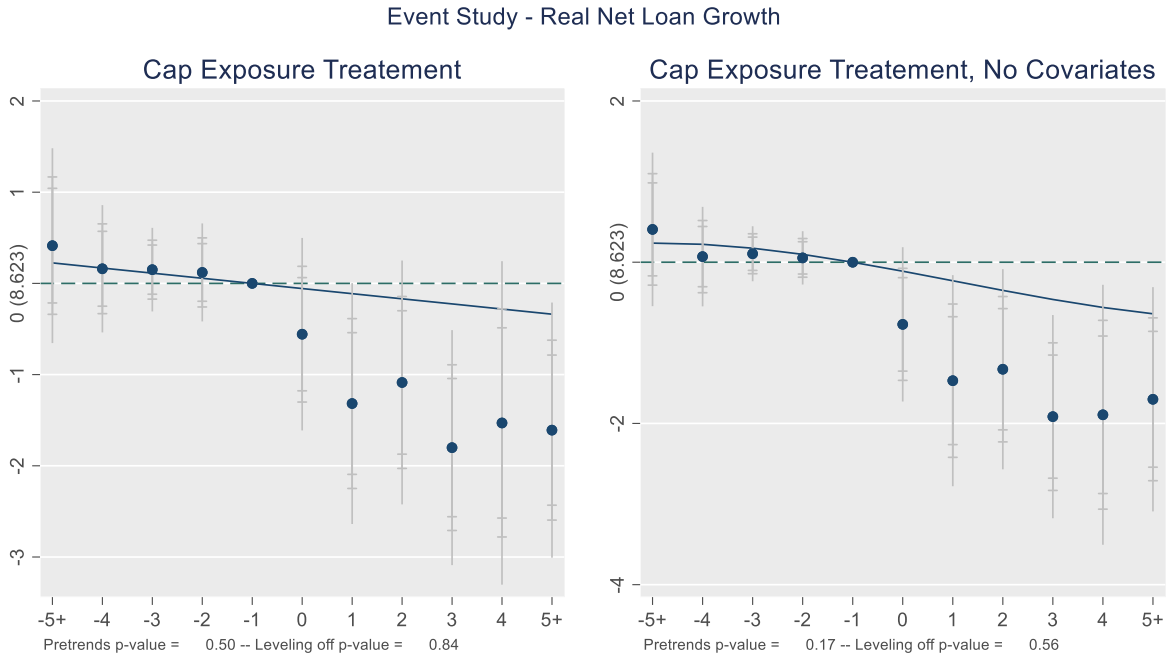
Economically, the results are meaningful. During the Global Financial Crisis (GFC), aggregate real net loans reduced by approximately 25% and the average reduction in cumulative loan growth across established non-banks was approximately 45%. For many non-bank lenders, the IRR thus had similar and potentially worse effects on loan volumes than the GFC.

The results support the findings from the DL modelling in the previous section that relative reductions in lending appear to be associated with the intensity of IRR bindingness. They also support the broad consensus in the literature that IRRs can have repressive effects on credit markets (Cubillos-Rocha et al., 2018; Romero, 2021; Aiba et al., 2020; Heng et al., 2021; and Roa et al., 2021). The confidence intervals around our point estimates remain large; given our small sample and the uncertainty associated with it, we should consider these findings as directional rather than precisely estimated point estimates.

[160] 15%=1.5*0.1, with 1.5 being the cumulative coefficient for at the end of five quarters in the right-hand event plot in Figure 28Figure 24)

[161] The 120% reduction in growth seems large but remember that it is a counterfactual comparison to how they would have grown otherwise, which can be <-100%. The IRR also likely placed several lenders on a path to eventual failure, which drives some of the low growth rates. Robustness analysis is provided in the next section. Note also that we are approximating cumulative growth rates by comparing log differences, which does become inaccurate with the large effects we are registering here and will overestimate the coefficient compared to the correct geometric growth rate interpretation.

Figure 28 – Baseline Event Study Results for IRR Implementation



Note: The figure provides coefficient estimates of $\sum_{m=-G}^{k} \hat{\beta}_m$ from equation 5.10. The left figure includes lender covariates, the right figure does not. The dependent variable for both is the log of outstanding real net loans and the point estimate coefficients follow the time path for the marginal effect slope parameters. A comparison group is formed from lenders pricing below the IRR in Q4 2012. Pointwise confidence intervals are provided at 90% and 95% (horizontal grey lines), and the 95 percent uniform confidence band (vertical grey line). Standard errors clustered at lender level. Covariates in the left graphic include lagged values of the capital asset ratio, loss allowance ratio, operating expense ratio, and lender age (see Table 14 for definitions). The approach follows Freyaldenhoven et al (2021). See Annex 14 – Event Study Implementation, for additional details. The Pretrends p-value statistic provides a Wald test for evidence of pretends.

The parallel trends assumption, pre-trends, group-level trends, and plausible confounds

Our results are contingent on the parallel trends assumption holding, which we believe is plausible. Visually, the lack of a strong pre-event trend in log outstanding loans in the estimates leading up to the implementation of the IRR in Figure 28 implies that such an assumption is plausible, even if not convincingly so. To more formally test for pretrends, we run a Wald test for evidence of pre-trends (H_0: $\delta_k = 0, -G - L_G - 1 \leq \mathrm{k} < -G$, where G is set at zero and L_G is set at 4). This tests whether there is evidence of pre-trends in the 4 periods before the cap was implemented. The outcome of the Wald test (provided at the bottom of each set of estimates in Figure 28) confirms we cannot reject the null of no pre-trends with a p-value of 50% (covariates included) and 17% (no covariates), although given the large confidence intervals around the pre-treatment point estimates the tests are weakly powered. But at the least it is plausible that there are no strong pre-trends, and this plausibility seems to broadly hold across other specifications in the sensitivity analysis.

Note, however, that not rejecting the null of no pre-trends, as we did above, does not necessarily mean that there is no confound or there is no pre-trend (Roth, 2021). As further discussed in Freyaldenhoven et al. (2021) p10, "any estimated event-time path for the outcome can be reconciled with a devil's-advocate model in which the policy has no effect on the outcome…[where] all of the dynamics of the estimated event-time path of the outcome variable are driven by an unmeasured confound." In such a case, the policy might plausibly have no effect on the outcome, even if the event path makes it seem so. We thus follow the recommendation to estimate a line representing the "most plausible" confound that is "statistically consistent with the estimated event-time path" of the loan growth rates and think through its plausibility (a linear devil's advocate confounder line, for example, might be concerning). The confound devil's advocate line in our results – a downward sloping line – is arguably a reasonable one. We hence cannot be overly certain that our results are not driven by unmodelled confounds or pre-trends, highlighting some of the challenges with low power and a small sample size inherent in our analysis.

Robustness and sensitivity

How stable are these results, and are they robust to sensible changes in modelling assumptions and treatment/control groupings? Does any individual lender drive the results? In this section we undertake a similar excise to our DL analysis and show that the event study results remain directionally consistent under other reasonable model specifications, as well as when removing outliers. Because of our small, noisy sample, and the fact that the policy treatment appears quite severe for some lenders, it is additionally important to check that the results appear stable under other reasonable modelling and data assumptions. Below, we test our results allowing for: differing control and treatment groups, modelling the IRR as a binary treatment, using very short and very long estimation windows, and modelling all policy variation (not just the imposition of the cap). The analysis below mirrors much of the robustness analysis from the DL modelling. Replications of the baseline results including linear group time trends are also consistent (Annex 15).

Redefining treatment and control groups

Many lenders in our sample hold a distribution of loans at different interest rates, rather than lending at a single simple average loan rate as we assume in our dataset and theoretical framework[162]. This

[162] In other words, they might use some risk/price interest rate differentiation at the borrower level.

means that some lenders lending at interest rates close to the implemented IRR might hold a mixture of loans priced both above and below the cap and might be misclassified[163].

To be conservative, we want to be quite sure that these lenders are not driving the core direction of our results, although using a continuous exposure setting compared to modelling a binary treatment already somewhat helps limit this because such lenders will already be receiving only a small exposure dosage. Nevertheless, as a robustness check, we try to remove the likeliness that lenders that have loans on both sides of the cap are driving the results by re-running equation 5.10 and arbitrarily removing 6 lenders pricing close to the cap (between -5% under and 20% over). More concretely, we adapt equation 5.10 such that $z_{i,t}$ is redefined:

$$z_{i,t} = \begin{cases} T.r_{i,\,t=2012q4} - r_{cap,t=2013q1}, & r_{i,\,t=2012q4} - r_{cap,t=2013q1} \geq 0.2 \\ removed, & 0.2 > r_{i,\,t=2012q4} - r_{capt,t=2013q1} \geq -0.05 \\ 0, & r_{i,\,t=2012q4} - r_{capt,t=2013q1} < -0.05 \end{cases}$$

Figure 29 (row 1) shows that, while the estimates are using less data and are less certain than the baseline case, the broad narrative remains.

Another implicit modelling assumption in our baseline analysis is that the line between whether a lender will ration their supply or not due to the IRR lies at $r_{i,\,t=2012q4} - r_{cap,t=2013q1} > 0$. However, this is an outcome of the simple theoretical framework and may not be practical. In reality, richer models will be at play. For example, lenders pricing close to the IRR might have other margins to help them manage the IRR imposition, like non-interest loan fees and commissions. A lender previously pricing at 50% when an IRR was imposed at 40% may be able to shift interest across loan-related income streams, limiting the effect of the IRR. At the same time, it is less clear that a lender pricing at 100% could recoup such a large loss in income by increasing loan fees.

To allow for such effects and others like it, we now arbitrarily redefine the "bound" and "unbound" groups for our analysis by the median IRR exposure in the market in Q4 2014 (approx., 35%), rather than the condition $r_{i,\,t=2012q4} - r_{cap,t=2013q1} > 0$. This means that we are shifting "marginally" bound lenders – approximately 8 of them – into the "unbound" control group. The is very similar to what we did with the DL modelling. In essence, it is a conservative shift if we believe that the IRR has a negative outcome for bound lender's lending, as we would be adding lenders to the control

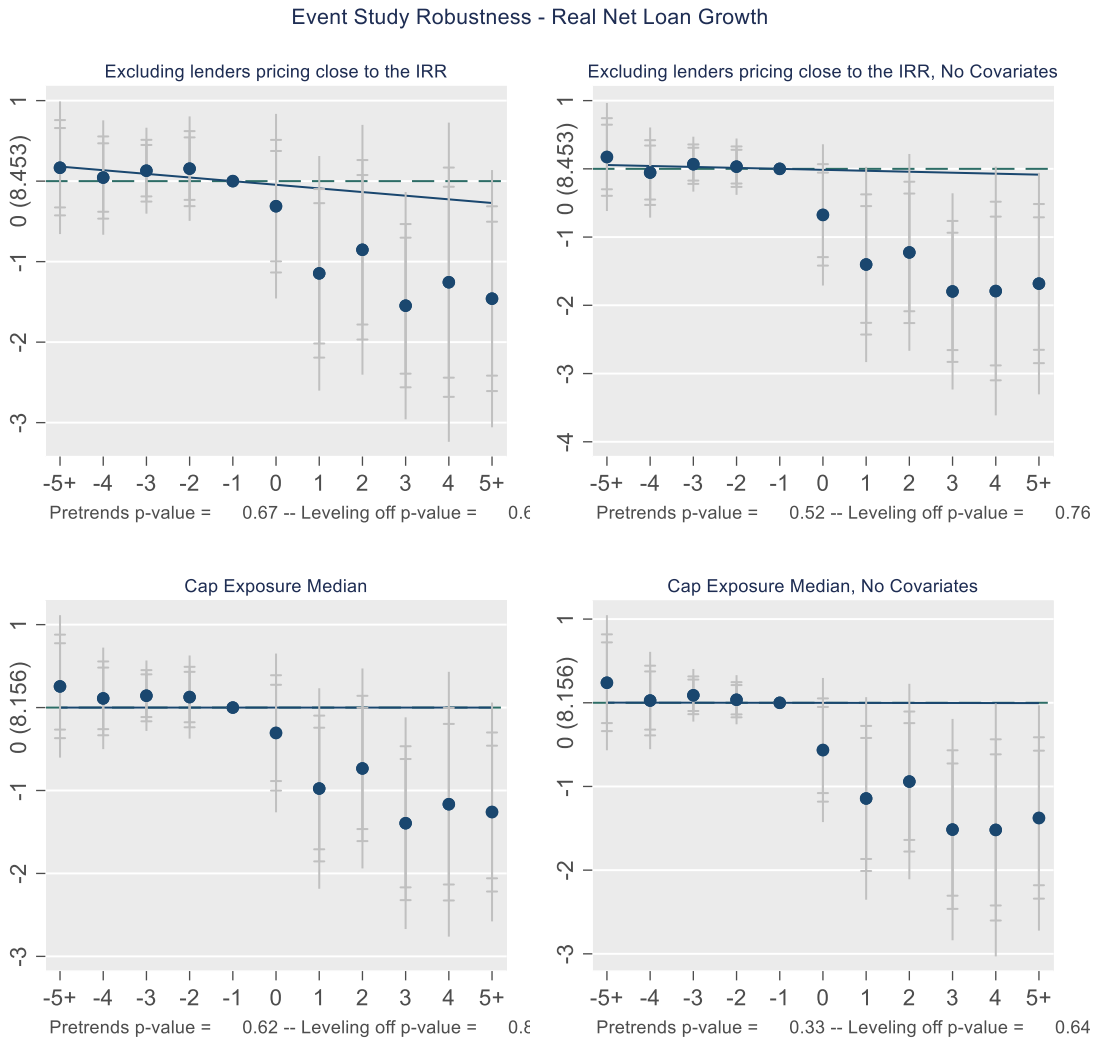
[163] We would be additionally concerned with any lenders whose loan portfolio exhibited an interest rate distribution that is highly skewed, as such lenders might be misclassified. These types of issues highlight the limitations of our lender-level analysis – loan level data would be of benefit here.

group who were more likely to have seen a negative lending growth post treatment. Specifically, we adjust $z_{i,t}$ in equation 5.10:

$$z_{i,t} = \begin{cases} T.r_{i,\,t=2012q4} - r_{cap,t=2013q1}, & r_{i,\,t=2012q4} - r_{cap,t=2013q1} > 0.35 \\ 0, & r_{i,\,t=2012q4} - r_{capt,t=2013q1} \leq 0.35 \end{cases}$$

Figure 29 (row 2) suggests that even if marginally bound lenders were misclassified in our baseline regressions, they don't appear to be driving the core results. The coefficient estimates using the median definition are smaller in magnitude, but they remain directionally consistent. These results also confirm our suspicion that much of the variation driving our results derives from the group of higher interest rate lenders who saw particularly poor growth performance post IRR implementation.

Figure 29 – Event Study Results with Alternative Group Definitions



Note: The figure provides coefficient estimates of $\sum_{m=-G}^{k} \hat{\beta}_m$ from equation 10. The dependent variable is the log of outstanding real net loans, and the y-axis coefficients capture the time path for the marginal slope parameters. Comparison groups are defined in the text. Pointwise confidence intervals provided are at 90% and 95% (horizontal lines), and the 95% uniform confidence band (vertical line). Standard errors clustered at lender level. LHS graphic includes lagged lender covariates, including the capital asset ratio, loss allowance ratio, operating expense ratio, and lender age (see Table 14 for variable definitions). RHS graphics excludes covariates. The event study regressions follow Freyaldenhoven et al (2021). See Annex 14 – Event Study Implementation, for additional details. The Pretrends p-value statistic provides a Wald test for evidence of pretends.

Modelling the IRR as a binary treatment

Our theoretical framework and previous results assume that a lender's reaction to an imposed IRR will be a function of the intensity of its exposure to the IRR. But what if this is not the case, or if our measures of exposure are misinformed or inaccurate? To check that our results are not overly sensitive to these assumptions we now estimate 5.10 as if the effect of the IRR is homogenous across all bound lenders (see Figure 30, row 1). Note that we will be trying to uncover a different estimand

in this case, the ATT(t) rather than the ACRT(t,z|z), which also requires slightly different identifying assumptions. Specifically, we model:

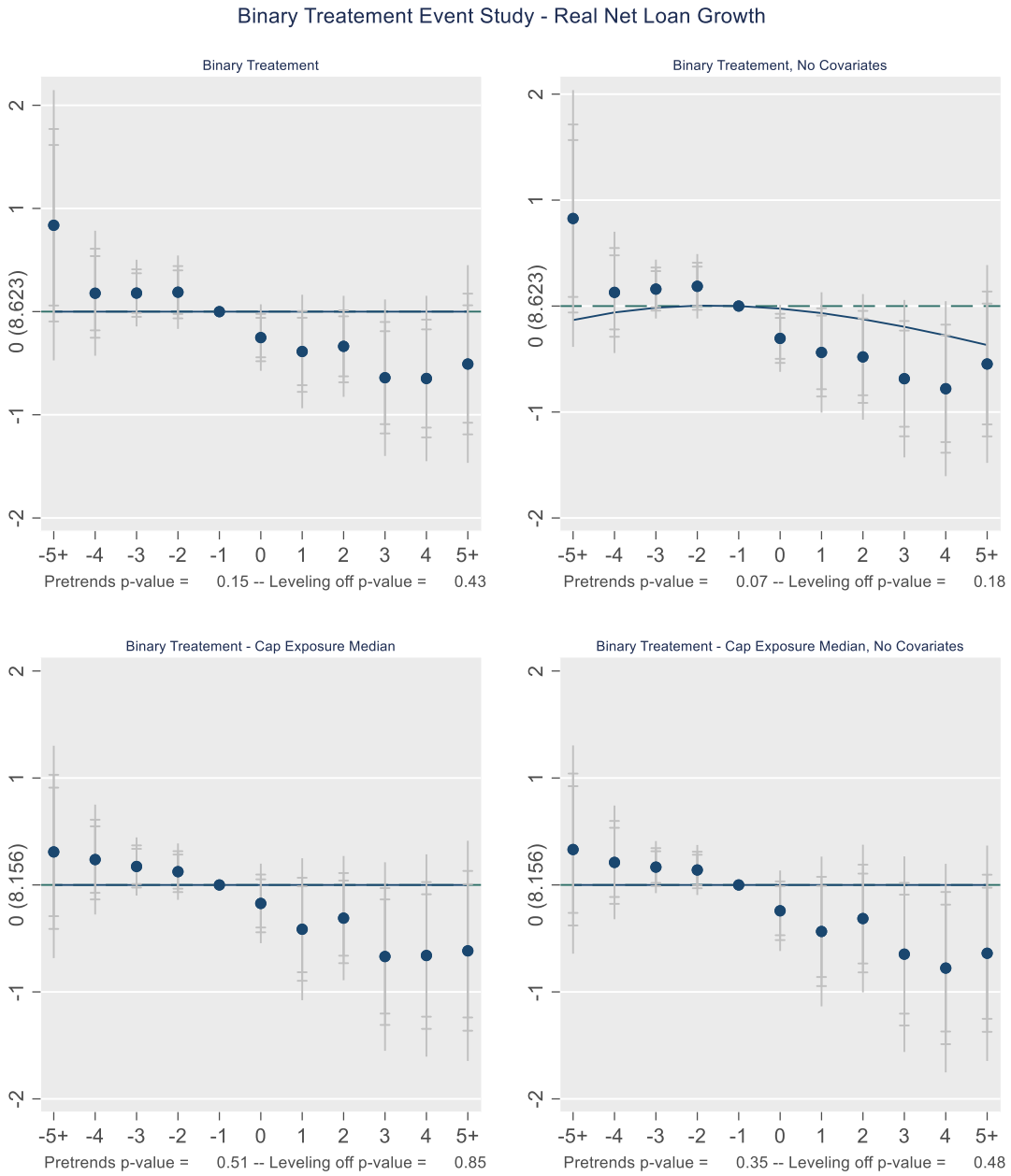
$$z_{i,t} = \begin{cases} T.1, & r_{i,\,t=2012q4} - r_{cap,t=2013q1} > 0 \\ 0, & r_{i,\,t=2012q4} - r_{capt,t=2013q1} \leq 0 \end{cases}$$

We also want to be sure that our results hold if we also assume both (i) that the treatment is binary, and (ii) that "marginally" bound lenders are not affected by the cap. To do so, we again assume that "marginally" is defined by the median exposure and hence estimate 1b (see Figure 30, row 1), but with:

$$z_{i,t} = \begin{cases} T.1, & r_{i,\,t=2012q4} - r_{cap,t=2013q1} > 0.35 \\ 0, & r_{i,\,t=2012q4} - r_{capt,t=2013q1} \leq 0.35 \end{cases}$$

Treating the IRR as a binary treatment and estimating the treatment effect on bound lenders provides broadly similar results to our previous findings (Figure 30). As we are estimating direct treatment effects on bound lenders (and not marginal treatment effects, as previously), the y-axis on Figure 30 can be interpreted directly as a growth effect. Across the specifications, bound lenders appear to reduce lending due to the IRR by approximately 50% on average compared to what they might have done otherwise, with a 95% confidence interval of approximately -10% and -115%. As with previous results, this is clearly an economically meaningful change. The end-point dummy variable binning i.e., the dummy variable capturing differences in growth for the fifth quarter before the IRR and before (-5+ period dummy in Figure 30), is picking up some substantial historical differences in growth between control and treatment, which is concerning for the credibility of the parallel trends assumption and a reminder of the short and unbalanced nature of our panel.

Figure 30 – Event Study Results with Binary Treatment and Binary Treatment and Median Grouping



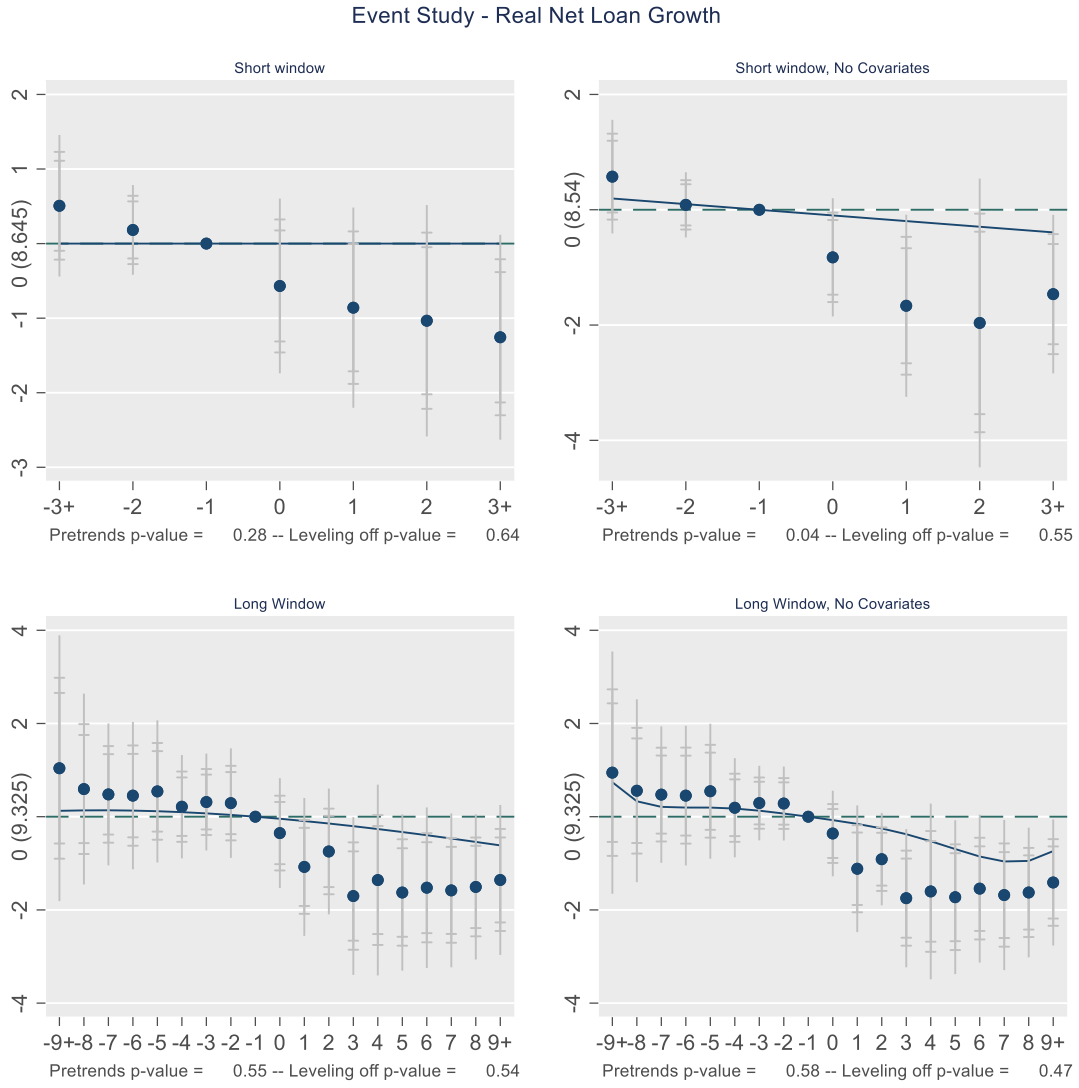
Note: The figure provides coefficient estimates of $\sum_{m=-G}^{k} \hat{\beta}_m$ from a binary treatment version of equation 5.10. The dependent variable is the log of outstanding real net loans. Comparison groups are defined in the text above. Pointwise confidence intervals are provided at 90% and 95% (horizontal lines), and the 95 percent uniform confidence band (vertical line). Standard errors clustered at lender level. LHS graphic includes lagged lender covariates, including the capital asset ratio, loss allowance ratio, operating expense ratio, and lender age (see Table 14 for variable definitions). RHS graphics excludes covariates. The event study regressions follow Freyaldenhoven et al (2021). See, Annex 14 for additional details. The Pretrends p-value statistic provides a Wald test for evidence of pretends. RHS graphics excludes covariates.

Event study estimation window

We estimate our baseline results from equation 5.10 (or, more precisely, 5.A1 in Annex 12) but for a very short (2 quarters) and a very long (8 quarters) estimation window. In other words, we re-run

equation 5.A1 but set L_G and M equal to 2 and 8 respectively. The samples across the two sets of estimates differ as not all lenders were present for the 18 periods of data over the period when the IRR was imposed. In any case, Figure 31 shows that the general baseline results appear robust, at least in terms of the directional policy narrative.

Figure 31 – Event Study Results with Short and Long Estimation Windows



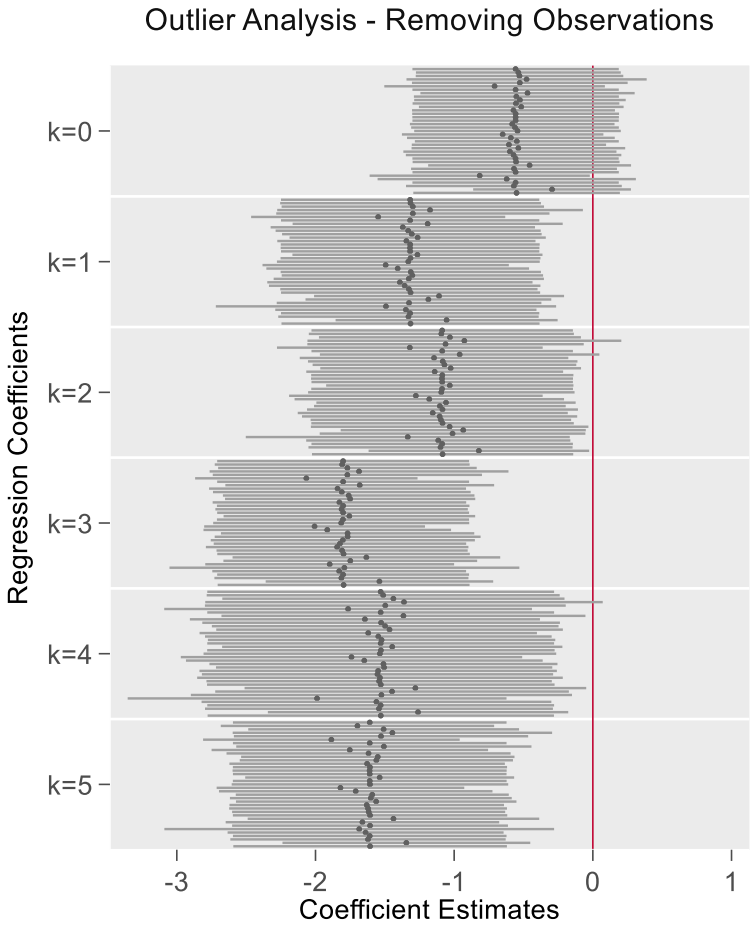
Note: The figure provides coefficient estimates of $\sum_{m=-G}^{k} \hat{\beta}_m$ from equation 5.A1 in, with different event horizon windows. The dependent variable is the log of outstanding real net loans. Comparison group based on simple IRR bindingness. Pointwise confidence intervals provided are at 90% and 95% (horizontal lines), and the 95 percent uniform confidence band (vertical line). Standard errors clustered at lender level. LHS graphic includes lagged lender covariates, including the capital asset ratio, loss allowance ratio, operating expense ratio, and lender age (see Table 14 for variable definitions). RHS graphics excludes covariates. The event study regressions follow Freyaldenhoven et al (2021). See Annex 14 – Event Study Implementation, for additional details. The Pretrends p-value statistic provides a Wald test for evidence of pretends. RHS graphic excludes covariates.

Outlier sensitivity

We undertake two pieces of analysis to examine whether our results are sensitive to outlier observations. First, we re-run equation 5.10 several times. For the first estimation, we use the full sample, but we then remove one lender per estimation (with replacement) allowing us to check whether a single lender's experience is driving our results. Figure 32 provides the results for each group of $\sum_{m=-G}^{k} \hat{\beta}_m$ estimates (from k=0 to k=5+). The coefficient estimates for each regression (i.e., for each regression, for each of which one lender has been removed) are shown with its confidence interval for each $\sum_{m=-G}^{k} \hat{\beta}_m$. Broadly speaking, for each $\sum_{m=0}^{k} \hat{\beta}_m$, the exclusion of an individual lender does not appear to drive the results, even if there is some variability in the point estimates.

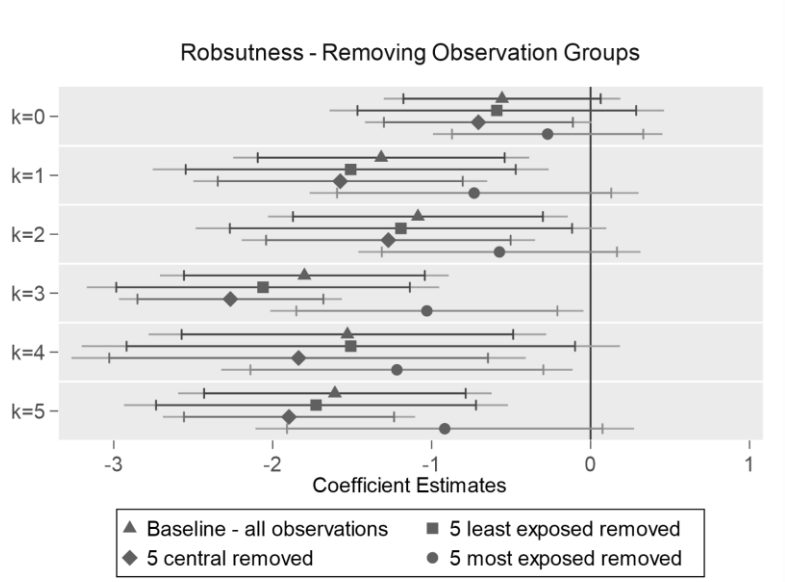
The second piece of analysis takes a similar approach to Figure 32, but instead of removing individual observations per regression, removes the five most bound, the five least bound, and the five most central observations (the "most central" observations include the median lender and the two closest observations either side). The idea is to check that our results are not being entirely driven by only a small set of extreme lenders. Predictably, with our small sample the estimates shift in magnitude and significance when we remove the most extreme observations (Figure 33). Indeed, removing the 5 most exposed lenders clearly reduces the magnitude and sometimes shifts the 95% confidence interval over the zero-coefficient line. Yet the underlying direction of the results seem reasonably stable.

Figure 32 – Event Study Results Removing Individual Lenders



Note: The figure provides coefficient estimates of $\sum_{m=0}^{k} \hat{\beta}_m$ from equation 5.10 where one lender is removed (with replacement). The results are first stacked by time period for each 6>k≥0. Note that there is only regression for each excluded lender, and coefficients are then stacked for each k. The dependent variable is the log of outstanding real net loans. Pointwise confidence intervals provided are at 95% (grey lines around solid point estimate). Standard errors clustered at lender level. Regressions included lagged covariates covering a lender's capital asset ratio, loss allowance ratio, operating expense ratio, and lender age (see Table 14 for variable definitions). The event study regressions follow Freyaldenhoven et al (2021). See Annex 14 – Event Study Implementation, for additional details.

Figure 33 – Event Study Results Removing Lender Groups



Note: The figure provides coefficient estimates of $\sum_{m=0}^{k} \hat{\beta}_m$ (where k≥0) on the x-axis from equation 5.10 but removes five lenders from different ends of the IRR exposure spectrum each time (with replacement), stacking the results for each k coefficient (results stacked on the y-axis). The dependent variable is the log of outstanding real net loans. Pointwise confidence intervals provided are at 95% (grey lines around solid point estimate). Standard errors clustered at lender level. Estimates include lagged lender covariates, comprised of the capital asset ratio, loss allowance ratio, operating expense ratio, and lender age (see Table 14 for variable definitions). The event study regressions follow Freyaldenhoven et al (2021). See Annex 14 – Event Study Implementation, for additional details.

Taken as a whole, this DID analysis arguably supports the findings in the DL modelling from earlier in the Chapter. While the estimands and counterfactuals between the two approaches are clearly different, there is a broad consensus that Zambia's IRR regime did likely result in lower lending for cap-exposed MFIs. While our Zambia sample is small and the parallel trends assumptions are based on groups that are clearly different, these DID results do appear robust to changing the grouping definitions and removing sets of observations that are the most likely to be outliers. The discussion at the end of the Chapter places these findings in a broader context.

The loan income shifting hypothesis and plastic policy effects

> *"...we now spread our costs in our pricing, rather than bundle all into our interest rate. Of course, this does raise issues of transparency.."*
>
> *".... If BOZ decided to impose a cap on fees, that would be a killer for us..."*
>
> *(Anonymous MFI quotes, cited in Financial Sector Deepening Africa (FSDA) and Davel, 2013 p.27)*

The Zambian Government's IRRs placed restrictions on the interest rate that lenders could charge, yet there remained significant scope for lenders to increase their fees and commission charges instead of higher interest rates. Some fees and commissions were captured within BOZ's definition of interest rates; however, many were not[164].

This section investigates whether the imposition of Zambia's caps influenced lenders' mix of interest vs fee and commission shares. The risk that lenders might switch income from interest income to fees in response to IRRs is important because it may undermine the potential effectiveness of the caps in achieving their policy goals (goals like correcting for market competition imperfections or reducing actual effective interest rates for borrowers). It may also result in reduced pricing transparency within the market, a negative outcome in-and-of itself (Ferrari et al, 2018).

Internationally, there is growing evidence that a shift towards fees and other non-interest loan income is a distinct risk when introducing IRRs. Evidence of such a shift draws from Kenya's IRRs and its banking system (Central Bank of Kenya, 2018 and Alper et al., 2020), from microfinance lenders in Cambodia (Heng et al., 2021) as well as in Nicaragua, South Africa, and Armenia (Helms and Reille, 2004). Unfortunately, much of this evidence is anecdotal rather than quantitative.

Our quantitative analysis of Zambia's IRRs finds that the imposition of the regime did likely result in lenders shifting interest income into fees. Our baseline results find that lenders' fee share of loan income increased by 20 percentage points five quarters after the imposition of the IRR on average. We did not find strong evidence that lenders shifted back to a lower fee regime once the caps were lifted. These are concerning findings for policymakers in Zambia because they suggest that (i) the IRR policy may have not been as successful in reducing the cost of credit as policymakers had hoped (at least in APR (Annual Percentage Rate terms) while still risking reduced pricing transparency in the market[165], and (ii) that the IRR may have had plastic effects on pricing in the market: once the IRR policy was removed, there is little evidence that non-interest fees and commissions were removed.

[164] Fees are regulated in Zambia via Statutory Instrument No. 179 of 1995, but the definitions carve out significant room for lenders to apply non-interest loan fees (carve-outs include, for example, fees relating to arrangement/renewal, mandatory insurance, disbursement charges, asset appraisals, legal charges, prepayment, etc.).
[165] Politically, reducing in the "shop window" interest cost of credit may still have been considered a success.

Theoretical Mechanisms

Our simple theoretical framework did not consider non-interest loan income and lenders' abilities to charge loan-based fees, and so we have no traceable theory precisely isolating how banks might adjust their fees post IRR. Nevertheless, it still makes sense to discuss how we might expect lenders to manage their non-interest income in response to the imposition of an IRR as well as the context they faced, even if without a formal model. We concentrate on a few possible theoretical mechanisms and hypotheses, below.

First, intuition suggests that we might expect lenders to shift to charging fees in the face of an IRR, and that such a shift might be slow but persistent. Consider a simple context in which an IRR is imposed on a lender for whom there is a one-off cost involved in changing its business pricing model from an interest-based model to an interest and fee model. Before the IRR was imposed, the lender simply charged interest[166]. But with the IRR imposed, it may make sense for the lender to incur the fixed cost in changing its business model and begin to charge non-interest fees in order to retain loan-based income/profitability[167]. While the cost of changing business models initially incentivised the lender to not charge fees, after the IRR is imposed the value of changing outweighs the fixed cost. Once the IRRs are removed, however, the cost to a lender of changing its business model *back* to only charging interest would disincentivize further change, creating persistence in lenders' fee shares of income. Changing business/pricing models might also take time, which means we should expect a lag in observing any change in income share. This would be in addition to lags associated with portfolio roll-over and the fact that lenders might also want to wait to ensure that policymakers were not going to reverse their IRR decision before making any investments in changing their business models. These inertial characteristics suggest we may need to be open to lags and dynamics in our modelling.

Another relevant question is whether lenders more severely bound by the IRRs will shift their income from interest to fees by more. On the one hand, lenders more bound by the IRR will face more pressure on their profit margins (per loan) from the IRR and so might be more inclined to try and shift more interest income into other fees and commissions. Yet lenders operating in high interest

[166] In Zambia, some lenders charged fees before the cap, but such lenders were in the minority and the fees were typically small.
[167] In this context, the costs of changing business/pricing model might be due to creating new loan origination processes and documentation.

rate markets might also find it harder and costlier to transfer such a large proportion of their loan income from interest into fees and commissions – a lender originally pricing at 100% APR with a 40% cap might find it challenging to offset 60% of interest income into non-interest fees, while a lender only looking to offset 10% might find it relatively easy. The cost economics of interest vs fees and the demand elasticity of interest vs fees remain unclear – we do not know borrowers' reactions to facing large dollar-denominated fees or other fee strategies, such as monthly transaction fees. So, while there may be a marginal effect of IRR exposure on income shares, it might be harder to justify restricting it to be linear.

Finally, when an IRR is imposed, price competition within the market may lead us to expect that lenders who were unbound by the cap (or less bound by it) may also increase their use of fees in lieu of interest income. If we assume that lenders compete for some borrowers (a reasonable but different assumption to the segmented market framework from earlier), it is plausible that bound lenders reducing their interest rates and increasing their fees (to offset the IRR) might result in other less bound lenders also doing so, at least on the margins where they compete. This is because, if borrowers do not perfectly calculate the new real cost of borrowing (APRs can be challenging to calculate), unbound lenders might need to defend their market position from bound lenders who are now using non-interest fees to lower their interest rates. Increasing their own fees might be the best means to do this, which may in turn result in a level shift in fees and commissions across the market. In such a setting it is less clear that we would expect only bound lenders to react to the cap by increasing their non-interest income share. The specific consequence of this argument for our empirical strategy is that a DID approach, whereby we separate bound and unbound lenders, may be questionable for our fee analysis as both groups are likely affected by the policy.

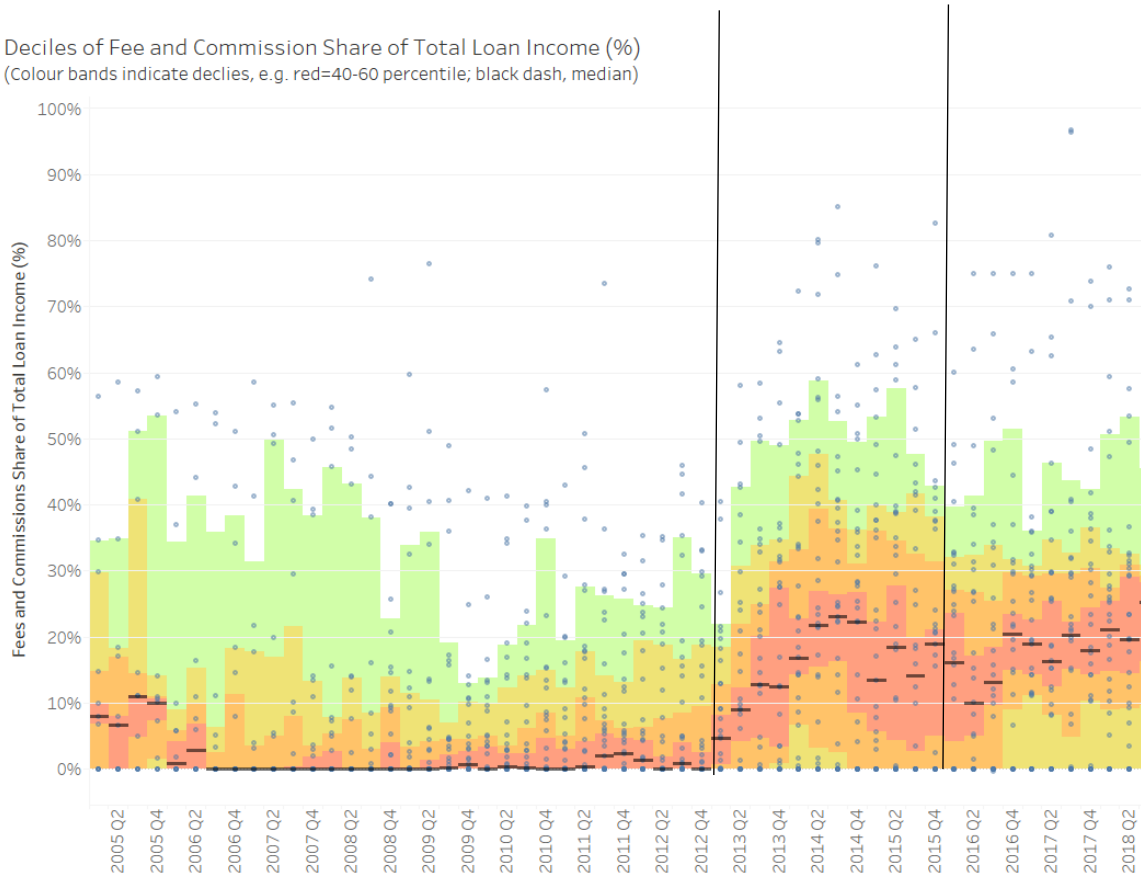
Empirical Approach and Quantitative results – IRR imposition

An informal analysis of Zambia's raw interest rate share data indicates that the implementation of the IRRs likely resulted in lenders moving towards business models with less reliance on interest income and more reliance on fees. Figure 34 illustrates the decile breakdown of lenders' fee and commission income as a percentage of total interest and fee income. The data reveals a gradual but notable increase in the post-Q1 2013 period towards a regime of higher fees and commissions. Furthermore, this trend appears to have continued even after the cap was lifted. Qualitative evidence

also supports this shift, for example the first anonymous quote at the beginning of this section, as well as an interview with an MFI undertaken by Siwale and Kimmitt (2017) in July 2015:

> *"All MFIs have increased fees and yet these fees were not this high before the cap…You see when these clients come to us to borrow all they ask for are interest rates, they don't ask about other fees? So us we just tell them the interest rates and they don't know that in some cases the fees could even be higher than interest charged" (MFI NDT 6, cited in Siwale and Kimmitt 2017, p.22)*

Figure 34 – Non-banks' Fee and Commission Loan Income Shares



Note: Each dot provides a fee share for a given MFI in a given quarter. The colour banks provide decile ranges across all lenders' fee shares in a given quarter. The red band in the middle shows the 40th-60th percentile range, the dark orange the 30th-40th and 60th-70th. The horizontal black line provides the median interest rate for a given quarter. The vertical black lines show the imposition and removal of the cap.

Source: BOZ Prudential Returns

We now more quantitatively examine whether the introduction of IRRs in Zambia was associated with a shift in lenders' fee structures. Our approach uses a distributed lag model and is thematically similar to equation (5.8), with the clear exception that we are using the percentage-point change in non-interest loan income shares as the dependent variable, rather than loan growth.

We begin by modelling the IRR as a binary treatment variable while allowing for bound and unbound lenders to have different reactions to the policy. Given the possible theoretical mechanisms underpinning income shifting and IRRs, this is likely a reasonable starting point. Specifically, we estimate:

$$\Delta Y_{it} = \alpha_i + \sum_{j=d}^{J} \beta^j . I . \Delta T + \sum_{k=d}^{K} \beta^k . (1 - I) . \Delta T + \sum_{l=1}^{L} \beta^l X_{it-1}^l + \sum_{m=1}^{M} \theta^m M_t^m + \varepsilon_{it} \quad (5.11)$$

$$I = \begin{cases} 1, & r_{i,\, t=2012q4} - r_{cap,t=2013q1} \geq 0 \\ 0, & r_{i,\, t=2012q4} - r_{capt,t=2013q1} < 0 \end{cases}$$

$$T = \begin{cases} 1, & t \geq 2013q1 \\ 0, & t \leq 2012q4 \end{cases}$$

Where Y_{it} is the percentage share of fees and commissions of a lender i's total loan-related income (loan interest revenue and fees and commissions) at time t, α_i is a lender fixed effect, I is a group indicator variable (coded a 1 for lenders bound by the cap when it was imposed and 0 for those that were not)[168]. T is the IRR indicator, indicating when a cap regime was in place or not. X_{it}^l includes several lender-level covariates, capturing firms' size (log of total lender assets), leverage (leverage ratio), cumulated losses (loss allowance), business model and efficiency (operating expense ratio), and age.

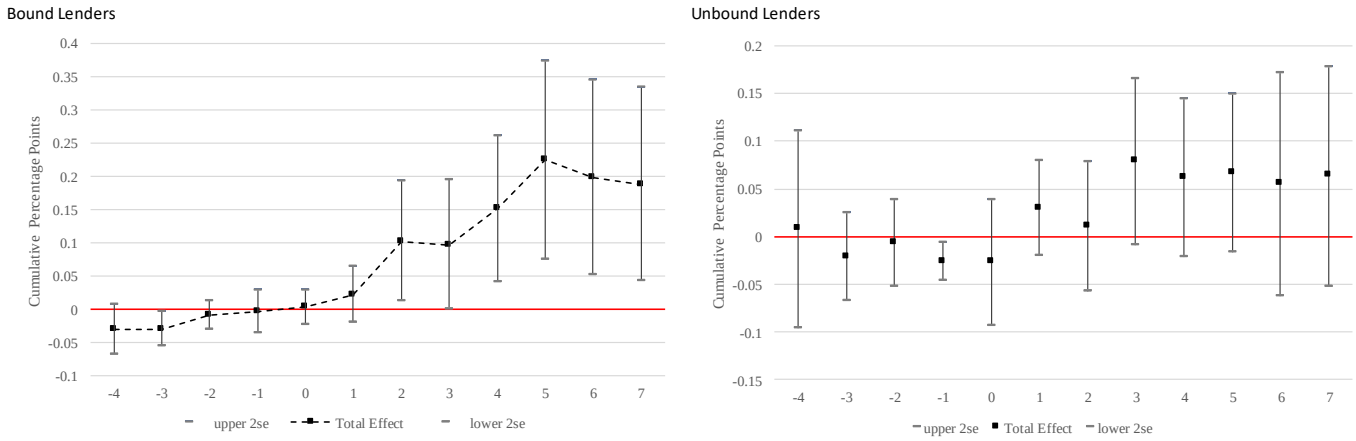
Estimating equation 5.11, we find that bound lenders increased their non-interest income share cumulatively by 20 percentage points within five quarters after the imposition of the IRR on average, with a 95% confidence interval of 5% and 35% (left-hand side, Figure 35)[169]. This constitutes a substantial change in the pricing structure for a given bound lender. For unbound lenders, we cannot reject the null hypothesis of no change at the 95% level (right hand side, Figure 35), although at less constrictive significance levels we estimate a level shift for them too, albeit a smaller one. As we suspected, there appears to have been some lag in lenders adjusting their pricing practices and business models in the wake of the cap, which might be explained by adjustment frictions in lenders'

[168] $Y_{it} = \frac{fees_{it} + commissions_{it}}{fees_{it} + commissions_{it} + total\ interest_{it}}$. We use total interest income in the denominator to minimize the effect of a lender shifting its portfolio away from loans towards securities, although this is not typically an issue for most MFIs who hold limited liquid assets.

[169] The estimates are relying on within-lender variation in IRRs and income shares only, although the results are robust to using TWFE.

changing business models. It could also be because lenders may also have waited to see whether the Government would keep the policy in place before taking action.

Figure 35 – Binary Effect Piecewise Distributed Lage Models for Bound and Unbound Lenders – Cap Implementation

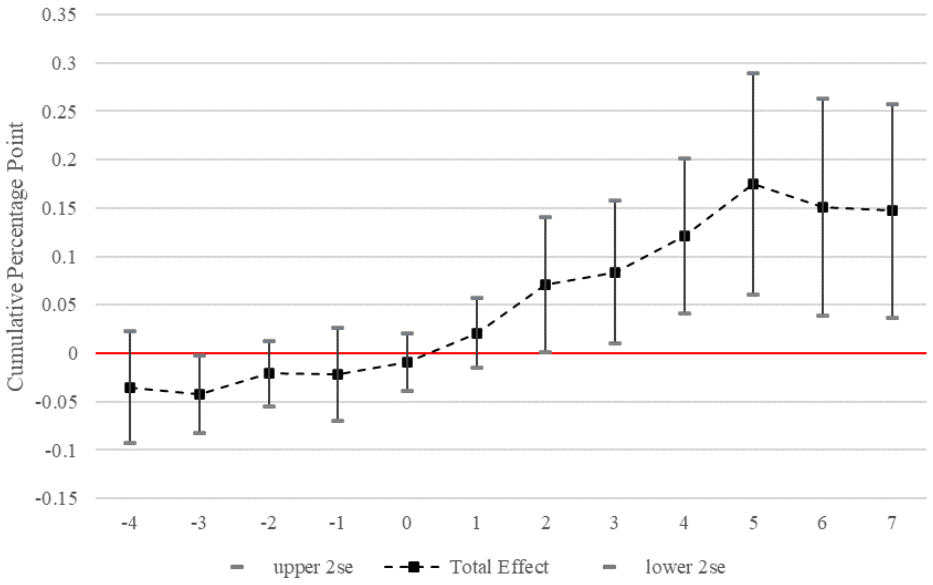


Note: The figure provides results from equation 5.11 regressing the percentage change in lenders fee share of loan income against the change in a group-specific binary treatment dummy capturing the imposition of the IRR policy. The LHS graphic provides the estimates for lenders bound by the IRR, $\sum_{j=d}^{J} \beta^j$, and the RHS graphic provides the estimates for lenders not directly bound by the IRR, $\sum_{j=d}^{J} \beta^j$. Controls include both lender-specific and macroeconomic control variables, as per Table 15. The regressions include lender fixed effects. Standard errors are robust and clustered at the lender level. The y-axis is given in decimals. Confidence interval at 95%.

The baseline results also appear to be robust to other modelling specifications. For example, if we assume that all lenders (bound and unbound) can be modelled as one group, we still find evidence of a change in income share, albeit with a smaller 15 percentage point mid-point cumulative effect of the cap (see Figure 36)[170].

[170] Specifically, we estimate: $\Delta Y_{it} = \alpha_i + \sum_{j=d}^{J} \beta^j . \Delta T + \sum_{l=1}^{L} \beta^l X_{it}^l + \sum_{m=1}^{M} \theta^m M_t^m + \varepsilon_{it}$. The approach essentially estimates a single dynamic lag structure and cumulate effect parameter for both groups combined.

Figure 36 – Binary Effect Distributed Lage Models for All Lenders – Cap Implementation



Note: The figure provides results from regressions using equation footnote 165. It regresses the percentage change in lenders fee share of l income against the change in a binary treatment dummy capturing the imposition of the IRR policy over different potential policy lag and l structures, providing estimates the IRR policy dummy, $\sum_{j=d}^{J} \beta^{j}$. Controls include both lender-specific and macroeconomic control variab as per Table 15. The regressions include lender fixed effects. Standard errors are robust and clustered at the lender level. The y-axis is gi in decimals. Confidence interval at 95%.

The results also appear to hold under various robustness tests similar to those used in the loan DL modelling.

Quantitative results – IRR removal

Compared to the imposition of the IRRs in Zambia, we find less evidence that lenders resorted back to an interest income-based model when the caps were loosened or removed. In fact, it is plausible that there is a plasticity in the how the IRR changed the structure of pricing in the market: fees increased post IRR imposition, but do not appear to have undergone an offsetting change post IRR removal.

To evaluate whether lenders reduced their income and fee share of income on IRR removal, we estimate a binary treatment, single parameter version of (5.11) for the loosening (and removal, estimated separately) of the cap in Q3 2014 (Q4 2015):

$$\Delta Y_{it} = \alpha_i + \sum_{j=d}^{J} \beta^{j}.\Delta T + \sum_{l=1}^{L} \beta^{l} X_{it}^{l} + \sum_{m=1}^{M} \theta^{m} M_{t}^{m} + \varepsilon_{it} \qquad (5.12)$$

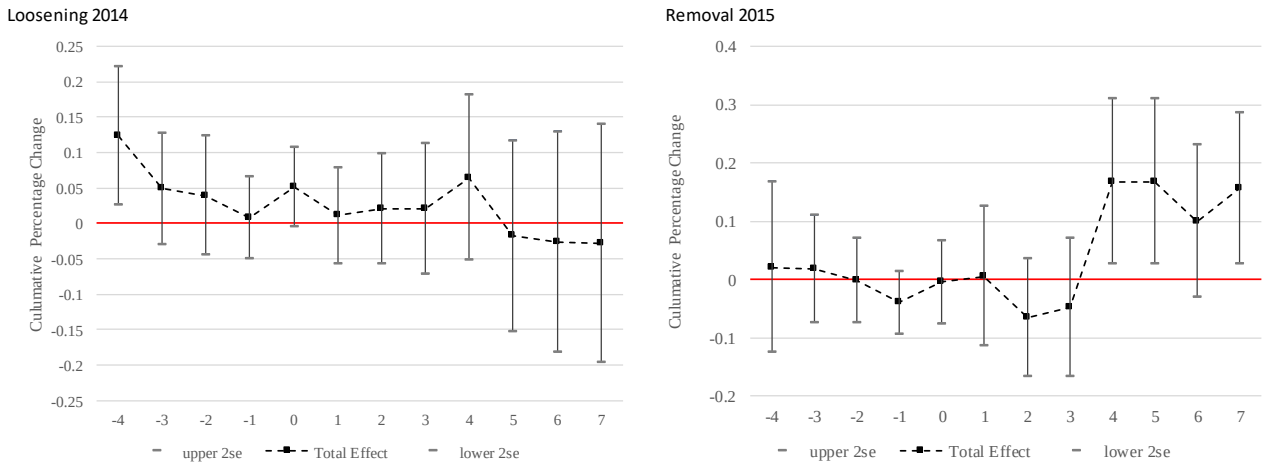
$$T = \begin{cases} 1, & t \geq 2014Q3 \ (or\ 2015Q4) \\ 0, & t < 2014Q3 \ (or\ 2015Q4) \end{cases}$$

We do not find convincing evidence that the loosening/removal of Zambia's IRR regime was associated with large changes in lenders' non-interest income shares. Our estimates of the cumulative effect of the IRR regime change, $\sum_{j=d}^{J} \hat{\beta}^j$, are neither clearly nor consistently negative post IRR removal. Figure 37 shows the estimates of the cumulative percentage change for a given lag/lead structure, spanning over an array of possible lag structures for the loosening in 2014 and removal in 2015 of the IRR. If lenders truly did unwind the increase in fee shares that we suspect were associated with the imposition of the IRR regime, we might have expected to see a stronger and more significant negative signal in these results.

Yet there are important caveats to these results and our interpretations. First, note that we cannot rule out large reductions in income fee shares post IRR removal given our sample and analysis – the 95% confidence interval point estimates for the removal of the cap still include an unwinding of the cumulative effect estimates of 15% to 20% (i.e., our estimates from the cap imposition)[171]. Such wide confidence intervals likely reflect the imperfect and noisy data capture we highlighted previously, as well as our small sample. The second caveat is that the removal of the IRR regime was almost certainly not exogenous to contemporaneous macroeconomic conditions, potentially biasing our result. While we argue that the imposition of the IRR was arguably exogenous, the same cannot be said for the removal of the IRRs. There is a significant increase in fee shares approximately a year after the caps were removed, which is a useful reminder of imperfect BOZ data capture (one institution recorded 100% of loan income as fees and commissions, a clear error), the challenging macroeconomic environment in mid-2016, and the noise in the sample.

[171] The adage "the absence of evidence is not evidence of absence" clearly applies here.

Figure 37 – Binary Effect Distributed Lage Models for All Lenders – Cap Loosening and Removal



Note: The figure provides results from regressions using the equation in footnote 5.12. It regresses the percentage change in lenders fee share of loan income against the change in a single binary treatment dummy for all lenders. Estimates provided are for the cumulative IRR policy dummy, $\sum_{j=d}^{J} \beta^{j}$, capturing the loosening (or removal, see RHS) of the IRR policy over different treatment lag and lead regressions. Controls include both lender-specific and macroeconomic control variables, as per Table 15. The regressions include lender fixed effects. Standard errors are robust and clustered at the lender level. The y-axis is given in decimals. Confidence interval at 95%.

What types of fees did MFIs increase in the wake of the IRR?

Our analysis so far has used financial statement data to calculate fee income shares and yields, providing an amalgamated picture of how IRRs increased non-banks use of non-interest pricing. But we should also be interested in the *types* of fees that were increased in the wake of the cap. While all non-interest fees arguably reduce transparency (due to the complexity of APR calculations), a switch to unpublished and hidden fees may do so even more.

While we cannot quantitatively answer all we would like about fee usage and the IRRs, we do find that, on average, MFIs use of some fee types increased. For example, MFIs use of "application fees" and "other fees" in the post IRR period appear higher compared to the market before the IRRs were imposed. It is also worth noting that "application fees" were exempt from contributing to MFIs' IRR-regulated interest rates. MFIs use of other typical fee types, such as "processing fees" and "insurance fees", remained reasonably static. This perhaps also made sense given that these fee types were not exempt from the IRR cap calculations.

Table 16 provides simple average MFI fee share data before and after the IRR regime, calculated using BOZ's Cost of Borrowing survey[172]. The table shows that "application fees" were 5.8% of

[172] The data is based on BOZ's Cost of Borrowing survey and captures five "transparent" fee types: Application fees; Processing and Admin Fees; Insurance Fees; Credit Bureau Charges; and Other Charges) as well as interest costs. See

total surveyed loan costs in the MFI cross section after the IRR, up from 0.6% in the year before the regime was implemented. Average MFI usage of "other fees", a data catch-all for other origination-based fees that may or may not be transparently provided to the borrower, also increased (from 0.2% of total loan costs to 2.1%, on average). This makes sense given that application fees were explicitly exempt from the IRR calculation and an easy option to shift interest income, while other fee types, such as processing fees, were not.

What we cannot directly capture here, however, are changes in the use of ex-post fees not captured in BOZ's survey (for example, late fees or monthly transaction fees). This is because we only have either consolidated financial statement data (which we used in the previous quantitative analysis), or the low-quality ex-ante fee data used in Table 16. We suspect it is plausible that less transparent, post-origination loan fees may have increased post IRR. This is because of the difference in magnitudes between the results from the financial statement data and the income survey data. In the previous section, we estimated that MFIs increased their fees share after the IRR was imposed by approximately 15% on average, and that this increase persisted (see Figure 36). This analysis used fee line items from MFIs' financial statements that include *all* fee types, including less transparent post-origination fees (and hence are captured in the fee shares calculations and results). At the same time, Table 16 shows that the approximate increase in average origination fees share using the BOZ interest rate survey was somewhat smaller in magnitude (≈7%). Several factors, including poor data collection and reconciliation issues, might drive this difference. Yet it is possible that less transparent post-origination fee types (those included in the financial statement data but not included in BOZ's Cost of Borrowing survey), may also have contributed to some of this difference. Having said that, the above is a speculative, back of the envelope analysis, and we would need additional data to understand the role of ex-post fees more precisely.

https://www.boz.zm/demonstration-of-the-cost-of-borrowing.htm. In 2012 and 2016 this data was not of high quality, meaning quantitative and statistical analysis was not possible. BOZ's survey is unfortunately the best fee data available, yet it does not address less transparent ex-post fee charges, such as late payment fees.

Table 16 – Fee Composition of Cost of Borrowing at Origination

Fee share component	Average fee share component of total surveyed cost of borrowing, across all MFIs in the year pre and post the IRR		Diff
	2012Q1-2012Q4	2015Q4-2016Q3	
Loan Application Fees	0.6%	5.8%	5.2%
Other Charges	0.2%	2.1%	1.9%
Processing and Administration Fees	10.5%	8.1%	-2.4%
Insurance Fees	2.9%	3.2%	0.2%
Credit Bureau Fees	0.0%	0.5%	0.5%

Note: This table provides average MFI fee shares of the cost of borrowing. Specifically, it provides the cross-MFIs average of each fee component share of the total surveyed cost of borrowing, as surveyed by BOZ in its Cost of Borrowing Survey. Fees shares are calculated as a percent of the "total cost of borrowing" (i.e., the total fees associated with borrowing and the total interest cost associated with a normalized 1000 Kwacha loan) and are averaged across the MFIs within each of the two time periods. Each MFI fee share (for each borrower (i), fee share type (j), and quarter (t)) are calculated as: $Fee\ share\ componenet_{ijt} = \frac{Fee\ componenet_{ijt}}{\sum_{j=1}^{J} Fee\ componenet_{ijt} + Interest_{it}}$. These are then averaged across MFIs and within the various time groups (the year before the IRR and the year after) for each fee type (j). The Diff column provides the simple arithmetic difference between the two periods' averages. The data quality is poor because of inconsistent MFI data completion and BOZ validation.

Summary – Fees, commissions, and the IRR

Taken together these results examining the imposition and removal of IRRs on fee and commission income shares are concerning. They indicate that the imposition of IRRs may have had unintended and long-term consequences for the structure of Zambia's non-bank credit market. We find a substantial shift in pricing from primarily interest income to fee income (between 15% to 20%) that was likely due to the cap, and we do not find strong evidence that income shares reverted on cap removal. Such asymmetric and plastic effects should be concerning for policymakers looking to implement IRRs in a market where fees are not already a large component of lenders' business models, especially given that a higher fee environment lowers transparency and potentially hides predatory pricing. These findings are some of the first quantitative results provided in the literature, adding weight to substantial anecdotal evidence (Central Bank of Kenya, 2018; and Alper et al., 2020; Heng et al., 2021; Helms and Reille, 2004). The results also indicate that, if a government does want to implement an IRR, they need to be mindful in how they design and implement it (for example how to include fees into any APR and interest rate measurements), as well as how they plan to monitor it.

Summary and Discussion - IRRs and Zambia's Microfinance sector

This Chapter has analysed the effect of Zambia's IRRs on its microfinance sector. We found evidence (using both DL and DID approaches) that lenders exposed to the cap likely reduced their lending after its implementation, and that lenders that were more bound likely reduced their lending by more. We also found that the IRRs were associated with lenders shifting their business models towards charging fees and commissions in lieu of interest, and that such effects may be plastic, i.e., not clearly reverting after the IRRs were removed. In a practical sense, Zambia's IRR appears to have had quite a severe effect on lenders, especially those operating in higher interest rate sections of Zambia's MFI market. Our results show that the imposition of the IRR had similar effects to the GFC for many lenders – a reduction in loan growth by at least 50% compared to what it might have otherwise been.

We have previously highlighted some of the contributions of this Chapter, such as being one of the first to provide quantitative estimates of (i) the effects of a cap regime in SSA, and (ii) income switching dynamics. However, the Chapter also uniquely highlights the plight of lenders further up the interest rate distribution, lenders and loan market types that are not always captured in other papers or larger cross-county MFI datasets[173]. Our results stress that policymakers should directly and carefully consider the cost structures and interest rate distribution within the market before imposing IRRs, as the effects may be heterogenous. While the political frustrations that led to IRRs in Zambia are understandable, our results also argue that in some cases there may be no easy and harmless solutions to structurally high interest rates. Unless there are clear signs and diagnostics suggesting strong market imperfections, there are risks in implementing strongly binding IRR policies and expecting both, lower rates, and improved market access[174].

There are some clear weaknesses in our Zambia case study. Econometric issues with our current analysis aside, we are not able to examine how IRRs may have led to firm failure because Zambia's cross-sectional sample was too small for statistical analysis (say, for example, in a binary logit/probit setup). A few firms *did* fail directly after the IRR, but their data was not used in our analysis as we relied only on those lenders who initially survived. From a policy perspective, the IRRs also may

[173] Roa et al. (2021) and Cubillos-Rocha et al. (2018), for example, use a binary treatment indicator, limiting their ability to capture heterogenous effects based on cap bindingness.
[174] The narrative might be framed differently if the objective of Zambia's IRR regime was to remove smaller higher-risk lenders from the market based on arguments of predatory pricing and usury. Yet lenders in the higher segments were not clearly making huge rents and the arguments in Zambia for the IRRs were based around deepening financial access.

have had other downsides which we could not capture. For one, they likely contributed towards an environment of policy uncertainty within the credit market and more broadly. Future research, perhaps qualitative, might be able to investigate how the IRRs impacted policy uncertainty in the market and how such uncertainty might impact future investment in the sector and growth in the market.

Another limitation of this chapter is that the analysis is necessarily county and context specific, i.e., external validity is intrinsically limited. Therefore, humility is required in extrapolating its results outside a small geographic or market-type domain. The next chapter in this section, however, builds a dataset of IRR implementation across several countries and experiences, with the goal of providing broader evidence on this topic.

Chapter 6 – An Exploratory Cross-country Analysis of MFI Interest Rate Regulations using the MiX

Introduction

In this chapter, we attempt to undertake a simple exploratory cross-country lender analysis of IRRs. Over the last five years, there has been a growing number of country-specific studies on interest rate restrictions (see, for example, Cubillos-Rocha et al. (2018); Romero (2021); Aiba et al. (2020); Heng et al. (2021); Roa et al. (2021); Fekrazad (2020)), Cuesta and Sepũlveda (2021) and Maderia (2019)) and our analysis of Zambia's regime in Chapter 5 contributes to this body of country-specific research.

To date, however, there has been limited quantitative firm-level cross-country analysis of the effects of IRRs, and such work has typically concentrated on banks rather than microfinance lenders (e.g., Calice et al. 2020). Country-specific studies no doubt have their strengths: they take a country's specific policy and institutional context seriously and provide internally valid IRR policy evaluations. Yet, such research is inherently limited given that the IRR regimes are often applied across the market (limiting both options for choosing econometric frameworks and policy variation). In addition, firms' reactions might be market context-specific, and datasets at the lender level might be limited in size.

In this chapter, we undertake likely the first cross-country panel analysis examining IRRs and MFI outcomes. We concentrate specifically on restrictions to MFI interest rates, either in the form of an interest rate regulation (such as in Zambia) or an enforced usury law[175]. We begin by constructing a new longitudinal database of IRRs covering the last decade or so, documenting an increasing prevalence of interest rate rules across countries. We then undertake descriptive analysis using the new data, recording associations between IRRs and lender outcomes.

Across different sources of variation in the dataset, we typically find a statistically significant and economically meaningful negative association between IRR regimes and MFI outcomes. For

[175] There are a few examples of non-enforced Usury laws, such as in Honduras, which were ignored.

example, we found that MFIs constrained by an IRR regime recorded loan portfolio growth that was 13.8 percentage points lower than when the same MFI was not constrained (but only at the 90% confidence level). Similarly, the growth in the number of loans was lower by 10.9 percentage points. Although not causal, these results are consistent with our Zambia case study in Chapter 5 and invite further research.

Note that the main contribution of this Chapter is the dataset itself. We build a new longitudinal dataset from secondary sources that tracks microfinance IRRs across countries and time[176]. As far as we know, this has not been attempted previously and does not exist in the literature. We then link this country-policy dataset to a recently released cross-country World Bank dataset on MFIs – The Microfinance Information Exchange (MiX) – and examine associations between interest rate restrictions on MFI lending and outcomes. Though this merged dataset is limited in some dimensions, it does allow us to analyse the relationship between IRRs and lender growth (by value and number), operational self-sufficiency, and average loan sizes. The data covers, at least in principle, some of the shift towards financial restrictions discussed in the introduction in Chapter 4. It also provides several different sources of IRR variation: the imposition of policy within a country; a change in the value of the IRR while imposed; variations in the size of policy across countries, and variation in the exposure of MFIs to the policy within a given country and across countries. There is also a large set of countries without an explicit IRR, which might be used in some empirical setups as a non-treated control group – another avenue for future research.

Data

Our data originates from broadly three sources: (i) an original database encoding MFI IRR regimes constructed for this paper; (ii) the World Bank Microfinance Information Exchange (MiX) database, which provides consistent and standardised (albeit imperfect) financial and operational data on MFIs across nearly 100 countries; and (iii) macroeconomic data from the World Bank's World Development Indicators (WDI) and the IMF's World Economic Outlook (WEO) and International Financial Statistics (IFS). Combined, these databases let us track multiple regime changes across countries and control for time-varying local macroeconomic context.

[176] We concentrate specifically on restrictions on MFI interest rates, either in the form of an interest rate regulation (such as in Zambia) or in the form of an implemented usury law.

The dataset comprises approximately 13,000 usable MFI-years of data across 2000 or more lenders and 90 countries. Of these lenders, approximately 15% always faced a cap regime, 10% faced a change in cap regime, and the remaining 75% were never bound. The average number of years for an MFI in the dataset is 6-7 (Table 17).

Challenges with cross-country analyses

To date, there has been little quantitative cross-country research on IRRs and MFIs, reflecting some undeniable challenges in undertaking such analysis. One of the key challenges is standardising cap regimes across countries, as each regime can be subtly different or applied differentially within the market. A question remains whether it is possible to meaningfully standardise caps over several dimensions, such as their scope, number of ceilings, type, methodology, benchmark, whether the caps are binding within a given market, and whether they also encompass fees in their calculation (Ferrari et al. 2018). Another problem is the potential endogeneity of the timing and severity of the IRRs, something that we could argue as being less severe in our Zambia case study. However, it may be far more problematic in other contexts (especially if a country's IRR was implemented in response to a financial crisis or was a part of a "package" of financial market interventions). Finally, a country's interest rate context may interact with the severity of any given cap regime. For example: an IRR of 40% might mean something quite different in a country where most MFIs lend at 20%, compared to a country where they lend at 50%. Further, we might expect the reaction to this cap to be a function of the market structure in each country (treatment effects may differ across perfect vs. imperfect/incomplete markets).

Next, we discuss each dataset component in detail.

Interest Rate Restriction database

We constructed an IRR policy database to track changes in rules across lower-income countries over time. The policy data coverage is designed to match that of the MFIs in the MiX (approximately 100 countries with strongly unbalanced coverage between 2000 to 2018). The IRR policy database covers several countries transitioning back to higher levels of control over interest rates, an unravelling of the financial market deregulation of the late 80s and 90s as discussed in Chapter 4.

Table 17, below, and Figure 48, show the sample of the interest rate rules across countries over time. As mentioned, the data suggests an increased trend for IRR implementation in countries from 2006

to 2018 (2018 is when we stopped collecting policy data). Some countries, for example, those in the West African Monetary Union, Namibia, China, Paraguay, Colombia, and Armenia, have had some form of IRR in place since the dataset began (around 2001). However, a substantial set also imposed an IRR during the period (for example, Zambia, India, Bangladesh, Ecuador, Bolivia, Cambodia, and others). There is also substantial variation in where the IRR attached.

Table 17 – IRR Database

Country	Year Imposed / Earliest Date Recorded	Initial Cap	Min. in dataset	Max. in Dataset
Namibia	2001	23%	17%	29%
Benin	2001	27%	24%	27%
Burkina Faso	2001	27%	24%	27%
Cote d'Ivoire (Ivory Coast)	2001	27%	24%	27%
Guinea-Bissau	2001	27%	27%	27%
Mali	2001	27%	24%	27%
Niger	2001	27%	24%	27%
Senegal	2001	27%	24%	27%
Togo	2001	27%	24%	27%
Armenia	2001	27%	23%	45%
Thailand	2001	28%	28%	28%
Myanmar (Burma)	2001	30%	17%	30%
Colombia	2001	41%	25%	55%
Paraguay	2001	53%	38%	91%
South Africa	2006	35%	31%	46%
Ecuador	2007	32%	29%	42%
Macedonia	2009	14%	13%	18%
Bangladesh	2010	28%	27%	30%
Chile	2011	44%	32%	56%
India	2011	26%	25%	28%
Kyrgyzstan	2013	35%	33%	36%
Zambia	2013	58%	43%	67%
Bolivia	2014	12%	12%	12%
El Salvador	2014	187%	93%	203%
Azerbaijan	2015	29%	29%	29%
Nepal	2016	19%	13%	44%
Cambodia	2017	18%	18%	18%
Zimbabwe	2017	313%	313%	313%
Sri Lanka	2018	35%	35%	35%

Source: Multiple secondary sources and Author's calculations

We mostly relied on secondary sources to construct the dataset[177]. If there were inconsistencies between secondary sources, we reverted to primary legislation if possible. When the path of an IRR

[177] The collation process involved desktop research for each country: collating official reports, country-research reports, microfinance market reports, etc. We tried to have at least two secondary sources confirming a rule or a change in rule

was still unclear from primary sources (because of a lack of availability/publishing or our inability to translate), the country was not included in the sample.

In some cases (like Egypt), an usury rate limitation in law could be coupled with strong evidence that the laws are not actually enforced (Hoxhaj, 2010 and USAID, 2009). Therefore, we followed de jure rules unless there was clear evidence that rules were enforced differently de facto.

For countries with variable or relative IRRs that change each period (for example, Chile, Colombia, and Nicaragua, to name a few), we manually calculated the binding rule for each period based on primary data if the time series was not available directly from the national regulator). Where countries updated the rule more frequently than annually, we used year-averages. Some countries had IRRs specific to certain lending types, groups, or priority areas, for example, in Vietnam (Vietnam Microfinance Working Group, 2013). If rules were simple or the MiX lenders clearly fit within one category, we used that category, for example, in El Salvador, where rate ceilings were based on loan sizes and borrower types. Otherwise, we dropped them from the sample.

Tracing the rules over time was not always possible. In some countries, the IRRs are based on poorly published directives that were not easily traceable solely with desk research. We translated several countries' primary legislation into English; however, it was sometimes challenging to cleanly identify the history of a country's IRR regimes (this was sometimes true even when the source material was in English). As a result, countries in such categories were dropped from the dataset. For example, we were unable to follow the details of the Russian cap regime clearly enough to include in our sample, and we could not be sure that the lenders in the database for China were subject to their IRR.

The dataset is in no way perfect. With the time constraints, it was not possible to cleanly document the multiple hundred secondary sources used. Additional time and funding might allow discussions with analysts within countries to confirm our classifications and data with primary sources.

MiX database and MFI Outcomes

The MiX Market database provides financial and operational data for a strongly unbalanced panel of microfinance organisations from across 90+ countries spanning the period from 2000 to 2018. The data was initially collected to promote financial services for underserved communities and

before allowing a country into the final dataset, as well as trying to find evidence that a rule did not apply in a given country.

provide performance and financial information on microfinance institutions. It was made public in 2020 and now sits on the World Bank website.

The MiX data has been used extensively in the microfinance academic literature (for example, by Cull et al. (2016). We use the MiX as the primary MFI data source for our analysis, relying on it for financial and operational data. It provides three outcome variables that we will use to analyse the IRRs: (i) growth in the value of an MFI's loan portfolio, (ii) growth in the number of an MFI's outstanding loans, and (iii) an MFI's operational sustainability. It also provides reasonable MFI controls.

As a dataset, the MiX has both positive and negative characteristics. The limitations of the MiX for quantitative policy analysis have been documented (for example, in (Morduch 2016)). One significant issue is that the MiX is self-reported and subject to substantial selection bias; it is certainly not representative or a random sample within a country. It is biased towards lenders willing and able to complete the survey each year, so it typically under-represents smaller, less successful, and less profit-oriented MFIs. Coverage also differs substantially across countries. Take the coverage of Zambia, for example. The MiX only captures two large non-bank lenders and misses most of the market. It especially misses the smaller lenders targeting higher interest rate market segments who we now know were more likely to be negatively affected by the IRR. If we take the hypothesis that binding IRRs are likely to reduce MFI lending, then we might expect analysis using MiX to be biased towards a null of no effect. On the other hand, the dataset provides by far the most extensive, consistent cross-country coverage available.

Some limited interpolation was required due to incomplete MiX data submissions. Interpolation is not an unusual step for adjusting for single period missing years in the MiX (see, for example, Baquero et al. (2012)). We linearly interpolated any individual missing observations for a lender in the original levels data, contingent on the lender having at least four other observations. In total, approximately 1.6% of datapoints were interpolated across MFIs. The results in this chapter are not sensitive to removing all interpolated data. Our baseline results also winsorise the dependent variables at the 1% level. We aim to sample MFIs operating in normal, going concern conditions, and the dataset does contain some abnormally high (and low) growth rates that we do not want to drive our results. We make note when the baseline results are materially different from when non-winsorised data are used (which is seldom).

All raw financial data was converted to real, local currency units. We use real, local currency units to minimize foreign exchange effects, a concern had we used a common currency (e.g., the MiX USD dataset). We drop all lenders with average loans sizes larger than USD 5000 to orient the sample towards development microfinance; however, few of the results (especially qualitatively) are sensitive to this decision or when applying other reasonable cut-offs).

Macroeconomic data

Macroeconomic data for country-level controls were sourced from the World Bank World Development Indicators (WDI), the IMF World Economic Outlook (WEO), and the IMF International Financial Statistics (IFS).

Summary statistics are provided in Table 18 below. Note that some covariates, especially the macroeconomic variables, can have quite a short coverage, so regressions can have far fewer observations. Annex 16 provides a correlation table and a full country list.

Table 18 – MFI IRR Descriptive Statistics

Variable	Full Sample								
	Obs	Mean	Std. dev.	Coef.Var	Min	p25	p50	p75	Max
Growth in Net Loans (real, local currency)	11,723	0.21	0.49	2.30	-1.77	0.02	0.16	0.35	2.73
Growth in Number of Outstanding Borrowers	12,036	0.17	0.37	2.12	-0.86	0.00	0.12	0.30	1.68
Operational Sustainability (Financial revenue)/(financial expense+opp exp+imparement)	14,039	1.13	0.38	0.34	0.14	0.98	1.11	1.28	2.71
Log Number of Active Borrowers	14,046	9.19	2.05	0.22	0.69	7.88	9.21	10.53	16.01
Loan Share of Assets (%)	14,014	0.75	0.20	0.26	0.15	0.65	0.78	0.87	1.37
Average Outstanding Balance (USD)	14,047	816	987	1	0	159	399	1080	4995
Impairment Ratio (impairment exp/total loans)	13,829	0.04	0.05	1.23	0.00	0.01	0.03	0.05	0.36
Operating Expense Ratio (opp exp/total assets)	14,009	0.18	0.14	0.80	0.02	0.09	0.13	0.22	0.83
CPI Inflation (%)	13,987	6.3	6.0	0.9	-9.8	3.2	5.2	8.2	135.1
Import Growth Rate (%)	13,590	7.8	13.0	1.7	-53.6	1.2	7.0	13.9	108.8
Export Growth Rate (%)	13,595	6.8	11.7	1.7	-69.1	0.1	5.9	12.0	123.6
Government Revenue (% GDP)	13,954	21.5	8.1	0.4	3.0	16.1	19.9	25.2	67.3
Net Government Borrowing (% GDP)	13,929	-2.7	3.7	-1.4	-35.4	-4.6	-2.6	-0.7	31.0
Current Account Deficit (% GDP)	14,019	-2.8	7.7	-2.8	-65.0	-5.8	-2.2	0.2	63.4
GDP Growth Rate (%)	14,016	5.3	3.5	0.7	-28.0	3.4	5.3	7.1	53.4
Bank Lending Rate (%)	11,466	14.7	8.8	0.6	2.6	9.5	12.9	17.9	82.3
Bank Branches (per 1000 people)	12,091	11.0	10.8	1.0	0.1	4.9	8.4	13.8	92.3
Private Credit (% GDP)	13,820	30.9	19.2	0.6	0.4	16.9	27.5	42.8	156.2
Real Interest Rate (%)	11,448	7.5	8.8	1.2	-39.1	2.6	5.7	10.8	60.9

Descriptive Stats	**Total**	**Always Under IRR**	**Never IRR**	**Change in IRR Regime**
No. lenders	2485	341	1,859	285
Total Obs.	14,039	1,651	9,534	2,854
Avg. no. years	6-7 years			
Total Treatment -MFI Years	3,089			

Results

We now provide basic descriptive and conditional statistics using the full dataset. For now, we are interested in simple associational results: do MFIs in countries with interest rate cap regimes grow faster or slower than those without such regimes? Do lenders operating under an IRR have lower operational sustainability than those not operating under one? Are more stringent IRR regimes associated with lower loan growth rates?

In this analysis we include lenders from all countries and with all types of IRR exposure (i.e., those that were never under an IRR, those that were always under an IRR, and those that saw a change in an IRR). We also explore several sources of variation within the data by using unit-fixed effects, common time-fixed effects (which are clearly less meaningful at the global level than the country-level), and a combination of both. These descriptive results are not causal and instead aim to capture simple average differences and correlations within the dataset. Note that unlike the event study

approach used in the Zambia case study, these comparisons are analysing growth rates and level changes in operational sustainability with respect to a simple binary IRR treatment dummy, rather than modelling dynamics.

Simple dummy variable IRR modelling

We first investigate whether lenders operating under an IRR regime grow their portfolios slower than those not operating under an IRR regime. To do so, we run versions of a simple regression with a binary IRR indicator variable:

$$\Delta Y_{ijt} = \alpha_{ij} + \beta_2 . Cap_{jt} + \boldsymbol{\beta X_{ijt-1}} + \boldsymbol{\beta M_{jt}} + \varepsilon_{ijt} \qquad (6.2)$$

Where ΔY_{ijt} represents the growth rate in the value of real net outstanding loans, the growth rate in the number of net outstanding loans, or the level of operational sustainability (the latter being defined as Financial Revenue / (Financial Expense + Net Impairment Loss + Operating Expense) and reflects how well a lender is able to cover its' expenses). Cap_{jt} is a dummy variable that takes the value of 1 when an IRR regime was in place in country j during year t and 0 otherwise. α_{ij} is an MFI-specific fixed effect that captures time-invariant MFI characteristics. X_{ijt-1} is a vector of lagged institution-specific covariates, including the lagged values of the number of active borrowers in a lender's portfolio, the loan share of a lender's portfolio, the average loan size in a lender's portfolio (in USD), and a lender's impairment ratio. We do not include a profitability measure because we want to allow profitability to vary with the IRR for the loan growth rate outcomes. M_{jt} is a vector of macroeconomic and country-level controls including the inflation rate, import growth rate, export growth rate, the number of bank branches per 1000 people, and the real interest rate. Government revenue, net government borrowing, the stock of government debt, the current account, and net private debt, all as a percentage of GDP, are also included. Together, these attempt to control for macroeconomic developments within a country, as well as its level of financial development[178].

Our results show that lending growth for lenders operating under an IRR was lower on average than that for lenders not operating under a cap regime, both as a simple unconditional average as well as conditional on covariates and fixed effects (Table 19). Lenders under IRRs grew their loan books (net of losses) by 5.8 percentage points lower than those not under IRRs (column 1). We also

[178] One limitation of the MiX dataset is that there is not enough data in country cross sections to use country-time fixed effects.

recorded lower growth rates when we examined only within-lender variation (-13.8 percentage points, column 3), or when using global time-fixed effects to control for common year effects or TWFE (-12.1 percentage points and -11.8 percentage points respectively, see columns 4 and 5)[179]. Columns 3,8, and 13 essentially throw out the variation between never-capped and always-capped institutions, concentrating purely on within country comparisons, while columns 4, 9, and 14 allow for common time-period shocks across countries, but does not control for time-invariant lender characteristics. We find similar results for the number of outstanding loans, but less consistently for the level of operational self-sustainability (column 6-10, Table 19).

While these initial associations between IRRs and credit market growth and sustainability might summon a pause for thought when it comes to implementing such a policy, associations are far from conclusive evidence. For one, there are several reasons why we might expect MFI growth within a country implementing an IRR to be higher or lower, independent of the implementation of IRRs. For example, countries with high structural lending interest rates and expensive lending environments (and hence, perhaps, low loan growth rates) might be more likely to see the implementation of IRRs as a political response to this environment (although time-invariant factors should be absorbed with the lender-fixed effects). Alternatively, IRRs might be more likely to be implemented when economic conditions are systematically poor (or good). Such dynamics and any other endogenous links between IRRs and lender reactions will still confound any causal interpretations. In our Zambia analysis, we believe that the imposition of the cap was exogenous, strengthening the argument that the cap was not imposed in response to economic events. However, in a cross-country analysis such assumptions are not plausible.

The covariate estimates from these initial MFI results appear reasonable, although these are not necessarily well-specified determinants-style regressions. There is evidence of size convergence, whereby smaller lenders tend to grow faster (row 2, column 2-10, Table 19). Larger lenders (by number of loans) are not clearly more operationally self-sufficient. Higher asset loan shares are also associated with lower loan growth and higher operational sustainability (row 4). Higher operating expenses are associated with lower operational sustainability, whereby a 10-percentage point increase in operating expenses is associated with an approximately 0.05 decrease in operational sustainability (row 5, columns 11-15). We included several country-level macroeconomic variables

[179] Note that for many countries the data is not deep enough to use county-time fixed effects, a serious shortcoming.

in the regressions to account for different types of common county-level shocks and trends, although many of these variables clearly move together (Table 44). Nevertheless, some associations in the results seem reasonable, but almost certainly need a more precise analysis to make any meaningful comment.

Table 19 – IRR regimes and Average MFI Average Outcomes

Notes: This table presents estimates regressing various MFI lending or sufficiency outcomes (the dependent variable) against a dummy variable indicating whether or not the bank was bound by an Interest Rate Cap in a given year. The regressions are approximately based on equation (6.2) in the text. Each column represents a regression using a different estimation approach (POLS = Pooled Ordinary Least Squares, Within = Lender Fixed Effects, Time = Time Fixed Effects and Country Fixed Effects, TWFE = Two-way Fixed Effects). Lender data is annual and was taken from the World Bank's MiX dataset. Lender-specific controls include the following variables: the log of the number of active borrowers in a lender's portfolio, the loan share, the average loan size, as well as a lender's impairment ratio, all lagged by one year and winsorised. Macro variables include the inflation rate, import growth rate, export growth rate, the number of bank branches per 1000 people, and the real interest rate, as well as government revenue, net government borrowing, the stock of government debt, the current account, and net private debt, all as a percent of GDP. Standard errors are clustered at the Country level and are provided in parentheses. Dependent variables winsorised at 1%. *** p < 0.01, ** p < 0.05, * p < 0.1.

Dependent Variable	Net Loan Growth (Percent)					No. of Active Borrowers Growth (Percent)					Op. Self Sufficiency				
	POLS	POLS	Within	Time	TWFE	POLS	POLS	Within	Time	TWFE	POLS	POLS	Within	Time	TWFE
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
IRR Regime Dummy	-0.058**	-0.022	-0.138*	-0.121*	-0.118*	-0.056***	-0.04	-0.109**	-0.092*	-0.088**	-0.021	-0.094***	-0.03	-0.063*	-0.066*
	[0.025]	[0.032]	[0.076]	[0.065]	[0.061]	[0.020]	[0.030]	[0.049]	[0.054]	[0.042]	[0.026]	[0.030]	[0.042]	[0.037]	[0.039]
Log Number of Active Borrowers		-0.020***	-0.202***	-0.054***	-0.204***		-0.023***	-0.233***	-0.110***	-0.239***		0.005	0.013	0.002	0.004
		[0.006]	[0.018]	[0.009]	[0.017]		[0.005]	[0.010]	[0.008]	[0.010]		[0.006]	[0.011]	[0.005]	[0.009]
Loan Share of Assets (%)		-0.258***	-0.423***	-0.481***	-0.482***		-0.125**	-0.079	-0.239***	-0.111**		0.435***	0.401***	0.415***	0.429***
		[0.049]	[0.078]	[0.046]	[0.067]		[0.048]	[0.066]	[0.047]	[0.050]		[0.057]	[0.048]	[0.038]	[0.042]
Ave. Loan Size (USD)		-0.021	-0.115***	-0.028**	-0.100***		0.021***	0.044**	0.054***	0.047***		-0.004	0.004	0.009	0.013
		[0.014]	[0.027]	[0.013]	[0.025]		[0.007]	[0.019]	[0.010]	[0.015]		[0.007]	[0.010]	[0.006]	[0.011]
Impairment Ratio (impairment exp/total loans)		-1.095***	-0.457	-0.686***	-0.568**		-1.213***	-1.094***	-0.967***	-1.041***		-0.560***	-0.478***	-0.542***	-0.567***
		[0.264]	[0.329]	[0.187]	[0.246]		[0.176]	[0.201]	[0.134]	[0.142]		[0.147]	[0.174]	[0.138]	[0.149]
Operating Expense Ratio (opp exp/total assets)		0.102	0.1030	0.272***	0.1380		0.022	-0.106	0.01	-0.139		-0.918***	-0.555***	-0.781***	-0.641***
		[0.111]	[0.234]	[0.105]	[0.179]		[0.073]	[0.178]	[0.109]	[0.138]		[0.100]	[0.040]	[0.042]	[0.048]
CPI Inflation (%)		-0.122	-0.395				0.057	-0.196				-0.004	-0.082		
		[0.409]	[0.542]				[0.150]	[0.151]				[0.168]	[0.121]		
Import Growth Rate (%)		0.335***	0.195***				0.084	0.006				0.083*	0.052		
		[0.072]	[0.064]				[0.063]	[0.057]				[0.046]	[0.036]		
Export Growth Rate (%)		-0.190**	-0.157				-0.063	-0.069				-0.0050	-0.085**		
		[0.083]	[0.106]				[0.054]	[0.056]				[0.066]	[0.040]		
Government Revenue (% GDP)		0.0380	-0.704*				-0.05	0.400				0.178	-0.151		
		[0.126]	[0.418]				[0.077]	[0.253]				[0.172]	[0.356]		
Net Government Borrowing (% GDP)		-0.113	1.426***				-0.371	-0.068				0.319	0.106		
		[0.497]	[0.341]				[0.404]	[0.285]				[0.441]	[0.313]		
Current Account Deficit (% GDP)		0.206	-0.539***				0.136	0.014				-0.085	0.055		
		[0.197]	[0.134]				[0.108]	[0.123]				[0.223]	[0.130]		
GDP Growth Rate (%)		0.783	-0.268				1.285***	0.656***				-0.035	0.644***		
		[0.724]	[0.686]				[0.348]	[0.219]				[0.278]	[0.165]		
Bank Lending Rate (%)		0.0220	0.296				-0.072	-0.039				0.111	-0.169		
		[0.263]	[0.535]				[0.146]	[0.194]				[0.178]	[0.190]		
Bank Branches (per 1000 people)		-0.324***	0.422				-0.283***	0.467*				0.086	-0.054		
		[0.122]	[0.360]				[0.101]	[0.243]				[0.085]	[0.217]		
Private Credit (% GDP)		-0.008	0.003				0.046	0.321**				-0.206**	-0.412**		
		[0.075]	[0.195]				[0.083]	[0.135]				[0.086]	[0.172]		
Real Interest Rate (%)		-0.025	0.211				-0.054	-0.038				-0.023	-0.043		
		[0.231]	[0.244]				[0.143]	[0.131]				[0.156]	[0.119]		
Constant	0.227***	0.614***	2.637***	1.498***	2.366***	0.184***	0.481***	2.246***	1.500***	2.346***	1.133***	0.985***	1.010***	0.663***	0.941***
	[0.020]	[0.104]	[0.216]	[0.102]	[0.160]	[0.018]	[0.062]	[0.084]	[0.067]	[0.103]	[0.016]	[0.075]	[0.127]	[0.049]	[0.068]
Observations	11723	7913	7913	10730	10730	12036	8008	8008	10919	10919	14039	8007	8007	10919	10919
R-Squared	0.003	0.054	0.171	0.133	0.160	0.004	0.090	0.333	0.3	0.332	0.001	0.205	0.098	0.089	0.091
Lender-Specific Controls	X	√	√	√	√	X	√	√	√	√	X	√	√	√	√
Lender Fixed Effect	X	X	√	X	√	X	X	√	X	√	X	X	√	X	√
Country Fixed Effects	X	X	X	√		X	X	X	√		X	X	X	√	
Time Fixed Effects	X	X	X	√	√	X	X	X	√	√	X	X	X	√	√
Macro Controls	X	√	√	X	X	X	√	√	X	X	X	√	√	X	X
Error Cluster Level	Country	Country	Country	Country	Country	Country	Country	Country	Country	Country	Country	Country	Country	Country	Country

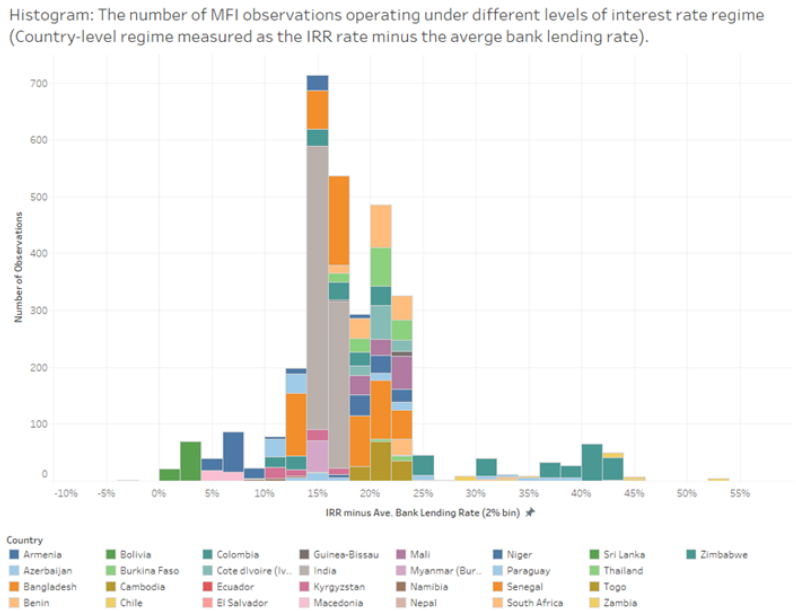
Country-wide IRR intensity approach

Next, we examine whether lenders operating in harsher cap regimes grew less quickly than those in a weaker regime or those with no regime. In essence, we want to model variation in IRRs across countries, as well as the background interest rate environments in each. To do this we change equation (6.2) to include a cap intensity interaction term, $\beta_3 . Cap_{jt} . (r_{cap,jt} - r_{bank,jt})$, which should allow for lenders in countries with high IRRs relative to the background interest rate environment to see larger or smaller effects. The intuition is that an IRR set at 30% might have quite different effects if it were imposed in a market where the bank lending rate was 5% versus if it was 30%. If we think of the average bank rate as a low-risk, base lending rate for the credit market, we are examining how far above this base lending rate the cap sits. We use the following regression:

$$\Delta Y_{ijt} = \beta_{ij} + \beta_2 . Cap_{jt} + \beta_3 . Cap_{jt} . (r_{cap,jt} - r_{bank,jt}) + \boldsymbol{\beta X_{ijt-1}} + \boldsymbol{\beta M_{jt}} + \varepsilon_{ijt} \qquad (6.3)$$

Where $r_{bank,jt}$ is the average bank lending rate for a MFI's given country. When estimating equation (6.3), we would expect β_3 to be negative if MFIs operating in countries with a more binding cap grow less quickly. Figure 38 provides the distribution of $r_{cap,jt} - r_{bank,jt}$ for MFIs facing IRRs, where lenders to the left face tighter regimes.

Figure 38 – Histogram of Country-Level IRR Intensity



Note: Figure provides the distribution of $r_{cap,jt} - r_{bank,jt}$ for MFIs facing IRRs, where lenders to the left face tighter regimes (MFIs are grouped by country, given that the measure is at the country level). Source: MiX, IMF WEO, WB WDI, Author's calculations.

We do not find evidence of an association between lenders operating under a more severe IRR and lower lending growth rates, either in terms of loan values or loan numbers (Table 20 columns 1-8). Remember, severe here is when the cap is low and close to the bank rate, so as $(r_{cap,jt} - r_{bank,jt})$ reduces, the cap is more binding. We still find some evidence of a level growth shift due to the presence of the cap (estimates are typically in the same range as in Table 19), yet do not find the β_3 coefficient from the interacted $r_{cap,jt} - r_{bank,jt}$ term from equation (6.3) to be significantly different to zero. This non-result is robust to using non-winsorised data. The economic size of the interaction estimates is also typically small. If the largest coefficient were significant (row 2, column 6 in Table 20) it would suggest that, given a cap was imposed, each 10 percentage point loosening in $r_{cap,jt} - r_{bank,jt}$ (i.e., a loosening of the cap regime) would increase the growth in the number of MFI loans by only 2 percentage points.

We do, however, find that a lender's operational sustainability (on average) appears to reduce less for lenders operating under a less binding IRR regime (although not in the pooled OLS cross-section, but that could be due to many other factors, such as lender and country effects). For each additional ten percentage points the IRR regime increased above the bank rate, operational sustainability was associated with an increase of 0.04 percentage points (column 2 rows 10-12 in Table 20). This is still reasonably small given the mean operational sustainably of 1.12 (std. dev of 0.34), but it does "undo" the cap's "effect" if the IRR is imposed approximately 30-40 percentage points above the bank rate. Note that only 5% of MFIs facing an IRR saw a cap that was 40 percentage points above the bank rate – most regimes were typically more constraining than that.

Note that data quality is a concern with this analysis. Several regions appear to have implausibly low IFS or WDI average bank lending rates, which might bias the results. In both the WAEMU and CEMAC, for example, the bank rate data seems very low. Low bank rates may reflect other policy settings (for example, IRRs on banks) or incorrect data entry. Yet the results still do not significantly change if African MFIs are removed from the dataset. To move forward, we need to be more precise with capturing the variation across MFIs within a country rather than only measuring IRR binding at the country level. We suggest this as an area for future research in the following discussion.

Table 20 – IRR regimes and Simple Country-Level Cap Intensity Interaction

Notes: This table presents estimates regressing various bank lending and sufficiency outcomes (the dependent variable) against a dummy variable indicating whether or not the bank was bound by an Interest Rate Cap in a given year and an interaction variable based on the difference between a country's IRR and its average bank lending rate. The regressions are approximately based on equation (6.3) in the text. Each column represents a regression using a different estimation approach (OLS = Ordinary Least Squares, Within = Bank Fixed Effects, Time = Time Fixed Effects and Country Fixed Effects, TWFE = Two-way Fixed Effects). Lender data is annual and was taken from the World Bank's MiX dataset. Controls consistent with Table 18. Robust standard errors are clustered at the Country level and are provided in parentheses. Data not winsorised. *** p < 0.01, ** p < 0.05, * p < 0.1.

Dependent Variable	Net Loan Growth (Percent)				No. of Active Borrowers Growth (Percent)				Op. Self Sufficiency			
	POLS	Within	Time	TWFE	POLS	Within	Time	TWFE	POLS	Within	Time	TWFE
IRR Dummy	-0.057	-0.156*	-0.125	-0.164*	-0.067*	-0.139**	-0.11	-0.149***	0.013	-0.095*	-0.130***	-0.101**
	[0.047]	[0.085]	[0.084]	[0.091]	[0.036]	[0.055]	[0.069]	[0.056]	[0.058]	[0.051]	[0.050]	[0.051]
IRR Dummy*(interest rate cap - bank lending rate)	-0.037	0.135	-0.17	0.113	0.087	0.197	0.033	0.143	-0.235	0.423**	0.391***	0.389**
	[0.159]	[0.254]	[0.215]	[0.263]	[0.125]	[0.138]	[0.139]	[0.180]	[0.259]	[0.184]	[0.143]	[0.168]
Observations	11030	7913	10127	7913	11333	8008	10313	8008	13235	8007	10313	8007
R-Squared	0.003	0.171	0.131	0.184	0.003	0.333	0.299	0.344	0.002	0.1	0.089	0.104
Bank-Specific Controls	X	√	√	√	X	√	√	√	X	√	√	√
Bank Fixed Effect	X	√	X	√	X	√	X	√	X	√	X	√
Country Fixed Effects	X	X	√		X	X	√		X	X	√	
Global Time Fixed Effects	X	X	√	√	X	X	√	√	X	X	√	√
Macro Controls	X	√	X	X	X	√	X	X	X	√	X	X

Discussion and Conclusions

This brief and exploratory Chapter analysed associations between IRRs and MFI outcomes across countries using a newly constructed dataset. We found evidence that MFI loan growth rates were lower when operating under an IRR, both in terms of the value of their portfolio as well as the number of outstanding loans. Further MFI operational sustainability was typically lower for MFIs operating under an IRR regime. The magnitudes of the associations were also economically meaningful and typically robust to the inclusion of covariates. MFIs constrained by an IRR regime recorded loan portfolio growth that was 13.8 percentage points lower than that of the same MFI when not constrained. Similarly, the growth in the number of outstanding loans was lower by 10.9 percentage points.

These results, although in no way causal, are directionally consistent with the findings from our Zambia case study in Chapter 5. Again, we see lower growth on average for MFIs when they were constrained, compared to when they were not. We did not necessarily expect to see this in the cross-country analysis given that our Zambia sample covers MFIs of all sizes and market segments, whereas the MiX sample skews towards typically larger lenders (who are prepared and have the capacity to upload detailed data).

We expected that tighter cap regimes (measured relative to the bank lending rate) might have induced a stronger and more negative association with MFI outcomes. Using a simple interaction model, however, we only find evidence for such an association for MFI's operational self-sufficiency – whereby more bound MFIs saw lower levels of operational self-sustainability.

To further the analysis in this Chapter, we have two (not mutually exclusive) suggestions for future research. The first would be to try modelling cap bindingness at the MFI level. While the MiX has limited usable interest rate data, it might be possible to use an MFI's interest yield data to better measure bindingness at the lender (rather than the country) level. Binary treatment approaches otherwise average over what might be substantial heterogeneity in MFI responses. As our IRR database quantitatively tracks the level of interest regulations over time, it should be possible to capture some form of IRR exposure at the MFI-level.

The second is to undertake an event study to analyze the dynamics surrounding the imposition of the IRR regimes. The descriptive statistics and correlations provided in this paper, especially those using within MFI variation, indicate some relationship between IRRs being in place and poorer microfinance outcomes. Yet the estimates average over dynamics and do not allow for causal inference. Additional research using an event study methodology could help provide more causal evidence regarding IRRs by building a more credible counterfactual of non-bound lenders. Table 21 shows there are approximately 150-200 MFIs across countries in the sample with the appropriate data dimensions that saw a change in cap regime and could be used in an event study. Note that while results would be clearly driven by a smaller set of country experiences, namely India, Bangladesh, Ecuador, and Bolivia, the sub-sample would still cover some 50 million affected borrowers[180].

[180] An event study analysis would not be without its own econometric challenges. Holding aside whether comparisons across countries is a valid approach to building a counterfactual in this case, there has been increasing skepticism in the last couple of years concerning difference-in-difference and event studies in a setting with staggered implementation or heterogeneous or dynamic treatment effects (Roth et al, 2022; Goodman-Bacon, 2018; Sun and Abraham, 2020), likleyt he setting with IRRs and this MIX dataset. Having a large set of non-treated MFIs is useful, but analysis would likely need to use estimators that adjust for these dynamic/heterogenous/staggered issues, such as Callaway and Sant'Anna (2020) or and Sun and Abraham (2021).

Table 21 – IRR Imposition Timing and Sample Size

	Time to IRR Imposition										
Country	**-5**	**-4**	**-3**	**-2**	**-1**	**0**	**1**	**2**	**3**	**4**	**5**
Azerbaijan	9	11	12	12	12	13	13	11	9		
Bangladesh	24	28	29	30	30	31	31	31	29	28	27
Bolivia	21	21	21	21	21	21	19	17	16	16	
Chile	3	3	3	4	4	5	4	4	4	3	3
Ecuador	17	22	23	23	36	41	40	36	34	32	31
El Salvador	11	11	11	11	11	11	8	2	1	1	
India	43	49	57	73	97	112	87	80	73	72	67
Kyrgyzstan	7	7	7	7	8	8	8	7	6	6	6
Macedonia	4	4	4	4	4	4	4	4	4	4	3
Nepal	8	8	8	8	8	8	6	6			
South Africa	1	1	1	2	3	3	2	2	2	2	2
Zambia		0	1	1	2	3	3	3	3	3	1
Grand Total	148	165	177	196	246	270	234	212	190	176	150

Note: Table shows the number of MFI observations for a given country in relation to the time (in years) until the imposition of the IRR regime in that country. For example, in the year the India imposed its IRR, 112 Indian MFIs recoded data in that given year.

Source: MiX and Author's calculations

Chapter 7 – Zambia's banking system and IRRs: Bank Margins and Fees

Introduction and Background

In Chapter 5, we found evidence that the IRR regime affected Zambia's non-banks. We found broad support for the idea that bound non-banks: (i) reduced lending in the wake of the caps being introduced and (ii) increased their use of fees and commission income. Yet Zambia's interest rate caps were also targeted (perhaps primarily) at the banking sector. When implemented in January 2013, the IRR on banks was set at 18.25%, or 9% plus the BOZ rate. Although the regime was less binding on the banking sector than the non-banking sector, there is evidence that the caps bound at least some lending in the banks' portfolios[181]. Hence, there is value in analysing more specifically how the IRRs may have affected Zambia's banking sector.

This chapter will empirically test for two potential impacts of IRRs on banks in Zambia. Using a detailed bank-level panel dataset, we will first test whether Zambia's interest rate caps were associated with changes in banks' costs of financial intermediation[182]. The efficiency of financial intermediation is important because it is a driver of the depth, breadth, and accessibility of the banking system (Calice and Zhou, 2018), which in turn can play an important part in supporting broader development processes (Rajan and Zingales, 1998; Levine and Zervos, 1998; Beck et al., 2000; Levine, 2004; and others). In Zambia, as with much of SSA, the role of an efficient banking system is perhaps even more important because domestic firms and households have few alternatives to banks[183]. If IRRs influenced the cost of intermediation – and there are some theoretical reasons why this might be so – this would clearly be relevant to policymakers[184].

[181] For example, BOZ's loan-level data on banks' largest 20 loan exposures shows that the cap likely bound at least some banks, especially those providing higher-risk Kwacha-denominated loans. In the year preceding the cap, the median bank's average interest yield was recorded in the 14-15% range with median Kwacha-denominated lending yields nearly twice as high as foreign-denominated corporate lending; hence it is plausible the cap constrained some loans.
[182] Financial intermediation is the process by which a financial system pools and transfers excess funds from deposits to borrowers, providing liquidity transformation, transforming/diversifying risks, and minimising information asymmetries between depositors and lenders (Claus and Grimes (2003) and Diamond (1996)).
[183] Large/international firms can and do acquire financing from abroad and at lower rates.
[184] Direct measurements of the efficiency of financial intermediation are difficult (Bernanke, 1983), thus a strategy used in much of the financial development literature is to use banks' margins as a proxy measure for the efficiency of financial intermediation. See, for example, Angbanzo (1997), Demirgüç-Kunt, and Huizinga (1999), Maudos and Fernandez (2004), Maudos and Solis (2009), Entrop et al. (2015).

The second effect we will examine is whether lending institutions offset IRRs restrictions by shifting income from regulated interest payments to other unregulated fees and commissions, an analysis similar to our work on Zambia's non-banking sector. As explained in Chapter 5, there is reason to believe that fees and commissions are a margin that lenders might exploit if faced with an IRR (see Central Bank of Kenya (2018), Ferrari et al. (2018), and Maimbo and Gallegos (2014)). As with the non-banking sector, such a structural shift may harm price transparency in the loan market, worsening borrower outcomes[185].

Examining Zambia's IRR regime will also require a detailed analysis of bank margins in Zambia, a piece of analysis which will itself contribute to the literature. Over the last few decades, banks' margins in Zambia have been relatively high, well above the SSA and low-income country averages[186]. There is clearly value in better understanding the determinants of these high margins, especially given that they likely contributed to elevated lending interest rates and thus may have partly underpinned the imposition of the IRR regime in the first place. In terms of contributions, this chapter will hence provide one of the first in-depth analyses of bank margin determinants in Zambia.[187].

From our analysis, we do not find evidence that the IRR regime either influenced bank margins in Zambia or that it likely impacted banks' fee/non-interest income shares. Additionally, we do not find evidence that exposure to the IRR, proxied by interest yields, mattered for banks' margin outcomes. We are wary not to interpret a null effect as evidence of no effect, especially given that our sample is small and perhaps underpowered for Zambia's IRR regime (implemented at 9% above the BOZ rate).

Our analysis of the determinants of bank margins in Zambia does, however, provide some clearer insight. As with most of the literature, we found that operating costs were a consistent and significant driver of banks' margins in Zambia, which suggests that structural market reforms to improve bank lending and efficiency may be a worthwhile policy option for reducing margins. Unlike Bologna

[185] Note that in this chapter we are not analysing the effects of the IRR regime on bank lending volumes. This is due to the huge, forced recapitalisations that banks began in 2012. The theoretical link between bank capitalisation and the cost of financial intermediation is less clear, thus we believe there is still scope for meaningful analysis for this paper.
[186] See data on Zambia in Calice and Zhao (2018)
[187] This analysis will contribute to bank margin literature in a SSA context (see, for example, analysis of bank margins in Uganda by Beck and Hesse (2009), in Malawi by Chirwa and Mlachila (2004); Kenya by Were and Wambu (2013), and across banks in SSA by Ahokpossi (2013) and Boutin-Dufresne (2013)) as well as using more detailed margin determinants than is often possible in a low-income setting.

(2018), we did not find that contractual maturity mismatch was a strong driver of Zambian banks' margins. We also found potential evidence of heterogeneity between bank types: larger more diversified banks' margins appear to be driven more by macroeconomic and liquidity factors, while smaller, less diversified banks appear to have their margins driven by scale and concentration.

The remainder of the chapter is as follows: it briefly reviews the literature on bank margin determinants that sits underneath our empirical analysis of Zambia's IRRs. We then undertake an informal theoretical framing to help outline how we might expect IRRs to influence bank markings. To understand the context of bank margins in Zambia, we outline recent historical trends in bank margins and undertake a simple bank net interest margin decomposition. Finally, we analyse (i) the determinants of bank margins in Zambia, and (ii) whether there is evidence that the IRR regime affected such margins.

Relevant Literature (bank margin determinants)

This chapter closely relates to, borrows from, and contributes to the empirical literature on bank margins and bank margin determinants. This is in addition to the empirical literature on interest rate caps that we have already discussed in detail in Chapter 4.

There is little previous research quantitatively examining the impact of interest rate caps on bank margins. This leaves limited direct research to draw upon. Yet, there is a substantial literature examining the determinants of bank margins, as well as studies extending such a 'bank margin determinants'-type framework to examining how other policy changes or events impact margins on which we build.

Hence, our empirical strategy draws from a lengthy literature that evaluates the factors underpinning the costs of intermediation, including bank-specific, industry structure, and macroeconomic factors (see Ho and Saunders (1981), Allen(1988), Angbanzo (1997), Demirgüç-Kunt and Huizinga (1999), Maudos and Fernandez (2004), Maudos and Solis (2009), Beck and Hesse (2009), Entrop et al. (2015), Calice and Zhao (2018)). These studies provide a baseline approach for our empirical specification as well as a theoretical basis to determine which variables might be relevant for our analysis (see Table 24 for a detailed discussion of the variables we use and where they relate to the literature). In addition to the interest rate cap analysis, this chapter also relates and contributes to the

literature on margin determinants in SSA, for example, Beck and Hesse (2009), Chirwa and Mlachila (2004), Were and Wambu (2013), Ahokpossi (2013), and Boutin-Dufresne (2013).

Finally, we draw from studies examining the influence of policy changes on bank margins. Claessens, Demirgüç-Kunt, and Huizinga (2001), for example, use a large cross-country panel to test the impact of changes in bank ownership (between majority foreign shareholders versus majority domestic shareholders) on banks' net margins and profitability. Beck and Hesse (2009) and Brock and Franken (2003) use bank-level country-specific panel data and dummy variables to isolate the association between bank margins and a policy or event. Beck and Hesse (2009), for example, found that the privatisation of one of Uganda's large banks did not likely impact margins in the market. Our analysis on IRRs also adds to this literature by allowing for potential bank-level (linear) heterogenous policy responses to an event.

Theoretical Framing (IRRs and bank margins)

This section examines the primary mechanisms through which the cap regime may have influenced banks' margins. This is distinct from the extensive theoretical and empirical literature examining the potential determinants of interest rate margins, which we will draw upon to help us construct/specify our margin regressions in a later section (including Table 24).

As with the theoretical framing for microfinance lenders' lending, there may be several plausible models underpinning banking margins under IRRs. Modelling bank margins can be slightly more complex than the simple microfinance lender model we used earlier because banks operate and make decisions conjointly in both loan and deposit markets.

We briefly consider three key mechanisms through which caps might directly impact margins and the cost of financial intermediation. To summarise, the first is to view caps as a correction for imperfect market competition in the banking sector based on market power or imperfect information. The second mechanism is the influence caps may have on banks' portfolio diversification efficiency, and whether this flows through to the cost of financial intermediation. And finally, whether banks insulate their margins from the cap by shifting income from loan revenue (interest income) to non-interest income (for example, fees and commissions).

This first mechanism - that caps can reduce bank margins - is supported analytically by analysis in Courakis (1984), which models the impact caps have on market interest rates and lending volumes in a classic loan and deposit market setting under assumptions of perfect competition and monopoly.

The high-level idea is that if financial providers make large margins and rents from uncompetitive behaviour and market power (the justification for caps by some policymakers), capping rates may place downward pressure on abnormal profits. In this case, the adjustment of margins downwards after the implementation of caps may indicate the success of the cap regime.

One key difference between Courakis (1984) and much of the work done in the US on usury laws is that Courakis (1984) directly models the deposit market in addition to the loan market[188]. Furthermore, he breaks away from the assumption of perfect competition. Within a loan and deposit market setting and assuming a monopoly market structure, a binding price cap will lower the interest rate the bank can charge, increasing the volume of loans it optimally supplies (to the point where the cap meets the bank's marginal cost of loans). To finance these additional loans, the bank entices more deposits by increasing its deposit rate. The combined impact of these effects is to reduce the bank's spreads and margins[189]. Courakis' (1984) analysis predicts that the impact of the caps on margins will be a function of the market structure of the banking system, the relative elasticities of deposit and lending markets, and where the Government placed the cap relative to the hypothetical monopoly or perfect competition price[190]. In this setting, we expect both, the interest yield (-) and interest cost (+) components of the margin calculation to adjust to the cap regime, reducing margins.

The applicability of Courakis' (1984) framework is contingent on the level of competition in Zambia's banking market. If there is strong evidence of imperfect competition, we would want to place more weight on this potential spread and margin-reducing dynamic. Simpasa (2010) applies the Panzer-Rosse approach to Zambia's banking sector, an approach that assesses the competitive climate in an industry. He finds that since liberalisation, the market has been characterised by oligopolistic competition, suggesting that the market context for a margin effect, as described in Courakis (1984), is potentially appropriate. One caveat is that competition in Zambia's

[188] See Ostas (1976), Vandenbrink (1980), and Chicago Federal Reserve (2001) for perfect, loan-market only frameworks.
[189] Under perfect competition, the volume of loans will decrease with the implementation of a cap, but the bank spread – already at zero because of the competitive pressures – will remain at zero.
[190] This is an important point, and one that carries over from our simple MFI model. If the IRR were placed so low that bank lending was not profitable, we might expect bank withdrawal. If the cap is placed extremely high, we might expect no effect.

banking market has been steadily increasing, and this has been the case since Simpasa's (2010) estimates.

The second mechanism is how IRRs may influence banks' abilities to diversify and optimise their portfolios, potentially impacting the cost of financial intermediation. Examining banks' decision-making processes from an asset management/portfolio optimisation perspective, interest caps may increase intermediation costs and bank margins by potentially adjusting downward the risk/return price point for loan types that are bound by the cap. Hart and Jaffee (1974) generalise the mean-variance approach to financial intermediation to account for multiple assets (and liabilities) and derive a bank's opportunity set. In this setting, constraining loan returns for higher-risk market segments (a possibility given the caps) may limit a bank's ability to construct efficient portfolios by limiting its capacity to diversify across loan assets and in relation to its deposit liabilities. This, in turn, could impact intermediation efficiency. Pyle (1971) shows that, under a mean-variance model of financial intermediation, intermediation is more likely the smaller the standard deviation of aggregated bank loan yields and the lower the correlation between loan and deposit yields. Constraining loan assets' risk/return characteristics with a cap may influence the likelihood and cost of intermediation.

Within these frameworks, the interest rate cap's influence on a bank's portfolio will depend on the covariance between the cap-exposed assets and the bank's other assets/deposits, and the bank's risk preferences. However, the analytical comparative statics of how the interest rate caps could work through a bank's portfolio optimisation are not explored here, so we are cautious in placing too much weight on the role of portfolio effects resulting from the cap. Moreover, collecting and meaningfully quantifying data on portfolio diversification effects is not easily possible with our dataset. Nevertheless, the anecdotal evidence is that credit supply from banks to SMEs reduced in the wake of the cap (World Bank, 2013), which may (although perhaps not materially) have had efficiency and intermediation consequences.

Another related example of a portfolio effect that may dampen the impact of the caps on margins is that banks could ration credit to lower-risk borrowers, de-risking their portfolios (Vandenbrink, 1982). They could do this by increasing the stringency of lending terms by, for example, raising collateral or down payment requirements, increasing minimum loan sizes, or increasing credit score thresholds. These actions help retain profitability at a lower interest rate by reducing expensive

default and collection costs, but for fewer disbursements. Exactly how this dynamic would feed through into intermediation margins is unclear and likely a function of competition, interest rate and deposit elasticities. It is possible that such effects might dampen the impact of the cap on bank margins, although they would be coupled with a reduction in both the riskiness and quantity of loan disbursements. Unfortunately, it is difficult to measure the decrease in the risk profile of disbursements, and it will be challenging to untangle this effect given our dataset. Linked to this risk/loan retrenchment, other investment opportunities for banks (such as tilting their balance sheets towards government securities or foreign deposits), may also insulate banks' margins.

Finally, if banks' lending is bound by caps, they may use non-price lending dimensions to protect their margins. This would mean that, even if there was imperfect competition in the bank lending market, caps may not have had a significant effect. One way that banks can protect their margins – and effectively hold the cost of intermediation constant – is to shift income that would have been collected from a loan's interest payment into other fees and charges (Maimbo and Gallegos, 2014 and Ferrari et al., 2018). These fees and charges are usually placed on the borrower, taking the form of registration, processing, administration, or legal fees. But banks could also increase fees on account holders and depositors, although this would be less likely if strong competition existed in the deposit market. The net effect of this income switching is that, although banks' margins may initially fall in the wake of a cap, they may not be easily shifted.

There are two other potential effects worth mentioning that we will not be considering directly. The first is how caps may reduce the incentives for firm entry and capital deployment or increase the likeliness of market exit. This may have flow-on effects on competition, leading to less competition and poorer current or future intermediation. Both the Ho and Saunders (1981) bank dealership and the Monti-Klein (Klein (1971) and Monti (1972)) industrial organisation models predict that such a reduction in competition would place upward pressure on margins. In addition, the National Commission on Consumer Finance (1972) found that rate ceilings reduced competition in the US consumer market. Given the likely long-term nature of this effect and our relatively short cap regime, we do not directly consider it and instead control for competition as a covariate in our regression framework. The second effect is how the cap regime may have created additional policy uncertainty in the banking sector, and how this in turn may have influenced banks' investment and lending decisions. As a first-order effect, banks could potentially price in such policy uncertainty in their

lending risk premiums, thus increasing margins. It is unclear how strong this effect likely is, and how we could measure it, so we do not consider it further.

Based on the above, what is the likely net effect of Zambia's interest rate cap regime on banks' margins? Given that we are uncertain of Zambia's market context and that multiple mechanisms may operate in opposing directions, making an a priori claim is challenging. Furthermore, the difficulty of measuring certain aspects of bank behaviour (for example, the efficiency of portfolio diversification and the dynamics of this flowing through to margins or disbursement risk assessments) makes empirically untangling each mechanism impossible. However, at the least, we should be able to test for significant net effects, and we should be able to test the income-switching mechanism directly.

Table 22 separates the loan margin ratio into its component parts to help clarify how the cap might affect each mechanism and thus influence the margin. It highlights one empirical challenge we are unavoidably going to face: that several margin components could be influenced by multiple mechanisms concurrently and hence we will not be able to disentangle them.

Dataset and Margin Measurement

Our dataset covers all licensed banks operating in Zambia between Q1 2005 and Q4 2018, with the number of banks ranging between 13 and 18 over the period. The average number of quarters per bank in the full sample is approximately 48. The dataset was constructed from raw monthly supervisory prudential returns provided by the Bank of Zambia. Balance sheet and income statement data were averaged and summed across each quarter respectively. The data covers simple accounting data as well as more detailed liquidity, interest rate, foreign exchange, and capital data. The variables used in our analysis are explained in Table 24 and descriptive statistics are provided in Annex 18.

Our raw data was recorded monthly, but we chose to collapse this into a quarterly series to smooth accounting recognition mismatches which we suspect may exist between income statement flows and the balance sheet stocks in the monthly data series. Using quarterly data also dampens the bias that stale data may introduce, whereby banks revalue their less-liquid assets only in anticipation of their own internal quarterly reporting[191]. It also accounts for the fact that some loans make interest payments less frequently than every month.

[191] We found large income jumps to be more likely in the final month of each quarter.

Over the time period, six banks entered the market, one departed, and there was one merger. During banks' entry and exit transitions, their financial ratios sometimes recorded extreme values that likely reflected abnormal operating conditions. We drop these observations and proceed with a clear caveat in mind: our results only apply to banks operating under normal, business-as-usual conditions. To adjust for the merger, we follow Gambacorta (2005) and combine the two banks and treat them as one timeseries.

There is debate in the literature regarding the appropriate margin to use for measuring the cost of financial intermediation in empirical analysis (see Brock and Rojas-Suarez (2000) pg. 120-121, for example). Our analysis uses a margin definition that best fits our context for evaluating IRRs. Our measure remains broadly consistent with the financial development literature, for example it does not deviate too far from Beck and Hesse (2009), Brock and Rojas-Suarez (2000), Chirwa and Mlachila (2004):

$$Loan\ Margin\ = \frac{Interest\ Received\ from\ loans\ + fees}{Performing\ Loans} - \frac{Interest\ paid\ to\ all\ sources}{All\ Interest\ Bearing\ Liabilities} \quad (7.1)$$

This margin captures only banks' interest on loan assets, excluding the effects of Government securities and other interest-bearing assets the bank may own. In addition, it includes deposit and non-deposit interest costs and liabilities on the financing side of the margin calculation. The margin takes this form because we do not want the primary driver of our measure to be either (i) the direct result of increasing Treasury yields (off the back of a loosening fiscal stance - a feature in our dataset), or (ii) the portfolio effects of banks switching their asset portfolios from loans towards Treasury Bills (a potentially relevant consequence of the cap regime, but not in terms of the cost of financial intermediation).

Although the loan margin in equation 7.1 is our preferred measure, as with Brock and Rojas-Suarez (2000), we will also present results for the Net Interest Margin (NIM) to facilitate comparability with the literature[192].

[192] $NIM = \frac{Net\ Interest\ Income\ from\ all\ source+fees}{Total\ Earning\ Assets}$. The issue with the NIM is that it can be sensitive to banks' sometimes arbitrary decisions regarding how to finance its non-interest-bearing assets. For instance, let's say two banks decide to purchase a new physical asset (for example, purchasing a new branch in a rural town). One uses deposits to finance the purchase and the other issues new equity. In an accounting sense, the bank that used deposit financing will now have a lower margin, reflecting less costly intermediation (it is lower because interest costs increased as a result of the deposit financing, but they would not, in an accounting sense, if the asset were equity financed). Yet this is not reflecting the bank's efficiency, it is merely reflecting the fact that a leveraged physical asset purchase results in a lower NIM. In the ideal world,

Because of the volatile operating environment for Zambian banks from early 2014 onwards, we provide results for both the full sample ending in 2018, as well as a sub-sample ending in Q1 2014. The worry is that the banks began to operate under either significant liquidity constraints, macroeconomic volatility, or both from Q2 2014 until 2017. In the shorter sub-sample this leaves just over a year for the cap to impact bank margins. The other issue with the extended sample is that BOZ's loosening of the cap in Q2 2014, as well as its removal in Q4 2015, was almost certainly endogenous to other macroeconomic and political events. In any case, we provide results for several subsamples.

Table 22 – Potential mechanisms via which a cap might influence banks' margins.

	$Loan\ Margin = \frac{Interest\ from\ loans}{Performing\ Loans} + \frac{Fees}{Performing\ Loans} - \frac{Interest\ to\ all\ sources}{Interest\ Bearing\ Liabilities}$			
Mechanism	Total Effect on Margin	Interest yield	Fees Yield	Financing Costs
Correcting for Non-Competitive Loan Market	Negative (-)	(-)		(+)
Portfolio Diversification Degeneration	Positive (+)	(+)		(-)
Income Switching	Unclear	(-)	(+)	
Reducing Higher-Risk Lending	Unclear	(+)/(-)		

Note: Table separates the loan margin ratio into its component parts and comments, based on the discussion in the theory section, on how the IRR might plausibly impact each component for a given bank. This table assumes the cap was binding on the bank, which may not be true for several of the banks.

Zambia Context and Margin Decompositions

Macroeconomic context and bank margins

We now outline several of the main trends and macroeconomic events in Zambia over the sample period, as well as how they influenced bank intermediation and bank margins. A more detailed history of Zambia's banking and financial history can be found in Chapter 2 of this thesis.

As Chapter 2 discusses, from 2005 Zambia's financial system was becoming increasingly stable. The volatility associated with the early 1990s liberalisation had significantly dampened and the pains associated with a banking crisis in the late 1990s had passed. Although bank margins were still high by international standards in 2005, it is within this increasingly stable macroeconomic context that

we would not want a margin measure that was overly sensitive to arbitrary financing decisions. In a low-income country context, using the NIM is likely to be more problematic because the cost of increasing banking services is expensive and does requires physical asset purchases. Zambia is a good example whereby a geographically disbursed deposit and lending base will require higher physical asset investments.

they began to decline: in the years leading up to the Global Financial Crisis (GFC) median bank margins reduced from 6.5% to below 5%, or almost 25%. Part of this decline can likely be linked to the 2005 HIPC write-off and lower inflation, both of which also helped drive down government yields. But increased competition and reduced operating costs in the market may also have played a role.

In addition to this general decline in margins from 2005, there were three other significant events that influenced Zambia's margins in our dataset. The first was the Global Financial Crisis, which impacted Zambia from 2008 onwards. At the onset of the crisis, Zambia's financial system had limited integration with the international financial markets, dampening the crisis' immediate impacts (BOZ, 2009). But the crisis eventually fed through into Zambia's economy, via reduced mineral revenues for the Government, a slowdown in private sector investment, and job losses in the extractives industry. Copper prices at the London Metal Exchange (LME) fell from highs of US$8,980 per tonne in July 2008 to as low as US $2,812 per tonne in December 2008 (BOZ, 2009). The Kwacha also declined as global capital shifted away from emerging and developing country risk (especially those countries with commodities exposures) and had a pass-through effect on inflation. Treasury yields increased contemporaneously. Clearly these rate changes had tangible effects on the financial sector, yet banks were able to pass on much, if not all, of these rate increases on to their lending rates and fees, without facing equivalently large increases in their financing costs. Indeed, it was this asymmetry that caused margins to increase during the GFC (see row two and three of Figure 39). Banks also tightened lending standards and reduced loan volumes, potentially pricing higher risk premia into their rates, or compensating for increasing loan write-offs. Figure 1 from Chapter 2 shows the portfolio shift away from loans towards securities and Figure 49, row 4 in Annex 19 shows the increase in provisioning. Bank lending margins did return to their decreasing trend after the crisis period, although the key impact on the banking sector was longer lived: non-performing loans spiked as corporate and household balance sheets deteriorated due to the slowdown, increasing from 7.2% in 2008 to 13.8% in September 2009. These elevated NPL levels persisted for several years; for several banks, accounting profitability took some time to recover (Figure 49). In summary, the GFC thus resulted in a positive short-term shock to Zambia's banks' spreads and lending-specific margins, a sizeable and persistent increase in non-performing loans, and a long tail of banks recording negative accounting profits (Figure 49).

The second event was the Bank of Zambia's overly intense tightening of liquidity and monetary conditions at the beginning of 2014, which created severe funding pressures for banks. This event was arguably less severe than the GFC or Zambia's inflation and FX crisis, although it is still visible in the bank margin data. In early March 2014, BOZ took several actions to reduce excess liquidity in the banking system and quell possible inflationary pressures caused by a 41.9% increase in civil service wages implemented by the Government in September 2013 (BOZ, 2014). This included increasing the statutory reserves requirement from 8% to 14%, increasing the BOZ policy rate half a percentage point from 9.75% to 10.25%, increasing the Overnight Lending Facility (OLF) margin from 2.5% above the BOZ rate to 6% above, and restricting OLF access to once per week (see Table 42 for a record of all changes in monetary policy settings during the period)[193]. Taken together, these actions constituted a considerable squeezing of liquidity. Almost immediately, liquidity tightened as banks turned to the interbank market and OLF for liquidity as they adjusted to the new reserve requirements[194]. The interbank lending rate reflected these pressures and jumped well outside the Monetary Policy Target Zone (the 10.25% BOZ rate +/- 2%), reaching over 20% for nearly two months. BOZ further increased the OLF margin to 10% above the BOZ rate in May, putting additional pressure on liquidity and the interbank market. Over June and July, the interbank rate sat above banks' recorded *lending* rates, and for one week sat 11% above the upper BOZ Monetary Policy target limit. In other words, the IRR cap (9% above the BOZ rate) was lower than the interbank lending rate.

This tight liquidity had an immediate impact on inflating bank's financing costs, although much of the increase appears to have been successfully passed onto borrowers, leaving lending margins largely intact (Column 1 Row 2 in Figure 39). BOZ's loosening of the interest caps in June may have helped facilitate this passthrough. In terms of tracing the liquidity crisis through the liability side of banks' balance sheets, bank non-deposit financing costs, for example interbank loans, jumped significantly in Q2 2014, which was very likely linked to the liquidity-crisis induced costs. Other deposit-based interest costs also increased over the period (especially time deposits), although it is

[193] The application of the statutory reserve requirement was also extended to Government deposits and the maintenance of statutory reserves was tightened from a weekly average to daily compliance.

[194] The margin we use in our regressions should remove the majority of the direct effects of the increase in reserve requirements on profitability. This is because by using only interest-bearing assets in the denominator, we are not penalising the banks' margins for holding non-interest-bearing central bank reserves. In saying this, the resulting tight liquidity conditions resulting from the change in the reserve requirements will still impact the cost of financing and liquidity indirectly through tighter liquidity conditions and financing (and we see this in the spike in the interbank lending rate and financing costs more broadly).

hard to clearly link these to liquidity conditions given that the Treasury bill rates had also increased significantly.

The final event was the foreign exchange and inflation crisis Zambia underwent in 2015-2016, which saw the Kwacha depreciate from K/USD 6.4 to K/USD 11.2, or 42%, over 2015. Much of the import price shock that followed flowed through to inflation (monthly inflation hit 6% in October), creating significant distress. Several structural factors lead to the crisis, including poor hydrology and power shortages, an increasing and increasingly dollarised debt stock (Zambia issued an additional USD 1.25B Eurobond at 9% in July), an uncontrolled budget deficit, and soft commodity prices (especially copper). The change in sentiment came to a head from July to October in 2015 with the Kwacha essentially decoupling from its previous fundamentals, appreciating by 38% (in direct FX terms), and topping the charts for the largest value depreciation in the world that year. Inflation passthrough in Zambia is extremely high, and so the FX crisis was coupled with a substantial inflation shock: October, November, and December saw monthly inflation rates of 6.2%, 5.0%, and 2.1% respectively. BOZ responded by tightening monetary policy, increasing reserve requirements to 18% and the policy rate to 15.5%[195]. Zambia's Eurobond-US T-bill spread reflected investors' concerns about Zambia's ability to finance Eurobond coupons and roll-over the placement, more than doubling from July (500 BPs) to the end of the year (more than 1250 BPs).

The banking system was placed under significant strain during the FX/inflation crisis. Dollarisation increased significantly as depositors tried to shift into foreign-denominated accounts. Nominal rates rose, and liquidity was extremely tight. The interbank rate rose well above 25%, as did Treasury yields (which had already been trending up since 2014 as the copper price slid). Many banks that were able to shift short-term security positions into offshore accounts did so (Figure 1 in Chapter 2). Bank lending weakened.

The initial effect of the crisis on bank margins in mid-late 2015 was mildly negative, as banks were able to pass at least some of the short-term liquidity costs onto borrowers[196]. As the crisis moved into 2016 and eventually subsided, however, bank margins increased. Bank lending rates increased as other nominal rates in the economy shifted up, but deposit financing rates did not increase nearly as

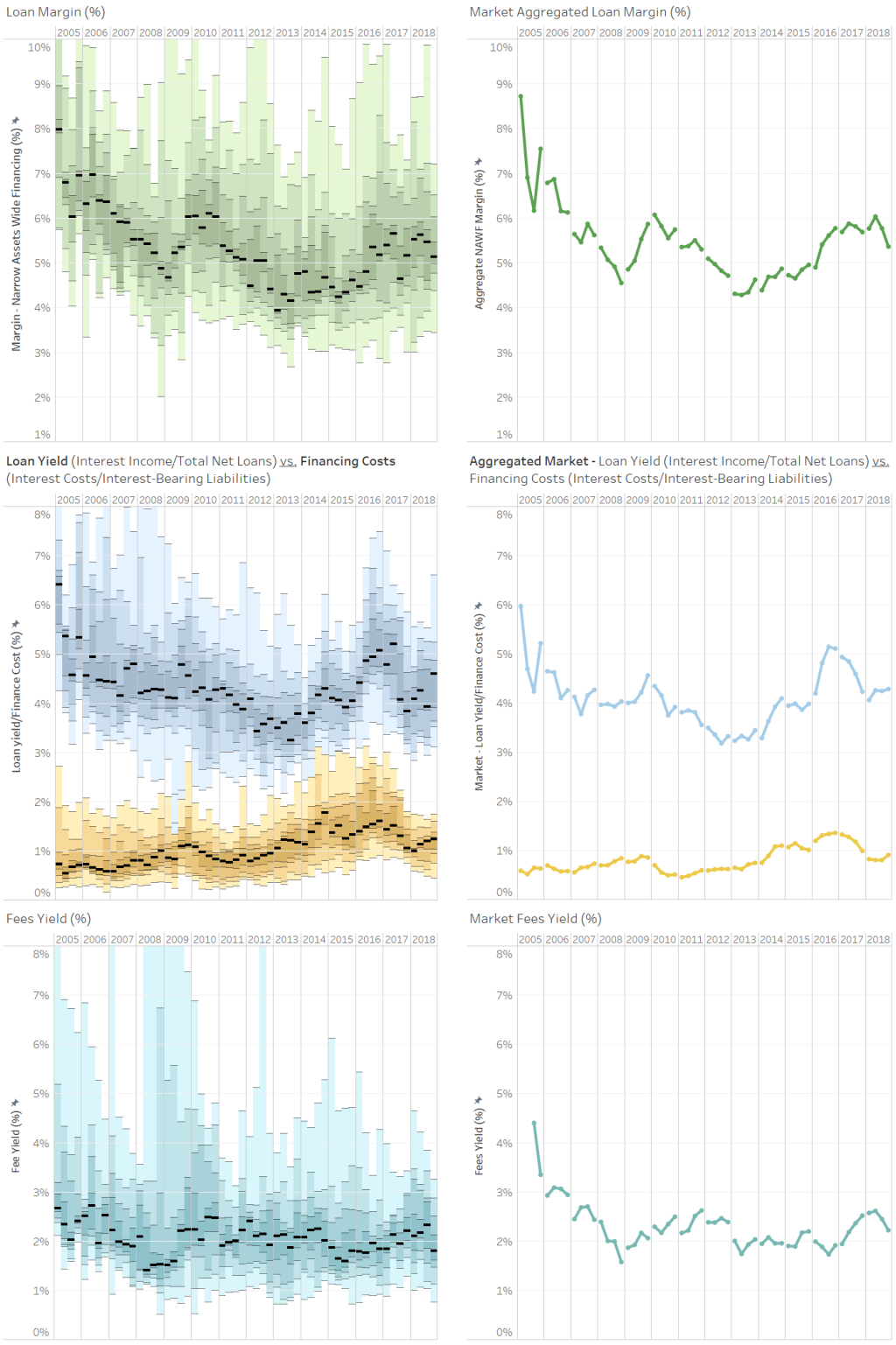
[195] All Zambia macroeconomic data in this section sources from BOZ's fortnightly statistical bulletin.
[196] Banks NIMs dropped much more severely than the narrower margin calculation we primarily use in the paper. Note that these margins are rates of return for a given unit of lending or asset, and do not trace 1:1 to profitability levels or volumes.

much as loan interest yields (row 2 Figure 39). Part of this could be a dollarisation allocation and valuation effect, in which the panic-induced deposits shift towards higher volumes of potentially more profitable foreign-denominated deposits[197]. It is also unclear whether banks operating during this period simply required large margins for providing banking in such a volatile environment or whether depositors and savers essentially paid the price for the macroeconomic uncertainty.

The acuteness of the crisis dampened over late 2016 and into 2017 as the copper price recovered back to above USD 6000, Zambia's trade balance stabilised, the FX-induced inflation impulse subsided, and the Kwacha itself mildly strengthened on the back of substantial BOZ intervention (the Kwacha strengthened by 25% in the first half of 2016)[198]. From early 2017, BOZ began to loosen monetary policy and the Government revamped its domestic deficit financing. However, this was essentially the beginning of the end, as the lack of quality government expenditure (World Bank, 2017), an increasingly large dollar denominated debt stock, and continued politically driven fiscal dominance meant Zambia's position began deteriorated, especially when the copper price began to fall in mid-2018, at which point our dataset ends.

[197] Deposits in dollars attracted almost no return and such transactions and accounts are subject to fees. Further, even Kwacha denominated deposits have a lower interest rate elasticity than bank lending.
[198] FX intervention calculated form BOZ Fortnightly Bulletin.

Figure 39 – Margin Components by Disaggregated Decile Bands (LHS) and for Market Aggregate (RHS)



Note: Figure provides data on key bank margin and margin components data. The left column graphics provide decile ranges (colored ban and medians (black horizontal lines) for bank margin and margin component data each quarter, based on the distribution of bank data in t quarter. The right column provides the aggregated market figure for the same variables. Variable definitions provided in Table 22.

Source: BOZ Prudential Returns and Author's calculations

Accounting decompositions

Before we turn to regression analysis, we use a simple accounting decomposition to help outline some of the main trends in the banking sector that might have impacted margins. This is a common exercise in the bank margins literature, for example in Calice and Zho (2018) and Boutin-Dufresne et al. (2014). The intuition behind a decomposition is to show how the net interest margins contributed to banks' broader profitability, and how they compared to (or covered) the banks' non-interest costs and other income streams. The decomposition hence links the banks' core intermediation revenues and costs to their wider financial performance and business operations. An interest margin records the outcome of the pressures a bank faced and the decisions it made in setting its lending and deposit rates, which must implicitly reflect other non-interest revenues and costs related to the banking business. Decompositions help outline these other aspects of the banking business, and how they have changed over time[199]. We undertake both an aggregate decomposition (Figure 40) as well as a bank-level decomposition (including the decile distribution or each NIM component – see Figure 49 in the Annex).

The simple accounting decomposition of Zambia's aggregated Net Interest Margin (NIM) highlights several trends. First, that there was a reduction in the banking system's NIM beginning in 2005, and that this occurred contemporaneously with a reduction in overhead costs and provisioning (Table 23 and Figure 40). Further, other income sources did not compensate, and as a result, profit rates (i.e., the measured return on assets) broadly reduced over the period leading up to the imposition of the caps. Banks profit rates were, at least in aggregate, not significantly higher than they had been in the past, and in fact profits and margins had broadly reduced over time. It also alludes to the fact that competition may have driven margins down in such a way that banks were not able to compensate via other revenue lines. Profitability, broadly speaking, was somewhat insulated by improvements in provisioning and reduced overhead costs.

While the Financial Crisis adversely affected bank profitability via high loan impairments and operating costs, core intermediation profitability was not immediately impacted[200]. As the crisis

[199] Take, for example, the case where a bank's overhead costs were acutely increasing. The decomposition will show how net interest margins or other revenue streams would need to adjust (or may have adjusted) for the bank to defend a certain level of profitability. The limitation of decompositions is that they merely measure a related set of financial stocks and flows and cannot concretely tell us about the factors that might be influencing these changes.

[200] There is one caveat: by using this definition of the NIM we are also including interest from non-loan interest-bearing assets, such as government securities. This means that actual loan lending intermediation costs may have been impacted, but we cannot tell this from our data because such portfolio reallocations will not be picked up.

progressed, however, NIMs do appear to have reduced, especially from mid-2010. Aggregate bank profitability appears to suffer during the two main periods of crisis (the last quarter of 2014 and 2015). Later, we discuss disaggregated results.

Equation (7.2) provides banks' accounting profit before tax and equation (7.3) provides the decomposition[201]. In Figure 40, the black line shows profitability, and the first green bar shows the NIM.

$$Profit_t = Int_t - int.expn_t + Non.int.inc_t - opp.costs_t - provisions_t - tax_t \qquad (7.2)$$

$$\frac{Int_t - int.expn_t}{TotAssets_t} = NIM_t = \frac{Profit_t - Non.int.inc_t + opp.costs_t + provisions_t + tax_t}{TotAssets_t} \qquad (7.3)$$

For comparability to other studies, we use the Net Interest Margin defined as the net interest income divided by total bank assets and do not adjust for reserve holdings etc. *Non.int.inc* is the non-interest income, *opp.costs* represents operating costs, and *tax* records tax payments. *Profit* records net profit after tax and p*rovisions* records the accounting write-down associated with loan losses for the period (i.e., the income statement entry for the recognition of increases in non-performing loans).

Table 23 – Bank Accounting Profit and NIM Decompositions

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
FX Gains and Losses	2.88%	2.14%	2.29%	2.05%	1.90%	1.61%	1.49%	1.31%	1.43%	1.25%	1.54%	1.30%	1.29%
FX Fees	0.35%	0.42%	0.38%	0.36%	0.27%	0.23%	0.26%	0.19%	0.16%	0.22%	0.24%	0.20%	0.23%
Non-interest Income	4.07%	3.68%	3.07%	3.46%	4.11%	3.05%	3.55%	3.14%	3.14%	3.05%	2.93%	3.37%	3.25%
Net Interest Margin	9.27%	8.60%	8.36%	8.80%	7.22%	6.80%	7.08%	6.77%	6.36%	6.22%	6.33%	6.95%	7.36%
Tax	-1.58%	-1.17%	-1.39%	-0.99%	-1.32%	-1.27%	-1.56%	-1.05%	-1.16%	-1.03%	-0.96%	-1.13%	-1.17%
Overheads	-9.29%	-8.82%	-9.60%	-10.06%	-9.60%	-7.96%	-8.11%	-7.54%	-7.30%	-7.45%	-7.75%	-7.87%	-8.18%
Provisions	-2.13%	-1.55%	-1.32%	-3.00%	-1.23%	-0.10%	-0.50%	-0.56%	-0.41%	-0.57%	-0.66%	-0.84%	-0.97%
After-tax Profit (ROA)	3.57%	3.30%	1.80%	0.60%	1.30%	2.40%	2.20%	2.30%	2.20%	1.70%	1.70%	2.00%	1.80%

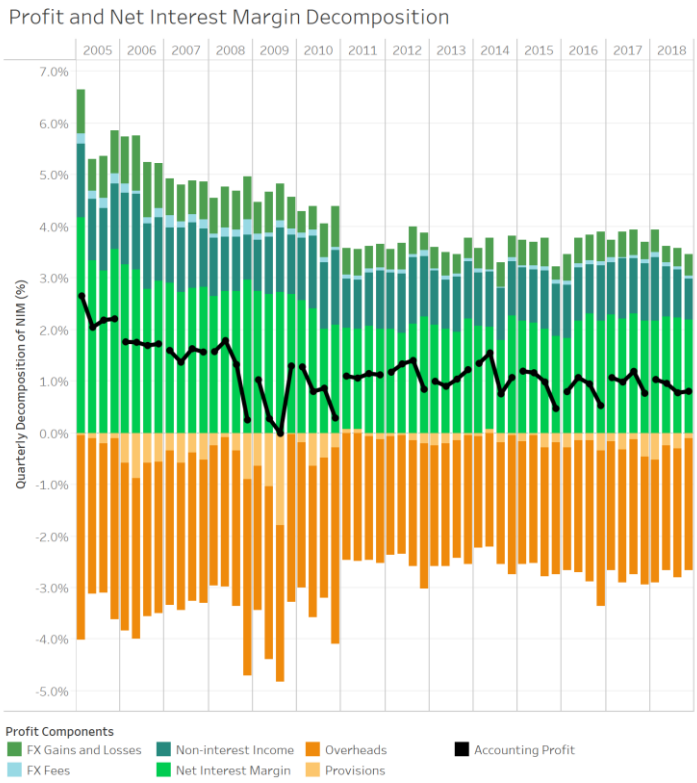
Note: Annual decomposition based on equation 7.3 and using BOZ prudential data.

How do Zambia's decompositions compare internationally? While Zambia's margins have reduced over the sample, they remain high. In a cross-country decomposition exercise, Calice and Zhou (2018) find that Zambia's NIM over 2005-2014 was 6.7%, 5th highest in a sample of 160. Lower-income countries and SSA countries have higher margins on average (4.47 and 5.03, respectively) but intermediation in Zambia is expensive even compared to these benchmarks[202]. Profits in Zambia are comparably higher than in higher-income countries, but they are consistent with other banks in

[201] (7.3) is re-arrangement of (7.2) with rebasing both sides of the equation with a banks' total earning assets. The decomposition shown below is not for individual banks but for the hypothetical aggregate bank.
[202] Average NIM in OECD countries and high-income countries during the period were 1.75% and 2.03%, respectively. Note that Figure 39 and Figure 40 data are not annualised.

the region and income group. Whether these higher profits reflect higher risk or rents is more complicated and a topic for further study.

Figure 40 – Bank Accounting Profit and NIM Decompositions



Note: Figure provides a quarterly decomposition of the banking sector's accounting profit and NIM, replying on a version of equation Costs and expenses are shown as negative. The NIM is shown in light green.

Source: BOZ Prudential Return

Finally, we disaggregate the decomposition. In Figure 49 in Annex 19, we use quarterly bank-level data to track the deciles of the banks' decomposition components. While our previous aggregated approach is likely a fair representation of the large, diversified banks in Zambia, it does not tell us much about the experiences of smaller banks in the sample. Figure 49 provides some intuition for how the disaggregated market has tracked. Several interesting points stand out; first, the trend on reducing NIMs from the aggregated data appears to be a broadly market-wide phenomenon, as is the reducing role of FX fees and FX gains. Secondly, the financial crisis created a long negative tail of poor profitability in the market which lasted several quarters and was associated with a spike in operating costs. The operating cost spike may be related to banks writing off large sections of their loan portfolios (as loans become non-performing, banks may be unable to reduce their overhead

costs in line with the loss-allowance write-off, creating a spike in the ratio). Non-interest income appears to be a stable revenue stream but with high levels of variation in the cross section, although this could simply be a result of variations in banks' accounting policies. Several banks appear to have been caught out by the FX volatility in late 2015 – losses in the final quarter of 2015 are very high. As macroeconomic conditions began to worsen in late 2018, we can see that bank NIMs began to increase again. In summary, this disaggregated data broadly reflects the trends in the aggregate market, but also shows us that we should be careful in any panel estimations as there is evidence of significant bank heterogeneity.

Empirical Approach

Determinants of bank margins in Zambia

The starting point for our analysis is to investigate the empirical determinants of bank margins in Zambia and benchmark the results to other countries in the region[203]. This analysis allows us to have some confidence that we are adequately modelling bank margins before we move to investigating IRRs. We build from the empirical bank margin literature and estimate various versions of:

$$Y_{it} = \alpha_i + \sum_{l=1}^{L} \rho^l Y_{it-l} + \sum_{j=1}^{J} \beta^j X_{it}^j + \sum_{k=1}^{K} \theta^k M_t^k + \varepsilon_{it} \qquad (7.4)$$

Where Y_{it} is an interest margin measure, $\sum_{l=1}^{L} \rho^l Y_{it-l}$ captures lagged dependent variables up to a lag length L., X_{it} is a vector of bank-specific margin determinants for bank i at time t, M_t is a vector of macroeconomic and regulatory measures, α_i is the bank-specific intercept, and ε_{it} is the residual error. Annex 18 provides summary statistics. Our primary approach is to use within-bank variation to calculate the β^j coefficients (i.e., equation 7.4). In some regressions, however, we replace $\alpha_i \; for \; \alpha_t$ and instead rely on cross-sectional variation in banks margins, providing a different source of variation for margin determinants.

[203] This will be the one of the first papers to examine bank margins in Zambia at the bank level. Further, it will be the first paper to use loan maturity as a margin determinant in Africa.

We include margin determinants from both theory and empirical margin literature in our regression (i.e., the components of X_{it}^{j} and M_{t}^{k}). Table 24 explains in detail which variables we include in our regressions and how they link back to the bank margins literature[204].

We will generally model margins dynamically as banks' margins almost certainly exhibit persistence. Margins take time to adjust to shocks and changes in bank strategies, and we will need to model such dynamics into our specifications. Part of this persistence is simply mechanical: banks' balance sheets are comprised, at least partly, of fixed-term asset and liability contracts which can take time and be costly to adjust. This is especially true for banks' lending books, where the spread on a loan may be locked in for the loan's duration; banks can't change loan terms or renege on term deposits at will, and we would expect this to be reflected in the dynamics underpinning their margin adjustments. If relationship banking is an important component of banks' business models, this may also factor into margin dynamics. These constraints on bank flexibility are partially offset by the fact that, as discussed in Chapter 2, much of the Zambian market is on floating rate loan agreements, which reset more frequently than the loan maturity. Nevertheless, we find strong evidence of serial correlation in the static models' error processes. Breusch–Godfrey Serial Correlation Lagrange Multiplier tests unambiguously reject the null of serially independent errors for all specifications. This serial correlation is likely an indication that the static model is mis-specified and that we should take dynamics into account (Beck, 2011; Hendry and Mizon, 1978). We suspect that contemporaneous correlation is also likely an issue as banks operate in a common market and are exposed to similar risks and shocks (and Breusch-Pagan tests for cross-sectional independence confirm this). Finally, time-invariant bank-specific characteristics are almost certainly correlated with our margin determinants[205].

We select the optimal lag length based on the BIC[206]. Ideally, we would have allowed for a more general ARDL specification (to capture variable-specific dynamics), but we are constrained by limited observations and instead assume a common response function across explanatory variables and shocks. To account for cross-sectional dependence and heteroskedasticity, we follow Beck and

[204] We include a variable for contractual maturity transformation in our regression, borrowing from the work of Bologna (2018) and Entrop et al. (2014), the first paper we have come across to do so in a low-income setting. Maturity transformation is a key function of a banking system, but there has been very little research into its links to bank profitability, especially in low-income countries.
[205] Consider a bank with easy access to parent-subsidised wholesale finance. This access would persistently dampen the interest cost component of the bank's margin, yet such access would likely be correlated with other time-varying margin determinants, such as management efficiency.
[206] This was almost always a single lag.

Katz (1995) and Beck (2001) and use Panel Corrected Standard Errors (PCSEs). A Feasible General Least Squares (FGLS) approach is also sometimes used for robustness checks. As with most of the bank panel literature, we find that bank-specific intercepts are consistently jointly significant, and so we include bank fixed effects[207].

Our general approach is broadly consistent with previous times-series–cross-section (TSCS) research in the bank margins literature, sitting most closely to Papavangjeli and Leka (2016) who use PCSE corrected standard errors for a panel of 12 years of data for 16 Albanian banks[208]. Beck and Katz (2009) and Wawro et al. (2011) find that FE-PCSE approaches are appropriate given our data dimensions and the strong likelihood of contemporaneous correlation[209]. In terms of the time-series properties of our data, Im, Pesaran, and Shin's (2003) panel unit root tests consistently rejected the hypothesis that all panels contain unit roots for the margin data and bank-level controls. Additional Fisher-type Augmented Dickey-Fuller tests (Choi, 2001) support this. ADF, ADF-GLS and Phillips Perron tests show that the macroeconomic variables are generally stationary, except for the Treasury Bill rate, which does not consistently reject the unit root null. We check that the results are not sensitive to using the Treasury Bill rate in first differences.

[207] The cost of this is that our estimates will be biased (Nickel, 1981), but given our longer time dimension we can afford to be less concerned. A GMM approach may be inappropriate given our high T low N data dimensions.
[208] Other examples of dynamic specifications include Chirwa and Mlachila (2004), who use monthly Malawian data and a type of FGLS transformation, Brock and Franken (2003), who use Hahn and Kuersteiner (2002)'s bias-corrected MLE with dynamic fixed effects in their work on Chile's banking sector. Finally, Beck and Hesse (2009) do not explicitly model dynamics in their analysis of Ugandan bank margins.
[209] Beck and Katz (1995, 1996) and Beck (2001) have shown that Feasible General Least Squares can be unreliable, given our data dimensions.

Table 24 – Variable Descriptions and Discussion

Variable Type	Theoretical Variable and Expected Coefficient	Proxies/Measurement and Intuition	Discussion Points
Bank Specific Variables	Operating costs	Cost Ratio, calculated as total operating costs divided by total assets.	• Banks with higher costs will need to cover these in their lending margin. A frequent finding in the theoretical literature (see Wong (1997) and Maudos and Fernandez de Guevara (2004), among many others) • The broad norm in the literature is to proxy bank operational costs as total operational costs as a proportion of total bank assets. Some papers use total loans in the denominator, for example, in Brock and Rojas Suarez (2000)).
	Liquidity Risk	Liquidity Ratio, calculated as total liquid assets divided by deposits.	• Liquidity risk is the risk that banks will not have the necessary liquid assets or lines of credit to meet depositors' withdrawals or new loan demands. Banks may need to tap emergency financing at a high cost in this case. Therefore, a higher proportion of liquid assets should reduce liquidity risk and reduce the liquidity premium in margins. • The liquidity ratio is consistent with Beck and Hesse (2009), Santiago and Fernandez (2007), Brock and Rojas Suarez (2000), among others.
	Bank Size	Deposit market share measured as the ratio of a bank's deposit portfolio to the market as a whole.	• We use the deposit market share as our measure of bank size. This is broadly consistent with Brock and Franken (2003), who use banks' total asset market share, and is consistent with Beck and Hesse (2009). • The relative size measure captures whether larger banks have more market power. In Zambia, this will more likely pick up the fact that in expanding their retail businesses, banks have had to increase their distribution networks, which is expensive.
	Forward-Looking Credit Risk	Risk-weighted Asset Ratio.	• Lenders that make riskier loans may charge higher margins to compensate for the risk. The risk-weighted asset ratio (risk-weighted assets over total assets) should act as a forward-looking measure of credit risk exposure (Entrop et al. 2015). We expect it to be positively related to net interest margins.

		• This is a blunter measure in Zambia than in more developed markets because Zambia's Basel II reporting requirements do not provide a granular view of risk.
Historic Credit Risk and Portfolio Losses	Non-performing loan ratio	• Banks that have historic losses might look to recoup capital by charging higher margins. To the extent that NPL ratios are stable, they may also reflect banks' current lending risk decisions. The literature typically finds that NPL ratios positively affect margins (Angbazo 1997, Maudos and Fernández de Guevara 2004; Maudos and Solís 2009, among others).
Maturity Transformation or Contractual Maturity Mismatch	Time-weighted loans minus time-weighted liabilities.	• Measures the misalignment between the expiration dates of loans and loan financing and captures banks' structural refinancing risk. Our measure is consistent with Bologna (2018), whereby higher margins are associated with higher structural maturity mismatches. • Entrop et al. (2015) extend the Ho and Saunders (1981) model to capture expected returns from maturity transformation. • We follow Bologna (2018) in constructing the maturity transformation variable as the difference between the contractual time-weighted maturity of banks' lending portfolios and the contractual time-weighted lending portfolios and their non-equity financing. Bologna (2018) finds that maturity transformation is a significant determinant of net interest margins. • Lending in Zambia is typically dominated by floating rate loans, which adjust rates each quarter (See Chapter 2 and Annex 2, although the larger banks in the Zambian do make some fixed-rate loans). Contractual loan maturities are longer than the interest rate resets, so exposures are typically passed on to borrowers. Given this, we might expect maturity to have a different effect on margins in Zambia than what has been found previously, at least until banks start making more fixed-rate loans.
USD Loan Share	Foreign-denominated loans/Total Loans	• Banks with more dollarised books may have different funding costs and lending rates to those operating primarily in Kwacha. Dollar-denominated loans shift interest rate risk onto borrowers.
Bank Risk Aversion	Capital buffer, calculated as capital in excess of the required minimum risk-weighted asset ratio divided by total risk-weighted assets	• In most forms of the bank dealership model, more risk-averse banks require a higher premium for accepting interest rate risk (Ho and Saunders, 1981). • There are no direct measures of risk aversion. Therefore, we follow Maudos and Fernández de Guevara (2004) 's suggestion and use the capital buffer as our primary measure of bank risk aversion. In addition, we use Bologna (2018) 's approach of using risk-weighted assets in the denominator. • Empirical papers often use capital adequacy or leverage ratios as proxies for a bank's level of risk aversion (see Angbazo (1997), Brock and Rojas Suarez, 2000; McShane and Sharpe, 1985; Maudos

			and Fernández de Guevara (2004)), but these can be artificially inflicted by accounting conventions and do not account for within-asset risk (Bologna (2018)
Policy and Institutional Variables	Reserve Requirements	Bank of Zambia Reserve Requirement Ratio (%)	• Reserve Requirements can increase the costs of liquidity. However, our margin variable excludes non-interest-bearing reserves in its calculation, so the direct "taxation" effect of Reserve Requirements has already been removed. So, Reserve Requirements only capture the secondary liquidity effects and broader monetary policy conditions. • Reserve Requirement data are taken from BOZ's fortnightly macroeconomic series.
	Industry Concentration	Herfindahl–Hirschman Index for loan market shares $HHI_t = \sum_{i=1}^{N} \left(\frac{Loan_{it}}{\sum_{i=1}^{N} Loan_{it}} \right)^2$	• The Herfindahl–Hirschman Index measures the competitive structure within the loan market. Bank dealership models support the role of market power in determining margins (Entrop et al., 2015), and much of the literature finds that margins do typically increase with bank concentration (Chirwa and Mlchila, 2004; Maudos and Fernández de Guevara, 2004; Maudos and Solís, 2009; Saunders and Schumacher, 2000; and Williams, 2007, etc.), although not all find that this relationship is robust (Demirgüç-Kunt et al. (2004), Claessens and Laeven (2004) and Beck and Hesse (2009)).
Macroeconomic Variables	Inflation	The quarterly inflation rate is calculated as the log difference between the quarter-end price level and the price level at the end of the previous quarter.	• Many papers in the literature control for inflation, including Demirgüç-Kunt et al. (2003), Beck and Hessee (2009), Entrop et al. (2015). • Inflation captures the effects of nominal contracting. As a result, information asymmetries may increase with inflation (Huybens and Smith, 1999). Further, banks may adjust loan and deposit rates asymmetrically when inflation is high.
	Business Cycle	Growth in Zambia's Production Index. The international copper price.	• Most studies use a quarterly GDP growth series as their business cycle measure. For Zambia, this data is unavailable. We use the international copper price as an imperfect measure of the business cycle. • Other studies use production indexes, such as Chirwa and Mlchila (2004) 's use of growth in the Malawian production index as a proxy for GDP growth, or Beck and Hesse (2009) who created an index based partly on production indexes to circumvent the lack of a quarterly GDP series in Uganda. Zambia does have an industrial production index, although this data has not been published since 2017 and coverage is short. We use it as a robustness check.

		• During downturns in the economic cycle, borrowers' creditworthiness and net worth reduce. As a result, banks may increase rates to match the higher risk (Bernanke and Gertler, 1989).
Exchange Rate	Average exchange rate over the quarter.	• A significant portion of Zambia's loan market is denominated in dollars. Therefore, substantial changes in the exchange rate will directly impact the relative sizes of banks' domestic and foreign assets, potentially impacting margins. Exchange rate changes could also affect the bank's net worth, the size of its NPL book, etc. In addition, these changes may cause banks to change their pricing. Finally, the currency will also impact borrowers' ability to service loans if their revenue streams and loan obligations are not balanced. • Like Zambia, Uganda's banks have significant dollar holdings. Beck and Hesse (2009) include the exchange rate in their specifications.
Treasury Rate	One-year Treasury Bill rate.	• Treasury bill rates are positively correlated with bank spreads (Ahokpossi, 2012 and Poghosyan, 2010). Therefore, we use the one-year Treasury Bill rate to approximately match the bank's loan maturities. • The Treasury Bill rate plays an important role in credit markets by setting the opportunity cost of banks' lending and providing a floor reference rate for institutional/long-term depositors providing liquidity to the banking system. However, exactly how the Treasury Bill rate impacts the margin is complicated because it may impact both lending and deposit rates. • The Government has engaged heavily in domestic lending in Zambia, and the Treasury rate has been quite volatile.
Interest Rate Risk	Quarterly standard deviation of the average weekly interbank lending rate.	• A high level of interest rate volatility makes it riskier for banks to match deposit supply and loan demand (as they arrive asymmetrically in time and require short-term financing). For this higher risk, banks theoretically charge a higher premium. There is a long history of using the volatility of interest rates as a measure of interest rate risk, linking back to Ho and Saunders (1981). Since then, many papers have used the standard deviation of interbank lending or other Treasury Bills or bond rates as a measure of interest rate volatility. For example, • Maudos and Fernández de Guevara (2004), Hawtrey and Liang (2008), Entrop et al. (2015), all use interest rate volatilities in their margin regressions.

*All variables are constructed from monthly data. Balance sheet or stock data are calculated as an average over the quarter. For example, fourth quarter balance sheet data will average the month-end balance sheet values from September-December. Income statement or flow variables are calculated as the sum over the quarter. For example, fourth quarter data will sum the monthly income statement data for October through December. This approach helps ensure that the income or expense flows generated from a period are best matched to the stock or balance sheet that generated them. Using quarter-end balance sheet data would bias ratios towards the balance sheet’s quarter-end position.

Examining the IRR regime

Our strategy for the analysis of IRRs builds from the empirical bank margin literature and then adapts it to model the effects of the IRR. Our approach is most similar to Bologna (2018), and our general empirical specification is as follows:

$$Y_{it} = \alpha_i + \sum_{l=1}^{L} \rho^l Y_{it-l} + \sum_{j=1}^{J} \beta^j X_{it}^j + \sum_{k=1}^{K} \theta^k M_t^k + \tau^a C_{2013} + \tau^b C_{2013}.E_i^1 + \tau^c C_{2014} + \tau^d C_{2014}.E_i^2 + \varepsilon_{it} \quad (7.5)$$

We create two policy dummy variables which act as indicator variables for the cap regime:

$$C_{2013} = \begin{cases} 1, & 2012,Q4 < t < 2014,Q2 \\ 0, & Otherwise \end{cases}$$

$$C_{2014} = \begin{cases} 1, & 2014,Q1 < t < 2015,Q4 \\ 0, & Otherwise \end{cases}$$

Y_{it}, X_{it}^j, M_t, α_i, and ε_{it} are the same as in equation 7.4. τ^a is a dummy parameter measuring the short-term level shift in the margin associated with the implementation of the initial cap rule (the BOZ rate plus 9%), τ^ccaptures the loosening of the cap regime beginning in Q2 2014. E_i^1 and E_i^2 are interest cap exposure variables, described in detail below.

We use two approaches to account for bank exposure heterogeneity. The first is to build a continuous "bank cap exposure" variable to capture variation in IRR "bindingness" across banks. By interacting this variable with an interest rate cap policy dummy, we can better account for bank exposure heterogeneity in our estimations. This approach is not dissimilar to our non-bank analysis. Still, it sits most closely in the bank margins literature to Bologna (2018), who looks at the effect of a maturity transformation regulation on banks' interest margins in Italy during the early 1990s. The second, less granular and more restrictive approach, is to break up the sample into two groups: diversified banks, which are less likely to be constrained by the cap, and small, undiversified banks, which are more likely to be constrained by the cap[210].

[210] Zambian banks do not necessarily have the same business models or operate in the same market segments. Large diversified commercial banks take liquid deposits, lend via large retail bases and dispersed branch networks, participate in the Government's Treasury auctions, and typically have access to international markets via their parents or shareholders. The smaller banks look more like investment companies, collecting large illiquid deposits and making a small number of large commercial loans. Some of our analysis hence breaks the market into two groups: Diversified Banks, which comprises

In terms of calculating the interest cap exposure variable, E_i^1 is calculated as the average interest yield for the three months before each bank's implementation of the cap regime change. This is our best approximation for a bank's pricing relative to the cap, as, unfortunately, the other bank-lending data we have is low quality and equally aggregated. The three-month period helps smooth out what could otherwise be quite noisy accounting data. We create two exposure variables: one for the introduction of the regime in Q1 2013 and one for the loosening of the regime in Q2 2014. These tell us where the bank's pricing sat, on average, before a regime change. Net outstanding loans are used in the denominator of the calculation because we do not want to bias, downwards, the interest rate estimate due to the presence of non-interest paying non-performing loans[211]. The two exposure variables are calculated as:

$$E_i^1 = \frac{\frac{\sum_{t=2012,12}^{t=2012,10} Interest\ Income_{it}}{3}}{\frac{\sum_{t=2012,12}^{t=2012,9} Net\ Outstanding\ Loans_{it}}{4}} \quad E_i^2 = \frac{\frac{\sum_{t=2014,5}^{t=2014,3} Interest\ Income_{it}}{3}}{\frac{\sum_{t=2014,5}^{t=2014,2} Net\ Outstanding\ Loans_{it}}{4}} \quad (7.51)$$

There is variation between banks' exposure indices (Table 25 and Figure 41). This approach assumes that each bank's average income yield accurately measures its loan pricing and exposure to the cap. As discussed later, this will not be accurate if banks price market segments very differently within their portfolios, as, in this case, we would ideally need to break the panel down further to the margin for each bank's market segments. Between Q4 2012 and Q1 2014, there was some shift in the interest yield distribution between the cap regime changes, although the shift was far from radical. Interestingly, the cap exposure variable for loosening the cap in Q2 2014 has a higher average value than before the cap was imposed, suggesting that the cap regime did not strongly bind banks.

the six largest banks, each having over 5% market share for both the loans and deposit markets (making up 80-90% of total market) and each lending to both individual and corporates via branch networks that extend outside of Lusaka; and, the Small Banks, each of which have less than 5% market share and typically lend to corporates only, and not necessarily servicing large branch networks. Simpasa and Pla (2016)'s analysis of banks' sectoral credit concentration in Zambia provide evidence that separating large from small banks in the Zambian market might make sense.

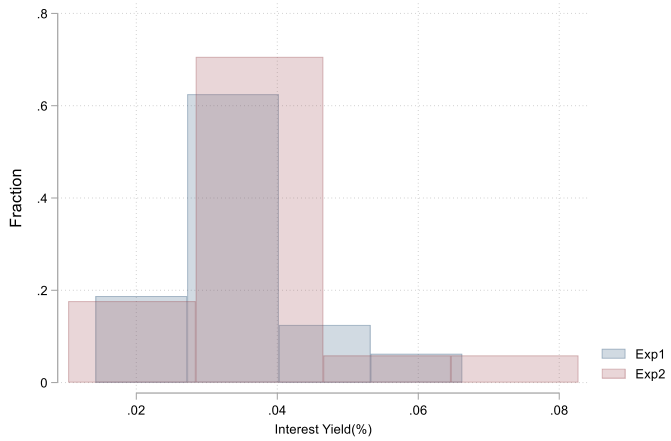
[211] There are other technical issues with using the yield as a measure of bank pricing; for example, when a loan becomes non-performing but has not yet been written off, it will underestimate the interest yield (this is because interest will not have been paid, but the loan will not yet have been provided for). Furthermore, the average yield will not do a good job at reflecting banks' exposure if they price multi-modally.

Table 25 – Exposure Variable

Summary Statistics

Variable	Obs	Mean	25th	Median	75th	Std.Dev	Coef.Var	Min	Max
Exposure Measure 1 (2013)	17	0.0361	0.0290	0.0363	0.0401	0.0121	0.336	0.0142	0.0662
Exposure Measure 2 (2014)	17	0.0369	0.0304	0.0380	0.0403	0.0155	0.419	0.0103	0.0827

Figure 41 – Exposure Variables Histogram



Note: This figure provides a histogram for banks' exposure indices both before the imposition of the cap regime (Exp1) and when it was loosened (Exp2). The exposure indices are measured as an interest yield, provided on the x-axis.

Source: BOZ prudential returns and Author's calculations.

As discussed, Zambia's cap regime is not a bad episode to study because the initial imposition of the cap does not appear endogenous to acute macroeconomic or political events, so we can be less concerned about the endogeneity issues clearly outlined in Ferrari et al. (2018) p.18.

A good question to ask is why there is a difference between our modelling of the IRR here as compared to our Chapter on Zambia's MFI. In this Chapter, we use a dummy variable to model the IRR regime and then an exposure interaction variable to capture the heterogeneity. However, in the MFI Chapter, we do not include a dummy to capture the IRR regime and directly model MFIs' cap bindingness. Why the difference? The primary reason is that the higher quality MFI interest rate data and more segmented market structure allowed us to be reasonably precise regarding how bound each lender was by the cap. In addition, the intuition of the simple model and a loan reaction that is plausibly linear with the level of MFI "boundness" allowed us to model the cap bindingness as the treatment itself. Unfortunately, we do not have access to high-quality interest rate data in this bank analysis. There is also the complication that banks can and do lend across sectors and borrower types, meaning that any use of an average interest rate might be misleading, and we cannot implicitly

assume market segregation across banks. The interest yield/exposure data for banks allows us to measure the intensity of the IRR regime for banks, but the level of the exposure data is not plausibly the treatment in itself.

Results

Background Results - Bank Margin Determinants in Zambia

Table 26 provides our regression results for bank margin determinants in Zambia using various versions of equation 7.4. In summary, the results tell a consistent story: operating costs are strongly associated with banks' margins in Zambia. However, we also find evidence of possible determinant heterogeneity between larger/diversified banks and smaller banks: smaller banks' margins appear driven by bank size and market structure variables. At the same time, liquidity holdings and macroeconomic factors determine large banks' margins. Across both bank types, there is strong evidence that margins are persistent.

We found the strongest and most consistent determinant of bank margins to be operating costs, a result which holds across both our margin calculation as well as the NIM. Operating costs capture the role of all non-interest costs in providing banking services. It reflects how inefficient or expensive it is to provide intermediation per unit of bank asset[212]. We find that for every percentage point that the operating cost ratio increases, banks' margins likely increase by between approximately 0.5-1 percentage points, depending on the exact regression and the source of variation (Table 26, row 2). The results argue that a significant proportion of the high intermediation costs in Zambia are likely linked to high underlying costs of operating and banking in Zambia. Considered in conjunction with our decomposition analysis, this high long-term operating cost passthrough argues that helping banks reduce their operating costs (wages, branch costs, loan enforcement and legal costs, etc.) may be one option to help shrink the costs of financial intermediation (and potentially reduce the costs of borrowing)[213]. One characteristic of the data driving the within-bank regression results is the reduction in overhead costs (and margins) in Zambia's banking system over the past decade due to

[212] Operating costs include the following: Salaries and benefits, Rent, Equipment, Education and Training, Audit and Legal Fees, Insurance, Fraud, and "Management Fees", sometimes used for transfer payments.

[213] What drives high absolute levels of operating costs? High fixed costs of lending for some borrower types, especially SMEs (Beck, 2013), operating a branch network in a large country with a low population density can be expensive, and legal processes can be slow in Zambia, increasing the costs of loan recovery and claims adjudication. Future work could look to use some of the detailed cost and geographic data in the prudential return to better understand a number of these drivers.

increased digitalisation banking and uptake of mobile banking. The time fixed effects regressions, which use cross-bank variation and produce more significant estimates for the smaller banks, suggest that banks with high operating costs pass this on to higher margins[214]. This finding of a strong link between operating costs and bank margins is consistent with much of the regional literature. Indeed, concentrating primarily on East Africa, our results are consistent with Beck and Hesse (2009) in Uganda; Chirwa and Mlachila (2004) in Malawi; Were and Wambu (2013) in Kenya, and Ahokpossi (2013) and Boutin-Dufresne (2013) who use a Sub-Saharan Africa panel. The other consistent determinant across bank types in our preferred bank fixed effects regressions is the bank-specific demand measure. The measure is inverted so lower demand in a sector that a bank has loans (measured as future entries into the NPL accounts) increases margins.[215] Finally, BOZ's reserve requirements, are also positively associated with higher margins for the several of the within-bank dynamic regressions [216]. The positive coefficients on reserve requirements – see row 11 in Table 26 – suggest that banks may pass on negative profitability associated with the liquidity effects high reserve requirements impose. However, as we discuss below, it can be a challenge to accurately disentangle how Zambia's macroeconomic environment fed through to margins more broadly.

The results also hint at possible heterogeneity between large/diversified and small banks. For example, margins for small banks appear to be consistently correlated/determined by bank size (measured by deposit market share), credit risk (measured RWA/Total Assets), Industry Concentration (measured by the Herfindahl Hirschman Index), and the price of copper. On the other hand, none of these variables are consistently significant for larger/diversified banks, which instead appear to have their margins driven by Liquidity, USD loans shares (in the cross-section), the level of the Kwacha, and inflation.

[214] A warning that these time-fixed effects results are perhaps more subject to omitted variable bias, given that our bank-level covariates may not capture all time-invariant bank characteristics.

[215] Annex 4 discusses this measure in detail, but it essentially measures a bank's loan exposure to sector-specific loan defaults, as they occur, based on future movements in the NPL accounts (the measure assumes loan defaults approximately measure loan demand conditions in a sector). For example, a bank that had a high exposure to farming would register a higher value if it held agricultural exposure while at that time agricultural loans that would end up in the aggregate NPL account (in 3 months) stopped paying interest etc. A bank holding 100% of its portfolio in agricultural loans while the average increase in future NPLs in the market associated with defaults in the current period increased by 5% we would expect margins to increase by 0.75 percentage points. This is a large change, but 5% is also a large number of future defaults to flow through the NPL 90-120 day account in a quarter.

[216] The results, although not convincingly consistent, indicate that 1 percentage point increase in reserve requirements is associated with a 0.16 percentage point to 0.2 percentage point increase in margin.

For the small banks, we suspect that the negative bank size coefficient results (row 4 in Table 24) reflect that some smaller banks may be prepared/required to forgo margin to capture higher deposit market shares. Either that or they face improving economies of scale as they increase in absolute size (several banks entered the sample and had strong market share growth while they became more established in the market)[217]. A one percentage point increase in deposit market share is associated with approximately a -0.6 percentage point (-0.215/(1-0.64)) level shift in margin. The positive coefficients on credit risk for the small banks (row 5) suggest that their loan margins reduced as they increased the share of risky loans/ assets in their portfolios. For small banks, this might reflect a size effect once again – more lending (rather than fixed-income securities) might be associated with positive economies of scale in lending. The industry concentration coeffects, measured by the HH index, which is a positive function of bank concentration, suggest that higher market concentration (in aggregate) is associated with high margins. This finding is consistent with some of the literature (for example, Chirwa and Mlachila (2004) in Malawi). However, it is interesting that only small banks are found to have industry concentration as a statistically significant determinant, possibly reflecting that for larger banks, changes in concentration might not equate to changes in competition. Future work might look to find better data and use other measures of concentration, for example, the Lerner Index as in Calice and Zhao (2018) or Maudos and de Guevara (2002). Finally, the copper price coefficient might suggest that small banks can charge higher rates when the commodities cycle is strong.

Large banks' margins appear to be driven more by liquidity and market conditions and less by size. Bank's holdings of liquid assets appear to be a significant margin determinant, ceteris paribus, with a ten percentage point increase in the ratio of liquid assets (excl. BOZ reserves) to deposits being associated with a level increase in margins by approximately 0.00275 percentage points (0.0076/(1-0.745)). Theoretically, banking dealership models typically argue that banks' margins should reduce as liquidity increases because banks have a better ability to meet short-term demands on their liquid assets (Nassar et al. 2014). But in Zambia's context, where liquid assets are already very high – the average liquidity variable is 44% among diversified banks – and "liquid" assets do not always have a deep tradeable market, the ratio might reflect broader macroeconomic and loan market conditions

[217] Note that the direction of these results holds when using an alternative absolute measures of bank size from the literature, for example the log of total assets.

as well as portfolio management decisions. For example, during Zambia's inflation and exchange rate shocks, banks switched to holding foreign deposits and domestic Treasury assets. If the marginal retrenched loan in this switch was higher risk (and more highly priced), we might expect to see a negative relationship between liquidity and margins. The exchange rate growth and inflation coefficients suggest that bank margins reduced during times of high inflation and exchange rate depreciation[218]. The negative exchange rate coefficients may reflect the fact that as the Kwacha falls, borrowers that have taken dollar-denominated loans yet earn in Kwacha, might start struggling to pay interest, which would reduce our margin ratio[219]. This negative exchange rate coefficient is a common finding in the literature (see Beck and Hesse (2009)), although direct comparisons across countries might be misleading due to varying levels of dollarisation. Many of these papers also find that inflation positively relates to bank margins, an outcome inconsistent with our results[220].

Finally, we do not find consistent support that Maturity Transformation is a determinant of bank margins in Zambia. This is in contrast to Bologna (2018), who does find a positive association between NIMs and contractual maturity transformation. One possible reason for this is that in Zambia, a very high proportion of lending is written with a short-term floating rate, meaning banks pass on to lenders a lot of the interest rate and market uncertainty that they might otherwise price in the margin.

In Annex 17 we provide results using banks' NIM rather than our margin definition. Remember that the NIM captures intermediation costs from all financing sources to all interest-generating uses and treats internal asset allocation and interest returns differently from our other margin calculations. Nevertheless, we again find strong evidence that operating margins drive margins. There are also clear signs of heterogeneity between large and small banks; however, small bank margins are more

[218] Our margin calculation should be insulated from most direct revaluation effects of changes in the Kwacha's value (even given the high dollarisation of Zambia's banking system). A depreciation in the Kwacha will mean dollar denominated loans will increase in value, as will their interest yields (in kwacha terms). The same happens on the interest income component of the margin calculation. So, at least as at a first order, foreign exchange effects should not mechanically impact margins. Some of the effect may be reflecting funding mismatches between dollar-denominated lending and financing, whereby currency depreciations effect lending or foreign financing unequally.

[219] A 10 percent depreciation in the kwacha is associated with an approximate 0.1 percentage point reduction in margin.

[220] For empirical results, see Beck and Hesse (2009), Demirgüç-Kunt and Huizinga (1999), or Rojas Suarez (2000), among others. Demirguc-Kunt and Huizinga (1999) argue that inflation might increase margins, as it can drive higher banking costs (requiring more transactions, branch assets, etc). Given that our inflation coefficients are specific to large banks and hold only in the full sample, we suspect that this result is picking up the broader effects of the Kwacha and inflation crisis in 2015/2016 and the high uncertainty at the time. These volatile periods provided the majority of the variation in inflation rates in our data. Different papers use various sets of macroeconomic controls, so it can also be a challenge to make like-for-like, ceteris paribus, comparisons.

driven by non-performing loans and risk aversion rather than bank size and credit risk. Large bank margins again appear driven by macroeconomic factors.

Finally, a caveat. These determinant results, especially the macroeconomic figures, should be seen as indicative rather than conclusive, a starting point for further research[221]. Further work might attempt to more precisely unfurl the relationships between bank margins, monetary policy (and some aggressive use of required reserves by BOZ), fiscal policy decisions (resulting in very high Treasury yields) and the macroeconomic environment, especially given the significant variation in the data and the instability of the late 2015 and 2019 periods. A VAR type approach may be one option for clarifying these relationships. Further, coupling such an approach with an accounting analysis of what appropriate commercial bank returns might look like in Zambia (including any transfer pricing to foreign headquarters shifting profit disbursements under the net income line) might help better ground the discussion of high interest rates and financial intermediation costs in Zambia.

[221] The margins analysis here was in part designed for our cap regime regressions, in that we wish to partial out as many of the macroeconomic factors driving margins as possible. This has likely led to somewhat oversaturated analysis, which is less helpful for determinant analysis.

Table 26 – Margin Determinant Results

Independent Variable		Full Sample						Pre-IRR Sample (end-2012Q4)						Large/Diversified Banks (Full Sample)						Small/Undiversified Banks (Full Sample)					
		Static			Dynamic			Static			Dynamic			Static			Dynamic			Static			Dynamic		
		OLS	FE	Time	OLS	FE	Time	OLS	FE	Time	FE	FE	Time	OLS	FE	Time	OLS	FE	Time	OLS	FE	Time	OLS	FE	Time
Bank-Specific Determinants	Lagged Margin				0.7436***	0.7437***	0.8214***				0.6901***	0.6611***	0.7878***				0.7089***	0.7451***	0.7534***				0.6826***	0.6374***	0.7816***
					[0.0623]	[0.0276]	[0.0212]				[0.0888]	[0.0645]	[0.0409]				[0.0246]	[0.0404]	[0.0458]				[0.0891]	[0.0385]	[0.0317]
	Operating Costs	1.0066***	0.438	0.9833**	0.284	0.3038***	0.2893***	1.0836**	0.641	1.1107**	0.5270*	0.6314***	0.4394***	0.4849*	0.5690*	0.337	0.2224*	0.2410***	0.1335**	0.8796**	0.344	0.8676**	0.320	0.3853***	0.3855***
		[0.3873]	[0.2869]	[0.4257]	[0.1632]	[0.0642]	[0.0646]	[0.4499]	[0.3840]	[0.4707]	[0.2899]	[0.0909]	[0.1223]	[0.2539]	[0.2661]	[0.2552]	[0.1131]	[0.0675]	[0.0615]	[0.3944]	[0.4132]	[0.4120]	[0.2669]	[0.0908]	[0.0877]
	Liquidity Ratio	-0.001	0.008	-0.003	0.006	0.0054**	0.001	-0.007	-0.004	-0.006	0.008	0.006	0.000	0.0216*	0.0038**	0.020	0.0075***	0.0076*	0.0097**	-0.008	0.011	-0.004	0.003	0.004	-0.003
		[0.0107]	[0.0070]	[0.0115]	[0.0036]	[0.0027]	[0.0021]	[0.0107]	[0.0104]	[0.0114]	[0.0059]	[0.0045]	[0.0034]	[0.0111]	[0.0015]	[0.0128]	[0.0015]	[0.0042]	[0.0039]	[0.0122]	[0.0132]	[0.0127]	[0.0057]	[0.0040]	[0.0030]
	Bank Size	-0.034	-0.142	-0.040	-0.026	-0.037	0.001	-0.0787*	-0.237	-0.0786*	-0.060	-0.057	-0.015	0.012	0.086	0.004	0.011	-0.002	-0.003	-0.4897***	-0.7023***	-0.5264***	-0.2047**	-0.2158**	-0.037
		[0.0302]	[0.1508]	[0.0329]	[0.0469]	[0.0233]	[0.0068]	[0.0426]	[0.1859]	[0.0434]	[0.0653]	[0.0832]	[0.0140]	[0.0241]	[0.0662]	[0.0321]	[0.0210]	[0.0140]	[0.0103]	[0.1280]	[0.1483]	[0.1657]	[0.0878]	[0.0900]	[0.0459]
	Credit Risk	-0.0510**	-0.0494**	-0.0474**	-0.0087**	-0.0058*	-0.004	-0.0786***	-0.058	-0.0798***	-0.020	-0.0148**	-0.0119**	-0.0156**	-0.0280**	0.004	-0.002	-0.001	0.004	-0.0839***	-0.0488*	-0.0845**	-0.0128*	-0.0102*	-0.006
		[0.0211]	[0.0203]	[0.0228]	[0.0039]	[0.0032]	[0.0027]	[0.0304]	[0.0348]	[0.0307]	[0.0121]	[0.0074]	[0.0055]	[0.0070]	[0.0087]	[0.0103]	[0.0027]	[0.0046]	[0.0039]	[0.0322]	[0.0229]	[0.0359]	[0.0067]	[0.0053]	[0.0057]
	Maturity Transformation	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.0003**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-0.001	0.000	0.000	0.000	0.000
		[0.0001]	[0.0002]	[0.0002]	[0.0001]	[0.0001]	[0.0000]	[0.0004]	[0.0003]	[0.0003]	[0.0001]	[0.0001]	[0.0001]	[0.0001]	[0.0001]	[0.0001]	[0.0000]	[0.0000]	[0.0000]	[0.0003]	[0.0004]	[0.0003]	[0.0001]	[0.0001]	[0.0001]
	USD Loan Share	-0.0110*	-0.015	-0.013	0.000	-0.001	0.001	0.002	-0.013	0.004	0.001	0.007	0.007	-0.0294***	-0.019	-0.0484***	-0.002	-0.002	-0.0091***	0.004	0.005	0.008	0.008	0.0097*	0.005
		[0.0065]	[0.0112]	[0.0081]	[0.0037]	[0.0030]	[0.0022]	[0.0094]	[0.0285]	[0.0099]	[0.0069]	[0.0090]	[0.0045]	[0.0098]	[0.0131]	[0.0136]	[0.0038]	[0.0034]	[0.0033]	[0.0153]	[0.0121]	[0.0176]	[0.0054]	[0.0057]	[0.0033]
	Non-performing Loans	0.024	0.011	0.019	0.004	0.004	0.005	0.059	0.011	0.060	-0.0175*	-0.0212**	0.000	0.0443***	0.013	0.016	0.006	0.005	-0.001	0.007	0.016	0.007	0.008	0.0108*	0.007
		[0.0250]	[0.0222]	[0.0238]	[0.0079]	[0.0042]	[0.0034]	[0.0410]	[0.0301]	[0.0431]	[0.0096]	[0.0090]	[0.0070]	[0.0136]	[0.0292]	[0.0113]	[0.0108]	[0.0063]	[0.0070]	[0.0291]	[0.0240]	[0.0323]	[0.0094]	[0.0065]	[0.0060]
	Risk Aversion	0.005	0.006	0.006	0.005	0.0054***	0.0048***	0.004	0.007	0.003	0.005	0.004	0.0051*	-0.0080**	-0.002	-0.004	-0.005	-0.005	-0.004	0.000	-0.002	-0.003	0.006	0.0059**	0.0072**
		[0.0078]	[0.0070]	[0.0078]	[0.0036]	[0.0017]	[0.0014]	[0.0109]	[0.0096]	[0.0112]	[0.0058]	[0.0033]	[0.0027]	[0.0039]	[0.0048]	[0.0029]	[0.0038]	[0.0033]	[0.0033]	[0.0096]	[0.0135]	[0.0099]	[0.0041]	[0.0030]	[0.0028]
	Demand Measure	0.2407***	0.2725***	0.2644***	0.0830**	0.0725***	0.029	0.3676***	0.3032*	0.3941***	0.0910**	0.0870*	0.037	0.1084**	0.1493**	0.0595*	0.0771***	0.0765*	0.046	0.394	0.399	0.6546***	0.070	0.083	0.010
		[0.0784]	[0.0810]	[0.0896]	[0.0343]	[0.0257]	[0.0295]	[0.1228]	[0.1561]	[0.1302]	[0.0386]	[0.0483]	[0.0413]	[0.0447]	[0.0408]	[0.0347]	[0.0177]	[0.0398]	[0.0488]	[0.2467]	[0.2438]	[0.2122]	[0.0935]	[0.0668]	[0.0767]
Macro and Institutional Determinants	Reserve Requirements	0.037	0.026		0.014	0.012		0.1278**	0.1501**		0.031	0.0403***		0.034	0.038		0.018	0.0149*		0.071	0.088		0.040	0.0396***	
		[0.0352]	[0.0324]		[0.0090]	[0.0108]		[0.0520]	[0.0593]		[0.0249]	[0.0148]		[0.0316]	[0.0284]		[0.0141]	[0.0082]		[0.0739]	[0.0706]		[0.0228]	[0.0152]	
	Industry Concentration	0.1644**	0.166		0.027	0.0272*		0.061	0.059		0.011	0.019		-0.022	-0.055		-0.021	-0.017		0.3219***	0.3562***		0.1038*	0.1222***	
		[0.0788]	[0.0961]		[0.0263]	[0.0144]		[0.0902]	[0.0900]		[0.0336]	[0.0355]		[0.0255]	[0.0484]		[0.0150]	[0.0108]		[0.0844]	[0.0903]		[0.0483]	[0.0272]	
	Copper Price	0.002	0.001		0.001	0.001		0.0030*	0.002		0.001	0.0010*		-0.001	-0.001		0.000	0.000		0.0056**	0.0044*		0.0023**	0.0021***	
		[0.0014]	[0.0012]		[0.0006]	[0.0005]		[0.0016]	[0.0015]		[0.0008]	[0.0005]		[0.0008]	[0.0008]		[0.0003]	[0.0003]		[0.0022]	[0.0022]		[0.0010]	[0.0007]	
	Inflation	0.007	0.023		0.017	0.021		0.0811**	0.084		0.045	0.0583*		-0.019	-0.028		-0.0247*	-0.0281*		0.059	0.052		0.050	0.0585**	
		[0.0249]	[0.0278]		[0.0188]	[0.0208]		[0.0400]	[0.0492]		[0.0474]	[0.0305]		[0.0144]	[0.0164]		[0.0116]	[0.0144]		[0.0377]	[0.0319]		[0.0300]	[0.0295]	
	Treasury Bill Rate	0.037	0.049		0.018	0.018		0.1083*	0.074		0.002	-0.004		-0.031	-0.038		0.003	0.006		0.1105*	0.074		0.014	0.017	
		[0.0395]	[0.0357]		[0.0175]	[0.0146]		[0.0590]	[0.0758]		[0.0188]	[0.0186]		[0.0347]	[0.0253]		[0.0165]	[0.0110]		[0.0588]	[0.0469]		[0.0356]	[0.0206]	
	Exchange Rate	-0.004	-0.008		-0.0092***	-0.0095**		0.003	-0.002		-0.0139*	-0.0133***		-0.0160***	-0.0162***		-0.0109**	-0.0106***		0.012	0.000		-0.006	-0.008	
		[0.0074]	[0.0060]		[0.0031]	[0.0042]		[0.0129]	[0.0113]		[0.0073]	[0.0044]		[0.0033]	[0.0030]		[0.0030]	[0.0028]		[0.0106]	[0.0084]		[0.0042]	[0.0057]	
	Interest Rate Volatility	0.000	0.000		0.000	0.000		-0.001	0.000		0.000	0.000		0.000	0.000		0.000	0.000		0.000	-0.001		0.000	0.000	
		[0.0004]	[0.0003]		[0.0002]	[0.0003]		[0.0006]	[0.0007]		[0.0003]	[0.0003]		[0.0002]	[0.0003]		[0.0002]	[0.0002]		[0.0006]	[0.0006]		[0.0003]	[0.0004]	
	Constant	0.010	0.027	0.0651***	-0.005	-0.001	0.001	0.013	0.040	0.0697***	0.005	-0.002	0.003	0.0564***	0.0615***	0.0625***	0.011	0.009	0.006	-0.026	-0.031	0.0750***	-0.025	-0.0300***	0.004
		[0.0219]	[0.0211]	[0.0153]	[0.0091]	[0.0074]	[0.0039]	[0.0231]	[0.0276]	[0.0159]	[0.0072]	[0.0127]	[0.0062]	[0.0121]	[0.0078]	[0.0107]	[0.0064]	[0.0062]	[0.0055]	[0.0429]	[0.0390]	[0.0240]	[0.0183]	[0.0103]	[0.0073]
	Observations	738	738	738	715	715	715	395	395	395	375	375	375	350	350	350	342	342	342	388	388	388	373	373	373
	R-Squared (Total)	0.246	0.278	0.278	0.693	0.845	0.852	0.253	0.287	0.264	0.651	0.829	0.850	0.304	0.349	0.477	0.715	0.917	0.922	0.383	0.436	0.433	0.708	0.836	0.840
	Bank Fixed Effects	x	√	x	x	√	x	x	√	x	x	√	x	x	√	x	x	√	x	x	√	x	x	√	x
	Time Fixed Effects	x	x	√	x	x	√	x	x	√	x	x	√	x	x	√	x	x	√	x	x	√	x	x	√
	Standard Errors	robust	robust	robust	robust	PCSE	PCSE	robust	robust	robust	robust	PCSE	PCSE	robust	robust	robust	robust	PCSE	PCSE	robust	robust	robust	robust	PCSE	PCSE

Note: Estimates from equation 7.4 regressing bank margins ($Margin = \frac{Interest\ Received\ from\ loans + fees}{Performing\ Loans} - \frac{Interest\ paid\ to\ all\ sources}{All\ Interest\ Bearing\ Liabilities}$) on bank-specific and common bank determinants. Different sub-samples and estimation approaches are provided in each column. Variables are defined in full in Table 24. The large Diversified Banks include the largest seven banks in the quarter before the IRR (ABC, Barclays, FNB, IZB, Stanbic, SCB, ZANACO), all of whom have over 5% of market share (by loans and deposits) and maintain large branch networks. Standard error type provided in final row. $^{***}p<0.01$, $^{**}p<0.05$, $^{*}p<0.10$.Table 24 – Variable Descriptions and Discussion

Interest Rate Caps and Bank Margins

We now turn to our primary research question and examine how bank margins (including fees) might be associated with Zambia's IRR regime. We begin by using a dummy variable approach to capture the simple level-effects of the cap. Specifically, we estimate (i) a dummy variable to capture the IRR up to its loosening in 2014, and (ii) dummy variables to capture the IRR over the whole dataset, including the cap's loosening and removal:

$$Y_{it} = \alpha_i + \sum_{l=1}^{L} \rho^l Y_{it-l} + \sum_{j=1}^{J} \beta^j X_{it}^j + \sum_{k=1}^{K} \theta^k M_t^k + \tau^a C_{2013} + \varepsilon_{it} \qquad (7.6)$$

and

$$Y_{it} = \alpha_i + \sum_{l=1}^{L} \rho^l Y_{it-l} + \sum_{j=1}^{J} \beta^j X_{it}^j + \sum_{k=1}^{K} \theta^k M_t^k + \tau^a C_{2013} + \tau^c C_{2014} + \varepsilon_{it} \qquad (7.7)$$

Equation 7.6 is essentially asking whether banks' average level of margins was different before the cap was implemented compared to when it was loosened in Q2 2014, ceteris paribus and including controlling for time-invariant bank fixed effects. Equation 7.7 is similar to equation 7.6 except that it compares margins during the two cap regimes to their pre-IRR and post-IRR values.

Using this simple levels dummy variable approach, we do not find overwhelming evidence that the initial imposition of the cap was associated with large level shifts in banks' margins. This holds across the pooled and the bank-type subsets in the pre-cap loosening sample (Table 27), which we feel are likely the most robust estimates due to the relative exogeneity of the initial cap regime and the more stable macroeconomic environment. Using the full sample provides some weak evidence that small banks may have faced downward pressure on their margins, which is consistent with our prior that small banks typically price higher and thus would be more likely to be exposed to and see an impact from the cap (Table 28). The small bank estimate implies that the initial cap regime was associated with a 0.78 percentage point long-run level reduction in loan margin and 0.33 percentage-point reduction in the NIM[222]. For the loan margin, this is an economically meaningful result.

[222] Note the long run effects for the margin and NIM were calculated as $long\ run = \frac{\tau^a C_{2013}}{(1-\rho)}$, or, for the Margin, 0.78% =0.0028/(1-0.649).

However, we are cautious to weigh the result too heavily, given the volatile operating conditions post Q2 2014 and the lack of corroborating effects in the constrained sample. As fiscal conditions deteriorated toward the end of the sample, bank NIMs also rose as banks found government securities profitable, which we might not be adequately controlling for and might make margins during the cap regime seem relatively reduced.

This weak or null non-rejection outcome is not necessarily surprising. First, given the low levels of cap exposure for many banks, the effect size may simply be too small (given our limited sample and the complexity of the banking system, picking up what might be small effects may just not be feasible). Secondly, if we believe that banks behaved according to the income-switching hypothesis and that interest income and fees are fungible, we wouldn't necessarily see a large change in the fee-inclusive margin. And finally, it could be that the large-diversified/small bank separation does not capture enough of the pricing variability amongst banks or, simply, the cap regime did not have an effect.

In terms of coherence with our margin determinants analysis, this IRR analysis we found that the coefficients for the margin determinants are reasonably consistent with our previous results (Table 26) and the literature.

Table 27 – Binary Treatment for Initial IRR Imposition and Bank Margins (Static and Dynamic)

Dependent Variable	Sample	Coefficient	Static		Dynamic (Short-run effect)	
			OLS	Bank FE	OLS	Bank FE
Margin	All Banks	τa, Cap2013	-0.0037	-0.0055	-0.0012	-0.0004
		[se]	[0.0042]	[0.0047]	[0.0015]	[0.0014]
		rho			0.7321***	0.7223***
	Diversified	τa, Cap2013	-0.0108**	-0.0109**	-0.0017	-0.0013
		[se]	[0.4648]	[0.3486]	[0.2424]	[0.0780]
		rho			0.006	0.0052
	Not Diversified	τa, Cap2013	-0.0043	-0.0042	-0.0021	0.0002
		[se]	[0.0079]	[0.0076]	[0.0029]	[0.0024]
		rho			0.6988***	0.6567***
NIM	All Banks	τa, Cap2013	-0.0015	-0.0011	-0.0016	-0.0013*
		[se]	[0.0021]	[0.0018]	[0.0012]	[0.0007]
		rho			0.5234***	0.4768***
	Diversified	τa, Cap2013	-0.0046**	-0.0038	-0.0014	-0.001
		[se]	[0.0019]	[0.0022]	[0.0012]	[0.0009]
		rho			0.5616***	0.5731***
	Not Diversified	τa, Cap2013	-0.0002	-0.0004	-0.002	-0.0006
		[se]	[0.0029]	[0.0027]	[0.0022]	[0.0011]
		rho			0.4554**	0.3248***

Note: Results from versions of equation 7.6. Each cell provides the policy coefficient estimate from a different regression, for example -0.0037 from row 1 column 1 uses bank margins as the dependent variable, the full bank sample, and a static OLS approach. Results are provided for two different dependent variables: the bank margin ($Margin = \frac{Interest\ Received\ from\ loans + fees}{Performing\ Loans} - \frac{Interest\ paid\ to\ all\ sources}{All\ Interest\ Bearing\ Liabilities}$) and the Net Interest Margin, as defined in equation 7.3 "Static" results do not include a lagged dependent variable, Y_{it-l} in the estimation. Covariates are included in all regressions and are consistent with the set used in Table 26. Only short run effects are shown for the dynamic estimates. Rho is the coefficient on the lagged dependent variable (either Margin or NIM). τa, Cap2013 is the estimate for a dummy variable taking the value of 1 between Q1 2013 to Q2 2014, and 0 otherwise. The large "Diversified" Banks include the largest seven banks, all of whom have over 5% of market share (by loans and deposits) and maintain large branch networks. The "Not Diversified" banks make up the remainder. Robust standard errors used for static regressions and PCSE for dynamic. $^{***}p<0.01$, $^{**}p<0.05$, $^{*}p<0.10$.

Table 28 – Binary Treatment for IRR Regime and Bank Margins (Static and Dynamic)

Dependent Variable	Sample	Coefficient	Static		Dynamic (Short-run effect)	
			OLS	Bank FE	OLS	Bank FE
Margin	All Banks	τa, Cap2013	-0.0061*	-0.0066**	-0.0021*	-0.0015
		[se]	[0.0032]	[0.0027]	[0.0010]	[0.0010]
		τc, Cap2014	-0.0044*	-0.0060**	-0.0014	-0.0017
		[se]	[0.0026]	[0.0023]	[0.0014]	[0.0012]
		rho			0.7435***	0.7455***
	Diversified	τa, Cap2013	-0.0103***	-0.0105***	-0.0021**	-0.0016
		[se]	[0.0018]	[0.0017]	[0.0008]	[0.0011]
		τc, Cap2014	-0.0093***	-0.0087***	-0.001	-0.0013
		[se]	[0.0023]	[0.0023]	[0.0018]	[0.0012]
		rho			0.6990***	0.7227***
	Not Diversified	τa, Cap2013	-0.0056	-0.0047	-0.0027	-0.0021
		[se]	[0.0057]	[0.0044]	[0.0019]	[0.0014]
		τc, Cap2014	-0.0007	-0.0062	-0.0029	-0.0028*
		[se]	[0.0047]	[0.0045]	[0.0022]	[0.0015]
		rho			0.6781***	0.6496***
NIM	All Banks	τa, Cap2013	-0.0025	-0.0023*	-0.0015***	-0.0016**
		[se]	[0.0016]	[0.0012]	[0.0005]	[0.0007]
		τc, Cap2014	-0.0026**	-0.0023*	-0.0008	-0.0011
		[se]	[0.0012]	[0.0012]	[0.0007]	[0.0008]
		rho			0.5946***	0.5700***
	Diversified	τa, Cap2013	-0.0035**	-0.0032**	-0.001	-0.0007
		[se]	[0.0016]	[0.0012]	[0.0009]	[0.0008]
		τc, Cap2014	-0.0025	-0.002	-0.0001	-0.0006
		[se]	[0.0016]	[0.0015]	[0.0010]	[0.0009]
		rho			0.5979***	0.6226***
	Not Diversified	τa, Cap2013	-0.0016	-0.0006	-0.0012	-0.0018**
		[se]	[0.0024]	[0.0020]	[0.0007]	[0.0009]
		τc, Cap2014	-0.002	-0.0028	-0.0017	-0.0018**
		[se]	[0.0021]	[0.0018]	[0.0010]	[0.0009]
		rho			0.5108***	0.4475***

Note: Results from versions of equation 7.7. Each set of τa, Cap2013 and τc, Cap2014 provides policy coefficient estimate from a different regression. For example, the estimates-0.0061 and -0.0044 from row 1 column 1 and row 2 column 1 are from the same static OLS regression using all banks and the "margin" as the dependent variable. Results are provided for two different dependent variables: the bank margin ($Margin = \frac{Interest\ Received\ from\ loans + fees}{Performing\ Loans} - \frac{Interest\ paid\ to\ all\ sources}{All\ Interest\ Bearing\ Liabilities}$) and the Net Interest Margin, as defined in equation 7.3. "Static" results do not include a lagged dependent variable, Y_{it-l} in the estimation. Covariates are included in all regressions and are consistent with the set used in Table 26. Only short run effects are shown for the dynamic estimates. Rho is the coefficient on the lagged dependent variable (either Margin or NIM). τa, Cap2013 is the estimate for a dummy variable taking the value of 1 between Q1 2013 to Q2 2014, and 0 otherwise. τa, Cap2014 is the estimate for a dummy variable taking the value of 1 between Q3 2014 to Q4 2015, and 0 otherwise. The large "Diversified" Banks include the largest seven banks, all of whom have over 5% of market share (by loans and deposits) and maintain large branch networks. The "Not Diversified" banks make up the remainder. Robust standard errors used for static regressions and PCSE for dynamic. ***p<0.01, **p<0.05, *p<0.10.

As a robustness check, we shift away from the margin levels approach described above and instead model the imposition of the cap and its association with respect to changes in bank margins. This approach sits closer to our work on Zambia's microfinance sector. Specifically, we estimate a distributed lag model with the change in bank margins as the dependent variable and model the imposition of the cap with a series of leading and lagged time dummy variables. Annex 20 shows that the results remain consistent. There appears to be little convincing evidence that the imposition

of the IRR regime was associated with reductions in large banks' margins or NIMs, although smaller banks may have seen a drop in margins in the periods post-IRR imposition.

Cap Exposure Results

We now examine the results from modelling variation in banks' exposure to the IRR regime. To better isolate banks' cap exposures, we add an interaction term between our exposure variable and the cap regime dummy variable to our regression equation (equations 7.8 and 7.9). The idea here is to allow a linear heterogeneous level shift in a bank's margins conditional on our measure of their exposure to the cap. This might capture more granularity in banks' pricing than the simple larger diversified/small categorisations used above, although it does impose that this relationship is linear in banks' exposure.

$$Y_{it} = \alpha_i + \sum_{l=1}^{L} \rho^l Y_{it-l} + \sum_{j=1}^{J} \beta^j X_{it}^j + \sum_{k=1}^{K} \theta^k M_t^k + \tau^a C_{2013} + \tau^b C_{2013}.E_i^1 + \varepsilon_{it} \quad (7.8)$$

$$Y_{it} = \alpha_i + \sum_{l=1}^{L} \rho^l Y_{it-l} + \sum_{j=1}^{J} \beta^j X_{it}^j + \sum_{k=1}^{K} \theta^k M_t^k + \tau^a C_{2013} + \tau^b C_{2013}.E_i^1 + \tau^c.C_{2014}.+\tau^d C_{2014}.E_i^2 + \varepsilon_{it} \quad (7.9)$$

Assuming that our exposure measure can capture variation in banks' pricing and that competition in Zambia's banking market structure is imperfect (using a Courakis' (1984)-type effect caps should reduce bank margins), we would expect that both τ^a and τ^b to be negative. This would mean that banks with more exposure to the cap would see a larger negative margin shift. τ^c and τ^d would have a similar sign (negative), albeit with a much smaller size because the IRR regime is weaker post-Q2 2014, ceteris paribus.

This exposure variable approach provides little or no evidence that the IRR regime may have influenced bank margins heterogeneously. Across most of our bank fixed effect specifications, our results do not find convincing evidence of an interaction effect. The one exception is for the short sample ending before the cap was loosened, where there is some evidence that bank margins reduced for less exposed lenders, which is arguably contrary to our theoretical priors (see Row 1-2, Column

4, in Table 29 and Table 30)[223]. The 25th percentile bank in the exposure distribution, with an exposure yield of 2.9%, for example, is estimated to have seen a level margin reduction of 0.3 percentage points associated with the cap regime; The 75th percentile bank, a level shift of upwards of 0.27 percentage points (see Table 30, which calculates the short-run and long-run marginal effects of the cap on loan margin). Yet this estimated effect does not hold when examining the full sample and using bank sub-samples. Small banks do register a small level effect on their margins, but not for their NIM, which is weakly consistent with Table 27. But the results do not find convincing evidence of a significant interaction effect.

These weak and null results, from both the simple dummy and the interaction approach, should be interpreted with caution. Lack of evidence is not evidence of absence, especially in underpowered regressions such as ours. That being said, there is a clear difference in qualitative outcomes between the strong effects of the IRRs on MFI lending, where we know that the cap strongly bound lenders, and these banking results where (i) we suspect banks may have been less bound, and (ii) banks have additional options within their balance sheets to switch between asset types.

In summary, so far, we have not found strong and consistent evidence that the cap regime impacted banks' margins, although there may be some limited effect for smaller banks. One potential explanation for the lack of an association, power issues aside, is that banks' income lines are fungible: interest can be interchanged with fees to sidestep the impacts of the cap. We examined this in Chapter 5 for MFIs, finding strong effects, and now in the next section, we test for the same effects in our banking sample.

[223] Note that smaller sample means the dummy variable and the exposure interaction effects are comparing bank margins during the initial 2013-2014 cap period to before the IRR was introduced.

Table 29 – IRR Regime, Exposure Interactions, and Bank Margins

Dependent Variable	Sample	Coefficient	Static		Dynamic (Short-run effect)	
			OLS	Bank FE	OLS	Bank FE
Margin	Short Sample (All banks)	τa, Cap2013	-0.0299***	-0.0234*	-0.0076	-0.0059*
		[se]	[0.0109]	[0.0120]	[0.0058]	[0.0030]
		τb, Interaction	0.7300***	0.5038	0.1814	0.1670**
		[se]	[0.2336]	[0.2950]	[0.1415]	[0.0849]
		rho			0.7263***	0.7155***
	Full Sample (All Banks)	τa, Cap2013	-0.0291**	-0.0215*	-0.0071*	-0.0043
		[se]	[0.0119]	[0.0116]	[0.0036]	[0.0030]
		τb, Interaction	0.6447**	0.4206	0.1416	0.081
		[se]	[0.3095]	[0.3245]	[0.0824]	[0.0870]
		τc, Cap2014	-0.0195**	-0.0122	-0.0012	-0.0018
		[se]	[0.0088]	[0.0093]	[0.0046]	[0.0028]
		τd, Interaction	0.4302*	0.1806	-0.003	0.0037
		[se]	[0.2296]	[0.2468]	[0.1148]	[0.0755]
		rho			0.7412***	0.7448***
	Diversified Banks (Full Time Peroid)	τa, Cap2013	-0.0277**	-0.0230*	-0.005	-0.0038
		[se]	[0.0138]	[0.0112]	[0.0050]	[0.0072]
		τb, Interaction	0.4902	0.3535	0.0802	0.0633
		[se]	[0.3448]	[0.2620]	[0.1606]	[0.2125]
		τc, Cap2014	0.0014	-0.0032	0.00000	-0.0013
		[se]	[0.0096]	[0.0114]	[0.0045]	[0.0065]
		τd, Interaction	-0.307	-0.1559	-0.0285	0.0025
		[se]	[0.3380]	[0.3956]	[0.1831]	[0.1997]
		rho			0.6972***	0.7214***
	Not Diversified Banks (Full Time Peroid)	τa, Cap2013	-0.0254*	-0.0269*	-0.0124*	-0.0074**
		[se]	[0.0154]	[0.0146]	[0.0056]	[0.0037]
		τb, Interaction	0.5507	0.6254	0.2726**	0.1531
		[se]	[0.4323]	[0.3987]	[0.1154]	[0.1156]
		τc, Cap2014	-0.0139	-0.0196*	-0.0069	-0.0068*
		[se]	[0.0123]	[0.0087]	[0.0065]	[0.0036]
		τd, Interaction	0.3659	0.3759*	0.1114	0.1131
		[se]	[0.2978]	[0.1716]	[0.1311]	[0.0991]
		rho			0.6671***	0.6442***
NIM	Short Sample (All banks)	τa, Cap2013	-0.0096	-0.0067	-0.0041	-0.0003
		[se]	[0.0088]	[0.0072]	[0.0034]	[0.0046]
		τb, Interaction	0.2264	0.1573	0.0729	-0.0294
		[se]	[0.2075]	[0.1685]	[0.0795]	[0.1217]
		rho			0.5184***	0.4777***
	Full Sample (All Banks)	τa, Cap2013	-0.0125	-0.0099	-0.0043	-0.0015
		[se]	[0.0087]	[0.0079]	[0.0031]	[0.0039]
		τb, Interaction	0.2799	0.2134	0.0786	-0.0055
		[se]	[0.2311]	[0.2126]	[0.0800]	[0.1131]
		τc, Cap2014	-0.001	0.001	0.0017	0.002
		[se]	[0.0023]	[0.0027]	[0.0017]	[0.0036]
		τd, Interaction	-0.0436	-0.0911	-0.0713*	-0.0879
		[se]	[0.0499]	[0.0679]	[0.0386]	[0.1041]
		rho			0.5887***	0.5710***
	Diversified Banks (Full Time Sample)	τa, Cap2013	-0.0104	-0.0077	-0.0016	0.0005
		[se]	[0.0122]	[0.0118]	[0.0055]	[0.0046]
		τb, Interaction	0.1921	0.1234	0.0174	-0.0346
		[se]	[0.3004]	[0.3155]	[0.1604]	[0.1326]
		τc, Cap2014	0.0100***	0.0076*	0.0035	0.0026
		[se]	[0.0021]	[0.0037]	[0.0033]	[0.0045]
		τd, Interaction	-0.3635***	-0.2796	-0.1042	-0.0937
		[se]	[0.0848]	[0.1528]	[0.1219]	[0.1321]
		rho			0.5945***	0.6185***
	Not Diversified Banks (Full Time Sample)	τa, Cap2013	-0.0099	-0.0103	-0.006	-0.002
		[se]	[0.0103]	[0.0088]	[0.0035]	[0.0046]
		τb, Interaction	0.2294	0.2682	0.134	0.0052
		[se]	[0.2751]	[0.2293]	[0.0863]	[0.1248]
		τc, Cap2014	0.0014	-0.0017	-0.0004	0.0008
		[se]	[0.0030]	[0.0038]	[0.0026]	[0.0040]
		τd, Interaction	-0.0944*	-0.0289	-0.0356	-0.0714
		[se]	[0.0502]	[0.0791]	[0.0491]	[0.1086]
		rho			0.4989***	0.4511***

Note: Results from versions of equation 7.8 and 7.9. Each set of τa, Cap2013, τb Interaction and τc, Cap2014, and τd Interaction (all within the same horizontal black lines) provide estimates are from the same regression. For example, the estimates -0.0291 to 0.4302 from row 3 column 1 to row 6 column 1 are from the same static OLS regression using all banks and the "margin" as the dependent variable. Results are provided for two different dependent variables: the bank margin ($Margin = \frac{Interest\ Received\ from\ loans + fees}{Performing\ Loans} - \frac{Interest\ paid\ to\ all\ sources}{All\ Interest\ Bearing\ Liabilities}$) and the Net Interest Margin, as defined in equation 7.3. "Static" results do not include a lagged dependent variable, Y_{it-l} in the estimation. Covariates are included in all regressions and are consistent with the set used in Table 26. Only short run effects are shown for the dynamic estimates. Rho is the coefficient on the lagged dependent variable (either Margin or NIM). τa, Cap2013 is the estimate for a dummy variable taking the value of 1 between Q1 2013 to Q2 2014, and 0 otherwise. τa, Cap2014 is the estimate for a dummy variable taking the value of 1 between Q3 2014 to Q4 2015, and 0 otherwise. The large "Diversified" Banks include the largest seven banks, all of whom have over 5% of market share (by loans and deposits) and maintain large branch networks. The "Not Diversified" banks make up the remainder. The "Short Sample" finishes in Q2 2014. Robust standard errors used for static regressions and PCSE for dynamic. ***p<0.01, **p<0.05, *p<0.10.

Table 30 – Short-Run and Long-Run Effects for All Banks and Short Sample

	Margin Exposure						
	0.01	0.02	0.03	**0.036**	0.04	0.05	0.06
Short-Run Effect	-0.0042	-0.0026	-0.0009	0.0001	0.0008	0.0025	0.0041
Lon-Run Effect	-0.0149	-0.0090	-0.0031	0.0004	0.0027	0.0086	0.0145

Note: Short run and long run effects estimates based on equation 7.8. Using the short sample ending in Q2 2014 (see column 4 rows 1 and 2 in Table 29) for all banks. Margin exposures (from equation 7.51) provided over range the 0.01 to 0.06. Short-run effects calculated as (τa, Cap2014)* E_i^1 and Long-Run effects calculated as (τa, Cap2014)* E_i^1/(1-rho).

Results – Banks and Interest Income Switching

As discussed in our work on MFIs in Zambia, one of the often-cited reasons not to introduce interest rate caps into a market is that, unless the caps are implemented with other strict constraints on banks' fees, the caps may have little impact on the final total cost of credit facing borrowers. World Bank researchers, for example, state: *"The effectiveness of caps is often undercut by the use of non-interest fees and commissions." – Ferrari et al. (2018) p.4.*

The possibility of banks switching income in response to the imposition of interest rate caps is important because it undermines the potential effectiveness of the caps in (i) correcting for market competition imperfections or (ii) reducing effective interest rates for borrowers. It can also lead to more opaque pricing within the market. Yet there has only been limited research analysing it.

In Zambia, the cap regulation was applied to banks' maximum effective annual lending rates. This rate captures some fees in its calculation yet still leaves significant room for income shifting[224]. In addition to loan fee shifting (for example, loan origination fees) we are also interested in whether banks shift interest income into non-loan revenue lines, an example of which might be increased account transaction charges or till charges.

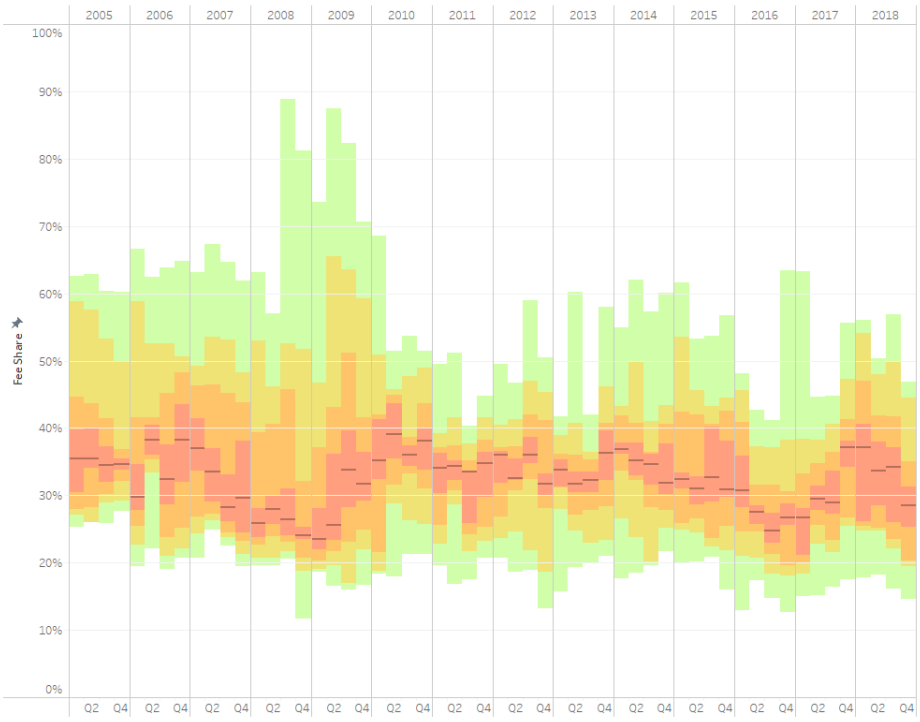
Before we begin a quantitative analysis, we first analyse the deciles of banks' fee and commission share of loan-related income over time [225]. For banks, the fee share of loan-related income reflects how much intermediation income is coming from fees versus interest. Unlike what we saw for Zambia's microfinance sector (see Figure 34), there is not a sharp kick-up in the distribution associated with the cap imposition in Q1 2013 for banks (Figure 42). Another interesting point is

[224] Fees are regulated in Zambia via Statutory Instrument No. 179 of 1995, but the definitions carve out significant room for lenders to apply non-interest loan fees.

[225] $Fee\ Share_{it} = \frac{Fees\ and\ commissions_{it}}{Loan\ Interest_{it} + Fees\ and\ commissions_{it}}$

that non-interest income has always been a component of banks' loan-related income streams, which has not been the case for MFIs[226].

Figure 42 – Deciles of Banks' Fee and Commission Shares of Total Loan Income



Note: Colour banks provide decile ranges across all lenders' fee shares in a given quarter. The red band in the middle shows the 40th-60th percentile range, the dark orange the 30th-40th and 60th-70th, etc. The horizontal black line provides the median interest rate for all banks in a given quarter.

Source: BOZ Prudential Returns and Author's calculations

To quantitatively test whether the cap is associated with income switching within each bank we use a similar specification to the bank margin regression. We will estimate:

$$Y_{it} = \alpha_i + \sum_{l=1}^{L} \rho^l Y_{it-l} + \sum_{j=1}^{J} \beta^j X_{it}^j + \sum_{k=1}^{K} \theta^k M_t^k + \tau^{af} C_{2013} + \tau^{cf} C_{2014} + \varepsilon_{it} \qquad (7.10)$$

Where Y_{it} is the fee and commission share of loan-related income or the fee yield[227]. Otherwise, the regression is similar to equation (7.7) whereby the X_{it}^j and M_t^k are aiming to control for other factors that might impact fee shares. We will use both the pre-Q2 2014 and the full sample. If income

[226] This historic higher non-income share makes sense given our fee and commission data captures deposit related income (e.g., account ownership and non-FX account transaction fees), which has always been part of banks' businesses models.

[227] $Fee\ Yield_{it} = \frac{Fees\ and\ commissions_{it}}{Total\ Loans_{it}}$

switching were an important dynamic, we would expect that τ^{af} (and potentially also τ^{cf} in the larger sample) would be positive and significant. τ^{af} and τ^{cf} are telling us whether the level of the fee share or fee yield was different for each bank, on average, during the IRR regime periods than when it was not, ceteris paribus.

We use fees share and fee yield ratios rather than raw Kwacha fees or fee growth rates because we want to isolate fee adjustments from changes in banks' portfolio compositions or from growth trends in their portfolios more broadly; the ratio approach should make our fee variables agnostic to changes in banks' asset mixes[228].

We do not find strong evidence for the income-switching hypothesis (Table 31). There is some evidence that smaller banks' fee income shares increased while the cap was in place but loosened (Table 31, column 4, estimate row 7), but economically, the effect is very small. We estimate that fee shares were 1-5 percentage points higher at a 90% confidence level during the loose regime compared to when the bank was under no regime, which is not a large impact given the median bank's fee loan income share is 32%.

It is also plausible that banks' income switching might be related to their exposure to the IRR, in which case we wish to model this. We investigate this Annex 21 but find no consistent evidence of the more exposed lenders shifting income to non-interest fees. The same applies when we examine whether the imposition of the cap in 2013 was related to a change in fee share.

These results then support the general narrative from the previous section on bank margins: the caps may not have bound banks strongly enough for us to have picked up effects, or, quite simply, the effects may not exist. Let's also consider the bank types that could respond to a binding cap by increasing their fees. These may not actually be the small domestic or non-diversified commercial/business banks (with limited branch footprints or transactions accounts and who do not lend in high loan numbers). So, in a sense, the banks which could most realistically increase their fees are possibly the least likely to have been bound by this cap regime, and those which may have been bound couldn't use fees as a revenue offset. Our findings do not disprove the narrative in the

[228] Otherwise, when banks shift their asset allocations (say, towards non-loan assets in response to the cap), we might see a drop in the absolute level of fees even though the rate of fees per loan (or deposit) remained unchanged or increased.

cap literature that income-switching is prevalent. Still, it does suggest that if caps are set high enough, policymakers may not expect large adjustments in bank fees.

Table 31 – IRR Regime and Fee and Commission Shares of Banks' Loan Income

Dependent Variable	Sample	Coefficient	Static		Dynamic (Short-run effect)	
			OLS	Bank FE	OLS	Bank FE
Fee Share	Short Sample (All Banks)	τa, Cap2013	-0.0072	-0.0131	-0.0099	-0.017
		[se]	[0.0204]	[0.0200]	[0.0171]	[0.0132]
		rho			0.3482**	0.2696***
	All Banks (Full Sample)	τa, Cap2013	0.02	0.0144	0.0066	-0.0008
		[se]	[0.0237]	[0.0171]	[0.0091]	[0.0085]
		τc, Cap2014	0.0755***	0.0303	0.0194*	0.0128
		[se]	[0.0215]	[0.0186]	[0.0096]	[0.0086]
		rho			0.4623***	0.4362***
	Diversified (Full Time Period Sample)	τa, Cap2013	-0.001	-0.0198	-0.0062	-0.0115
		[se]	[0.0164]	[0.0173]	[0.0105]	[0.0114]
		τc, Cap2014	-0.0076	-0.0259	-0.0069	-0.0082
		[se]	[0.0205]	[0.0231]	[0.0131]	[0.0112]
		rho			0.4246*	0.3820***
	Not Diversified (Full Time Period Sample)	τa, Cap2013	0.013	0.0233	0.0129	0.0052
		[se]	[0.0298]	[0.0226]	[0.0121]	[0.0090]
		τc, Cap2014	0.1129***	0.0552*	0.0335*	0.0183*
		[se]	[0.0275]	[0.0259]	[0.0149]	[0.0098]
		rho			0.3767**	0.3481***
Fee Yield	Short Sample (All Banks)	τa, Cap2013	-0.0025	-0.004	-0.0033	-0.0031**
		[se]	[0.0047]	[0.0057]	[0.0025]	[0.0015]
		rho			0.6981***	0.7307***
	All Banks (Full Sample)	τa, Cap2013	-0.0004	-0.0006	-0.0006	-0.0001
		[se]	[0.0027]	[0.0021]	[0.0010]	[0.0008]
		τc, Cap2014	0.0054**	0.0018	0.0008	0.0002
		[se]	[0.0024]	[0.0021]	[0.0012]	[0.0009]
		rho			0.6876***	0.7091***
	Diversified (Full Time Period Sample)	τa, Cap2013	-0.0029*	-0.0047**	-0.0011**	-0.0011
		[se]	[0.0015]	[0.0018]	[0.0004]	[0.0008]
		τc, Cap2014	-0.0039**	-0.0052**	-0.0007	-0.0007
		[se]	[0.0019]	[0.0017]	[0.0012]	[0.0008]
		rho			0.6682***	0.6655***
	Not Diversified (Full Time Period Sample)	τa, Cap2013	-0.003	-0.002	-0.001	-0.001
		[se]	[0.0043]	[0.0033]	[0.0014]	[0.0012]
		τc, Cap2014	0.0120***	0.0043	0.0014	0.0004
		[se]	[0.0035]	[0.0040]	[0.0019]	[0.0012]
		rho			0.6335***	0.6424***

Note: Results from versions of equation 7.10. Each set of τa, Cap2013 and τc, Cap2014 provides policy coefficient estimate from a different regression. For example, the estimates-0.02 and 0.0755 from row 2 column 1 and row 3 column 1 are from the same static OLS regression using all banks and the "fee share" as the dependent variable. Results are provided for two different dependent variables: bank fee shares ($Fee\ Share_{it} = \frac{Fees\ and\ commissions_{it}}{Loan\ Interest_{it}+Fees\ and\ commissions_{it}}$) and bank fee yields ($Fee\ Yield_{it} = \frac{Fees\ and\ commissions_{it}}{Total\ Loans_{it}}$). "Static" results do not include a lagged dependent variable in the estimation. Covariates are included in all regressions and are consistent with the set used in Table 25. Only short run effects are shown for the dynamic estimates. Rho is the coefficient on the lagged dependent variable. τa, Cap2013 is the estimate for a dummy variable taking the value of 1 between Q1 2013 to Q2 2014, and 0 otherwise. τa, Cap2014 is the estimate for a dummy variable taking the value of 1 between Q3 2014 to Q4 2015, and 0 otherwise. The large "Diversified" Banks include the largest seven banks, all of whom have over 5% of market share (by loans and deposits) and maintain large branch networks. The "Not Diversified" banks make up the remainder. Standard errors robust used for static regressions and PCSE for dynamic. ***p<0.01, **p<0.05, *p<0.10.

Discussion and Conclusion

This Chapter analyses whether Zambia's IRR regime influenced banks' margins and interest income pricing structure. Across almost all the analyses, we did not find convincing and consistent evidence that the IRRs either led to lower/higher margins or resulted in banks shifting income into or away from other non-interest revenue lines. There was also little evidence that the more exposed banks reacted to the IRR regime more severely. Consequently, we cannot reject the null that the IRR regime had no effect.

There are several reasons we have not been able to identify an effect, a very plausible one being that our analyses were simply underpowered. Not only is our sample small, but if the significant mass of banks' loan portfolios were priced below the cap, we perhaps should not have expected the effect of the IRR regime to be very strong. Initially set at 9% above the BOZ rate, the IRR regime may have been high enough not to shift bank margins drastically. When Zambia's banking system went through stress in late 2014, and the IRR might have begun to bind, the cap was loosened to 16% above the BOZ rate, limiting the chances that loans were not written because of the cap. The effects may have been quite different in Kenya, where the cap was set at 4% above the base rate.

One issue with our null results is that the results may invite quite different interpretations depending on one's view on the structure and level of competition within Zambia's credit market. For policymakers who suspect that Zambia's banking sector is uncompetitive and extracting rents, it is possible that a tighter cap may have been needed to reduce banks' margins and gouging. Yet for those who suspect that Zambia's credit market is already competitive and instead worry about the adverse effects a cap might have on bank profitability and loan risk rationing, the cap attaching at 9% above the BOZ rate may have been high enough to avoid the worst of these outcomes[229]. It is also possible that IRRs both (i) reduced bank portfolio efficiency because of a restricted risk lending set, placing upward pressure on margins, as well as (ii) putting downward pressure on margins because banks were operating in a non-competitive loan market, all adding up to a negligible net effect. Unfortunately, our analysis and results cannot distinguish between any of these[230].

[229] There is, perhaps, some weak evidence to support caution on the effects of the caps on small banks, even with a 9% cap.
[230] Further, without having better interest rate data and an analysis of how the IRRs may have influenced lending volumes and loan risk, we have not touched upon a significant part of the IRR story.

Even with these null results, this Chapter still contributes to the IRR literature. The discussion in the Section Conclusion below helps place this result within the context of the rest of our analysis on interest rate caps. However, we are now able to be somewhat confident that Zambia's banks did not drastically adjust their pricing and fee practices for a cap regime imposed at 9% above the BOZ rate (although, again, our low power meant that we did not precisely estimate an effect around zero and hence claim a small effect).

IRR analysis aside, the Chapter has also offered some detailed analysis of bank margin determinants in Zambia, the first to do so. We did not find evidence that banks' maturity structures were an important margin determinant in Zambia, the first piece of analysis to examine this question in SSA[231]. In addition to finding that some factors driving large and small banks may differ, we also found that operating costs were likely a key determinant for interest rate margins, confirming a frequent result in the literature. If operating costs strongly drive margins, it suggests that structural market reforms aimed at improving bank lending and efficiency (such as improving the credit information environment, bankruptcy proceedings, and collateral formality and availability) may be a worthwhile policy option for reducing interest rates and bank margins, even if they are slower, less visible, and a less politically popular option than implementing IRRs.

Section Conclusion

IRRs are a controversial policy and for good reason. At the outset, policy objectives for IRRs can range from limiting usury, a common justification nowadays, to structural interest rate setting, which was the case for many countries in the 1970s and 1980s. In some low-income contexts, setting interest rate policy navigates an additional political tension: market-determined rates remain elevated and above what might be considered fair, and this is often within a context of historic financial exclusion from formal credit institutions. The theory for IRRs is also not unambiguous; policy outcomes can potentially be a function of the competitive context, the information environment, the cost structures in the financial system, banks' asset optimisations, and the list goes on.

[231] We suspect this may be due to the high proportion of floating rate loans in the market.

This thesis uses Zambia's IRR regime to try and place some empirical bounds – albeit small ones – on lender outcomes under an IRR. Zambia's regime was well suited for observational empirical analysis for two reasons: first, it offered some valuable dimensions of variation, including the fact that IRRs were applied at different rates to banks and non-banks, and that the non-bank market was characterised by a wide distribution of interest rates, business models, and market segments; and second, because there was also little evidence that the imposition of the IRR was expected, plausible identification was possible.

One unifying theme across all our results is that IRR "bindingness" matters, perhaps unsurprisingly. For example, in Chapter 5, which analysed Zambia's microfinance sector, we found that lenders who were previously pricing above the IRR saw much worse lending outcomes in the following periods, both compared to their own previous experiences (the DL modelling) or compared to unbound lenders (the event study approach). Additionally, the outcomes could be pretty severe for bound lenders: All the lenders who failed in the quarter after the cap were heavily bound, and we find that MFIs pricing above the cap by 35%, approximately the median, reduced their lending by 53% compared to those who were not bound. In our analysis of Zambia's banks' margins in Chapter 7, on the other hand, we didn't find strong effects on margins and loan income pricing. This null finding could be because of the offsetting effects of the cap or, more likely, that the bulk of bank lending was made below the IRR.

In theory, there are reasonable justifications for IRRs in alleviating market inefficiencies and imperfections[232]. Yet, these need to be weighed against the risks of misdiagnosing the market context and accepting that a common rate cap regime – a popular choice – may leave some sections of the credit market unviable[233]. All lenders, even if operating as a monopoly, have a lower bound failure interest rate/price based on the cost structures of lending to a given market segment/borrower. A cap set significantly below such a failure rate will almost certainly curtail such lenders or markets, independent of the market context. This likely occurred in the higher interest rate segments of Zambia's MFI market due to the IRR, as outlined in Chapter 5.

[232] There can also clearly be justifications based on usury.
[233] Again, the policy caveat suggests that if policymakers want to remove high-risk lending based on usury arguments, this might be the desired outcome and caps might work well at achieving this.

The challenge in low-income markets is that lending to some borrowers is inherently expensive. Only in some contexts will price controls likely result in the desired improvement in market access. It can be challenging for policymakers to differentiate high-interest rates resulting from pricing power versus those simply reflecting expensive loan production functions. In Zambia, several MFIs that saw significant loan retrenchment were NGOs, which were unlikely to have been using strongly predatory pricing. It is unclear whether BOZ undertook borrower-level market and cost analysis before implementing the cap on MFIs. Remember, too, that the counterfactual outcome for borrowers losing out on the extensive margin due to cap-induced rationing might not be that they don't end up borrowing but instead that they borrow in an unregulated and informal market. Lastly, a comment about credit market development: we do not have a strong understanding of the processes involved in building successful and competitive lending organisations that can help the financial market grow towards lower interest rates and increased credit access. However, we suspect that the policy uncertainty associated with the imposition and removal of the IRRs in Zambia (and the fact that several firms failed and many lost significant investor capital) is unlikely to have supported broader market development.

Finally, our results also highlight that imposing an IRR may result in unintended consequences. We find strong and robust evidence that non-banks shifted a significant proportion of their income into non-interest fees after the IRR was imposed; after five quarters, lenders increased income share towards non-interest fees by 20 percentage points on average. This is a worrying result if we believe non-interest fees can reduce pricing transparency. Even more concerning is that the effect may have been plastic: when the IRR regime was removed, the effect appears to have remained, reducing transparency in the market thereafter. We did not find such effects for banks, although they faced a relatively less severe cap regime and were already using fees and commissions before the IRR was implemented.

There are several important pieces of analysis relating to IRRs that we did not undertake in this thesis. We have not analysed how the IRR may have impacted bank lending volumes, nor how the IRR may have encouraged lenders, but especially banks, to shift the composition of their portfolios. If the IRR did bind at least some portion of their portfolios, it is plausible that they may have rationed risk rather than, or in addition to, lending. For example, anecdotal evidence from some market

participants suggests that banks shifted lending away from medium and small enterprise lending towards salary-backed loans in the wake of the cap regime (FSDA and Davel, 2013). Furthermore, with government yields in Zambia being relatively attractive, the cap may have facilitated a portfolio shift towards securities. Future analysis with better data could try to pin down these portfolio shifts more precisely. BOZ did not have lender-borrower-linked data in our sample period. However, future research on IRRs in low-income settings should (and has, see Aiba et al. (2020)) begun to use more granular data.

Chapter 8 – Conclusion and Discussion

This thesis has provided two case studies of financial regulation in Zambia: one case examined absolute capital requirements, and the other interest rate restrictions.

The imposition of both regulations reflected long-held perceptions in Zambia that banks and the financial system had not played the role they should have in supporting the country's development. In Chapter 3 we examined the imposition of absolute capital requirements that resulted in most banks in Zambia being forced to raise capital. We found that, although the forced equity injections did improve bank capitalisation, there was no clear association with an increase in bank lending, at least in the short term. Instead, banks appear to have increased their holdings of Treasury securities. Chapter 5 provides the first in-depth quantitative analysis of Zambia's IRRs, implemented in 2013. We found that Zambia's IRR regime likely had a detrimental and quite severe impact on non-bank lending, especially for lenders operating in higher interest rate market segments. We also found a shift towards non-banks charging additional fees, in addition to interest, to help mitigate the effects of the IRR. Using an international MFI dataset and a new dataset on cross-country IRR regimes, Chapter 6 suggests that MFIs facing an IRR reduced their lending (both in terms of loan volumes and loan numbers) and were less sustainable, weakly corroborating the findings from Chapter 5. Finally, Chapter 7 examines whether Zambia's IRR regime impacted banks' costs of financial intermediation and whether banks shifted from charging regulated interest payments to other unregulated fees and commissions. We do not find evidence for either, which we suspect is because too little bank lending was meaningfully exposed to the cap. Taken together, the chapters on IRRs suggest that policymakers must take cap bindingness seriously when diagnosing problems and designing any IRR policy.

In the rest of this Chapter, we look across both policies to discuss financial policymaking in Zambia and consider any broader learnings from this thesis. We concentrate on three areas: (i) policy challenges in encouraging credit market growth in a low-income, small-loan market economy, (ii) the unintended consequences of Zambia's regulations, and (iii) the limitations and possible extensions of the analysis in this thesis.

Policy challenges in encouraging credit market growth in a low-income, small-loan market economy

In the wake of Zambia's market-constraining financial reforms in the early 1970s, Harvey (1973) wrote on various challenges in setting financial policy in Zambia post-independence. One of the main contentions at the time (and one that has arguably been relevant since) was why lending was so low and whether foreign banks were specifically excluding Zambian borrowers. At the end of his paper, Harvey (1973) explores what he thought lay behind the poor credit access for Zambians and the expensive cost of borrowing in such markets:

> *"Many Zambians find it hard to get credit, not because they are Zambian, but because their enterprises are high risk and small. The risk reduces repayment rates, and the smallness makes for very high administrative costs. These separate problems can perhaps be tackled by policies which would normally be included under the credit label - for example, higher interest rates and administrative subsidies. But the problems of small businesses, and of small farms in particular, raise much broader questions." (Harvey 1973, p.391)*

Some of the key results in this thesis suggest that this perspective – that credit access is limited because of the expensive, higher-risk, and small-sized nature of lending – may still be relevant in Zambia some fifty years later. Borrowing rates remain expensive in Zambia to this day. Much like the market-constraining regulations in the 1970s, neither the absolute capital requirements documented in Chapter 3, nor the interest rate caps examined in Chapter 5, appear to have improved credit access, at least in the short term. Indeed, in the case of the IRRs, the effect was potentially the opposite.

Under some market structures, models, or contexts, similar financial policies might encourage credit and market growth. Indeed, in a broader development strategy sense, the historical record provides several examples of financially restrictive policies supporting market growth, for instance, in Korea (Cho, 2002; Demetriades and Luintel, 2001). Yet for the context of the regulations examined in this thesis in Zambia, we suspect such models, structures, or development strategies were not likely in place, and/or our time horizons were too short. Instead, we suspect that the lack of an effect (or evidence of a reverse effect) may have resulted from policymakers' challenges with diagnosis and

policy calibration or simply that other historical, political, or economic objectives were more strongly weighted.[235]

Take the case and context of the absolute capital regulations, for example. Zambian banks were very liquid before the policy was implemented, yet at that point, they still lent only a small portion of their assets, even though they had the liquidity to lend more[236]. In this context, it is not initially apparent why or how holding additional capital might make the marginal loan more attractive for banks or encourage them to lend more when they already had the opportunity and capacity to do so. Do not forget that the absolute capital rule was also implemented during "fair-weather conditions", not due to a banking or macroeconomic crisis, when lending might have been naturally less. We hence suspect that the constraint on lending was at least partly due to a lack of perceived bankable projects (and the structural challenges of the market) rather than absolute capital being a binding constraint on lending[237].

Zambia's interest rate regime in the non-bank market provides an even starker example. The lending retrenchment and firm failures for non-banks operating at interest rates over the imposed cap suggest that high-risk lending was simply not economically feasible given a 42% cap[238,239]. Any imperfections that might have existed in these markets and perhaps justified a price cap intervention were likely second order to the high-cost structures of providing small loans to borrowers in Zambia[240]. Our back-of-the-envelope profit decompositions in Box 1 show that, more-or-less, lenders pricing at higher interest rates faced breakeven interest rates that sat above where the cap was imposed, and they could not realistically adjust to it. This is something that policymakers might have identified with better market diagnostics before the policy was implemented.

[235] Even if there were imperfections in some parts of Zambia's credit market, say pricing power, for example, setting a cap regime to appropriately address this is another issue. In this thesis we saw how a single MFI cap rate was at odds with the cost structure realities of most of the market.
[236] Across the system, banks held approximately 50% of their total balance sheets in liquid non-loan assets at the end of 2011.
[237] As we discuss, the absolute capital regulations did achieve the BOZ's goal of increasing capitalisation rates and, with hindsight, did help banks navigate what was to be a tumultuous second half of 2010s.
[238] If BOZ's goal was to remove lending at high interest rates based on arguments of usury, the caps were arguably successful.
[239] Note that there was some gradation in the level of the cap imposed, however the cap on MFIs at 42% was the most significant.
[240] Note that the high lending cost in Zambia is likely because incomes can be low, collateral can be limited, lending can be geographically diverse (see Allen et al. 2011), and market infrastructure is lacking.

The analysis in this thesis thus highlights some subtler points for Zambian policymakers. First, it may be more difficult to indirectly coerce financial institutions into expanding their outreach in what has become a fundamentally a market-based financial system. Lenders are unlikely to make additional loans if it is not profitable for them to do so; expensive borrowers will remain expensive, independent of an interest rate cap or capital injection. This also likely makes sense if we believe that the binding constraint on bank lending is not their willingness to lend but a lack of what they see as bankable projects, given their cost economics, which is a more challenging margin to change and may require different policies.

One clear set of policy options would be to support market infrastructure to try and lower the cost of doing risky lending and limit market information imperfections. Such policies include improving credit infrastructure, such as credit data systems, insolvency resolution, and collateral registries[241]. These are perhaps less public and slower-acting policies, but they might more directly address structural constraints on growth. Another would be to directly subsidise high-risk lending groups or more directly support priority sector lending[242]. These may be appropriate if we believe that low incomes and small loan sizes are a constraint. We have not analysed such policies in this thesis; their effects will be context dependent. Yet, they are likely important options that should be considered in line with forced equity injections or interest rate restrictions.

In the Zambia sample, we also saw severe macroeconomic and financial market volatility, which harmed credit market growth and was at least partly driven by policymakers' decisions. The build-up of debt in Zambia from 2012 onwards was unsustainable, and the eventual default in 2020 has seriously harmed the economy, including the financial and credit markets. When considering what matters or has mattered for financial market growth, we must keep the macroeconomic environment at the forefront. Indeed, even if the policies evaluated in this thesis positively affected Zambia's credit markets, there is a good argument that they would likely have been relatively inconsequential compared to losses seen as a result of macroeconomic instability.

[241] In Zambia progress has more recently been made towards improving parts of this infrastructure. In 2017, BOZ and the IFC launched a collateral registry and in 2018 BOZ implemented a Credit Reporting Act. Since 2006 Zambia has had enabling legislation for a credit bureau, yet as recently as 2016 there were data challenges with banks and other lenders only strategically uploading data to the credit bureau and enforcement was not clearly a priority for BOZ.

[242] Zambia does have polices in this area, for example The Zambia Credit Guarantee Scheme (ZCGS), introduced in 2017. Over the years the IFC has also issued loans and bonds to Zambian banks and non-banks to facilitate MSME lending (Zanaco, 2010; Stanbic, 2019; Bayport, 2016; and others).

Financial regulation and unintended consequences

Whether Zambia's capital and IRR policies achieved their direct objectives or not, this thesis provides a reminder that unintended consequences are a risk when setting financial policy without a deep context-relevant empirical literature to guide its design. With recent international trends in IRR adoption, this is a particularly pertinent point. For both policies evaluated in this thesis, we do find evidence of unintended consequences, or at least outcomes that were likely not aligned with the Government or BOZ's initial objectives.

First, in the case of the interest rate restrictions, we saw a significant shift towards a mixed interest and fees pricing model in the microfinance market. Lenders bound by the IRR increased their non-interest income share cumulatively by 20 percentage points within five quarters after the imposition of the IRR, on average. Moreover, this effect was likely plastic, not returning when the IRR was removed. A higher fee market structure can reduce pricing transparency within the market, a negative outcome in and of itself (Ferrari et al. 2018).

Second, in the case of the absolute capital restrictions, we saw evidence of large banks appearing to have had much stronger asset growth reactions in response to the recapitalisations. The forced recapitalisations were associated with larger banks expanding further (across asset types and deposit financing), potentially taking on additional market share in the Zambian banking system. While we don't show that this growth is at the expense of smaller domestic banks or that market share has shifted, this risk likely influences market dynamics and future growth.

These results provide a brief reminder that policy may affect other outcomes than those we consider as our objectives. If anything, it advocates for additional case studies and analysis to support future policy decisions[243]. Before this thesis analysed non-bank fees in Zambia, there was very little quantitative analysis to highlight the risks to market pricing transparency under an IRR. Yet, our findings also lack external validity as our fee results are contingent on Zambia's market not having a strong history of fees pre-IRR, which is not the case for many countries.

The thesis perhaps also strengthens the need to get the initial policy diagnostics correct. The outcome trade-off between expected and unexpected consequences may still make good economics in some

[243] This is especially the case if the policy risks having plastic effects.

contexts. For example, some reduction in pricing transparency in the MFI market might be justified if monopoly pricing truly was going to be constrained by an IRR. Equally, the increase in capitalisation levels in Zambia may have justified a shift in market share towards larger banks. But both require getting the first-order effects right.

Thesis limitations and extensions

This thesis is primarily comprised of what we hope are contextual, reasonable, and careful empirical case studies examining understudied but important financial regulations in Zambia. Yet such analysis inherently comes with limitations that we now discuss below.

Most obvious, perhaps, is that observational studies such as these face challenges providing defendable identification. While there are strengths of observational analysis (they can examine market-wide and systematically important policies, as well as potentially endogenous market reactions), identification is always a challenge.

In the banking and finance development literature, identification often leans upon detailed datasets and clever econometric strategies. One example is Khwaja and Mian (2008)'s approach using borrower-fixed effects. Another is Agarwal et al. (2019) and Abuka et al. (2019)'s use of detailed credit registry data. Unfortunately, in Zambia (and still many other lower-income contexts) such data is not always available. Moreover, the lending types we are interested in do not always lend themselves to such approaches (for example, a borrower fixed effects approach relies on syndicated lending, while most personal and MSME borrowers will only borrow meaningfully from one lender). Frequently time-series data at the loan level is also unavailable because credit bureau data is low quality or non-existent, or the central banks only process aggregate accounting data. Nevertheless, these policies (like the IRRs and the absolute capital requirements) are important and must be analysed, despite the identification limitations.

The identification underpinning the key results in this thesis typically relied on assumptions based on panel fixed effects or some form of DID or event study. While such approaches have clear shortfalls and are a clear limitation of this thesis, there were several mitigating factors that support our analysis.

For the panel fixed effects analysis, the fact that the policy history, politics, and public record all support the plausible exogeneity of the policy impositions is important. In Chapter 3, the legacy nature of banks' lower capitalisation in 2012 lessens the likelihood of loan demand endogeneity. The "fair-weather" economic conditions during both policies strengthen our claim. The analysis in this thesis would read quite differently if the policies were implemented in response to a macroeconomic, capital, or political shock. We also partly address some of the endogeneity risks associated with bank-panel work by building datasets from detailed prudential data. The detailed covariate data help with the conditional mean error assumption and we are also able to use forward-looking NPL data to proxy for bank-specific demand shocks – see Annex 4. For the Zambia-specific DID IRR analysis, the lender-specific interest rate data allows us to build a reasonable, albeit imperfect, counterfactual group. Nevertheless, there still needs to be a sense of humility in our interpretations and any claims concerning causality.

Another limitation is policy interactions across our case studies, especially in the banking sector. In Chapter 3, we analyse the effect of the absolute capital requirements and banks' forced recapitalisations on capitalisation and asset outcomes. Yet by the beginning of 2013, when only 20% of bank recapitalisations by value had occurred, BOZ introduced the IRR regime, limiting bank lending to 18.25%[244]. Now, we suspect that most bank lending was not bound by the cap, and we are careful to analyse how the IRR regime affected the cost of bank financial intermediation, an outcome intuitively less sensitive to capital levels than bank lending. Yet, it would be naive to think that an IRR that might have bound banks' higher-risk lending would not also have reduced their likelihood to expand their lending post a capital injection. While we did not find evidence that the capital injections were associated with banks expanding their lending, this needs to be considered within a context whereby banks' lending options may have been constrained in higher-risk segments because of the IRR. Our finding that banks seem to have parked their additional capital in Treasury assets may make additional sense in this case, especially given the relatively short-run nature of our estimates and rising treasury yields leading up to late 2015. When we think about these policy interactions, they make our analysis in Chapter 3 additionally contingent on Zambia's specific context. In a financial policy context, we also get a feel for just how repressive or interventionalist

[244] 62% of recapitalisation by value occurred during 2013.

the regime may have been in 2012/2013: the Government first required banks to hold more capital. Then it constrained them from lending at high rates[245].

Finally, this thesis has taken a partial equilibrium view of the credit market, yet, in some cases, understanding aggregate effects may also be useful for policymakers. Both theorical frameworks in this thesis are partial equilibrium, as are the empirical bank and non-bank panel analyses. For the non-bank IRR analysis, a partial equilibrium approach is likely the main appropriate analysis frame given the degree of market segmentation across lenders. For the bank capital analysis, however, policymakers may very well be interested in the aggregate effects of the policy in addition to the bank-level estimates. For example, they may wish to know whether the recapitalisations were associated with additional holdings of Treasury securities or deposits in aggregate, or whether banks traded the same stock of Treasury assets or deposits within themselves. Bank-level analysis cannot provide a clear picture of this. We were tempted to attempt a synthetic control approach to compare Zambia's aggregate bank outcomes to its regional peers over the recapitalisation period (see Abadie et al. (2010)). Such an approach might help us better understand how the capitalisation regime impacted the banking sector, with a counterfactual based on other countries. Unfortunately, aggregate cross-country capitalisation data has surprisingly weak coverage in SSA, with many countries only having time series starting in 2010. This analysis would thus require additional data collection and cross-country capital accounting standardisation, pushing it past the scope of this thesis. But it is an option for future research and would contribute a nice extension to Chapter 3.

[245] In fairness, only a small proportion of loans was likely written by banks above the IRRs, yet the polices do reflect a concern from Government with how the credit market was operating.

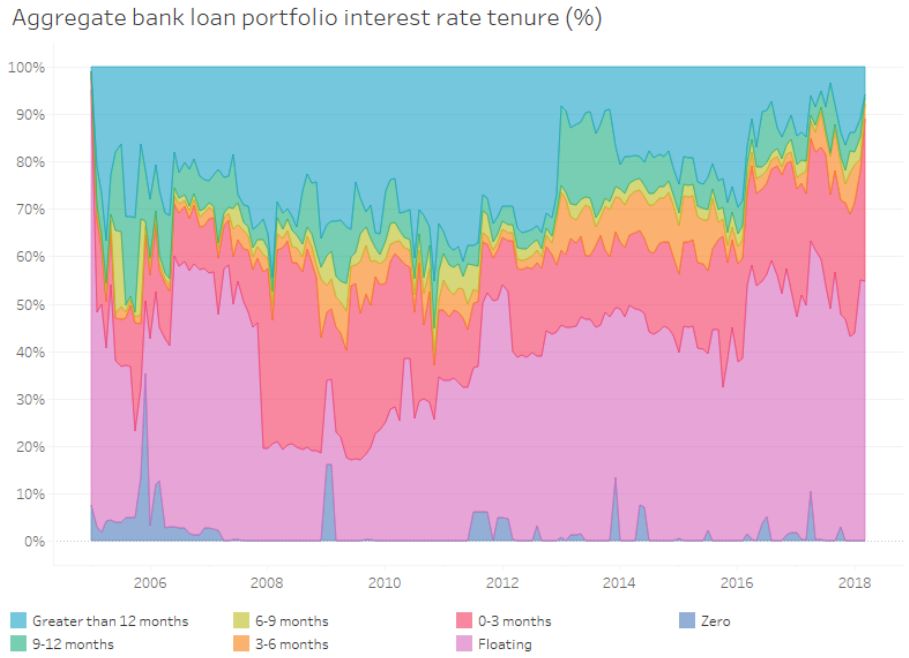
Annex 1 – Statistics on Zambia

Variable	2000	2010	2019
Pop (million)	10.4M	13.6M	17.8M
GDP const. (USD billion, 2010)	9.8B	20.2B	29.6B
GDP/cap (PPP)	1990	3125	3470
Agricultural share emp. (%)	70.1%	64.3%	48%
Inflation (%)	33%	14%	8%
Export share GDP (%)	24%	37%	35%
Copper share of exports (%)	51%	48.5%	48%
Tax Revenue (% GDP)	17%	12.8%	16.5%
External Govt Debt (% GNI)	168%	22%	120%

Source: World Bank WDI

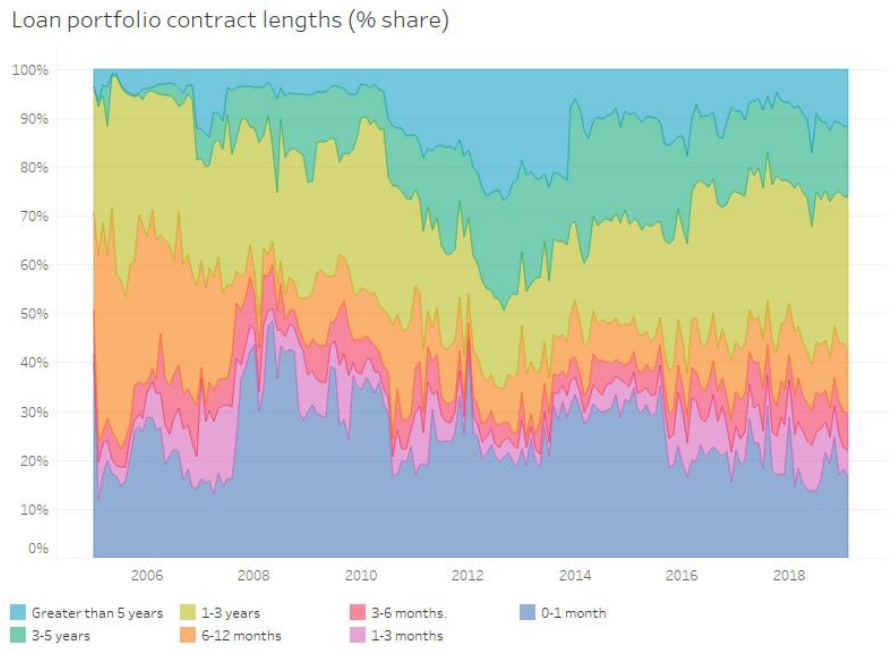
Annex 2 - Interest rate and Maturity Data

Figure 43 – Bank Loan Interest Rate Duration



Note: Figure provides the percent share of interest-rate tenures for the banking sector's aggregate loan portfolio in Zambia over time.

Figure 44 – Bank Loan Contractual Maturity



Note: Figure provides the share of contractual loan tenures for the Zambia's aggregate banking sector's loan portfolio over time.

Source: BOZ Prudential Returns

Annex 3 – Simple Theoretical Framework Extensions – Alternative Assumptions

This annex adjusts the baseline theoretical framework from the main body of the text for different market assumptions. The results complete the cells in Table 4.

Alternative (i) – Constant returns to borrower investments

In this scenario, we assume that the bank is a monopoly in its market segment, however we assume that the borrower has constant returns to investment/borrowing.

$$\max_{L_{ij}} \quad \pi_{ij} = (\bar{r}_{ij}).L_{ij} - r_{L_{ij}}.L_{ij} \tag{1}$$

The first-order condition for firms is:

$$\frac{d\pi_{ij}}{dL_{ij}} = \bar{r}_{ij} \; - \; r_{L_{ij}} = 0$$

Each firm will only accept a loan rate such that it will not make a loss. Profit per unit of loan is constant, but banks know that firms will accept a loan offer rate such that:

$$r_{L_{ij0}} = \bar{r}_{ij}$$

Solving out the bank's maximisation problem (and substitution in the deposit market constraint, the borrower's constraint, and the balance sheet identity, we find:

$$L_{ij} *= \frac{d_1}{2}.(\bar{r}_{ij} - \phi_{i1} + \phi_{i2}\bar{r}_{ij} \; - \overline{r_{D0}} - 2\tau) + \overline{E_{\iota 0}} + \frac{d_0}{2}$$

$$\frac{\partial L_{ij}}{\partial E_i} = 1$$

For the deposit market, the equilibrium is:

$$\frac{d_1.(\bar{r}_{ij} - \alpha E_{it} - \phi_{i1} + \phi_{i2}\bar{r}_{ij} - \overline{r_{D0}} - 2\tau) + d_0}{2} = D_{i0} *$$

$$\frac{\partial D_{ij}}{\partial E_i} = 0$$

So for firms with non-reducing returns to borrowing, we expect the loan market to expand 1:1 with the equity injection and for there to be no deposit offset. This is not surprising given the firms horizontal loan demand curve and the fact that they make zero economic profit: banks will want to use the new equity to finance their "on the margin" loans, whereas in the previous examples they would lose profit if all investment was passed on 1:1 to new loans post equity investment.

Alternative (ii) – Optimal leverage ratio

In this scenario we assume that the bank – for some exogenous and unmodelled reason – wishes to stay on its previous leverage ratio. This is a crude example because optimal leverage is almost certainly a function of the prevailing market structure and cost curves, and banks would unlikely stick to a specific leverage ratio if it were going to be expensive or very unprofitable to obtain. So, this scenario is better thought of as a boundary case.

We introduce the following simple leverage constraint into the baseline bank optimisation:

$$c = \frac{L}{E} = \frac{D + E}{E}$$

The Lagrangian is hence:

$$\mathcal{L}(L, D, \lambda) = L_{ij0}(\bar{r}_{ij} - \alpha.2L_{ji} - (\phi_{i1} - \phi_{i2}\bar{r}_{ij})) - \left[\frac{L_{ijt} - \overline{E_{i0}} - d_0}{d_1} + \overline{r_{D0}}\right].(L_{ij} - E_i) - \tau(L_{ij} - E_{it} + L_{ij}) + \lambda(c - \frac{L}{L - (L - E)})$$

Assuming that the bank is constrained by the ratio post injection, we find that:

$$\frac{dL * (constrained)}{dE} = c \qquad c > 0$$

$$\frac{dD * (constrained)}{dE} = c - 1 \quad c > 0$$

Predictably, this is a much stronger result than the previous scenarios. Banks will increase their deposits c-1 times for each unit of additional equity and expand their balance sheets by c times (in this case into lending, the only modelled asset). This is a quantitatively large result, but it is less plausible for a small market like Zambia's where there is a limit on individual savings (deposits) and lending quality does likely decline quite quickly – it is unlikely profitable for banks to expand their size by E.c post equity injection.

Alternative (iii) – Competitive Loan Market

In this section we assume that there is more than one bank in each market and instead the loan market clears where loan demand equals loan supply.

Demand for loans can be given by:

$$r_{L_{ij}} = \bar{r}_{ij} - \alpha.2L_{ij}$$

The bank's supply curve can be given by:

$$r_{Lij} = \phi_{i1} - \phi_{i2}\bar{r}_{ij} \; - \frac{-2.L_{ij} + 2\overline{E_{i0}} + d_0}{d_1} + \overline{r_{D0}} + 2\tau$$

Setting these equal and solving for the loan market outcome gives:

$$L_{ij} *= \frac{1}{\alpha.2} .\bar{r}_{ij} - \frac{1}{\alpha.2} . \frac{d_1\left(\phi_{i1} - \phi_{i2}\bar{r}_{ij} \; - \left(\frac{2\overline{E_{i0}} + d_0}{d_1}\right) + \overline{r_{D0}} + 2\tau\right) + \frac{1}{\alpha} .\bar{r}_{ij}}{(d_1 + \frac{1}{\alpha})}$$

$$\frac{\partial L_{ij*}}{\partial E_{i0}} = \frac{\partial}{\partial E_{i0}} \frac{E}{1 + d_1\alpha} = \frac{1}{1 + d_1\alpha} > 0$$

For the deposit market:

$$D_{i0} * \frac{d_1 . \left(\frac{d_1\left(\phi_{i1} - \phi_{i2}\bar{r}_{ij} - \left(\frac{2\overline{E_{i0}} + d_0}{d_1}\right) + \overline{r_{D0}} + 2\tau\right) + \frac{1}{\alpha} . \bar{r}_{ij}}{(d_1 + \frac{1}{\alpha})} - \phi_{i1} + \phi_{i2}\bar{r}_{ij} - \overline{r_{D0}} - 2\tau\right) + d_0}{2}$$

$$\frac{\partial D_{ij}}{\partial E_{i0}} = \frac{\partial}{\partial E_{i0}} . -\frac{\alpha d_1 E_{i0}}{\alpha d_1 + 1} = -\frac{\alpha d_1}{\alpha d_1 + 1} = \frac{1}{\alpha d_1 + 1} - 1 < 0$$

This is a similar result to the monopoly case, but the offset will be more negative.

Annex 4 – Sector-specific firm shocks and demand controls based on bank sector exposures and non-performing loan accounts

This annex discusses and constructs sector-specific demand controls based on bank sector exposures and non-performing loan accounts.

Banks in our dataset can have quite different loan sector concentrations, and each sector can potentially be exposed to its own shocks. A relevant example in Zambia of such a sector-specific shock was the poor rainy seasons in 2018-2019, which resulted in disrupted farm yields and load shedding (due to low water levels in Lake Kariba). Both farmers and manufacturing firms suffered, but banks' heavy lending to these sectors was also likely impacted, resulting in them writing down loans and collateral and suffering losses. As with the analysis of economy-wide shocks, banks that were exposed to these sectors potentially saw both a reduction in loan demand, while at the same time experiencing reduced equity and loan supply. Unfortunately, in this case, time fixed effects will not absorb such exposure heterogeneity that varies both across banks and over time.

Fortunately, there are some strategies in the empirical bank capital literature on how to address these sector-level shocks. Aiyar et al. (2014) create a control to capture sectorial variation in loan demand (i.e. sector-specific loan demand shocks) by creating a bank-specific measure of demand, calculated as the sum-product of each bank's sector exposure weights and employment growth in that particular sector. Fang et al. (2020) use a similar approach in their analysis of Peru's capital regime, but with two alterations: first, they use sectorial GDP growth instead of sectorial employment growth, and second, they use the bank's sector exposures at the beginning of their sample to weight the sectorial GDP growth, rather than a time-varying measure of bank loan sector exposure weights.

As mentioned previously, Zambian macroeconomic data series for our sample do not allow for such an approach. While we have data on each bank's sector lending exposure, employment growth data for the sectors does not exist. Instead, we propose alternative controls based on new additions to banks' non-performing loan accounts.

Our main bank-level demand control is calculated thus,

$$z_{it} = \sum_{q} z_{iqt}$$

$$where, \qquad z_{iqt} = \frac{\sum_{k=3}^{5} s_{iqt} . \frac{NPL_{q,t+k}^{90-119}}{L_{qt}}}{3}$$

where z_{it} is the demand measure for bank i at time t, s_{iqt} is the bank's lending exposure in sector q at time t (i.e. the share of that specific sector in its total loan portfolio), and $NPL_{q,t+k,q,t+k+1}^{90-120}$ measures the absolute growth in the market aggregate of the non-performing loans stock in the 90-120 day category for sector q, over the three months before the loan entered/was recorded as a non-performing loan ("NPL"). There are 12 economic sectors. The main assumption/idea is that appropriately recognised changes in recently incurred aggregate sector NPLs are an acceptable measure of sector-wide firm shocks, and hence can help control for variation in banks' sector-specific loan demand. Because the NPLs are only recognised 90 days after loans default, future NPL data may provide information about the demand conditions now. Growth in the 90-120-day category is averaged over three months to help smooth recognition and timing issues. The deciles of the variable for banks over time appears to pick up the key events in the lending market (Figure 14, where each coloured bar represents a decile). The weakness of this approach is that borrowers likely draw down on reserves before they fail to pay their bank back when they suffer a shock, but we are only picking up the timing of the latter. It also assumes that the market-wide change in NPLs in a given sector is an appropriate measure for all borrowers in that sector.

Graphically, a summary of the timing assumptions is shown below. Figure 45 shows the deciles of the measure across banks. Figure 46 shows the sum of these individual bank measures by sector.

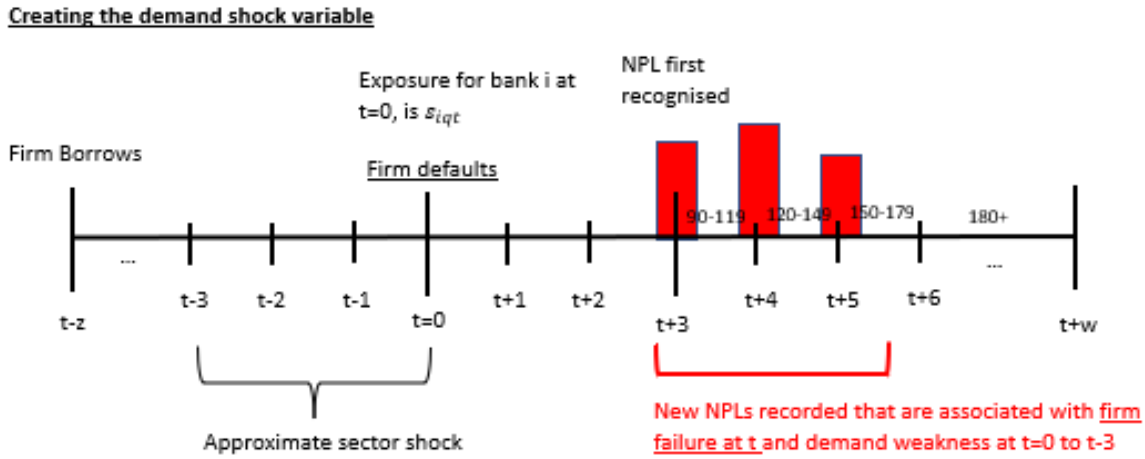
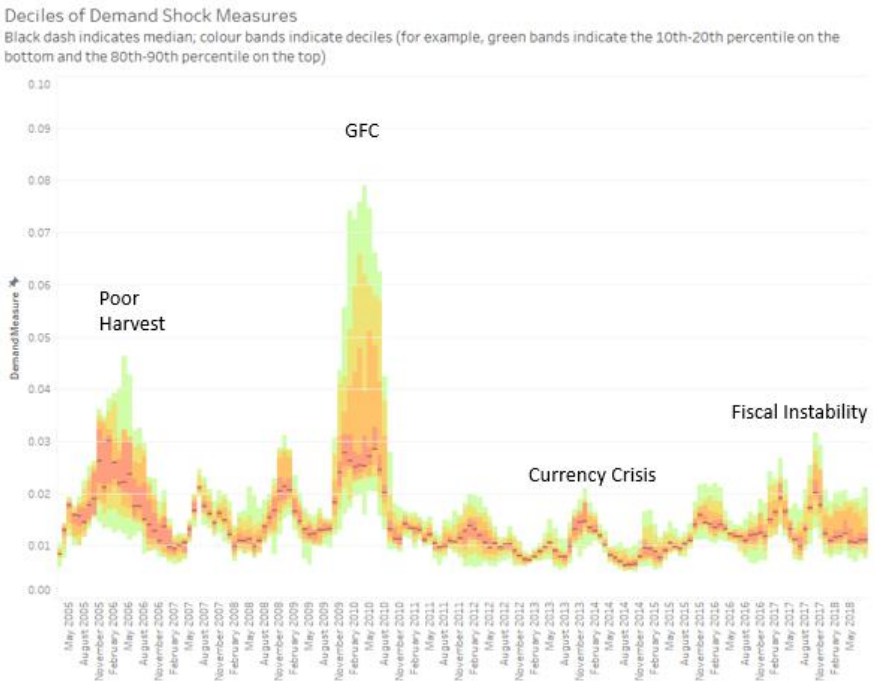


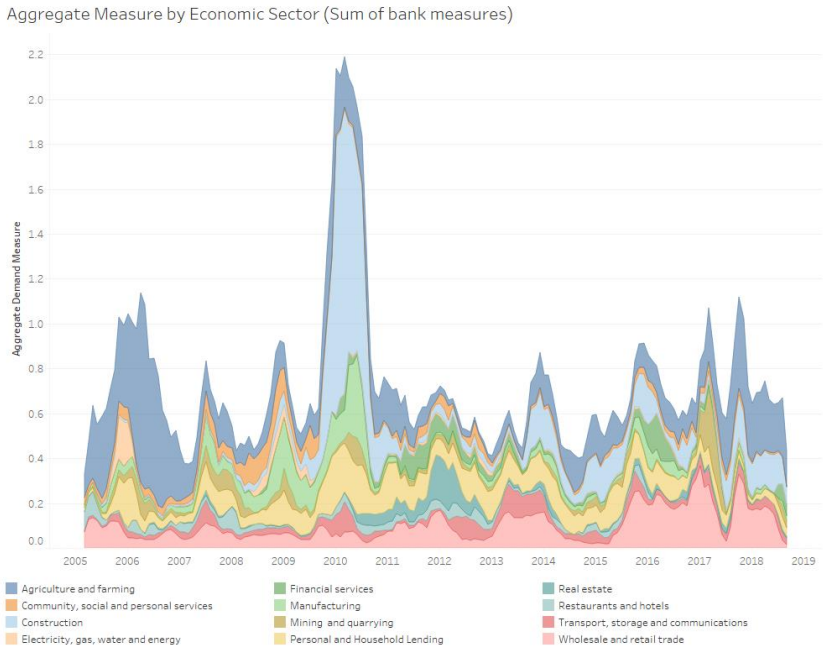
Figure 45: Decile Ranges of Demand Shock Measure



Note: The horizontal black line provides the median Demand Shock Measure for Zambia's banks in a given quarter. The colour bands provide decile ranges for banks' Demand Shock Measures. The red band in the middle, for example, shows the 40th-60th percentile range; the dark orange, the 30th-40th and 60th-70th percentiles, etc.

Source: BOZ Prudential Data

Figure 46



Note: The figure shows the aggregate of all banks' Demand Shock Measures over time, but separated by economic sector. Source: BOZ Prudential return.

Annex 5 - Calculating and using adjusted lending flows and making foreign exchange adjustments

Empirical studies often use the outstanding stock of bank loans as a measure of bank lending. In many cases, this is a practical constraint: the stock of loans is an easily available measure from published financial statements. As argued in Bridges et al. (2014), however, the stock of loans can be an imperfect measure of bank lending and decision-making, because it may reflect changes in write-offs and reclassifications, exchange rate effects (to both the outstanding loan portfolio and the magnitude of recorded non-performing loans), interest-rate revaluations, and more[246]. In an economic sense, we want to capture how exogenous changes in capital influence banks' decisions to lend (or take on deposits), and not other short-term factors outside of banks' immediate control. We are also able to see in the simple econometric model how loan write-offs may induce a contemporaneous fall in a bank's capital and its outstanding loan portfolio, inducing a spurious correlation and bias.

We follow Bridges et al. (2014), where possible, and construct adjusted lending flows, reflecting, as close as possible, loan portfolio 'transactions'. We aim to calculate:

$$\Delta L_{it} = \underbrace{L_{it} - L_{it-1} - (writeoff_{it} - recoveries_{it}) + fx.adj_{it} + other.adj_{it}}$$

Loan measure used in this paper

Where $L_{it}$is the outstanding loan portfolio measured in the balance sheet for bank i at time t, $writeoff_{it} - recoveries_{it}$ captures bank-specific net loan write-offs (which will induce a spurious correlation between capital and lending if left unadjusted), $fx.adj_{it}$, discussed in detail below, attempts to isolate and adjust for foreign exchange gains and losses on the loan portfolio. The foreign exchange adjustment is particularly important given the high level of dollarisation in the Zambian banking market.

[246] Take the situation whereby a bank cycles through writing off and then re-recognising a loan in its portfolio. This would result in loan growth/decline related not to the bank's lending decision-making (although they may be linked) and instead to accounting decisions.

Unfortunately, we are not able to make all the adjustments that we would like. For example, we cannot capture non-foreign exchange valuation adjustments (e.g., interest rate revaluations), as done in Bridges et al. (2014), because the BOZ data return aggregates all asset class revaluations into a single adjustment entry. We can also only make the net write-off and foreign exchange adjustment for loan analysis at the bank level, rather than the loan classification level (i.e. individually for the corporate/household loan books) or for other assets types. So, we will not be able to use "adjusted" variables for all our regressions.

Currency Exposure Adjustments

As discussed in the text, Zambia's banking system is highly dollarised, meaning that changes in exchange rates may lead to mismeasurement of bank's true lending or deposit growth. As such, what might appear initially to be lending growth/contraction by banks may potentially be driven by currency movements.

This annex outlines an attempt to make some simple adjustments for bank loan and deposit portfolios with foreign exchange exposures. The goal is to isolate changes in the outstanding portfolio that are directly linked to foreign exchange revaluations such that they are not confused in the analysis with those that are related to the banks' actively undertaking new lending or deposit-taking transactions. These adjustments are especially necessary in Zambia because of the coupling of a highly dollarised banking system and the relatively high volatility of the Kwacha.

The Bank of Zambia does did not collect loan disbursement data for our sample period, which we would otherwise be able to use to isolate new lending form exchange rate effects. So, we need to try and back out foreign exchange effects from banks' new lending given the dataset we have. The results will be imperfect, but likely still more accurate than not making any adjustments.

We use a simple example of a bank that makes both domestic and foreign lending, in two periods, to help frame our analysis. Banks begin with a portfolio of domestic and foreign assets, which we can measure directly and in the local currency:

$$L_1^T = L_1^D + L_1^F . f_1$$

Where L_1^T is the total loan portfolio denominated in the local currency, L_1^D is the outstanding lending stock at time 1, L_1^F is the outstanding foreign portfolio loans denominated in foreign currency, and f_1 is the spot exchange rate at time 1.

For the next observation in the dataset, which we can also measure, banks hold a total portfolio such that:

$$L_2^T = L_2^D + L_2^F . f_2$$

Between each period, we assume banks increase (or reduce) their foreign-denominated lending by some amount, captured by L_t^F. When the bank enters such transactions, the exchange rate is f_t. We can measure neither L_t^F or f_t directly in our dataset. The value of the portfolio immediately after the transaction is given by:

$$L_t^T = L_1^D + L_1^F . f_1 . \left(\frac{f_t}{f_1}\right) + L_t^F . f_t$$

Up to the point when the new transaction is made, the foreign-denominated portfolio will appreciate/depreciate by a factor $\left(\frac{f_t}{f_1}\right)$.

At the end of the period, the value of the portfolio is given by:

$$L_2^T = L_2^D + L_1^F . f_1 . \left(\frac{f_t}{f_1}\right) . \left(\frac{f_2}{f_t}\right) + L_t^F . f_t . \left(\frac{f_2}{f_t}\right)$$

Which also clearly simplifies to:

$$L_2^T = L_2^D + (L_1^F + L_t^F) . f_2$$

Note, the value of the new lending, L_t^F, can also appreciate/depreciate before the end of the quarter. Rearranging and isolating L_t^F.

$$L_2^T - L_2^D = (L_1^F + L_t^F) . f_2$$
$$\frac{L_2^T - L_2^D}{f_2} = (L_1^F + L_t^F)$$
$$\frac{L_2^T - L_2^D}{f_2} - L_1^F = L_t^F$$

We now need to back-out the foreign exchange adjustment for our loan variable. We calculate the change in the total portfolio, and then rearrange to isolate any foreign exchange effects:

$$\Delta L_2^T = L_2^D - L_1^D + L_1^F . f_1 . \left(\frac{f_t}{f_1}\right) . \left(\frac{f_2}{f_t}\right) + L_t^F . f_t . \left(\frac{f_2}{f_t}\right) - L_1^F . f_1$$

$$\Delta L_2^T = [L_2^D - L_1^D] + L_1^F . f_1 \left(\frac{f_2}{f_1} - 1\right) + L_t^F . f_t . \left(\frac{f_2}{f_t}\right)$$

$$\Delta L_2^T = [L_2^D - L_1^D] + L_1^F . f_1 \left(\frac{f_2}{f_1} - 1\right) + L_t^F . f_t . \left(\frac{f_2}{f_t} - 1\right) + L_t^F . f_t$$

$$\Delta L_2^T = [L_2^D - L_1^D] + L_1^F . (f_2 - f_1) + L_t^F . (f_2 - f_t) + L_t^F . f_t$$

The exchange rate adjustment is hence:

$$L_1^F . (f_2 - f_1) + L_t^F . (f_2 - f_t) \quad (1)$$

And the true lending change is:

$$\Delta L_2^T = [L_2^D - L_1^D] + L_t^F . f_t$$

The challenge we face is that we do not have direct data on either L_t^F or f_t. However, as mentioned above, we may be able to make assumptions on L_t^F or f_t which will allow us to make a sensible foreign exchange appreciation adjustment.

Given the data we do have and our assumptions concerning new foreign lending, we can calculate:

$$L_t^F = \frac{L_2^T - L_2^D}{f_2} - L_1^F$$

The above expression only requires data that we observe. We need to make a strong assumption regarding f_t because we do not know when or at what rate banks made (or withdrew) additional foreign-denominated loans. The best assumption we can probably make is to assume that the bank made transactions at the average foreign exchange rate over the period.

Which can then be used with the assumption that f_t is the average daily exchange rate for the period, to calculate:

$$\Delta L_2^T = [L_2^D - L_1^D] + (\frac{L_2^T - L_2^D}{f_2} - L_1^F) . f_{ave}$$

Because our raw balance sheet data is at a monthly frequency, we are able to make adjustments at a monthly frequency and then aggregate them up to the quarterly dataset used in the final analysis and improve precision.

The final exchange rate adjustment in measured variables is hence:

$$L_1^F . (f_2 - f_1) + \left(\frac{L_2^T - L_2^D}{f_2} - L_1^F\right)(f_2 - f_{ave}) \quad (2)$$

Annex 6 – Bank capitalisation and bank recapitalisations – winsorised and TWFE

Table 32 - Risk Weighted Assets and Bank Recapitalisation (TWFE)

Notes: This table presents estimates regressing banks' risk weighted asset ratios (the dependent variable, in percentage points) against a recapitalisation variable and its various lags. The recapitalisation variable is the ratio of equity raised by a bank in deficit under the absolute minimum capital rule in a quarter divided by the bank's total assets at the beginning of that quarter. The "Cumulative Effect" row calculates the sum of all the estimate lag coefficients included in that regression. The "P-value (H0: Cumulative Effect=0)" column is testing the hypothesis that the linear combination of the regulatory deficit variable and its lags are significantly different to zero. The regressions are approximately based on equation (10) in the text. Each column represents a regression using a different estimation approach (OLS = Ordinary Least Squares, Within = Bank Fixed Effects, Time = Time Fixed Effects, TWFE = Two-way Fixed Effects) and a different lag length (for example, (0,4) indicates that lags from 0 to 4 are included). Data is quarterly and was constructed from the Bank of Zambia's monthly Prudential Return. Bank-Specific controls include the following variables (for two lags, including t-1 and t-2): log of total bank assets, return on assets, liquid assets as a percent of total assets, RWAR ratio, a sector-specific loan measure, provisions for the quarter as a percent of beginning total assets. Macro variables include the following (for two lags, including t-1 and t-2): the statutory reserve ratio, quarterly growth in the exchange rate, the average inflation rate over the quarter, and the average Treasury Bill Yield (365 day) over the quarter. Driscoll Kraay or Robust Errors were used and are provided in parentheses. Data are not winsorised. *** p < 0.01, ** p < 0.05, * p < 0.1.

Dependent Variable	Change in Risk-Weighted Asset Ratio (Percentage Points)											
Lags (Beg, End)	(0,0)		(0,2)		(0,3)		(0,4)		(0,5)		(0,6)	
Regression Type	OLS	TWFE	OLS	TWFE	OLS	TWFE	OLS	TWFE	OLS	TWFE	OLS	TWFE
Recapitalisation Variable	1.7473***	1.7217***	1.6823***	1.7400***	1.6732***	1.7287***	1.7081***	1.6555***	1.6416***	1.6949***	1.6392***	1.6929***
	[0.3914]	[0.4027]	[0.4134]	[0.1312]	[0.4124]	[0.4134]	[0.4183]	[0.4135]	[0.4057]	[0.4085]	[0.4106]	[0.4104]
Recapitalisation Variable (t-1)			-0.2277	-0.2353*	-0.2458	-0.2528	-0.2564	-0.2607	-0.2582	-0.2593	-0.2688	-0.2604
			[0.2240]	[0.1315]	[0.2315]	[0.2252]	[0.2261]	[0.2293]	[0.2262]	[0.2250]	[0.2234]	[0.2254]
Recapitalisation Variable (t-2)			0.0262	0.021	0.0311	0.0275	0.02	0.0141	0.0065	0.0106	0	-0.0019
			[0.0688]	[0.1319]	[0.0744]	[0.0816]	[0.0794]	[0.0751]	[0.0763]	[0.0813]	[0.0758]	[0.0833]
Recapitalisation Variable (t-3)					-0.1511	-0.1492	-0.1542	-0.1666	-0.1659	-0.1556	-0.1692	-0.1567
					[0.1350]	[0.1280]	[0.1195]	[0.1246]	[0.1256]	[0.1220]	[0.1242]	[0.1190]
Recapitalisation Variable (t-4)							-0.0392	-0.0276	-0.0518	-0.055	-0.0619	-0.0722
							[0.0956]	[0.1052]	[0.1050]	[0.0963]	[0.0976]	[0.0848]
Recapitalisation Variable (t-5)									0.0903	0.035	0.0938	0.0432
									[0.0735]	[0.0649]	[0.0690]	[0.0640]
Recapitalisation Variable (t-6)											-0.0935	-0.1222
											[0.0839]	[0.0882]
Cumulative Effect	**1.747**	**1.722**	**1.481**	**1.526**	**1.307**	**1.354**	**1.278**	**1.215**	**1.262**	**1.271**	**1.140**	**1.123**
P-value (H0: Cumulative Effect=0)	**0.000**	**0.001**	**0.000**	**0.000**	**0.001**	**0.000**	**0.001**	**0.003**	**0.006**	**0.002**	**0.011**	**0.003**
Observations	746	712	714	712	698	696	680	682	666	664	650	648
R-Squared	0.249	0.297	0.278	0.3	0.287	0.308	0.326	0.311	0.311	0.327	0.314	0.33
Bank-Specific Controls	X	√	X	√	X	√	X	√	X	√	X	√
Bank Fixed Effect	X	√	X	√	X	√	X	√	X	√	X	√
Time Fixed Effects	√	√	X	√	√	√	√	√	√	√	√	√
Macro Controls	X	X	X	X	X	X	X	X	X	X	X	X

Annex 7 – Alternative Shock Variables (Equity Growth and RWAR shock) and winsorised Data

Table 33 - Asset Growth and Bank Recapitalisation using equity growth recapitalisation variable

Notes: This table presents estimates from regressing the growth rates for different types of bank assets (the dependent variable, in percent) against a recapitalisation variable and its various lags. The recapitalisation variable is the growth rate in equity (but specifically only equity raised by a bank in deficit under the absolute minimum capital rule in a quarter). The "Cumulative Effect" row calculates the sum of all the estimate lag coefficients included in that regression. The "P-value (H0: Cumulative Effect=0)" column is testing the hypothesis that the linear combination of the regulatory deficit variable and its lags are significantly different to zero. The regressions are approximately based on equation (11) in the text. Each column represents a regression using a different estimation approach (OLS = Ordinary Least Squares, Within = Bank Fixed Effects, Time = Time Fixed Effects, TWFE = Two-way Fixed Effects) and a different asset type (the set includes the following: Total Assets, Bank Loans, Treasury Assets, and Other Assets). Data is quarterly and was constructed from the Bank of Zambia's monthly Prudential Return. Bank-Specific controls include the following variables (for two lags, including t-1 and t-2): log of total bank assets, return on assets, liquid assets as a percent of total assets, RWAR ratio, a sector-specific loan measure, provisions for the quarter as a percent of beginning total assets. Macro variables include the following (for two lags, including t-1 and t-2): the statutory reserve ratio, quarterly growth in the exchange rate, the average inflation rate over the quarter, and the average Treasury Bill Yield (365 day) over the quarter. Driscoll Kraay or Robust Errors were used and are provided in parentheses. Data are not winsorised. *** p < 0.01, ** p < 0.05, * p < 0.1.

Dependent Variable	Percentage Growth in Asset Type (Percent)											
	Total Assets			Loans			Treasury Assets			Other Assets		
Regression Type	OLS	Within	TWFE	OLS	Within	TWFE	OLS	Within	TWFE	OLS	Within	TWFE
Recapitalisation Variable	0.0608***	0.0667***	0.0653***	-0.030	-0.030	-0.0387*	0.1643**	0.1600**	0.1680**	0.0710***	0.0857***	0.0780***
	[0.0081]	[0.0060]	[0.0065]	[0.0193]	[0.0193]	[0.0203]	[0.0725]	[0.0729]	[0.0793]	[0.0239]	[0.0189]	[0.0167]
Recapitalisation Variable (t-1)	-0.0072	-0.0115	-0.0169	-0.0422	-0.0422	-0.0371	0.0944	0.11	0.1039	-0.0147	-0.0175	-0.0217
	[0.0117]	[0.0122]	[0.0115]	[0.0373]	[0.0373]	[0.0377]	[0.0719]	[0.0828]	[0.0936]	[0.0260]	[0.0259]	[0.0276]
Recapitalisation Variable (t-2)	0.0076	0.0043	0.0033	0.0349	0.0349	0.0261	-0.0107	-0.0029	0.0046	0.0335	0.0206	0.0182
	[0.0128]	[0.0119]	[0.0110]	[0.0328]	[0.0328]	[0.0325]	[0.0136]	[0.0176]	[0.0148]	[0.0220]	[0.0250]	[0.0280]
Recapitalisation Variable (t-3)	-0.0055	-0.009	-0.0083	0.0076	0.0076	0.0008	-0.0440**	-0.0468	-0.0393	-0.0115	-0.0134	-0.0027
	[0.0085]	[0.0063]	[0.0069]	[0.0191]	[0.0191]	[0.0201]	[0.0168]	[0.0332]	[0.0283]	[0.0110]	[0.0111]	[0.0119]
Recapitalisation Variable (t-4)	-0.0002	-0.004	-0.0007	0.0281	0.0281	0.0193	-0.0052	0.0099	0.0098	-0.0152*	-0.0409	-0.0373
	[0.0086]	[0.0050]	[0.0048]	[0.0280]	[0.0280]	[0.0300]	[0.0196]	[0.0109]	[0.0113]	[0.0081]	[0.0247]	[0.0270]
Recapitalisation Variable (t-5)	-0.0088	-0.0210*	-0.0217**	-0.0116	-0.0116	-0.0105	0.0142	0.0093	0.0029	-0.0257*	-0.0474**	-0.0407**
	[0.0053]	[0.0106]	[0.0104]	[0.0093]	[0.0093]	[0.0072]	[0.0178]	[0.0159]	[0.0180]	[0.0137]	[0.0208]	[0.0193]
Cumulative Effect	**0.047**	**0.026**	**0.021**	**-0.013**	**-0.013**	**-0.04**	**0.213**	**0.239**	**0.25**	**0.037**	**-0.013**	**-0.006**
P-value (Cumulative effect=0)	**0.160**	**0.346**	**0.401**	**0.859**	**0.859**	**0.617**	**0.016**	**0.004**	**0.013**	**0.438**	**0.817**	**0.912**
Observations	666	664	664	664	664	664	645	643	643	666	664	664
R-Squared	0.04	0.26	0.34	0.09	0.09	0.16	0.07	0.18	0.24	0.02	0.15	0.24
Bank-Specific Controls	X	√	√	X	√	√	X	√	√	X	√	√
Bank Fixed Effect	X	√	√	X	√	√	X	√	√	X	√	√
Time Fixed Effects	X	X	√	X	X	√	X	X	√	X	X	√
Macro Controls	X	√	X	X	√	X	X	√	X	X	√	X

Note: Equity Growth Recapitalisation Variable calculated as $\frac{\Delta def_{t-k}^{i}}{E_{t-k}}$

Table 34 - Asset Growth and Bank Recapitalisation using RWAR growth recapitalisation variable

Notes: This table presents estimates from regressing the growth rates for different types of bank assets (the dependent variable, in percent) against a recapitalisation variable and its various lags. The recapitalisation variable is the growth rate in the RWAR (but specifically only for equity raised by a bank in deficit under the absolute minimum capital rule in a quarter). The "Cumulative Effect" row calculates the sum of all the estimate lag coefficients included in that regression. The "P-value (H0: Cumulative Effect=0)" column is testing the hypothesis that the linear combination of the regulatory deficit variable and its lags are significantly different to zero. The regressions are approximately based on equation (11) in the text. Each column represents a regression using a different estimation approach (OLS = Ordinary Least Squares, Within = Bank Fixed Effects, Time = Time Fixed Effects, TWFE = Two-way Fixed Effects) and a different asset type (the set includes the following: Total Assets, Bank Loans, Treasury Assets, and Other Assets). Data is quarterly and was constructed from the Bank of Zambia's monthly Prudential Return. Bank-Specific controls include the following variables (for two lags, including t-1 and t-2): log of total bank assets, return on assets, liquid assets as a percent of total assets, RWAR ratio, a sector-specific loan measure, provisions for the quarter as a percent of beginning total assets. Macro variables include the following (for two lags, including t-1 and t-2): the statutory reserve ratio, quarterly growth in the exchange rate, the average inflation rate over the quarter, and the average Treasury Bill Yield (365 day) over the quarter. Driscoll Kraay or Robust Errors were used and are provided in parentheses. Data are not winsorised. *** p < 0.01, ** p < 0.05, * p < 0.1.

Dependent Variable	Percentage Growth in Asset Type (Percent)											
	Total Assets			Loans			Treasury Assets			Other Assets		
Regression Type	OLS	Within	TWFE	OLS	Within	TWFE	OLS	Within	TWFE	OLS	Within	TWFE
Recapitalisation Variable	0.3027***	0.3252***	0.3192***	-0.1376**	-0.1376**	-0.1649**	0.8960***	0.8871***	0.9587***	0.3746**	0.4382**	0.4159***
	[0.0411]	[0.0440]	[0.0419]	[0.0616]	[0.0616]	[0.0753]	[0.2482]	[0.2351]	[0.2602]	[0.1688]	[0.1679]	[0.1543]
Recapitalisation Variable (t-1)	-0.0028	0.0192	0.0004	-0.2905	-0.2905	-0.2372	0.285	0.3538*	0.2759	-0.0634	-0.053	-0.0537
	[0.0510]	[0.0686]	[0.0702]	[0.1949]	[0.1949]	[0.1892]	[0.1930]	[0.1962]	[0.1841]	[0.1463]	[0.1945]	[0.1868]
Recapitalisation Variable (t-2)	0.0012	-0.0125	-0.014	0.2231	0.2231	0.1985	-0.0437	0.0271	0.0805	0.0578	-0.0182	-0.0365
	[0.0533]	[0.0545]	[0.0560]	[0.1387]	[0.1387]	[0.1340]	[0.0503]	[0.0661]	[0.0680]	[0.1174]	[0.1353]	[0.1462]
Recapitalisation Variable (t-3)	0.0149	-0.0038	0.0022	0.0729	0.0729	0.0404	-0.2875**	-0.2034	-0.2120**	0.0618	0.0067	0.0582
	[0.0375]	[0.0309]	[0.0322]	[0.0681]	[0.0681]	[0.0737]	[0.1148]	[0.1239]	[0.0990]	[0.0430]	[0.0566]	[0.0542]
Recapitalisation Variable (t-4)	-0.0266	-0.0229	-0.009	0.1361	0.1361	0.0918	0.1942	0.2870*	0.3092**	-0.1425***	-0.2233**	-0.2120**
	[0.0390]	[0.0234]	[0.0307]	[0.1117]	[0.1117]	[0.1124]	[0.1889]	[0.1556]	[0.1486]	[0.0341]	[0.0930]	[0.0924]
Recapitalisation Variable (t-5)	-0.0098	-0.0703	-0.0572	-0.0274	-0.0274	-0.0257	0.1136	0.1262	0.1097	-0.136	-0.2820**	-0.2251*
	[0.0226]	[0.0484]	[0.0527]	[0.0512]	[0.0512]	[0.0385]	[0.0932]	[0.0913]	[0.1043]	[0.0931]	[0.1241]	[0.1244]
Cumulative Effect	**0.280**	**0.235**	**0.242**	**-0.023**	**-0.023**	**-0.10**	**1.158**	**1.478**	**1.522**	**0.152**	**-0.131**	**-0.053**
P-value (Cumulative effect=0)	**0.008**	**0.036**	**0.046**	**0.927**	**0.927**	**0.731**	**0.004**	**0.000**	**0.001**	**0.547**	**0.701**	**0.878**
Observations	666	664	664	664	664	664	645	643	643	666	664	664
R-Squared	0.05	0.27	0.34	0.1	0.1	0.16	0.1	0.21	0.27	0.02	0.16	0.25
Bank-Specific Controls	X	√	√	X	√	√	X	√	√	X	√	√
Bank Fixed Effect	X	√	√	X	√	√	X	√	√	X	√	√
Time Fixed Effects	X	X	√	X	X	√	X	X	√	X	X	√
Macro Controls	X	√	X	X	√	X	X	√	X	X	√	X

Note: RWAR growth recapitalisation variable as $\frac{\Delta def_{t-k}^{i}}{RWA_{t-k}}$

Annex 8 – Bank Categorisations

Table 35 - Bank Types

Number of Banks in each category	Banks By Size	Bank Status
Large (or Foreign)	9	9
Small (or Domestic)	8	8
Total	**17**	**17**

Note: Table provides bank classifications. “Bank Size” refers to “Large” or “Small”; “Bank Status” refers to “Foreign” or “Domestic” banks. “Bank Status” is determined by the recapitalisation requirement (i.e., the lower or the higher capital requirement based on domestic ownership percentage). Bank Size type is determined by the median size of banks’ total asset market shares.

Table 36 – Bank Classifications

Bank	**Loan Mrkt. Share, % (Dec, 2012)**	**Asset Mrkt. Share, % (Dec, 2012)**	**Size**	**Bank Status (Ownership)**
Zanaco	16.5	16.8	L	D
Barclays	18.6	15	L	F
Stanchart	14	15	L	F
Stanbic	17.3	14.7	L	F
Bank of China	0.9	5.8	L	F
Citibank	2.7	5.4	L	F
Indo Zambia	5	5	L	F
Investrust	4.7	3.7	L	D
FNBZ	4.2	3.5	L	F
ABC (merged with Finance Bank)	4.1	2.8	S	F
ECO	1.2	1.7	S	F
First Alliance	1.3	1.1	S	D
Cavmont Capital	1.2	1.1	S	D
Access	0.5	0.9	S	D
UBA	0.3	0.5	S	D
AB Bank	0.4	0.4	S	D
FCB	0.1	0.1	S	D
Removed from Sample or Merged				
Intermarket (Failed)	0.9	0.9	S	D
Finance Bank (Domestic, Merged with ABC)	5.9	5.5	L	D

Note: L = Large; S = Small; D = Domestic; F = Foreign. The “Ownership” category is based on whether the bank has 51% domestic ownership (and defines the capital requirement the bank was subject to and recapitalized towards). Some international banks attained more than 51% domestic ownership to remain under the domestic capital regime (e.g., UBA).

Source: BOZ Annual Report 2012.

Annex 9 – Heterogeneous Bank Reactions to the Recapitalisation

Table 37 – Heterogeneous Bank Reactions to the Recapitalisation Across Asset Types (winsorised)

Notes: The regressions in this table are based on equation (12) in the text and the approach is loosely borrowed from Fang et al. (2021). This table presents estimates from regressing the growth rates for different types of bank assets over different periods (the dependent variable, in percent) against a recapitalisation variable. The recapitalisation variable is the ratio of equity raised by a bank in deficit under the absolute minimum capital rule in a quarter divided by the bank's total assets at the beginning of that quarter. Columns show different dependent variables (and are asset-type and growth period-specific). The main rows containing coefficients provide the beta on the recapitalisaion variable for the different bank types. The "Test of equality (p-value)" row provides the p-value from testing the hypothesis H0:B1=B(Type), which tests whether the asset growth betas for the different types of banks are equal. Bolded p-values are significant at 10%. Each column represents a regression using a different asset type (the set includes the following: Total Assets, Bank Loans, Treasury Assets, and Other Assets) and a different growth period (for example, (0,4) is estimating the growth rate for the period t=0 to t=4, i.e. beginning at the start of the period in which the recapitalisation occurred, covering up until the beginning of the fourth quarter). Data is quarterly and was constructed from the Bank of Zambia's monthly Prudential Return. Bank-Specific controls include the following variables (with one lag): log of total bank assets, return on assets, liquid assets as a percent of total assets, RWAR ratio, a sector-specific loan measure, provisions for the quarter as a percent of beginning total assets. Driscoll Kraay Standard Errors were used and are provided in parentheses. Data are winsorised. *** p < 0.01, ** p < 0.05, * p < 0.1.

	Dependent Variable	**Total Assets**					**Total Loans**					**Treasury Securities**					**Other Assets**				
Group	Growth (t,t+r)	(0,1)	(0,2)	(0,3)	(0,4)	(0,6)	(0,1)	(0,2)	(0,3)	(0,4)	(0,6)	(0,1)	(0,2)	(0,3)	(0,4)	(0,6)	(0,1)	(0,2)	(0,3)	(0,4)	(0,6)
Large	Small	0.6305***	0.5056***	0.8061***	0.5167*	0.1678	-0.3340***	-0.2901	-0.1138	0.0656	0.5642	0.512	2.2919***	2.2562***	1.3767*	-1.2598**	1.0459***	0.3335	1.0487***	0.9058***	-0.7522
		[0.0763]	[0.1468]	[0.2519]	[0.2696]	[0.3632]	[0.1145]	[0.2311]	[0.1378]	[0.2340]	[0.4866]	[0.3930]	[0.5142]	[0.6681]	[0.7355]	[0.5937]	[0.2432]	[0.3182]	[0.2442]	[0.3226]	[0.5688]
	Large	1.1503***	1.7176***	1.8510***	2.4678***	2.2889***	0.2657	0.5869**	1.6628***	1.8693***	2.0454***	2.7239***	4.1618***	4.7503***	3.8833***	4.8765***	1.5631***	1.4660***	0.6451	2.3659***	2.2948***
		[0.1305]	[0.2411]	[0.4037]	[0.5469]	[0.5562]	[0.2035]	[0.2895]	[0.5713]	[0.6940]	[0.7259]	[0.4256]	[0.7097]	[1.1372]	[1.3624]	[1.4605]	[0.2929]	[0.4117]	[0.5119]	[0.6750]	[0.8486]
	Test of equality (p -value)	**0.005**	**0.000**	**0.005**	**0.001**	**0.003**	**0.004**	**0.046**	**0.002**	**0.010**	**0.06**	**0.002**	**0.046**	**0.048**	**0.094**	**0.001**	**0.159**	**0.049**	0.476	**0.122**	**0.014**
	Obs.	729	713	697	681	649	729	713	678	646	598	704	689	673	658	628	729	713	697	681	649
	R2	0.35	0.4	0.44	0.45	0.51	0.27	0.33	0.37	0.39	0.46	0.23	0.34	0.35	0.37	0.39	0.23	0.25	0.29	0.30	0.35
	Bank-Specific Controls	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√
	Bank Fixed Effect	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√
	Time Fixed Effects	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√
	Macro Controls	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Ownership	Domestic	0.6118***	0.5798***	0.7911***	0.5994*	0.1041	-0.3656***	-0.1816	-0.1352	-0.0128	0.3803	0.7556**	2.1395***	2.1150***	1.6879	-1.4710*	0.8548**	0.5029	1.1501***	1.1688***	-0.8009
		[0.0796]	[0.1893]	[0.2495]	[0.3352]	[0.4054]	[0.0914]	[0.2613]	[0.1752]	[0.2897]	[0.5181]	[0.3491]	[0.5060]	[0.5777]	[1.0320]	[0.8119]	[0.3298]	[0.3797]	[0.2242]	[0.2947]	[0.5124]
	Foreign	0.9497***	1.0358***	1.4074***	1.4383***	1.4361***	0.0524	-0.0062	0.9067	1.2067*	1.6871***	1.3288	3.5574***	3.8461***	2.2499***	2.3896	1.6779***	0.6466	0.6425	1.2298**	0.9991
		[0.1317]	[0.3072]	[0.1908]	[0.5085]	[0.5062]	[0.1930]	[0.2838]	[0.6259]	[0.6848]	[0.5627]	[0.9525]	[0.5307]	[0.6215]	[0.7554]	[1.5901]	[0.2836]	[0.5866]	[0.5299]	[0.5383]	[0.7609]
	Test of equality (p -value)	**0.007**	0.210	**0.014**	0.194	**0.035**	**0.067**	0.652	**0.083**	**0.090**	0.108	0.572	**0.071**	**0.040**	0.671	**0.041**	**0.043**	0.833	0.328	0.927	**0.086**
	Obs.	729	713	697	681	649	729	713	678	646	598	704	689	673	658	628	729	713	697	681	649
	R2	0.35	0.39	0.44	0.45	0.51	0.27	0.32	0.37	0.39	0.46	0.22	0.34	0.35	0.36	0.39	0.23	0.25	0.29	0.30	0.35
	Bank-Specific Controls	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√
	Bank Fixed Effect	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√
	Time Fixed Effects	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√
	Macro Controls	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X

Table 38 - Heterogeneous Bank Reactions to the Recapitalisation Across Asset Types (No Bank-level controls)

Notes: The regressions in this table are based on equation (12) in the text and the approach is loosely borrowed from Fang et al. (2021). This table presents estimates from regressing the growth rates for different types of bank assets over different periods (the dependent variable, in percent) against a recapitalisation variable. The recapitalisation variable is the ratio of equity raised by a bank in deficit under the absolute minimum capital rule in a quarter divided by the bank's total assets at the beginning of that quarter. Columns show different dependent variables (and are asset-type and growth period-specific). The main rows containing coefficients provide the beta on the recapitalisaion variable for the different bank types. The "Test of equality (p-value)" row provides the p-value from testing the hypothesis H0:B1=B(Type), which tests whether the asset growth betas for the different types of banks are equal. Bolded p-values are significant at 10%. Each column represents a regression using a different asset type (the set includes the following: Total Assets, Bank Loans, Treasury Assets, and Other Assets) and a different growth period (for example, (0,4) is estimating the growth rate for the period t=0 to t=4, i.e. beginning at the start of the period in which the recapitalisation occurred, covering up until the beginning of the fourth quarter). Data is quarterly and was constructed from the Bank of Zambia's monthly Prudential Return. No Bank-Specific controls. Robust Standard Errors were used and are provided in parentheses. Data are not winsorised. *** p < 0.01, ** p < 0.05, * p < 0.1.

	Dependent Variable: Growth Rate for Each Asset Type	**Total Assets**					**Total Loans**					**Treasury Securities**					**Other Assets**				
Group	**Growth Period (t,t+r)**	(0,1)	(0,2)	(0,3)	(0,4)	(0,6)	(0,1)	(0,2)	(0,3)	(0,4)	(0,6)	(0,1)	(0,2)	(0,3)	(0,4)	(0,6)	(0,1)	(0,2)	(0,3)	(0,4)	(0,6)
Large	Small	0.6886***	0.5005	0.6038	0.1676	-0.1513	-1.1434*	-0.8861*	-0.2788	0.1704	0.7447	1.1478	1.9229***	2.8938***	-0.2458	-2.8739	1.1989*	0.6147	0.7278**	0.3037	-1.5237***
		[0.1375]	[0.4506]	[0.4872]	[0.6219]	[0.7830]	[0.6278]	[0.4186]	[0.5024]	[0.6362]	[1.2181]	[0.6798]	[0.2014]	[0.8862]	[1.6304]	[2.3775]	[0.6718]	[0.7449]	[0.3229]	[0.6929]	[0.4939]
	Large	1.6545***	2.2527***	1.9079***	2.7740***	2.2702**	0.6146	-0.4784	-0.0838	0.7724	0.4092	5.1056***	6.8202***	7.0878***	6.9631**	7.9559*	1.9451***	1.7673*	0.6224	2.9168**	2.1721
		[0.1635]	[0.4221]	[0.5225]	[0.5151]	[0.8221]	[1.2409]	[1.0483]	[0.8533]	[0.7160]	[1.8339]	[1.7218]	[1.8164]	[2.3703]	[3.1390]	[4.3353]	[0.5112]	[0.8614]	[0.6239]	[1.1082]	[1.5910]
	Test of equality (p -value)	**0**	**0.016**	0.124	**0.014**	**0.079**	0.303	0.725	0.856	0.566	0.886	**0.032**	**0.017**	0.121	0.115	**0.092**	0.374	0.353	0.898	**0.082**	**0.042**
	Obs.	746	730	714	698	666	746	727	679	647	599	721	706	690	675	645	746	730	714	698	666
	R2	0.17	0.19	0.18	0.17	0.15	0.09	0.14	0.14	0.14	0.13	0.12	0.15	0.19	0.11	0.11	0.14	0.14	0.13	0.13	0.14
	Bank-Specific Controls	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
	Bank Fixed Effect	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√
	Time Fixed Effects	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√
	Macro Controls	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Ownership	Domestic	0.6895***	0.6444	0.7252	0.4576	0.1626	-1.3567	-0.7624	-0.2698	0.1989	0.6242	1.6633***	1.6067***	2.7055**	-0.4739	-3.7482	0.952	0.9164	0.9619**	0.8399	-1.2378*
		[0.1536]	[0.5262]	[0.5990]	[0.7012]	[0.9072]	[0.8506]	[0.4992]	[0.6107]	[0.7567]	[1.5370]	[0.4409]	[0.2313]	[1.1320]	[2.1287]	[3.2593]	[0.7472]	[0.8205]	[0.3727]	[0.6636]	[0.6911]
	Foreign	1.2168***	1.1974	1.0973**	1.0665	0.6567	0.2148	-0.8901	-0.1874	0.4521	0.7577	2.4765	5.1149***	5.5002***	4.0793	4.3064	2.0642***	0.691	0.238	0.752	0.0299
		[0.2914]	[0.6923]	[0.4942]	[0.9890]	[0.9165]	[0.9734]	[0.7403]	[0.5987]	[0.5316]	[1.2672]	[2.0602]	[1.6530]	[1.7682]	[2.5328]	[3.2813]	[0.2923]	[0.7718]	[0.3085]	[1.2559]	[1.2903]
	Test of equality (p -value)	**0.09**	0.545	0.671	0.642	0.723	0.352	0.899	0.931	0.793	0.951	0.689	**0.041**	0.210	0.266	0.169	0.141	0.847	0.216	0.954	0.43
	Obs.	746	730	714	698	666	746	727	679	647	599	721	706	690	675	645	746	730	714	698	666
	R2	0.17	0.18	0.17	0.16	0.15	0.09	0.14	0.14	0.14	0.13	0.11	0.15	0.19	0.11	0.11	0.14	0.13	0.13	0.13	0.13
	Bank-Specific Controls	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
	Bank Fixed Effect	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√
	Time Fixed Effects	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√
	Macro Controls	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X

Annex 10 – Simple Robustness Tests

Given our small sample and volatile data, robustness tests are clearly necessary. In this section we undertake a bare-bones minimum of these.

We take three simple approaches for now: for investigating outliers, we rerun all our main baseline specifications but drop one bank at a time (Table 39); for covariate sensitivity, we run each baseline regression removing each covariate in order from each regression (Table 40); and finally, to investigate our choice of lag structure in the distributed lag models, we run each model from a three-period forwards lag structure through to effect to a seven quarter lag structure. Each series is provided both winsorised and unwinsorised. Colour coding is indicative only, showing how estimates sit relative to the mean of each variable (row) – white implies close to the mean, a strong red or blue a deviation. The colour coding does not imply the same absolute or percentage deviation from the mean across rows.

Broadly speaking, the headline results – an asset-equity growth effect, strong growth in treasury securities, a strong demand deposit response – do hold. We are slightly concerned that some institutions shift the magnitude of the results, although the broad narrative – likely all we can claim in this thesis – does hold. Given our dataset, the small number of banks, and the less-than-stable behaviour of some banks, some sensitivity is a distinct possibility. There is also a question whether some of the more volatile banks in the sample are operating as Business As Usual (the population we would like to sample), and whether they should be included in the sample.

Results seem robust to control variables being dropped, except for the exclusion of the lagged total assets and lagged provisions, which reduce the magnitude of the estimates. On dynamics, some results are reliant on longer-term effects, however, it is challenging to disentangle whether this is a feature (it takes banks time to originate and execute loans) or a bug (we are picking up other general trends that we are not controlling for given our limited set of macro variables).

In summary, assuming that the baseline specification we have used is defensible, the results seem reasonably robust. Furthermore, our results using the alternative Fang et al. (2020) specification paint a similar story to our main results. Finally, using a different definition for the key recapitalisation

variable – in this case replacing equity growth relative to total assets to equity growth with - we typically find that our results remain significant.

Table 39 – Robustness: Dropping an institution

Notes: This table examines the sensitivity of the results in Tables 10 and 11 form the main text. It presents estimates from regressing the two RWAR difference variables and the asset growth rates for different types of bank assets against a recapitalisation variable and 5 of its lags. Importantly, each column of this table represents uses a sample that has had one institution dropped. The recapitalisation variable is the ratio of equity raised by a bank in deficit under the absolute minimum capital rule in a quarter divided by the bank's total assets at the beginning of that quarter. The "Cumulative Effect" row calculates the sum of all the estimate lag coefficients for that given dependent variable for a given dropped institution. The "P-value (H0: Cumulative Effect=0)" column is testing the hypothesis that the linear combination of the regulatory deficit variable and its lags are significantly different to zero. The regressions are approximately based on equations (10) and (11) in the text. Data is quarterly and was constructed from the Bank of Zambia's monthly Prudential Return. Bank-Specific controls include the following variables (for two lags, including t-1 and t-2): log of total bank assets, return on assets, liquid assets as a percent of total assets, RWAR ratio, a sector-specific loan measure, provisions for the quarter as a percent of beginning total assets. Macro variables include the following (for two lags, including t-1 and t-2): the statutory reserve ratio, quarterly growth in the exchange rate, the average inflation rate over the quarter, and the average Treasury Bill Yield (365 day) over the quarter. Bank fixed effects were included. Time fixed effects are not included. Driscoll Kraay Standard Errors were used and are provided in parentheses. *** p < 0.01, ** p < 0.05, * p < 0.1.

Dependent Variable		Institution Dropped																
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Difference in RWAR (non-winsor)	Cumulative Effect	2.003	2.034	2.022	2.038	1.964	1.969	2.000	1.984	1.512	2.051	2.067	2.066	2.117	2.088	2.023	2.114	2.087
	p-value (H0: Cumulative = 0)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Difference in RWAR (winsor)	Cumulative Effect	0.858	0.875	0.871	0.893	0.819	0.846	0.870	0.830	0.921	0.887	0.909	0.922	0.913	0.927	0.780	0.845	0.860
	p-value (H0: Cumulative = 0)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Total Assets (non-winsor)	Cumulative Effect	0.798	0.682	0.794	0.772	0.667	0.792	0.775	0.821	0.860	0.593	0.699	0.602	0.713	0.865	1.071	0.984	0.655
	p-value (H0: Cumulative = 0)	0.068	0.111	0.077	0.088	0.138	0.070	0.076	0.063	0.043	0.166	0.105	0.184	0.098	0.072	0.022	0.046	0.138
Total Assets (winsor)	Cumulative Effect	0.642	0.562	0.647	0.619	0.526	0.638	0.626	0.684	0.683	0.477	0.589	0.473	0.605	0.649	0.986	0.753	0.527
	p-value (H0: Cumulative = 0)	0.048	0.082	0.053	0.064	0.105	0.050	0.055	0.042	0.043	0.124	0.073	0.171	0.062	0.071	0.004	0.045	0.116
Loan Assets (non-winsor)	Cumulative Effect	0.706	0.726	0.685	0.703	0.646	0.629	0.617	0.711	0.390	0.544	0.215	0.676	0.428	0.334	1.394	0.448	0.617
	p-value (H0: Cumulative = 0)	0.353	0.350	0.367	0.372	0.384	0.394	0.391	0.347	0.534	0.481	0.762	0.409	0.553	0.641	0.018	0.624	0.433
Loan Assets (winsor)	Cumulative Effect	0.571	0.591	0.585	0.552	0.581	0.609	0.574	0.632	0.587	0.541	0.443	0.470	0.457	0.362	1.132	0.536	0.527
	p-value (H0: Cumulative = 0)	0.138	0.135	0.131	0.159	0.119	0.107	0.133	0.101	0.193	0.184	0.277	0.256	0.223	0.285	0.002	0.291	0.201
Treasury Securities (non-winsor)	Cumulative Effect	4.207	4.214	4.289	4.113	4.187	4.197	4.164	4.265	3.628	4.086	4.492	4.005	4.416	4.525	4.823	2.740	4.164
	p-value (H0: Cumulative = 0)	0.001	0.001	0.001	0.002	0.001	0.001	0.001	0.001	0.001	0.003	0.001	0.002	0.002	0.001	0.003	0.006	0.001
Treasury Securities (winsor)	Cumulative Effect	1.169	1.241	1.271	1.133	1.113	1.171	1.165	1.212	1.004	0.988	1.321	1.215	1.188	1.471	1.061	0.885	1.165
	p-value (H0: Cumulative = 0)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Other Assets (non-winsor)	Cumulative Effect	0.020	-0.060	-0.037	0.005	-0.219	-0.077	-0.027	0.044	0.091	-0.252	-0.143	-0.316	-0.150	0.172	-0.154	1.012	-0.357
	p-value (H0: Cumulative = 0)	0.986	0.957	0.974	0.997	0.839	0.943	0.980	0.968	0.937	0.825	0.895	0.783	0.894	0.885	0.909	0.225	0.753
Other Assets (winsor)	Cumulative Effect	0.538	0.514	0.518	0.536	0.346	0.452	0.501	0.591	0.630	0.326	0.401	0.246	0.448	0.554	0.628	1.111	0.281
	p-value (H0: Cumulative = 0)	0.452	0.487	0.473	0.459	0.617	0.517	0.479	0.409	0.382	0.653	0.571	0.754	0.532	0.469	0.405	0.146	0.702

Note: Colour coding is indicative only, showing how estimates sit relative to the mean of each variable (row) – white implies close to the mean, a strong red(higher) or blue (lower), a deviation. The colour coding does not imply the same absolute or percentage deviation from the mean across rows.

Table 40 – Robustness: Dropping a variable

Notes: This table examines the sensitivity of the results in Tables 10 and 11 form the main text. It presents estimates from regressing the RWAR difference variable and the asset growth rates for different types of bank assets against a recapitalisation variable and 5 of its lags. Importantly, each column of this table uses a regression that has had one covariate variable dropped, and each row pair provides estimates for a different dependent variable. The recapitalisation variable is the ratio of equity raised by a bank in deficit under the absolute minimum capital rule in a quarter divided by the bank's total assets at the beginning of that quarter. The "Cumulative Effect" row calculates the sum of all the estimate lag coefficients included in that regression (5 lags) for a given dependent variable. The "P-value (H0: Cumulative Effect=0)" column is testing the hypothesis that the linear combination of the regulatory deficit variable and its lags are significantly different to zero. The regressions are approximately based on equations (10) and (11) in the text. Data is quarterly and was constructed from the Bank of Zambia's monthly Prudential Return. Bank-Specific controls include the following variables (for two lags, including t-1 and t-2): log of total bank assets, return on assets, liquid assets as a percent of total assets, RWAR ratio, a sector-specific loan measure, provisions for the quarter as a percent of beginning total assets. Macro variables include the following (for two lags, including t-1 and t-2): the statutory reserve ratio, quarterly growth in the exchange rate, the average inflation rate over the quarter, and the average Treasury Bill Yield (365 day) over the quarter. Bank fixed effects are included. Time fixed effects are not. Driscoll Kraay Standard Errors were used and are provided in parentheses. *** p < 0.01, ** p < 0.05, * p < 0.1.

		Variable Dropped								
		Demand Measure	Bank Size	Bank Profitability	RWA Ratio	Provisions	TBill	Required Stat Ratio	Inflation	FX Growth
Difference in RWAR (non-winsor)	Cumulative Effect	2.04	2.06	2.03	1.06	1.98	2.01	1.92	2.01	2.03
	p-value (H0: Cumulative = 0)	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00
Difference in RWAR (winsor)	Cumulative Effect	0.89	0.96	0.89	0.30	0.88	0.88	0.86	0.87	0.88
	p-value (H0: Cumulative = 0)	0.00	0.00	0.00	0.06	0.00	0.00	0.00	0.00	0.00
Total Assets (non-winsor)	Cumulative Effect	0.71	0.33	0.79	1.25	0.72	0.76	0.69	0.99	0.93
	p-value (H0: Cumulative = 0)	0.11	0.38	0.08	0.01	0.10	0.08	0.13	0.02	0.03
Total Assets (winsor)	Cumulative Effect	0.60	0.31	0.64	0.96	0.57	0.61	0.56	0.76	0.69
	p-value (H0: Cumulative = 0)	0.08	0.25	0.05	0.01	0.07	0.06	0.10	0.02	0.03
Loan Assets (non-winsor)	Cumulative Effect	0.61	0.70	0.58	0.57	0.45	0.57	0.11	0.98	0.90
	p-value (H0: Cumulative = 0)	0.42	0.38	0.41	0.41	0.51	0.42	0.86	0.19	0.24
Loan Assets (winsor)	Cumulative Effect	0.54	0.59	0.56	0.68	0.46	0.56	0.56	0.77	0.69
	p-value (H0: Cumulative = 0)	0.16	0.09	0.15	0.07	0.21	0.14	0.13	0.05	0.09
Treasury Securities (non-winsor)	Cumulative Effect	4.45	4.16	4.04	3.90	4.21	4.18	3.54	4.32	4.52
	p-value (H0: Cumulative = 0)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Treasury Securities (winsor)	Cumulative Effect	1.38	1.24	1.10	1.22	1.12	1.20	0.94	1.29	1.44
	p-value (H0: Cumulative = 0)	0.01	0.03	0.04	0.03	0.06	0.03	0.09	0.04	0.01
Other Assets (non-winsor)	Cumulative Effect	-0.15	-0.84	0.09	1.04	-0.02	-0.06	-0.05	0.14	0.02
	p-value (H0: Cumulative = 0)	0.89	0.44	0.93	0.35	0.99	0.95	0.96	0.89	0.99
Other Assets (winsor)	Cumulative Effect	0.37	-0.20	0.59	1.29	0.54	0.48	0.57	0.56	0.44
	p-value (H0: Cumulative = 0)	0.61	0.77	0.41	0.10	0.45	0.51	0.37	0.41	0.52

Table 41 – Robustness: Different lead and lag structures

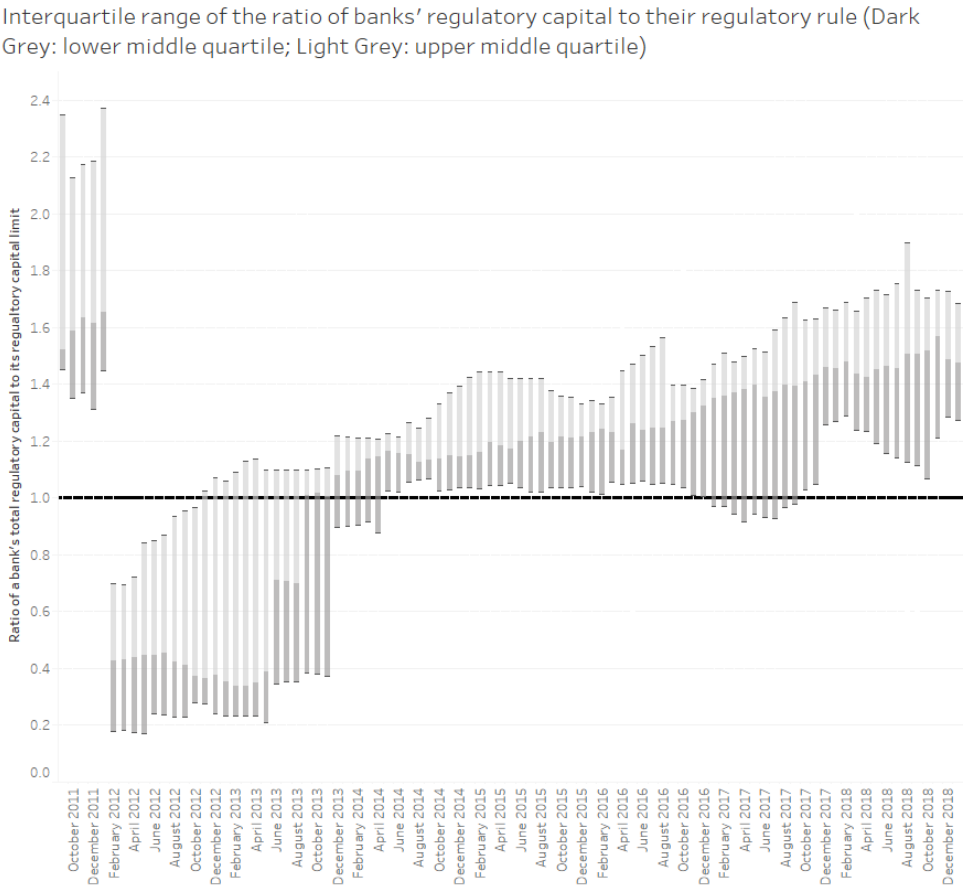
Notes: This table examines the sensitivity of the results in Tables 10 and 11 form the main text. It presents estimates from regressing the RWAR difference variable and the asset growth rate variables for different types of bank assets against a recapitalisation variable and its lags. Importantly, each column of this table uses a regression with a different number of recapitalisation policy variable leads/lags. The number of leads/lags is provided as the number at the top of the column. The table provides the Cumulative Effect parameters for the given lag structure and its p-value for a given dependent variable. The recapitalisation variable is the ratio of equity raised by a bank in deficit under the absolute minimum capital rule in a quarter divided by the bank's total assets at the beginning of that quarter. The "Cumulative Effect" row calculates the sum of all the estimate lag coefficients included in that regression. The "P-value (H0: Cumulative Effect=0)" column is testing the hypothesis that the linear combination of the regulatory deficit variable and its lags are significantly different to zero. The regressions are approximately based on equations (10) and (11) in the text. Data is quarterly and was constructed from the Bank of Zambia's monthly Prudential Return. Bank-Specific controls include the following variables (for two lags, including t-1 and t-2): log of total bank assets, return on assets, liquid assets as a percent of total assets, RWAR ratio, a sector-specific loan measure, provisions for the quarter as a percent of beginning total assets. Macro variables include the following (for two lags, including t-1 and t-2): the statutory reserve ratio, quarterly growth in the exchange rate, the average inflation rate over the quarter, and the average Treasury Bill Yield (365 day) over the quarter. Bank fixed effects are included. Time fixed effects are not. Driscoll Kraay Standard Errors were used and are provided in parentheses. *** p < 0.01, ** p < 0.05, * p < 0.1.

		Lags										
		(-3)	(-2)	(-1)	0	1	2	3	4	5	6	7
Difference in RWAR (non-winsor)	Cumulative Effect	-0.49	-0.19	0.00	1.56	1.79	2.02	2.15	1.99	2.14	2.18	2.29
	p-value (H0: Cumulative = 0)	0.04	0.21	0.97	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Difference in RWAR (winsor)	Cumulative	-0.14	-0.01	0.04	0.59	0.75	0.90	0.87	0.91	1.02	1.02	1.01
	p-value (H0: Cumulative = 0)	0.36	0.94	0.70	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Total Assets (non-winsor)	Cumulative	-0.57	-0.32	-0.02	0.90	0.82	0.71	0.57	0.79	0.62	0.91	0.89
	p-value (H0: Cumulative = 0)	0.18	0.16	0.93	0.00	0.00	0.07	0.20	0.03	0.15	0.10	0.11
Total Assets (winsor)	Cumulative	-0.49	-0.38	-0.11	0.78	0.76	0.71	0.64	0.74	0.54	0.76	0.78
	p-value (H0: Cumulative = 0)	0.03	0.00	0.34	0.00	0.00	0.01	0.01	0.01	0.07	0.04	0.07
Loan Assets (non-winsor)	Cumulative	-1.88	-1.24	-0.97	-0.62	-1.00	-1.46	-1.41	-0.38	-0.49	-0.05	-0.93
	p-value (H0: Cumulative = 0)	0.01	0.02	0.00	0.02	0.02	0.06	0.08	0.60	0.51	0.95	0.33
Loan Assets (winsor)	Cumulative	-0.73	-0.61	-0.59	-0.18	-0.15	-0.16	0.02	0.50	0.44	0.89	0.38
	p-value (H0: Cumulative = 0)	0.01	0.00	0.00	0.11	0.39	0.69	0.96	0.23	0.30	0.09	0.52
Treasury Securities (non-winsor)	Cumulative	-1.26	-0.55	0.26	1.84	3.03	4.31	2.94	4.16	5.48	6.12	5.99
	p-value (H0: Cumulative = 0)	0.07	0.37	0.48	0.08	0.01	0.01	0.08	0.01	0.01	0.00	0.00
Treasury Securities (winsor)	Cumulative	-0.47	-0.31	0.21	0.89	1.41	1.95	1.12	1.36	1.53	1.94	1.83
	p-value (H0: Cumulative = 0)	0.29	0.49	0.38	0.04	0.00	0.00	0.05	0.02	0.04	0.04	0.10
Other Assets (non-winsor)	Cumulative	-1.20	-0.35	-0.06	1.52	1.29	0.98	1.02	0.49	-0.46	0.50	1.76
	p-value (H0: Cumulative = 0)	0.02	0.38	0.83	0.00	0.00	0.20	0.20	0.50	0.65	0.66	0.10
Other Assets (winsor)	Cumulative	-0.95	-0.25	-0.07	1.32	1.04	0.64	0.71	0.50	-0.16	-0.03	0.54
	p-value (H0: Cumulative = 0)	0.02	0.51	0.72	0.00	0.00	0.30	0.22	0.36	0.78	0.97	0.43

Annex 11 – Additional Figures

Note: Figure 47 Figure 47provides the interquartile range of the ratio of banks' capital to their regulatory capital requirement (either the relative or the absolute rule, whichever was higher) per quarter; a ratio of 1 for the ratio means that the bank is just compliant with the regulatory rules, but with no capital buffer. This is represented by the black horizontal line at 1 on the y-axis. The dark grey bars represent the lower-middle quartile per quarter; the light grey bars, the upper-middle. Clearly, after December 2012 the interquartile range dropped from having all banks being compliant (both grey bars were above the ratio of 1 line) to all banks in the interquartile range not holding enough capital.

Figure 47: Interquartile Range of Bank Capital to Regulatory Rule Ratios



Source: BOZ prudential returns

Table 42 - Monetary policy settings

		Monetary Policy Tool (Changes)			Supervision Tool
	Date	Policy Rate	Statutory Reserves	Overnight Lending Facility (OLF) and other instruments	Capital Adequacy
2006	June				
2007	September		14% - 8%		Max. (2B, 10% RWA, 5% Tier1) - Max. (12B, 10% RWA, 5% Tier1)
2009	November			OLF: Introduced	
2010	January		8% - 5%		
2011	February			OLF: 6% - 4%	
2012	January				Max. (12B, 10% RWA, 5% Tier1) - Max. (504B or 104B, 10% RWA, 5% Tier1) -
	April	Policy Rate Introduced		OLF: Set at 2.5% above BOZ rate	
2013	January			Interest Rate Caps Implemented (9% margin above BOZ Rate)	
	February		5% - 8%		
2014	March	9.75%-10.25%	8% - 14%	OLF: 2.5% - 6.0% Access limited to once per week and Failure to pay treated as default.	
	April	10.25% - 12.0%			
	May			OLF: 6.0% - 10%	
	June		Daily Compliance	Interest Rate Caps Loosened	
	November	12.0% - 12.5%			
2015	March			OLF: 9.5% - 6%	
	April		14% - 18%		
	November	12.5% - 15.5%		OLF: 6% - 10% Interest Rate Caps Removed	
2016	November			OLF access restrictions removed - automatic rollover of intraday credit into intraday repo facility	
2017	February	15.5% - 14.0%	18% - 15.5%	OLF: 10% - 6%	
	March	14.0% - 12.5%	15.5% - 12.5%		
	August	12.5% - 11.0%	12.5% - 9.5%		
	November	11.0% - 10.25%	9.5% - 8%		
2018	February	10.25% - 9.75%	8% - 5%		

Source: BOZ website (https://www.boz.zm/laws-and-regulation.htm) and IMF (2017)

Annex 12 – Descriptive Statistics

Variable	Variable Type	N	mean	cv	min	p25	p50	p75	max
Difference in RWAR	Equity Shock Variables	774	-0.006	-17.990	-0.992	-0.031	-0.004	0.023	1.298
Difference in Leverage Ratio		774	-0.002	-24.002	-0.228	-0.021	0.000	0.017	0.267
Recapitalisation/Total Assets (t-1)		774	0.006	5.323	0.000	0.000	0.000	0.000	0.388
Recapitalisation/Risk Weighted Assets (t-1)		774	0.015	5.991	0.000	0.000	0.000	0.000	1.584
Total Assets (USD, m)	Context	774	349.09	0.96	5.17	71.34	237.85	544.59	1243.95
Log (Total Bank Assets)	Bank-Specific Control Variables	774	13.97	0.09	10.18	12.96	14.07	14.99	16.50
Return on Assets		774	0.001	13.628	-0.157	0.000	0.002	0.004	0.180
Liquid Asset Ratio		774	0.613	0.312	0.080	0.499	0.596	0.720	1.000
RWAR		774	0.327	0.665	0.009	0.180	0.245	0.423	1.648
Sector Demand Measure		758	0.015	0.772	0.001	0.009	0.012	0.017	0.194
Provisions		774	0.002	2.899	-0.024	0.000	0.000	0.002	0.068
Treasury Bill Yield	Macro	774	15.728	0.322	5.520	11.558	15.346	19.043	26.305
Required Statutory Reserves		774	9.971	0.480	5.000	5.000	8.000	14.000	18.000
Annualized Inflation Rate		774	9.537	0.439	6.242	6.832	7.811	11.039	22.297
Growth in the Inflation Rate		774	0.025	0.106	-0.230	-0.036	0.009	0.048	0.508
Growth in Total Bank Assets	Dependent Variables	774	0.043	3.197	-0.697	-0.035	0.026	0.109	0.657
Growth in Total Bank Loans (Adj)		774	0.075	3.853	-0.904	-0.048	0.032	0.120	3.024
Growth in Treasury Assets		774	0.081	5.147	-1.000	-0.099	0.017	0.160	4.614
Growth in Other Bank Assets		774	0.067	4.374	-0.653	-0.115	0.019	0.193	2.384
Growth in Total Bank Financing		774	0.056	3.528	-0.789	-0.044	0.022	0.129	2.431
Growth in Total Deposit Financing		774	0.056	3.380	-0.847	-0.047	0.028	0.133	1.578

Annex 13 – MFI Balance Tables

Table 43 – Pre-IRR MFI Balance Tests

Variable	All Pre-IRR data			Two Years Pre-IRR			Quarter Before IRR		
	Non-Bound Lenders	Bound Lenders	Difference	Non-Bound Lenders	Bound Lenders	Difference	Non-Bound Lenders	Bound Lenders	Difference
Growth in real net loans (%)	0.055	0.039	-0.016	0.088	0.054	-0.034	0.157	-0.004	-0.161**
	(0.215)	(0.269)	(0.022)	(0.210)	(0.219)	(0.033)	(0.209)	(0.158)	(0.073)
Chg. in Fees and commissions share of income (%-point)	-0.006	0.001	0.007	0.001	0.001	0.000	-0.012	0.007	0.018
	(0.065)	(0.050)	(0.005)	(0.055)	(0.070)	(0.010)	(0.013)	(0.061)	(0.024)
Log total lender assets	10.628	8.575	-2.053***	11.170	8.748	-2.423***	11.636	8.944	-2.692***
	(1.458)	(1.978)	(0.157)	(1.278)	(2.123)	(0.294)	(1.122)	(2.185)	(0.861)
Fees and commissions share of income (%)	0.168	0.041	-0.127***	0.129	0.057	-0.073***	0.147	0.056	-0.090*
	(0.193)	(0.079)	(0.010)	(0.150)	(0.105)	(0.018)	(0.186)	(0.097)	(0.051)
Leverage ratio (Total Regulatory Capital/Total Assets)	0.076	0.443	0.367***	0.292	0.546	0.254***	0.162	0.499	0.338**
	(0.509)	(0.358)	(0.034)	(0.449)	(0.362)	(0.057)	(0.216)	(0.324)	(0.130)
Loss allowance (Recognized losses/Gross Loans)	0.081	0.149	0.068***	0.092	0.140	0.048**	0.074	0.105	0.031
	(0.115)	(0.185)	(0.014)	(0.109)	(0.157)	(0.022)	(0.095)	(0.125)	(0.051)
Operating expense ratio (Non-interest Expense/Total Assets)	0.131	0.223	0.092***	0.085	0.215	0.130***	0.074	0.317	0.243*
	(0.181)	(0.333)	(0.025)	(0.056)	(0.227)	(0.030)	(0.052)	(0.360)	(0.138)
Age of firm	13.109	10.371	-2.738***	18.375	13.971	-4.404***	20.250	14.655	-5.595
	(9.156)	(7.321)	(0.650)	(10.690)	(8.005)	(1.247)	(11.732)	(9.363)	(3.947)
Observations	201	536	737	64	206	270	8	29	37

Note: The table compares the two pre-IRR subsamples (bound and non-bound lenders) according to their covariate variables. This table provides simple averages for the baseline period, defined over three timespans: all pre-IRR data; two-years pre-IRR; and for the single quarter before the IRR was implemented. Numbers in brackets provide standard errors. The Difference column provides the results from a difference in means test. $^{***}p<0.01$, $^{**}p<0.05$, $^{*}p<0.10$.

Annex 14 – Event Study Implementation

Below, we try to be more precise about what we are estimating when we follow Freyaldenhoven et al. (2021). Specifically, we estimate:

$$Y_{i,t} = \delta_{M+L_M} z_{i,t-M-L_M} + \delta_{-G-L_G-1}\left(1 - z_{i,t+G+L_G}\right) + \sum_{k=-G-L_G}^{M+L_M-1} \delta_k \,.\, \Delta z_{i,t-k} + \alpha_i + \gamma_t + X'_{it}\psi + \varepsilon_{it} \qquad (5.A1)$$

$$z_{i,t} = \begin{cases} T.r_{i,\,t=2012q4} - r_{cap,t=2013q1}, & r_{i,\,t=2012q4} - r_{cap,t=2013q1} > 0 \\ 0, & r_{i,\,t=2012q4} - r_{capt,t=2013q1} \leq 0 \end{cases}$$

$$T = \begin{cases} 1, & t \geq 2013q1 \\ 0, & t \leq 2012q4 \end{cases}$$

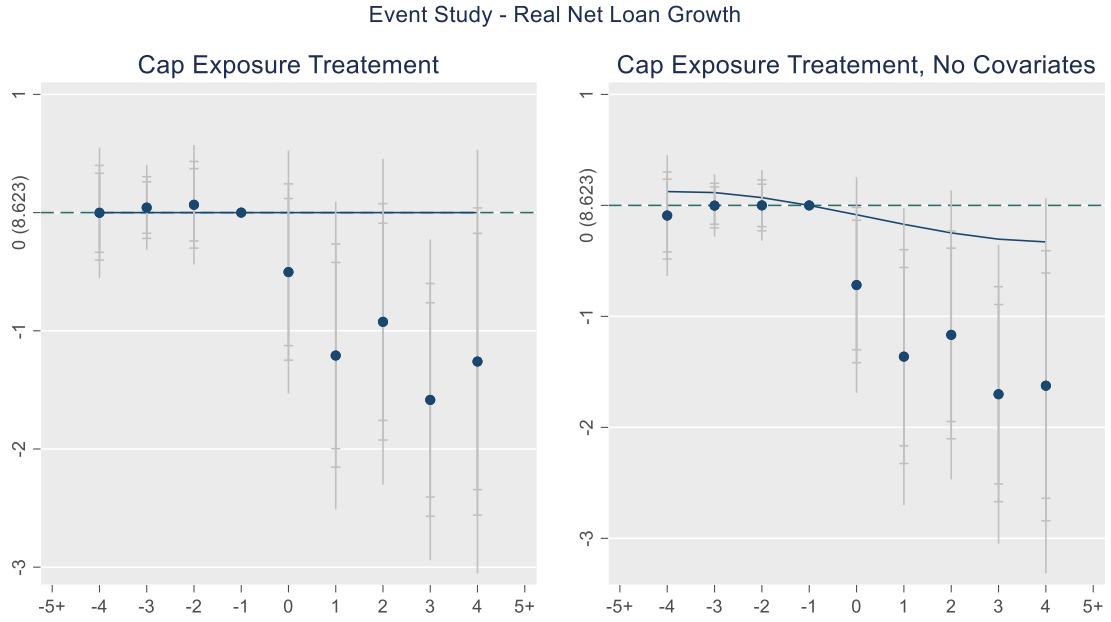
Where Y_{it}, T, and $z_{i,t}$ are discussed in the text. $\Delta z_{i,t-k}$ is an indicator for whether a lender was exposed to the IRR exactly k periods before period t[247] .G and M define the period over which we assume the outcome can be affected by the treatment, and L_M and L_G are additional time periods for which we estimate policy effects that allow us to test for anticipatory effects and parallel trends. $\delta_{M+L_M} z_{i,t-M-L_M}$ and $\delta_{-G-L_G-1}\left(1 - z_{i,t+G+L_G}\right)$ are terms to capture the effects of the policy outside the estimation horizon (in other DID setups, these are analogous to "binned" distant time periods), whereby $z_{i,t-M-L_M}$ and $\left(1 - z_{i,t+G+L_G}\right)$ are policy indicator variables for these early and late periods. As with much of the literature, Freyaldenhoven et al. (2021) normalise on the time period before a policy change. $\alpha_i, \gamma_t, X'_{it}, and\ \varepsilon_{it}$ are as before. For most of the analysis we set G at zero and L_G and M equal to 4 (i.e., we do not allow for anticipation effects, and allow for 5 quarters for the policy to affect loans).

The coefficient plots in this chapter span the set of slope coefficient effect estimates of δ_k from equation (5.1A): $\left\{\left(k, \hat{\delta}_k\right)\right\}_{k=-G-L_G-1}^{k=M+L_M}$. When multiplied by a given IRR exposure, they provide estimates of the cumulative growth effect of the IRR's imposition on a lender's outstanding loans.

In the text we say we plot $\sum_{m=-G}^{k} \hat{\beta}_m$ for different horizons k, but more formally we are actually plotting $\left\{\left(k, \hat{\delta}_k\right)\right\}_{k=-G-L_G-1}^{k=M+L_M}$, with the estimates of $\hat{\delta}_k$ providing the percent growth effect of the IRR implementation per percentage point of cap bindingness over the event window.

[247] k=0 defines the period for a lender in which the IRR was implemented. In our specific, more constrained policy case, k will be the same for all bound lenders because all lenders bound by the IRR were bound at the same time. With a different definition of $z_{i,t}$, 1b could be used to investigate additional forms of policy variation.

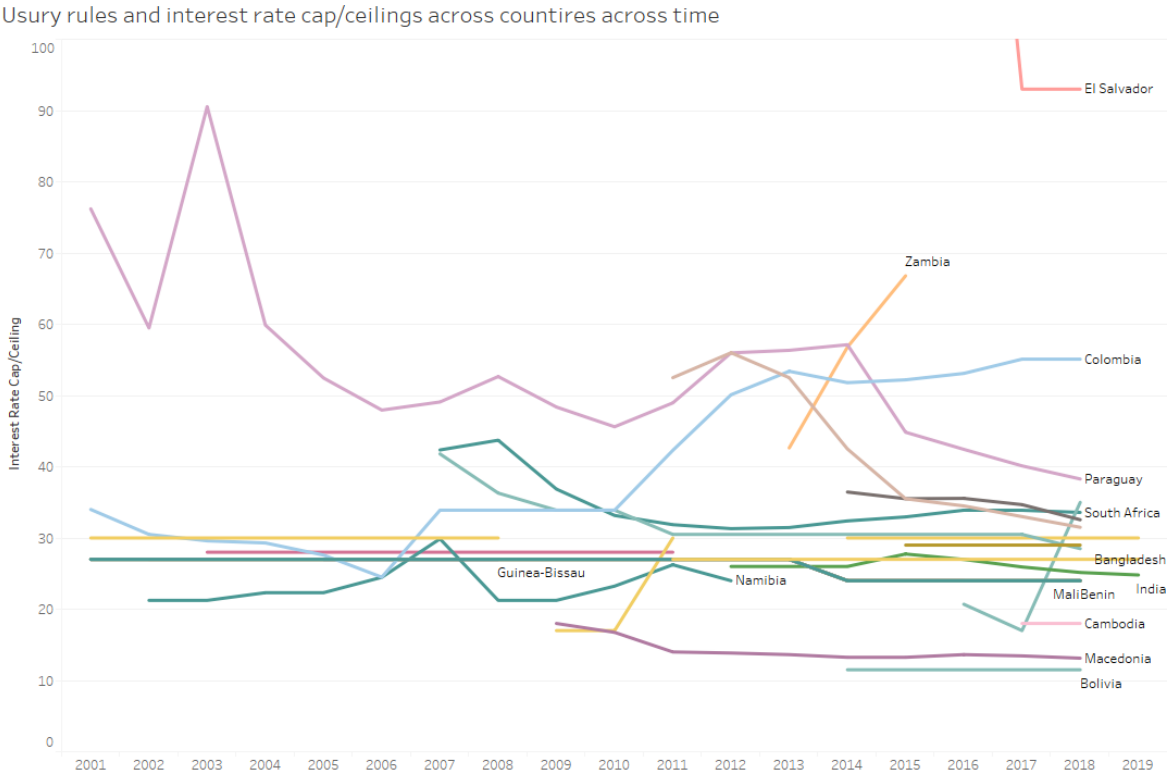
Annex 15 – Robustness, Linear Group Time Trends Included



Note: The figure provides coefficient estimates of $\sum_{m=-G}^{k} \hat{\beta}_m$ from equation 5.10, including group linear time trends. The left figure includes lender covariates, the right figure does not. The dependent variable for both is the log of outstanding real net loans and the point estimate coefficients follow the time path for the marginal effect slope parameters. A comparison group is formed from lenders pricing below the IRR in Q4 2012. Pointwise confidence intervals are provided at 90% and 95% (horizontal grey lines), and the 95 percent uniform confidence band (vertical grey line). Standard errors clustered at lender level. Covariates in the left graphic include lagged values of the capital asset ratio, loss allowance ratio, operating expense ratio, and lender age (see Table 14 for definitions). The approach follows Freyaldenhoven et al (2021). See Annex 14 – Event Study Implementation, for additional details. The Pretrends p-value statistic provides a Wald test for evidence of pretends.

Annex 16 – International Interest Rate Restrictions and Descriptive Statistics

Figure 48 – MFI IRR Regimes Around the World



Source: Author's calculation

Table 44 – IRR Sample Correlations

	Growth in net loans	Growth in number of outstanding borrowers	Operational sustainability	Log number of active borrowers	Loan share of assets (%)	Average outstanding balance (USD)	Impairment Ratio	Operating expense ratio
Growth in net loans (real, local currency)	1							
Growth in number of outstanding borrowers	0.5541*	1						
Operational sustainability (Financial revenue/(financial expense+opp exp+imparement)	0.0321*	0.0263*	1					
Log number of active borrowers	0.017	0.0279*	0.1079*	1				
Loan share of assets (%)	0.0602*	0.0353*	0.2167*	0.1956*	1			
Average outstanding balance (USD)	-0.0599*	-0.0989*	0.0994*	-0.1401*	0.0923*	1		
Impairment Ratio (imparements exp/total loans)	-0.2369*	-0.1918*	-0.1803*	-0.0177*	-0.0569*	0.0432*	1	
Operating expense ratio (opp exp/total assets)	-0.0701*	-0.0611*	-0.4182*	-0.1653*	-0.0500*	-0.2422*	0.1511*	1
CPI Inflation (%)	-0.0025	0.0534*	-0.0118	-0.0331*	-0.1051*	-0.1491*	0.0011	0.0338*
Import Growth Rate (%)	0.0881*	0.0828*	0.0215*	-0.0594*	0.0043	-0.0281*	-0.0493*	-0.0287*
Export Growth Rate (%)	0.0545*	0.0678*	0.0079	-0.0461*	0.005	-0.0218*	-0.0605*	-0.0263*
Government Revenue (% GDP)	-0.0318*	-0.0433*	0.0724*	-0.1261*	0.1747*	0.3826*	-0.0132	-0.0002
Net Government Borrowing (% GDP)	0.0345*	0.0079	0.0421*	-0.1832*	-0.0205*	0.1732*	0.0358*	0.0551*
Current Account Deficit (% GDP)	0.0436*	0.0473*	-0.0081	0.0655*	-0.0007	-0.0842*	0.0351*	-0.0017
GDP Growth Rate (%)	0.1125*	0.1380*	0.0532*	0.0221*	0.0590*	-0.0892*	-0.1146*	-0.0969*
Bank Lending Rate (%)	0.0166	0.0024	0.0301*	-0.1323*	-0.0239*	0.0741*	0.0352*	0.1117*
Bank Branches (per 1000 people)	-0.0801*	-0.0884*	0.0718*	-0.0606*	0.1633*	0.3000*	-0.0746*	-0.0560*
Private Credit (% GDP)	-0.0488*	-0.0789*	0.0342*	0.1539*	0.2556*	0.1014*	-0.0790*	-0.1920*
Real Interest Rate (%)	0.0036	-0.0265*	0.0185*	-0.0619*	0.0143	0.1304*	0.0341*	0.0590*

*p<0.05

	CPI Inflation (%)	Import Growth Rate (%)	Export Growth Rate (%)	Government Revenue (% GDP)	Net Government Borrowing (% GDP)	Current Account Deficit (% GDP)	GDP Growth Rate (%)	Bank Lending Rate (%)	Bank Branches (per 1000 people)	Private Credit (% GDP)	Real Interest Rate (%)
CPI Inflation (%)	1										
Import Growth Rate (%)	0.0395*	1									
Export Growth Rate (%)	-0.0422*	0.4150*	1								
Government Revenue (% GDP)	-0.1188*	-0.0077	-0.0413*	1							
Net Government Borrowing (% GDP)	-0.0205*	0.1199*	0.0122	0.2439*	1						
Current Account Deficit (% GDP)	0.0116	-0.0754*	0.0119	-0.0395*	0.2212*	1					
GDP Growth Rate (%)	0.0475*	0.4148*	0.4002*	-0.0779*	0.0946*	0.0656*	1				
Bank Lending Rate (%)	0.2500*	-0.0174	-0.0582*	0.0820*	0.0915*	-0.0405*	-0.0722*	1			
Bank Branches (per 1000 people)	-0.1479*	-0.0862*	-0.0324*	0.3667*	-0.0196*	-0.1563*	-0.1545*	-0.0826*	1		
Private Credit (% GDP)	-0.1550*	-0.0447*	0.0129	0.1913*	-0.1701*	-0.0260*	0.0200*	-0.2125*	0.4216*	1	
Real Interest Rate (%)	-0.1849*	-0.1176*	-0.0938*	0.0786*	0.0517*	-0.0263*	-0.1103*	0.7383*	-0.0067	-0.0723*	1

Table 45 – Country List

	Country	IRR Obs.	Total Obs.
Africa	Angola	0	13
	Burundi	0	74
	Central African Republic	0	8
	Comoros	0	4
	Congo, Democratic Republic of th	0	103
	Congo, Republic of the	0	33
	Ethiopia	0	155
	Gambia, The	0	9
	Ghana	0	257
	Guinea	0	35
	Kenya	0	158
	Liberia	0	16
	Madagascar	0	142
	Malawi	0	65
	Mozambique	0	96
	Nigeria	0	159
	Rwanda	0	124
	Sierra Leone	0	49
	Swaziland	0	4
	Tanzania	0	116
	Uganda	0	181
	Benin	166	166
	Burkina Faso	87	87
	Cote d'Ivoire (Ivory Coast)	83	83
	Guinea-Bissau	11	11
	Mali	131	131
	Namibia	9	9
	Niger	74	74
	Senegal	156	156
	Togo	119	119
	South Africa	25	60
	Zambia	8	57
	Zimbabwe	1	24

		IRR Obs.	Total Obs.
East Asia and the Pacific	China, People's Republic of	0	151
	East Timor	0	28
	Fiji	0	9
	Indonesia	0	314
	Laos	0	126
	Malaysia	0	3
	Papua New Guinea	0	55
	Philippines	0	699
	Samoa	0	18
	Solomon Islands	0	5
	Tonga	0	7
	Myanmar (Burma)	46	46
	Thailand	9	9
	Cambodia	25	242
Eastern Europe and Central Asia	Albania	0	59
	Bosnia and Herzegovina	0	191
	Bulgaria	0	117
	Georgia	0	135
	Kazakhstan	0	151
	Kosovo	0	100
	Moldova	0	42
	Mongolia	0	76
	Montenegro	0	29
	Romania	0	24
	Serbia	0	45
	Tajikistan	0	266
	Turkey	0	16
	Ukraine	0	22
	Armenia	124	126
	Macedonia	27	48
	Kyrgyzstan	47	220
	Azerbaijan	37	211

		IRR Obs.	Total Obs.
Latin America and The Caribbean	Argentina	0	120
	Brazil	0	237
	Costa Rica	0	114
	Dominican Republic	0	112
	Guatemala	0	218
	Guyana	0	7
	Haiti	0	72
	Honduras	0	234
	Jamaica	0	14
	Mexico	0	607
	Nicaragua	0	112
	Panama	0	36
	Peru	0	762
	Paraguay	76	78
	Colombia	330	346
	Ecuador	424	585
	Chile	20	50
	Bolivia	48	280
	El Salvador	22	159
Middle East and North Africa	Egypt	0	147
	Iraq	0	52
	Jordan	0	84
	Lebanon	0	47
	Morocco	0	131
	Palestine	0	85
	Yemen	0	54
South Asia	Afghanistan	0	127
	Pakistan	0	393
	Bangladesh	265	600
	India	697	1,352
	Nepal	20	303
	Sri Lanka	2	153

Note: IRR Obs. provides the total number of MFI years for which a country had a regime in place. Total Obs. provides the total number of MFI-year observations in the dataset for that country.

Annex 17 – NIM Determinants Regression

Table 46 – NIM Determinants Regression

Independent Variable		Full Sample						Pre-IRR Sample (end-2012Q4)						Large/Diversified Banks (Full Sample)						Small/Undiversified Banks (Full Sample)					
		Static			Dynamic			Static			Dynamic			Static			Dynamic			Static			Dynamic		
		OLS	FE	Time	OLS	FE	Time	OLS	FE	Time	OLS	FE	Time	OLS	FE	Time	OLS	FE	Time	OLS	FE	Time	OLS	FE	Time
Bank-Specific Determinants	Lagged NIM				0.5900***	0.5680***	0.7176***				0.4815***	0.4174***	0.6659***				0.6024***	0.6338***	0.7441***				0.5180***	0.4494***	0.6194***
					[0.0957]	[0.0346]	[0.0280]				[0.1055]	[0.0603]	[0.0365]				[0.0618]	[0.0415]	[0.0376]				[0.1533]	[0.0482]	[0.0478]
	Operating Costs	0.5430***	0.3124*	0.5187***	0.2770*	0.3335***	0.2540***	0.4601***	0.356	0.4622**	0.372	0.5305***	0.3011***	0.4785*	0.359	0.4517*	0.2529**	0.2656***	0.2306***	0.4099**	0.208	0.4199*	0.249	0.3637***	0.3138***
		[0.1523]	[0.1697]	[0.1673]	[0.1369]	[0.0527]	[0.0391]	[0.1702]	[0.2555]	[0.1808]	[0.2560]	[0.0677]	[0.0578]	[0.2477]	[0.1973]	[0.2510]	[0.0999]	[0.0513]	[0.0485]	[0.1978]	[0.2307]	[0.2252]	[0.1981]	[0.0631]	[0.0569]
	Liquidity Ratio	-0.003	-0.006	-0.006	0.000	0.000	0.000	-0.003	-0.001	-0.004	0.0064**	0.005	-0.001	0.003	-0.002	0.000	0.002	0.002	0.003	-0.005	-0.008	-0.008	-0.002	-0.003	-0.0034**
		[0.0050]	[0.0042]	[0.0050]	[0.0018]	[0.0016]	[0.0012]	[0.0065]	[0.0060]	[0.0069]	[0.0022]	[0.0037]	[0.0020]	[0.0071]	[0.0048]	[0.0075]	[0.0027]	[0.0028]	[0.0024]	[0.0059]	[0.0058]	[0.0060]	[0.0024]	[0.0021]	[0.0016]
	Bank Size	0.0805***	0.064	0.0751***	0.035	0.023	0.0232***	0.0688***	0.000	0.0679***	0.004	0.000	0.0207***	0.0675***	0.054	0.0604***	0.014	0.013	0.0145*	-0.1315***	-0.017	-0.1107*	0.007	0.013	0.000
		[0.0167]	[0.0436]	[0.0170]	[0.0253]	[0.0153]	[0.0037]	[0.0167]	[0.0448]	[0.0166]	[0.0308]	[0.0188]	[0.0057]	[0.0128]	[0.0452]	[0.0179]	[0.0240]	[0.0120]	[0.0077]	[0.0458]	[0.0476]	[0.0590]	[0.0322]	[0.0243]	[0.0219]
	Credit Risk	-0.005	-0.005	-0.002	0.001	0.001	0.000	-0.007	-0.001	-0.008	0.000	0.000	-0.002	-0.003	0.012	0.004	0.0084*	0.0079***	0.002	-0.0161**	-0.010	-0.012	-0.002	-0.002	-0.002
		[0.0064]	[0.0069]	[0.0070]	[0.0028]	[0.0017]	[0.0011]	[0.0092]	[0.0124]	[0.0096]	[0.0074]	[0.0049]	[0.0021]	[0.0056]	[0.0084]	[0.0039]	[0.0042]	[0.0030]	[0.0028]	[0.0081]	[0.0088]	[0.0117]	[0.0037]	[0.0026]	[0.0029]
	Maturity Transformation	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-0.0003*	0.000	0.000	0.000	0.000	-0.0001*
		[0.0001]	[0.0001]	[0.0001]	[0.0000]	[0.0000]	[0.0000]	[0.0002]	[0.0001]	[0.0002]	[0.0000]	[0.0001]	[0.0000]	[0.0001]	[0.0001]	[0.0001]	[0.0000]	[0.0000]	[0.0000]	[0.0002]	[0.0002]	[0.0002]	[0.0001]	[0.0000]	[0.0000]
	USD Loan Share	-0.0102**	0.004	-0.0125***	0.003	0.003	-0.0017*	-0.0120**	-0.005	-0.0122**	-0.003	0.001	-0.002	-0.0201***	-0.0075*	-0.0276***	-0.001	-0.001	-0.0052**	-0.002	0.0149**	-0.003	0.008	0.0096***	0.000
		[0.0044]	[0.0050]	[0.0044]	[0.0027]	[0.0016]	[0.0009]	[0.0057]	[0.0116]	[0.0058]	[0.0047]	[0.0040]	[0.0016]	[0.0039]	[0.0032]	[0.0051]	[0.0018]	[0.0024]	[0.0023]	[0.0061]	[0.0050]	[0.0066]	[0.0053]	[0.0030]	[0.0015]
	Non-performing Loans	0.002	-0.009	-0.002	-0.0098**	-0.0094***	-0.0054**	0.0234*	0.005	0.0244*	-0.015	-0.0200***	-0.004	0.013	0.029	0.002	0.012	0.0097**	-0.003	-0.010	-0.0224**	-0.012	-0.0187***	-0.0148***	-0.0070**
		[0.0096]	[0.0092]	[0.0103]	[0.0041]	[0.0031]	[0.0023]	[0.0136]	[0.0102]	[0.0142]	[0.0101]	[0.0069]	[0.0052]	[0.0193]	[0.0203]	[0.0200]	[0.0091]	[0.0049]	[0.0047]	[0.0119]	[0.0082]	[0.0135]	[0.0028]	[0.0041]	[0.0035]
	Risk Aversion	0.0124**	0.0130**	0.0138**	0.0072**	0.0080***	0.0050***	0.012	0.004	0.012	0.004	0.0052*	0.0054***	0.005	0.003	0.005	-0.001	-0.001	0.000	0.008	0.0138*	0.010	0.0106***	0.0109***	0.0067***
		[0.0059]	[0.0045]	[0.0059]	[0.0030]	[0.0014]	[0.0012]	[0.0095]	[0.0056]	[0.0098]	[0.0027]	[0.0027]	[0.0017]	[0.0047]	[0.0044]	[0.0043]	[0.0037]	[0.0022]	[0.0021]	[0.0067]	[0.0072]	[0.0061]	[0.0028]	[0.0019]	[0.0016]
	Demand Measure	0.0679*	0.1103***	0.0982**	0.0388**	0.0364*	0.022	0.0864*	0.078	0.1058*	0.038	0.044	0.021	0.0397***	0.030	0.024	0.014	0.013	0.018	0.072	0.085	0.2300**	-0.001	0.037	0.017
		[0.0387]	[0.0272]	[0.0483]	[0.0160]	[0.0203]	[0.0193]	[0.0488]	[0.0505]	[0.0587]	[0.0245]	[0.0318]	[0.0257]	[0.0143]	[0.0174]	[0.0337]	[0.0150]	[0.0266]	[0.0320]	[0.1081]	[0.0811]	[0.1120]	[0.0467]	[0.0343]	[0.0447]
Macroeconomic and Institutional Determinants	Reserve Requirements	0.016	0.018		0.008	0.010		0.0776**	0.0732**		0.0321*	0.0353***		0.016	0.024		0.014	0.010		0.0391*	0.030		0.011	0.0185**	
		[0.0136]	[0.0159]		[0.0087]	[0.0073]		[0.0314]	[0.0259]		[0.0178]	[0.0086]		[0.0191]	[0.0170]		[0.0097]	[0.0064]		[0.0223]	[0.0278]		[0.0182]	[0.0092]	
	Industry Concentration	0.0567*	0.040		0.008	0.0203**		0.020	0.010		-0.007	0.012		0.012	0.005		-0.004	-0.005		0.1146***	0.0710**		0.029	0.0517***	
		[0.0304]	[0.0308]		[0.0175]	[0.0088]		[0.0494]	[0.0451]		[0.0218]	[0.0163]		[0.0262]	[0.0357]		[0.0164]	[0.0083]		[0.0360]	[0.0276]		[0.0290]	[0.0123]	
	Copper Price	0.000	0.000		0.000	0.000		0.000	0.000		0.000	0.000		-0.0011*	-0.001		-0.0006*	-0.0007***		0.001	0.001		0.0008**	0.0008**	
		[0.0005]	[0.0004]		[0.0003]	[0.0003]		[0.0006]	[0.0006]		[0.0003]	[0.0002]		[0.0006]	[0.0005]		[0.0003]	[0.0003]		[0.0008]	[0.0006]		[0.0003]	[0.0004]	
	Inflation	-0.010	-0.015		-0.007	-0.004		0.032	0.029		0.014	0.023		-0.0264**	-0.0386**		-0.0334**	-0.0355***		0.014	0.012		0.019	0.021	
		[0.0130]	[0.0147]		[0.0150]	[0.0120]		[0.0373]	[0.0428]		[0.0370]	[0.0201]		[0.0130]	[0.0131]		[0.0110]	[0.0119]		[0.0187]	[0.0216]		[0.0261]	[0.0151]	
	Treasury Bill Rate	-0.016	-0.011		-0.001	0.001		0.018	0.011		0.011	0.005		-0.039	-0.0472*		-0.015	-0.011		0.000	-0.004		0.000	0.004	
		[0.0215]	[0.0169]		[0.0090]	[0.0094]		[0.0276]	[0.0301]		[0.0182]	[0.0129]		[0.0242]	[0.0230]		[0.0149]	[0.0088]		[0.0284]	[0.0219]		[0.0113]	[0.0121]	
	Exchange Rate	-0.0040*	-0.0059**		-0.0068***	-0.0056**		-0.002	-0.004		-0.0080*	-0.0068***		-0.0072***	-0.0066**		-0.0077***	-0.0082***		0.000	-0.004		-0.005	-0.002	
		[0.0023]	[0.0024]		[0.0022]	[0.0025]		[0.0045]	[0.0041]		[0.0039]	[0.0022]		[0.0027]	[0.0023]		[0.0013]	[0.0023]		[0.0036]	[0.0037]		[0.0036]	[0.0030]	
	Interest Rate Volatility	0.000	0.000		-0.0003*	0.000		0.000	0.000		0.000	-0.0004*		0.000	0.000		0.000	0.000		0.000	0.000		0.000	-0.0003*	
		[0.0002]	[0.0002]		[0.0001]	[0.0002]		[0.0002]	[0.0003]		[0.0002]	[0.0002]		[0.0002]	[0.0002]		[0.0002]	[0.0002]		[0.0002]	[0.0002]		[0.0002]	[0.0002]	
	Constant	0.012	0.0158*	0.0325***	0.000	-0.002	0.001	0.008	0.015	0.0335***	0.004	0.000	0.004	0.0334***	0.0268**	0.0346***	0.007	0.007	0.003	0.006	0.008	0.0500***	-0.004	-0.0134**	0.007
		[0.0099]	[0.0082]	[0.0083]	[0.0028]	[0.0046]	[0.0029]	[0.0121]	[0.0147]	[0.0078]	[0.0045]	[0.0067]	[0.0037]	[0.0095]	[0.0106]	[0.0104]	[0.0052]	[0.0047]	[0.0037]	[0.0162]	[0.0134]	[0.0130]	[0.0058]	[0.0056]	[0.0057]
	Observations	738	738	738	715	715	715	395	395	395	375	375	375	350	350	350	342	342	342	388	388	388	373	373	373
	R-Squared (Total)	0.218	0.264	0.291	0.568	0.900	0.884	0.211	0.230	0.248	0.472	0.886	0.888	0.291	0.367	0.412	0.636	0.925	0.908	0.270	0.349	0.354	0.587	0.908	0.903
	Bank Fixed Effects	x	√	x	x	√	x	x	√	x	x	√	x	x	√	x	x	√	x	x	√	x	x	√	x
	Time Fixed Effects	x	x	√	x	x	√	x	x	√	x	x	√	x	x	√	x	x	√	x	x	√	x	x	√
	Standard Errors	robust	robust	robust	robust	PCSE	PCSE	robust	robust	robust	robust	PCSE	PCSE	robust	robust	robust	robust	PCSE	PCSE	robust	robust	robust	robust	PCSE	PCSE

Note: Estimates from equation 7.4 regressing bank margins ($Margin = \frac{Interest\ Received\ from\ loans + fees}{Performing\ Loans} - \frac{Interest\ paid\ to\ all\ sources}{All\ Interest-Bearing\ Liabilities}$1) on bank-specific and common bank determinants. Different sub-samples and estimation approaches are used in each column. Variables defined in Table 24. The Large Diversified Banks include the largest seven banks, all of whom have over 5% of market share (by loans and deposits) and maintain large branch networks. ***p<0.01, **p<0.05, *p<0.10.

Annex 18 – Descriptive statistics

Variable Correlations

	Loan Margin	Margin excl fees	NIM	NIM excl fees	Fee Yield	Credit Risk	Liquidity Ratio 1	Risk Aversion	Operating Costs	Bank Size 1	Maturity Transformation	Industry Concentration	Required Reserve Ratio	Inflation Quarterly	Production Index Growth	Copper Price	USD Kwacha	Treasury Rate	Interest Rate Volatility
Loan Margin	1																		
Margin excl fees	0.5995*	1																	
NIM	0.2568*	0.6278*	1																
NIM excl fees	0.2865*	0.6853*	0.9195*	1															
Fee Yield	0.7848*	0.0538	-0.1035*	-0.1066*	1														
Credit Risk	-0.2234*	-0.0083	-0.1274*	-0.0719*	-0.1023*	1													
Liquidity Ratio 1	0.0773*	0.1092*	0.2904*	0.2546*	0.0522	-0.1723*	1												
Risk Aversion	0.0771*	0.0165	0.2895*	0.2304*	0.0212	-0.4659*	0.5629*	1											
Operating Costs	0.3098*	0.3218*	0.0588	0.0889*	0.2700*	0.3365*	-0.0716*	-0.2886*	1										
Bank Size 1	-0.0893*	0.0171	0.2551*	0.2644*	-0.0647	0.0737*	-0.2471*	-0.3566*	-0.0311	1									
Maturity Transformation	0.0173	-0.1391*	-0.2153*	-0.1953*	0.0820*	-0.2321*	-0.3948*	-0.3077*	-0.0437	0.4170*	1								
Industry Concentration	0.2276*	0.2846*	0.3993*	0.3530*	0.0774*	-0.0987*	0.1318*	0.0282	0.1539*	0.1706*	-0.0856*	1							
Required Reserve Ratio	-0.0473	-0.0476	0.0276	0.0201	0.0008	0.0537	-0.0516	0.1189*	-0.0675	0.0092	-0.0842*	-0.1059*	1						
Inflation Quarterly	-0.0143	-0.0048	0.0285	0.0076	-0.0114	0.0658	-0.0528	-0.0327	0.0074	0.0198	0.0051	0.1024*	0.2449*	1					
Production Index Growth	0.1035*	0.1093*	0.1330*	0.0934*	0.0382	-0.0503	0.0048	-0.0043	0.0489	0.0467	-0.0078	0.2594*	-0.1376*	0.0391	1				
Copper Price	-0.0145	0.0077	-0.1163*	-0.1024*	-0.019	0.0021	0.0647	-0.1101*	0.0136	-0.0358	0.0501	-0.0940*	-0.6217*	-0.2810*	-0.0843*	1			
USD Kwacha	-0.0878*	-0.0882*	-0.0795*	-0.0634	-0.0413	0.0718*	-0.0054	-0.0174	-0.0325	-0.028	0.0054	-0.1258*	0.0461	-0.0411	-0.2857*	-0.1483*	1		
Treasury Rate	-0.1362*	-0.1714*	-0.1503*	-0.1256*	-0.0464	0.1186*	-0.1729*	0.0495	-0.1005*	-0.0582	-0.0047	-0.4049*	0.6588*	0.1603*	-0.1412*	-0.6649*	0.0921*	1	
Interest Rate Volatility	0.036	0.0709*	0.1618*	0.1111*	0.0163	-0.0281	0.0831*	0.0623	0.0324	0.0459	-0.0722*	0.1856*	0.3314*	0.2422*	0.2185*	-0.2801*	-0.2615*	0.0643	1

*Note: * shows significane at a 5% signifincal level or smaller*

Summary Statistics

Variable	Measurement	Obs	Mean	Median	Std.Dev	Coef.Var	Min	Max
Loan Margin		775	0.06	0.05	0.02	0.44	0.01	0.20
Margin excl fees		775	0.03	0.03	0.01	0.42	-0.01	0.09
NIM		775	0.09	0.09	0.04	0.40	-0.06	0.29
NIM excl fees		775	0.02	0.02	0.01	0.43	-0.05	0.08
Fee Yield		775	0.02	0.02	0.01	0.67	-0.01	0.09
Credit Risk		774	0.55	0.56	0.17	0.30	0.15	1.28
Liquidity Ratio 1		775	0.46	0.40	0.23	0.49	0.07	1.45
Risk Aversion		774	0.38	0.27	0.34	0.88	-0.78	2.01
Operating Costs		775	0.02	0.02	0.01	0.45	0.00	0.06
Bank Size 1		775	0.07	0.03	0.08	1.15	0.00	0.40
Maturity Transformation		765	0.55	0.33	11.31	20.40	-28.46	45.43
Industry Concentration		55	0.16	0.14	0.04	0.24	0.12	0.23
Required Reserve Ratio		55	10.01	8.00	4.78	0.48	5.00	18.00
Inflation Quarterly		55	0.023	0.021	0.018	0.816	-0.001	0.130
Production Index Growth		49	0.01	0.01	0.02	1.75	-0.04	0.05
Copper Price		55	6.56	6.79	1.45	0.22	3.06	9.53
USD Kwacha		55	0.02	0.01	0.10	5.91	-0.26	0.41
Treasury Rate		55	0.16	0.15	0.05	0.33	0.05	0.28
Interest Rate Volatility		55	1.56	0.84	1.73	1.11	0.04	7.26

Summary Statistics By Bank Type

	Large Diversified Banks			Small or Corporate Banks		
	Obs.	Mean	Median	Obs.	Mean	Median
Margin	361	0.053	0.050	414	0.056	0.051
Margin excl fees	361	0.033	0.031	414	0.030	0.030
NIM	361	0.099	0.096	414	0.082	0.083
NIM excl fees	361	0.026	0.025	414	0.021	0.021
Fee Yield	361	0.017	0.015	414	0.018	0.015
Credit Risk	361	0.570	0.556	413	0.538	0.573
Liquidity Ratio 1	361	0.436	0.394	414	0.481	0.411
Risk Aversion	361	0.297	0.216	413	0.461	0.393
Operating Costs	361	0.022	0.021	414	0.021	0.021
Bank Size 1	361	0.131	0.144	414	0.017	0.014
Maturity Transformation	357	4.425	5.953	408	-2.832	-3.087

Annex 19 – Profit decomposition

Figure 49 –Profit and NIM Bank-Level Decomposition Components



Note: The horizontal black line provides the median margin component for banks in a given quarter. The colour bands provide decile ranges for banks' components. The red band in the middle, for example, shows the 40th-60th percentile range; the dark orange, the 30th-40th and 60th-70th percentiles, etc. Each margin component is from equation 7.3, but at the bank level. Source: BOZ Prudential Returns and Author's calculations.

Annex 20 – Fees, IRR and DL Difference Modelling

Another approach to investigating the imposition of the IRR is to estimate a distributed lag model, concentrating on the relationship between IRR imposition and changes in banks' margins:

$$\Delta Y_{it} = \alpha_i + \sum_{j=d}^{J} \beta^j . \Delta T + \sum_{l=1}^{L} \beta^l \Delta X_{it}^l + \sum_{m=1}^{M} \theta^m \Delta M_t^m + \varepsilon_{it} \qquad (7.\text{A})$$

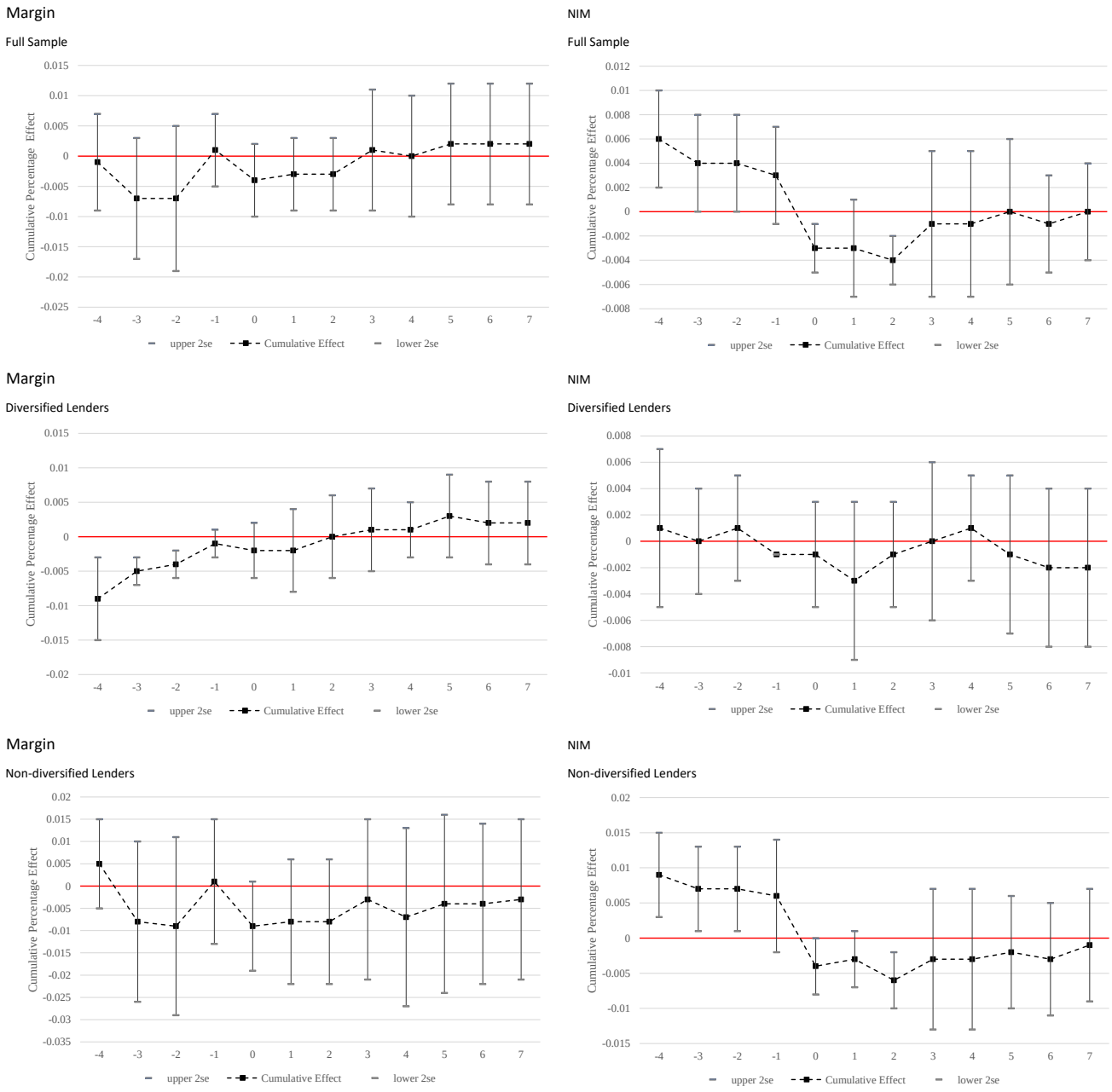
$$T = \begin{cases} 1, & t \geq 2013q1 \\ 0, & t < 2012q4 \end{cases}$$

Where Y_{it} is the percentage point change in banks' margins at time t, α_i is a lender fixed effect. T is the IRR indicator, indicating when a cap regime was in place or not. X_{it}^l and M_t^m are as in equation 7.6, but differenced.

In these regressions we are trying, as best possible given our small sample, to pick up any significant changes in margins that we might expect to see if the IRR regime had a strong short-term effect on margin levels.

As we can see in Figure 50 (which shows the cumulated effect of the imposition of the regime on margins and calculated from equation 7.A as $\sum_{j=\mathrm{d}}^{J} \widehat{\beta^j}$), this approach supports our primary margin levels findings in the text: there is little evidence that large banks' margins were directly affected by the cap, but smaller banks registered some negative change in NIM in the immediate periods after the cap was imposed.

Figure 50 – Distributed Lag apporach to IRRs and Bank Margins.



Note: Figures show the cumulated effect of the imposition of the regime on margins, calculated from equation 7.A as $\sum_{j=\mathrm{d}}^{J} \widehat{\beta^{j}}$, spanning over various lag structures, whereby each x-axis period provides an estimate from its own regression for each j. Controls include both lender-specific macroeconomic control variables, as per Table 24. The regressions include lender fixed effects. LHS figures use the "Margin" definition as per equation 7.1. RHS graphics use the NIM. Row 1 is for the full sample, and rows 2 and 3 are the sub-sample sets for Diversified and Non-diversified banks, respectively. Standard errors are robust and clustered at the lender level. Confidence interval at 95%.

Annex 21 – Fee Yield IRR Interactions

		Coefficient	Static		Dynamic (Short-run effect)	
			OLS	Bank FE	OLS	Bank FE
Fee Loan Income Share	Short Sample (All banks)	τa, Cap2013	-0.0415	-0.0367	-0.0187	-0.0318
		[se]	[0.0589]	[0.0532]	[0.0367]	[0.0289]
		τb, Interaction	0.2923	0.4113	-0.0029	0.2229
		[se]	[1.4813]	[1.1968]	[0.7604]	[1.0387]
		rho			0.3813**	0.2797***
	Full Sample (All Banks)	τa, Cap2013	-0.0084	0.022	0.0118	0.0216
		[se]	[0.0539]	[0.0436]	[0.0261]	[0.0280]
		τb, Interaction	0.0879	-0.4857	-0.3538	-0.7329
		[se]	[1.5658]	[1.0571]	[0.6414]	[0.8539]
		τc, Cap2014	0.1082	0.0316	0.01	0.0109
		[se]	[0.0700]	[0.0384]	[0.0207]	[0.0346]
		τd, Interaction	-1.579	-0.3246	0.0308	-0.0974
		[se]	[2.1034]	[0.8258]	[0.4070]	[0.9799]
		rho			0.4630***	0.4172***
	Diversified Banks (Full Time Peroid)	τa, Cap2013	-0.3144*	-0.2333	-0.0948	-0.0877*
		[se]	[0.1676]	[0.1775]	[0.0701]	[0.0479]
		τb, Interaction	8.9728**	6.1159	2.6517	2.4423*
		[se]	[4.2723]	[4.3899]	[1.7687]	[1.4428]
		τc, Cap2014	-0.1603	-0.0483	-0.0479	-0.0496
		[se]	[0.1492]	[0.1490]	[0.0618]	[0.0468]
		τd, Interaction	4.4577	0.5934	1.3415	1.3472
		[se]	[4.0515]	[3.7375]	[1.6578]	[1.4244]
		rho			0.6526***	0.6613***
	Not Diversified Banks (Full Time Peroid)	τa, Cap2013	-0.0286	0.0528	0.0378	0.0275
		[se]	[0.0540]	[0.0415]	[0.0305]	[0.0296]
		τb, Interaction	0.2557	-1.0262	-0.8478	-0.8353
		[se]	[1.3597]	[1.0756]	[0.7246]	[0.9294]
		τc, Cap2014	0.1242*	0.0633	0.0393	0.0268
		[se]	[0.0671]	[0.0373]	[0.0278]	[0.0408]
		τd, Interaction	-1.2941	-0.3872	-0.2788	-0.1438
		[se]	[1.9513]	[0.7296]	[0.4517]	[1.1222]
		rho			0.2779	0.2059***
Fee Yield	Short Sample (All banks)	τa, Cap2013	-0.0152**	-0.0073	-0.0022	-0.0033
		[se]	[0.0071]	[0.0061]	[0.0026]	[0.0026]
		τb, Interaction	0.3568**	0.1781	0.0126	0.0431
		[se]	[0.1496]	[0.1279]	[0.0586]	[0.0675]
		rho			0.6713***	0.7083***
	Full Sample (All Banks)	τa, Cap2013	-0.0161***	-0.0079	-0.0023	-0.0014
		[se]	[0.0053]	[0.0058]	[0.0020]	[0.0019]
		τb, Interaction	0.3411**	0.1484	0.0208	0.0122
		[se]	[0.1426]	[0.1391]	[0.0440]	[0.0483]
		τc, Cap2014	-0.0035	-0.0034	-0.0007	0.0002
		[se]	[0.0079]	[0.0060]	[0.0021]	[0.0020]
		τd, Interaction	0.1887	0.1146	0.0322	0.0013
		[se]	[0.2504]	[0.1709]	[0.0490]	[0.0509]
		rho			0.6247***	0.6500***
	Diversified Banks (Full Sample)	τa, Cap2013	-0.0335**	-0.0249	-0.0084	-0.0073*
		[se]	[0.0144]	[0.0156]	[0.0056]	[0.0043]
		τb, Interaction	0.8753**	0.5778	0.2141	0.1838
		[se]	[0.3564]	[0.3776]	[0.1467]	[0.1303]
		τc, Cap2014	-0.0138	-0.0073	-0.0054	-0.0047
		[se]	[0.0140]	[0.0155]	[0.0065]	[0.0039]
		τd, Interaction	0.2952	0.0652	0.1387	0.1183
		[se]	[0.3812]	[0.4238]	[0.2001]	[0.1228]
		rho			0.6874***	0.6827***
	Not Diversified Banks (Full Sample)	τa, Cap2013	-0.0193**	-0.0112	-0.0038	-0.0027
		[se]	[0.0080]	[0.0085]	[0.0028]	[0.0025]
		τb, Interaction	0.3680**	0.2642	0.0591	0.0413
		[se]	[0.1793]	[0.2056]	[0.0587]	[0.0668]
		τc, Cap2014	0.0014	-0.0052	-0.0014	-0.0009
		[se]	[0.0098]	[0.0064]	[0.0034]	[0.0031]
		τd, Interaction	0.1903	0.2813	0.086	0.0574
		[se]	[0.2984]	[0.1862]	[0.0789]	[0.0764]
		rho			0.5765***	0.6176***

Note:: Results from versions of equation 7.8 and 7.9 adjusted for fees (and not margins). Each set of τa, Cap2013, τb Interaction and τc, Cap2014, and τd Interaction (all within the same horizontal black lines) provide estimates are from the same regression. For example, the estimates -0.0084 to -1.579 from row 3 column 1 to row 6 column 1 are from the same static OLS regression using all banks and the "fee share" or fee "yield" as the dependent variable. I total the table provides results from 32 different regressions Results are provided for two different dependent variables: $Fee\ Share_{it} = \frac{Fees\ and\ commissions_{it}}{Loan\ Interest_{it}+Fees\ and\ commissions_{it}}$ and $Fee\ Yield_{it} = \frac{Fees\ and\ commissions_{it}}{Total\ Loans_{it}}$). "Static" results do not include a lagged dependent variable, Y_{it-l} in the estimation. Covariates are included in all regressions and are consistent with the set used in Table 26. Only short run effects are shown for the dynamic estimates. Rho is the coefficient on the lagged dependent variable (Share or Yield). τa, Cap2013 is the estimate for a dummy variable taking the value of 1 between Q1 2013 to Q2 2014, and 0 otherwise. τa, Cap2014 is the estimate for a dummy variable taking the value of 1 between Q3 2014 to Q4 2015, and 0 otherwise. The large "Diversified" Banks include the largest seven banks, all of whom have over 5% of market share (by loans and deposits) and maintain large branch networks. The "Not Diversified" banks make up the remainder. The "Short Sample" finishes in Q2 2014. Robust standard errors used for static regressions and PCSE for dynamic. ***p<0.01, **p<0.05, *p<0.10.

References

Abadie, A., Diamond, A. and Hainmueller, A.J. (2010) 'Synthetic control methods for comparative case studies: Estimating the effect of California's Tobacco control program', *Journal of the American Statistical Association*, 105(490), pp. 493–505. Available at: https://doi.org/10.1198/jasa.2009.ap08746.

Abuka, C. *et al.* (2019) 'Monetary policy and bank lending in developing countries: Loan applications, rates, and real effects', *Journal of Development Economics*, 139(March), pp. 185–202. Available at: https://doi.org/10.1016/j.jdeveco.2019.03.004.

Adam, C. (1995) 'Fiscal adjustment, financial liberalization, and the dynamics of inflation: Some evidence from Zambia', *World Development*, 23(5), pp. 735–750. Available at: https://doi.org/10.1016/0305-750X(95)00012-2.

Adam, C.S. and Simpasa, A.M. (2010) 'The Economics of the Copper Price Boom in Zambia', *Zambia, Mining, and Neoliberalism*, pp. 59–90. Available at: https://doi.org/10.1057/9780230115590_3.

Admati, A.R. (2015) 'The compelling case for stronger and more effective leverage regulation in banking', *Journal of Legal Studies*, 43(S2), pp. S35–S61. Available at: https://doi.org/10.1086/677557.

Admati, A.R. (2016) 'The Missed Opportunity and Challenge of Capital Regulation', *National Institute Economic Review*, 235(1), pp. R4–R14. Available at: https://doi.org/10.1177/002795011623500110.

Admati, A.R. *et al.* (2013) 'Fallacies, Irrelevant Facts, and Myths in the Discussion of Capital Regulation: Why Bank Equity is Not Socially Expensive', *SSRN Electronic Journal* [Preprint]. Available at: https://doi.org/10.2139/ssrn.2349739.

Admati, A.R. *et al.* (2018) 'The Leverage Ratchet Effect', *Journal of Finance*, 73(1), pp. 145–198. Available at: https://doi.org/10.1111/jofi.12588.

Ahokpossi, C. (2013) 'Determinants of Bank Interest Margins in Sub-Saharan Africa', *IMF Working Papers*, 13(34), p. 1. Available at: https://doi.org/10.5089/9781475551136.001.

Aiba, D. *et al.* (2020a) 'The Impact of Interest Rate Cap Policy on the Lending Behavior of Microfinance Institutions in Cambodia: Evidence from millions of observations in the Credit Registry Database The Impact of Interest Rate Cap Policy on the Lending Behavior of Microfinance', (224).

Aiba, D. *et al.* (2020b) 'The Impact of Interest Rate Cap Policy on the Lending Behavior of Microfinance Institutions in Cambodia: Evidence from millions of observations in the

Credit Registry Database The Impact of Interest Rate Cap Policy on the Lending Behavior of Microfinance', pp. 1–51.

Ailola, D. (1988) *The Regulation of Commercial Banks in Zambia and Their Role in Development*. Warwick.

Aiyar, S. *et al.* (2014) 'The international transmission of bank capital requirements: Evidence from the UK', *Journal of Financial Economics*, 113(3), pp. 368–382. Available at: https://doi.org/10.1016/j.jfineco.2014.05.003.

Aiyar, S., Calomiris, C.W. and Wieladek, T. (2014) 'Does macro-prudential regulation leak? Evidence from a UK policy experiment', *Journal of Money, Credit and Banking*, 46(SUPPL.1), pp. 181–214. Available at: https://doi.org/10.1111/jmcb.12086.

Aiyar, S., Calomiris, C.W. and Wieladek, T. (2015) 'Bank Capital Regulation: Theory, Empirics, and Policy', *IMF Economic Review*, 63(4), pp. 955–983. Available at: https://doi.org/10.1057/imfer.2015.18.

Alessie, R., Weber, G. and Hochguertel, S. (2005) 'Consumer credit: Evidence from Italian micro data', *Journal of the European Economic Association*, 3(1), pp. 144–178. Available at: https://doi.org/10.1162/1542476053295340.

Allen, F. *et al.* (2012) 'Resolving the African Financial Development Gap':

Allen, L. (1988) 'The Determinants of Bank Interest Margins: A Note', *The Journal of Financial and Quantitative Analysis*, 23(2), pp. 231–235.

Alper, E. *et al.* (2020) 'Do interest rate controls work? Evidence from Kenya', *Review of Development Economics*, (July 2019), pp. 1–17. Available at: https://doi.org/10.1111/rode.12675.

Andrianova, S. *et al.* (2015) 'Why Do African Banks Lend So Little?', *Oxford Bulletin of Economics and Statistics*, 77(3), pp. 339–359. Available at: https://doi.org/10.1111/obes.12067.

Angbazo, L. (1997) 'Commercial bank net interest margins, default risk, interest-rate risk, and off-balance sheet banking', *Journal of Banking and Finance*, 21(1), pp. 55–87. Available at: https://doi.org/10.1016/S0378-4266(96)00025-8.

Bahaj, S. and Malherbe, F. (2020) 'The Forced Safety Effect: How Higher Capital Requirements Can Increase Bank Lending', *Journal of Finance*, 75(6), pp. 3013–3053. Available at: https://doi.org/10.1111/jofi.12958.

Bailey, D. and Katz, J.N. (2011) 'Implementing Panel-Corrected Standard Errors in R : The pcse Package', *Journal of Statistical Software*, 42(Code Snippet 1). Available at: https://doi.org/10.18637/jss.v042.c01.

Banerjee, A., Karlan, D. and Zinman, J. (2015) 'Six Randomized Evaluations of Microcredit: Introduction and Further Steps', *American Economic Journal: Applied Economics*, 7(1), pp. 1–21. Available at: https://doi.org/10.1257/app.20140287.

Bank of Zambia (2009) *Bank of Zambia Annual Report, 2009*.

Bank of Zambia (2012) *Bank of Zambia Annual Report 2012*, pp. 1–103.

Bassett, W., Demiralp, S. and Lloyd, N. (2014) 'Government Support of Banks and Bank Lending', *SSRN Electronic Journal* [Preprint]. Available at: https://doi.org/10.2139/ssrn.2483966.

Beck, N. (2001) 'Time-Series-Cross-Section Data: What Have We Learned in the Past Few Years?', *Annual Review of Political Science*, 4(1), pp. 271–293. Available at: https://doi.org/10.1146/annurev.polisci.4.1.271.

Beck, N. and Katz, J.N. (1996) 'Nuisance vs. Substance: Specifying and Estimating Time-Series-Cross-Section Models', *Political Analysis*, 6(1996), pp. 1–36.

Beck, N. and Katz, J.N. (2011) 'Modelling Dynamics in Time-Series-Cross-Section Political Economy Data', *SSRN* [Preprint]. Available at: https://doi.org/10.1146/annurev-polisci-071510-103222.

Beck, T. (2009) 'The Econometrics of Finance and Growth', in T.C. Mills and K. Patterson (eds) *Palgrave Handbook of Econometrics: Volume 2: Applied Econometrics*. London: Palgrave Macmillan UK, pp. 1180–1209. Available at: https://doi.org/10.1057/9780230244405_25.

Beck, T. (2011) *Financing Africa: through the crisis and beyond*. Washington DC: World Bank.

Beck, T. (2012) 'The Role of Finance in Economic Development: Benefits, Risks, and Politics', in D.C. Mueller (ed.) *The Oxford Handbook of Capitalism*. 1st edn. Oxford University Press, pp. 161–203. Available at: https://doi.org/10.1093/oxfordhb/9780195391176.013.0007.

Beck, T. (2013) 'Bank Financing for SMEs – Lessons from the Literature', *National Institute Economic Review*, 225, pp. R23–R38. Available at: https://doi.org/10.1177/002795011322500105.

Beck, T. (2015) 'Cross-Border Banking and Financial Deepening: The African Experience', *Journal of African Economies*, 24(suppl 1), pp. i32–i45. Available at: https://doi.org/10.1093/jae/eju028.

Beck, T. (2016) 'Financial Inclusion – measuring progress and progress in measuring'.

Beck, T. and Cull, R. (2013) *Banking in Africa*. The World Bank (Policy Research Working Papers). Available at: https://doi.org/10.1596/1813-9450-6684.

Beck, T. and Hesse, H. (2009) 'Why are interest spreads So high in Uganda?', *Journal of Development Economics*, 88(2), pp. 192–204. Available at: https://doi.org/10.1016/j.jdeveco.2008.07.004.

Beck, T. and Maimbo, S.M. (2013) *Financial Sector Development in Africa - Opportunities and Challenges*, *The World Bank*.

Beck, T., Da-Rocha-Lopes, S. and Silva, A.F. (2021) 'Sharing the Pain? Credit Supply and Real Effects of Bank Bail-ins', *The Review of Financial Studies*. Edited by F. Cornelli, 34(4), pp. 1747–1788. Available at: https://doi.org/10.1093/rfs/hhaa067.

Beck, T., Jones, E. and Knaack, P. (2019) *Mind the Gap: Making Basel Standards Work for Developing Countries*. Paper submitted to International Financial Architecture for Stability and Development/ Crypto-Assets and Fintech. Available at: https://t20japan.org/wp-content/uploads/2019/03/t20-japan-tf2-13-mind-the-gap-basel-standards-work-for-developing-countries.pdf.

Beck, T., Levine, R. and Levkov, A. (2010) 'Big bad banks? The winners and losers from bank deregulation in the United States', *Journal of Finance*, 65(5), pp. 1637–1667. Available at: https://doi.org/10.1111/j.1540-6261.2010.01589.x.

Beck, T., Ross, L. and Loayza, N. (2000) 'Financial Intermediation and Growth: Causality and Causes', *Journal of Monetary Economics*, 46, pp. 31–77.

Berger, A.N. (1995) 'The Relationship between Capital and Earnings in Banking Author', *Journal of Money, Credit and Banking*, 27(2), pp. 432–456.

Berger, A.N. and Udell, G.F. (1994) 'Did Risk-Based Capital Allocate Bank Credit and Cause a " Credit Crunch " in the United States?', *Journal of Money, Credit and Banking*, 26(3), pp. 585–628.

Berger, A.N., Makaew, T. and Roman, R.A. (2019) 'Do Business Borrowers Benefit from Bank Bailouts?: The Effects of TARP on Loan Contract Terms', *Financial Management*, 48(2), pp. 575–639. Available at: https://doi.org/10.1111/fima.12222.

Berger, A.N., Roman, R.A. and Sedunov, J. (2020) 'Did TARP reduce or increase systemic risk? The effects of government aid on financial system stability', *Journal of Financial Intermediation*, 43. Available at: https://doi.org/10.1016/j.jfi.2019.01.002.

Bernanke, B. and Gertler, M. (1989) 'Agency Costs, Net Worth, and Business Fluctuations', *American Economic Review*, 79(1), pp. 14–31.

Bernanke, B.S. and Lown, C.S. (1991) 'The Credit Crunch', *Brookings Papers on Economic Activity*, 1991(2), p. 205. Available at: https://doi.org/10.2307/2534592.

Berrospide, J.M. and Edge, R.M. (2010) 'The effects of bank capital on lending: What do we know, and what does it mean?', *International Journal of Central Banking*, 6(4), pp. 5–54. Available at: https://doi.org/10.17016/feds.2010.44.

Bhattacharya, S., Boot, A. and Thakor, A. (1998) 'The Economics of Bank Regulation', *Journal of Money, Credit and Banking*, 30(4), pp. 745–770.

BIS (2010) *An Assessment of the Long-term Economic Impact of Stronger Capital and Liquidity Requirements, Bank for International Settlements*, pp. 1–69.

Bologna, P. (2018) 'Banks' maturity transformation: risk, reward, and policy', *European Systematic Risk Board Working Paper*, 63(2018). Available at: https://doi.org/10.1093/jnci/djv313.

Boutin-Dufresne, F. *et al.* (2014) 'Benchmarking Banking Sector Efficiency Across Regional Blocks in Sub-Saharan Africa: What Room for Policy?', *IMF Working Papers*, 13(51), p. 1. Available at: https://doi.org/10.5089/9781557753304.001.

Bradley, M.G., Jarrel, G.A. and Kim, E.H. (1984) 'On the Existence of an Optimal Capital Structure: Theory and Evidence', *The Journal of Finance*, 39(3), pp. 857–878. Available at: https://doi.org/10.1111/j.1540-6261.1984.tb03680.x.

Brei, M., Gambacorta, L. and Peter, G. von (2013) 'Rescue packages and bank lending', *Journal of Banking and Finance*, 37(2), pp. 490–505. Available at: https://doi.org/10.1016/j.jbankfin.2012.09.010.

Breusch, T.S. and Pagan, A.R. (1979) 'A Simple Test for Heteroscedasticity and Random Coefficient Variation', *Econometrica*, 47(5), pp. 1287–1294.

Breusch, T.S. and Pagan, A.R. (1980) 'The Lagrange Multiplier Test and its Applications to Model Specification in Econometrics', *The Review of Economic Studies*, 47(1), p. 239. Available at: https://doi.org/10.2307/2297111.

Brock, P.L. and Suarez, L.R. (2000) 'Understanding the Behaviour of bank spreads in Latin America', *Journal of Development Economics*, 63(1), pp. 113–134. Available at: https://doi.org/10.1016/S0304-3878(00)00102-4.

Brownbridge, M. (1998a) 'Financial distress in local banks in Kenya, Nigeria, Uganda and Zambia: causes and implications for regulatory policy', *Development Policy Review*, 16(2), pp. 173–188. Available at: https://doi.org/10.1111/1467-7679.00057.

Brownbridge, M. (1998b) 'The causes of financial distress in local banks in Africa and implications for prudential policy', *UNCTAD Discussion Papers* [Preprint], (March). Available at: https://doi.org/10.1111/1467-7679.00057.

Caballero, R.J., Hoshi, T. and Kashyap, A.K. (2008) 'Zombie lending and depressed restructuring in Japan', *American Economic Review*, 98(5), pp. 1943–1977. Available at: https://doi.org/10.1257/aer.98.5.1943.

Calani, M., García, P. and Oda, D. (2010) 'Supply and demand identification in the credit market', *Central Bank of Chile Working Papers*, 2010(571).

Calice, P. (2010) 'A Preliminary Assessment of the Implications of Financial Regulatory Reform for African Countries'.

Calice, P., Kalan, F.D. and Masetti, O. (2020) 'Interest Rate Repression: A New Database', (October). Available at: http://elibrary.worldbank.org/doi/book/10.1596/1813-9450-9457.

Callaway, B. and Sant'Anna, P.H.C. (2020) 'Difference-in-Differences with multiple time periods', *Journal of Econometrics* [Preprint], (December). Available at: https://doi.org/10.1016/j.jeconom.2020.12.001.

Caprio, J. and Honohan, P. (1999) 'Restoring banking stability: Beyond supervised capital requirements', *Journal of Economic Perspectives*, 13(1), pp. 43–64. Available at: https://doi.org/10.1111/j.1813-6982.2000.tb01159.x.

Central Bank of Nigeria (2013) 'Bank Intermediation in Nigeria: Growth, Competition, and Performance of The Banking Industry, 1990 – 2010'. Central Bank of Nigeria.

Chaisemartin, C. de and D'Haultfœuille, X. (2020) 'Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects', *American Economic Review*, 110(9), pp. 2964–2996. Available at: https://doi.org/10.1257/aer.20181169.

Chaisemartin, C. de, D'Haultfœuille, X., and (2017) 'Two-Way Fixed Effects and Differences-in-Differences with Heterogeneous Treatment Effects: A Survey', *AJR. American journal of roentgenology*, 186(2), pp. 227–236.

Cheesman, N. and Hinfelaar, M. (2010) 'Parties, Platforms, and Political Mobilization: The Zambian Presidential Election of 2008Nic Cheeseman and Marja Hinfelaar Published by : Oxford University Press on behalf of The Royal African Society Stable URL : https://www.jstor.org/stable/40388446 REF', *African Affairs*, 109(434), pp. 51–76. Available at: https://doi.org/10.1093/afraf7adp070.

Chirwa, E.W. and Mlachila, M. (2004) 'Financial Reforms and Interest Rate Spreads in the Commercial Banking System in Malawi', *IMF Working Papers*, 02(6), p. 1. Available at: https://doi.org/10.5089/9781451842302.001.

Chiuri, M.C., Ferri, G. and Majnoni, G. (2002) 'The macroeconomic impact of bank capital requirements in emerging economies: Past evidence to assess the future', *Journal of Banking and Finance*, 26(5), pp. 881–904. Available at: https://doi.org/10.1016/S0378-4266(01)00267-9.

Cho, Y.J. (2002) *Financial Repression, Liberalization, Crisis and Restructuring: Lessons of Korea's Financial Sector Policies*. Available at: www.adbi.org.

Choi, I. (2001) 'Unit root tests for panel data', *Journal of International Money and Finance*, 20(2), pp. 249–272. Available at: https://doi.org/10.1016/S0261-5606(00)00048-6.

Chu, Y., Zhang, D. and Zhao, Y.E. (2019) 'Bank Capital and Lending: Evidence from Syndicated Loans', *Journal of Financial and Quantitative Analysis*, 54(2), pp. 667–694. Available at: https://doi.org/10.1017/S0022109018000698.

Claessens, S. and Horen, N. van (2012) 'Foreign Banks: Trends, Impact and Financial Stability; by Stijn Claessens and Neeltje van Horen; IMF Working Paper 12/10; January 1, 2012', *IMF Working Paper*, 10, pp. 1–36.

Claessens, S. and Laeven, L. (2004) 'What Drives Bank Competition? Some International Evidence', *Journal of Money, Credit and Banking*, 36(3), pp. 563–583.

Claessens, S., Demirgüç-Kunt, A. and Huizinga, H. (2001) 'How does foreign entry affect domestic banking markets?', *Journal of Banking and Finance*, 25(5), pp. 891–911. Available at: https://doi.org/10.1016/S0378-4266(00)00102-3.

Coco, G. and Meza, D.D.E. (2009) 'In Defense of Usury Laws', *Journal of Money, Credit and Banking*, 41(8), pp. 1691–1703.

Cohen, B.H. and Scatigna, M. (2016) *Banks and capital requirements: Channels of adjustment, Journal of Banking and Finance*. Available at: https://doi.org/10.1016/j.jbankfin.2015.09.022.

Collier, P. and Bates, R.H. (1995) 'The Politics and Economics of Policy Reform in Zambia', *Journal of African Economies*, 4(1), pp. 115–143.

Cook, L.D. (2011) 'Were the Nigerian Banking Reforms of 2005 a Success ... and for the Poor?', *National Bureau of Economic Research*, III(September), pp. 157–180.

Courakis, A.S. (1984) 'Constraints on Bank Choices and Financial Repression in Less Developed Countries', *Oxford Bulletin of Economics and Statistics*, 46(4), pp. 341–370. Available at: https://doi.org/10.1111/j.1468-0084.1984.mp46004005.x.

Crick, Wilfred (1965) *Commonwealth banking systems*. Oxford, Clarendon Press.

Cubillos-rocha, J.S. *et al.* (2021) 'Effects of interest rate caps on credit access', *Journal of Regulatory Economics* [Preprint], (December). Available at: https://doi.org/10.1007/s11149-021-09437-0.

Cuesta, J.I. and Sepulveda, A. (2017) 'Price Regulation in Credit Markets: A Trade-off between Consumer Protection and Credit Access'.

Cull, R., Demirgüc-Kunt, A. and Morduch, J. (2007) 'Financial Performance and Outreach: A Global Analysis of Leading Microbanks', *The Economic Journal*, 117(February). Available at: http://www.mixmarket.org.

Cull, R., Demirgüç-Kunt, A. and Morduch, J. (2009) 'Microfinance meets the market', *Journal of Economic Perspectives*, 23(1), pp. 167–192. Available at: https://doi.org/10.1257/jep.23.1.167.

Dang, V.A., Kim, M. and Shin, Y. (2015) 'In search of robust methods for dynamic panel data models in empirical corporate finance', *Journal of Banking and Finance*, 53, pp. 84–98. Available at: https://doi.org/10.1016/j.jbankfin.2014.12.009.

Dasgupta, K. and Mason, B.J. (2020) 'The effect of interest rate caps on bankruptcy : Synthetic control evidence from recent payday lending bans', *Journal of Banking and Finance*, 119, p. 105917. Available at: https://doi.org/10.1016/j.jbankfin.2020.105917.

Degryse, H. *et al.* (2019) 'Identifying credit supply shocks with bank-firm data: Methods and applications', *Journal of Financial Intermediation*, 40(April), p. 100813. Available at: https://doi.org/10.1016/j.jfi.2019.01.004.

Demetriades, P.O. and Luintel, K.B. (2001) 'Financial restraints in the South Korean miracle', *Journal of Development Economics*, 64(2), pp. 459–479. Available at: https://doi.org/10.1016/S0304-3878(00)00146-2.

Demirguc-Kunt, A. and Huizinga, H. (1999) 'Determinants of Commerical Bank Interest Margins and Profitability', *The World Bank Economic Review*, 13(2 (2013)), pp. 1–408. Available at: https://doi.org/10.1093/wber/13.2.379.

Dermine, J. (1986) ‘Deposit rates, credit rates and bank capital. The Klein-Monti Model Revisited’, *Journal of Banking and Finance*, 10(1), pp. 99–114. Available at: https://doi.org/10.1016/0378-4266(86)90022-1.

Detragiache, E., Gupta, P. and Tressel, T. (2006) ‘Foreign Banks in Poor Countries: Theory and Evidence’, *IMF Working Papers*, 06(18), p. 1. Available at: https://doi.org/10.5089/9781451862782.001.

Diamond, D.W. (1996) ‘Financial intermediation as delegated monitoring: A simple example’, *Federal Reserve Bank of Richmond Economic Quarterly*, 82(3), pp. 51–66.

Dixon, R., Ritchie, J. and Siwale, J. (2007) ‘Loan officers and loan “delinquency” in Microfinance: A Zambian case’, *Accounting Forum*, 31(1), pp. 47–71. Available at: https://doi.org/10.1016/j.accfor.2006.11.005.

Driscoll, J.C. and Kraay, A.C. (1998) ‘Consistent covariance matrix estimation with spatially dependent panel data’, *Review of Economics and Statistics*, 80(4), pp. 549–559. Available at: https://doi.org/10.1162/003465398557825.

Duchin, R. and Sosyura, D. (2014) ‘Safer ratios, riskier portfolios: Banks’ response to government aid’, *Journal of Financial Economics*, 113(1), pp. 1–28. Available at: https://doi.org/10.1016/j.jfineco.2014.03.005.

Elliott, G., Rothenberg, T.J. and Stock, J.H. (1996) ‘Efficient Tests for an Autoregressive Unit Root’, *Econometrica*, 64(4), pp. 813–836. Available at: https://doi.org/10.2307/2171846.

Entrop, O. *et al.* (2015) ‘Determinants of bank interest margins: Impact of maturity transformation’, *Journal of Banking and Finance*, 54, pp. 1–19. Available at: https://doi.org/10.1016/j.jbankfin.2014.12.001.

Faherty, M., Mccarthy, O. and Byrne, N. (2017) *Interest Rate Restrictions on Credit for Low-income Borrowers on behalf of the Social Finance Foundation*.

Fang, X. *et al.* (2018) ‘The Effects of Higher Bank Capital Requirements on Credit in Peru’, *IMF Working Papers*, 18(222), p. 1. Available at: https://doi.org/10.5089/9781484378366.001.

Fang, X. *et al.* (2020) ‘Bank capital requirements and lending in emerging markets: The role of bank characteristics and economic conditions’, *Journal of Banking and Finance*, p. 105806. Available at: https://doi.org/10.1016/j.jbankfin.2020.105806.

Federal Reserve Bank of Chicago. (2001) *Controlling Interest: Are Ceilings On Interest Rates a Good Idea?* Federal Reserve Bank of Chicago.

Fekrazad, A. (2020) ‘Impacts of interest rate caps on the payday loan market : Evidence from Rhode Island’, *Journal of Banking and Finance*, 113, p. 105750. Available at: https://doi.org/10.1016/j.jbankfin.2020.105750.

Ferrari, A., Masetti, O. and Ren, J. (2018) ‘Interest Rate Caps: The Theory and The Practice’, *World Bank Policy Research Working Papers*, 8398(April), p. 42. Available at: https://doi.org/doi:10.1596/1813-9450-8398.

Feyen, E. and Huertas, I.Z. (2020) *Bank Lending Rates and Spreads in EMDEs Evolution, Drivers, and Policies*. Available at: http://www.worldbank.org/prwp.

Finance, N.C. on C. (1972) *Consumer Credit in the United States: Report by United States., 1972*.

Finmark Trust (2009) *FinScope Zambia 2009*.

Finmark Trust (2015) *FinScope Zambia 2015*.

Flamini, V., Mcdonald, C. and Schumacher, L. (2009) ‘The Determinants of Commercial Bank Profitability in Sub-Saharan Africa’, *IMF Working Paper* [Preprint].

Francis, W.B. and Osborne, M. (2012) ‘Capital requirements and bank behaviour in the UK: Are there lessons for international capital standards?’, *Journal of Banking and Finance*, 36(3), pp. 803–816. Available at: https://doi.org/10.1016/j.jbankfin.2011.09.011.

Francis, W.B. and Osbourne, M. (2010) ‘On the Behavior and Determinants of Risk-Based Capital Ratios: Revisiting the Evidence from UK Banking Institutions’, *International Review of Finance*, 10(4), pp. 485–518. Available at: https://doi.org/10.1111/j.1468-2443.2010.01112.x.

Freixas, X. and Rochet, J.-C. (2008) *Microeconomics of Banking, Second Edition*. 2nd edn. The MIT Press.

Freyaldenhoven, S. *et al.* (2021) ‘Visualization, Identification, and Estimation in the Linear Panel Event-Study Design’, *SSRN Electronic Journal* [Preprint]. Available at: https://doi.org/10.2139/ssrn.3909601.

Freyaldenhoven, S., Hansen, C. and Shapiro, J.M. (2019) ‘Pre-event trends in the panel event-study design’, *American Economic Review*, 109(9), pp. 3307–3338. Available at: https://doi.org/10.1257/aer.20180609.

Fuchs, M. *et al.* (2014) *Making cross-border banking work for Africa*. Available at: http://documents.worldbank.org/curated/en/2014/01/19761778/making-cross-border-banking-work-africa.

Furlong, F.T. and Kwan, S. (2006) 'Sources of Bank Charter Value', *Technical report, Federal Reserve Bank of San Francisco*, pp. 1–40.

Galbis, V. (1993) 'High Real Interest Rates Under Financial Liberalization: Is there a Problem?', *IMF Working Papers*, p. 1. Available at: https://doi.org/10.5089/9781451842364.001.

Galenianos, M. and Gavazza, A. (2022) 'Regulatory Interventions in Consumer Financial Markets: The Case of Credit Cards', *Journal of the European Economic Association*, 20(5), pp. 1897–1932. Available at: https://doi.org/10.1093/jeea/jvac016.

Gambacorta, L. (2005) 'Inside the bank lending channel', *European Economic Review*, 49(7), pp. 1737–1759. Available at: https://doi.org/10.1016/j.euroecorev.2004.05.004.

Gambacorta, L. and Marques-Ibanez, D. (2011) 'BIS Working Papers The bank lending channel: Lessons from the crisis', *Economic Policy*, 26.

Gambacorta, L., Mistrulli, P.E. and Requirements, C. (2014) 'Bank Capital and Lending Behaviour : Empirical Evidence for Italy', (May).

Ghosh, P., Mookherjee, D. and Ray, D. (2000) 'Credit Rationing in Developing Countries', *A Reader in Development Economics* [Preprint], (December 1999).

Gobat, J., Yanase, M. and Maloney, J. (2014) 'The Net Stable Funding Ratio: Impact and Issues for Consideration', *IMF Working Papers*, 14(106), p. 1. Available at: https://doi.org/10.5089/9781498346498.001.

Godfrey, L.G. (1978) 'Testing Against General Autoregressive and Moving Average Error Models when the Regressors Include Lagged Dependent Variables', *Econometrica*, 46(6), pp. 1293–1301.

Gondwe, M. (2013) *Developing economic and banking opportunities in Zambia*. BIS, pp. 1–4.

Goodman-Bacon, A. (2021) 'Difference-in-differences with variation in treatment timing', *Journal of Econometrics* [Preprint]. Available at: https://doi.org/10.1016/j.jeconom.2021.03.014.

Gropp, R. *et al.* (2019a) 'Bank Response To Higher Capital Requirements: Evidence form a Quasi-Natural Experiment'.

Gropp, R. *et al.* (2019b) 'Banks response to higher capital requirements: Evidence from a quasi-natural experiment', *Review of Financial Studies*, 32(1), pp. 266–299. Available at: https://doi.org/10.1093/rfs/hhy052.

Gulhati, R. (1991) *Impasse in Zambia, Public Administration and Development*. Available at: https://doi.org/10.1002/pad.4230110311.

Haan, W.J. den, Sumner, S.W. and Yamashiro, G.M. (2007) 'Bank loan portfolios and the monetary transmission mechanism', *Journal of Monetary Economics*, 54(3), pp. 904–924. Available at: https://doi.org/10.1016/j.jmoneco.2006.01.008.

Haan, W.J.D. *et al.* (2009) 'Bank Loan Portfolios and the Canadian Monetary Transmission Mechanism', 42(3), pp. 1150–1175.

Haan, W.J.D. *et al.* (2010) 'Bank Loan Components and the Time-varying Effects of Monetary Policy Shocks Published by : Wiley on behalf of The London School of Economics and Political Science and The Suntory and Toyota International Centres for Economics and Related Disciplines', 78(312), pp. 593–617.

Hahn, J. and Kuersteiner, G. (2002) 'Asymptotically Unbiased Inference for a Dynamic Panel With Fixed Effects When Both n and T are Large', *Econometrica*, 70(4), pp. 1639–1657. Available at: https://doi.org/10.1086/658342.

Hancock, D. and Wilcox, J.A. (1994) 'Bank Capital and the Credit Crunch: The Roles of Risk-Weighted and Unweighted Capital Regulations', *Real Estate Economics*, 22(1), pp. 59–94. Available at: https://doi.org/10.1111/1540-6229.00626.

Hancock, D., Laing, A.J. and Wilcox, J.A. (1995) 'Bank capital shocks: Dynamic effects on securities, loans, and capital', *Journal of Banking and Finance*, 19(3–4), pp. 661–677. Available at: https://doi.org/10.1016/0378-4266(94)00147-U.

Hart, O.D. and Jaffee, D.M. (1974) 'On the application of portfolio theory to depository financial intermediaries', *Review of Economic Studies*, 41(1), pp. 129–147. Available at: https://doi.org/10.2307/2296404.

Harvey, C. (1973) 'The Control of Credit in Zambia', *The Journal of Modern African Studies*, 11(3), pp. 383–392. Available at: https://doi.org/10.1017/S0022278X00008491.

Hashimzade, N., Thornton, M. and Baltagi, B.H. (2013) 'Dynamic panel data models', *Handbook of Research Methods and Applications in Empirical Macroeconomics*, pp. 229–248. Available at: https://doi.org/10.4337/9780857931023.00016.

Hawtrey, K. and Liang, H. (2008) 'Bank interest margins in OECD countries', *North American Journal of Economics and Finance*, 19(3), pp. 249–260. Available at: https://doi.org/10.1016/j.najef.2008.07.003.

Hendry, D.F. and Mizon, G.E. (1978) 'Serial Correlation as a Convenient Simplification, Not a Nuisance: A Comment on a Study of the Demand for Money by the Bank of England', *The Economic Journal*, 88(351), p. 549. Available at: https://doi.org/10.2307/2232053.

Heng, D. (2015) 'Impact of the New Financial Services Law in Bolivia on Financial Stability and Inclusion', *IMF Working Paper* [Preprint], (267).

Heng, D., Chea, S. and Heng, B. (2021) 'Impacts of Interest Rate Cap on Financial Inclusion in Cambodia', *IMF Working Papers*, 2021(107), p. 1. Available at: https://doi.org/10.5089/9781513582634.001.

Ho, T. and Anthony, S. (1981) 'The Determinants of Bank Interest Margins: Theory and Empirical Evidence', *The Journal of Financial and Quantitative Analysis*, 16(4), pp. 581–600.

Hoechle, D. (2007) 'Robust standard errors for panel regressions with cross-sectional dependence'.

Honohan, P. and Beck, T. (2007) *Making Finance Work for Africa*. The World Bank. Available at: https://doi.org/10.1596/978-0-8213-6909-8.

Huybens, E. and Smith, B.D. (1999) *Inflation, financial markets and long-run real activity*, *Journal of Monetary Economics*, pp. 283–315.

Im, K.S., Pesaran, M.H. and Shin, Y. (2003) 'Testing for unit roots in heterogeneous panels', *Journal of Econometrics*, 115(1), pp. 53–74. Available at: https://doi.org/10.1016/S0304-4076(03)00092-7.

IMF (1998) Report of the Group of Independent Persons Appointed to Conduct an Evaluation of Certain Aspects of the Enhanced Structural Adjustment Facility - Country Profiles.

IMF (2017) Financial System Stability Assessment for Zambia, *IMF Country Report*, p. 42.

IMF. (2004) Zambia: 2003 Article IV Consultation and Ex Post Assessment of Performance Under Fund-Supported Programs–Staff Reports; Staff Statement; Public Information Notice on the Executive Board Discussion; and Statement by the Executive Director for Zambia, *IMF Staff Country Reports*, p. i. Available at: https://doi.org/10.5089/9781451921601.002.

IMF. (2006) Zambia: 2005 Article IV Consultation, Third Review Under the Three-Year Arrangement Under the Poverty Reduction and Growth Facility, Request for Waiver and Modification of Performance Criterion, and Financing Assurances Review*: Staff Report; Public Inform*, *IMF Staff Country Reports*, p. i. Available at: https://doi.org/10.5089/9781451841275.002.

IMF. (2012) 'Zambia: 2012 Article IV Consultation', *IMF Staff Country Reports*, 13(40), p. 1. Available at: https://doi.org/10.5089/9781616356613.002.

IMF. (2016) ‘Nigeria: 2016 Article IV Consultation- Press Release; Staff Report; and Statement by the Executive Director for Nigeria’, *IMF Staff Country Reports*, 16(101), p. 1. Available at: https://doi.org/10.5089/9781484320037.002.

IMF. (2017) ‘Zambia: Selected Issues; IMF Country Report No. 17/328; October 3, 2017’, (17), pp. 0–62.

IMF. (2019) *Zambia - 2019 Article IV Consultation-Press Release; Staff Report; and Statement*, *IMF Staff Country Reports*, p. 1. Available at: https://doi.org/10.5089/9781498316972.002.

Intermediation, J.F. *et al.* (2013) ‘Capital ratios and bank lending: A matched bank approach’, *Journal of Financial Intermediation*, 22(4), pp. 663–687. Available at: https://doi.org/10.1016/j.jfi.2013.06.003.

Jadlow, J.M. (1977) ‘Adam Smith on Usury Laws’, *The Journal of Finance*, 32(4), pp. 1195–1200.

Jiménez, G. *et al.* (2012) ‘Credit supply and monetary policy: Identifying the bank balance-sheet channel with loan applications’, *American Economic Review*, 102(5), pp. 2301–2326. Available at: https://doi.org/10.1257/aer.102.5.2301.

Jiménez, G. *et al.* (2017) ‘Do Demand or Supply Factors Drive Bank Credit, In Good and Crisis Times?’, *SSRN* [Preprint], (April 2017). Available at: https://doi.org/10.2139/ssrn.1980139.

Jones, E. (2020) *The Political Economy of Bank Regulation in Developing Countries: Risk and Reputation*. Edited by O.U. Press. Oxford University Press. Available at: https://doi.org/10.1063/1.2914415.

Jones, E. and Knaack, P. (2019) ‘Global Financial Regulation: Shortcomings and Reform Options’, *Global Policy*, 10(2), pp. 193–206. Available at: https://doi.org/10.1111/1758-5899.12656.

Jones, E. and Zeitz, A.O. (2017) ‘The Limits of Globalizing Basel Banking Standards’, *Journal of Financial Regulation*, 3(1), pp. 89–124. Available at: https://doi.org/10.1093/jfr/fjx001.

Jones, E. and Zeitz, A.O. (2019) ‘Regulatory Convergence in the Financial Periphery: How Interdependence Shapes Regulators’ Decisions’, *International Studies Quarterly*, pp. 1–15. Available at: https://doi.org/10.1093/isq/sqz068.

Jones, E., Beck, T. and Knaack, P. (2018a) *Mind the Gap: Making Basel Standards Work for Developing Countries*. Blavatnik School of Government.

Jones, E., Beck, T. and Knaack, P. (2018b) *Room to Manoeuvre: How Developing Countries Can Tailor Basel Standards*. Global Economic Governance and Blavatnik School of Government.

Kashyap, A.K. and Stein, J.C. (2000) 'What Do a Million Observations on Banks Say about the Transmission of Monetary Policy?', *The American Economic Review*, 90(3), pp. 407–428.

Katz, J.N. (1995) 'What to do (and not to do) with Time-Series Cross-Section Data', *The American Political Science Review*, 89(3), pp. 634–647.

Kaunda, K.D. (1970) *Take Up the Challenge*. Zambia Information Services.

Khwaja, A.I. and Mian, A. (2008) 'Tracing the impact of bank liquidity shocks: Evidence from an emerging market', *American Economic Review*, 98(4), pp. 1413–1442. Available at: https://doi.org/10.1257/aer.98.4.1413.

King, M.R. (2013) 'The Basel III Net Stable Funding Ratio and bank net interest margins', *Journal of Banking and Finance*, 37(11), pp. 4144–4156. Available at: https://doi.org/10.1016/j.jbankfin.2013.07.017.

Klein, M. (1971) 'A Theory of the Banking Firm', *Journal of Money, Credit and Banking*, 3(2), pp. 205–218.

Kropko, J. and Kubinec, R. (2020) 'Interpretation and identification of within-unit and cross-sectional variation in panel data models', *PLoS ONE*, 15(4), pp. 1–22. Available at: https://doi.org/10.1371/journal.pone.0231349.

Larmer, M. (2005) 'Reaction & Resistance to Neo-Liberalism in Zambia', *Review of African Political Economy*, 32(103), pp. 29–45.

Larmer, M. and Fraser, A. (2007) 'Of cabbages and King Cobra: Populist politics and Zambia's 2006 election', *African Affairs*, 106(425), pp. 611–637. Available at: https://doi.org/10.1093/afraf/adm058.

Levine, R. and Zervos, S. (1998) 'Stock markets and Economic Growth: Stock Markets versus Banks ?', *The American Economic Review*, 88(3), pp. 537–558.

Li, L. (2013) 'TARP funds distribution and bank loan supply', *Journal of Banking and Finance*, 37(12), pp. 4777–4792. Available at: https://doi.org/10.1016/j.jbankfin.2013.08.009.

Madeira, C. (2019) 'The impact of interest rate ceilings on households' credit access: evidence from a 2013 Chilean legislation'.

Maderia, C. (2015) 'The Impact of Interest Rate Ceilings on households'credit access', *International Journal of Social Work*, 2(2), p. 10. Available at: https://doi.org/10.5296/ijsw.v2i2.7953.

Maimbo, S.M. (2002a) 'Explaining bank regulatory failure in Zambia', *Journal of International Development*, 14(2), pp. 229–248. Available at: https://doi.org/10.1002/jid.876.

Maimbo, S.M. (2002b) 'Explaining bank regulatory failure in Zambia', *Journal of International Development*, 14(2), pp. 229–248. Available at: https://doi.org/10.1002/jid.876.

Maimbo, S.M. and Gallegos, C.A.H. (2014) 'Interest Rate Caps around the World: Still Popular, but a Blunt Instrument', *World Bank Policy Research Working Paper*, 7070(October), pp. 1–37.

Maudos, J. and Guevara, J.F. de (2004) 'Factors explaining the interest margin in the banking sectors of the European Union', *Journal of Banking and Finance*, 28(9), pp. 2259–2281. Available at: https://doi.org/10.1016/j.jbankfin.2003.09.004.

Maudos, J. and Solís, L. (2009) 'The determinants of net interest income in the Mexican banking system: An integrated model', *Journal of Banking and Finance*, 33(10), pp. 1920–1931. Available at: https://doi.org/10.1016/j.jbankfin.2009.04.012.

Mbate, M. (2013) 'Domestic Debt, Private Sector Credit and Economic Growth in Sub-Saharan Africa', *African Development Review*, 25(4), pp. 434–446. Available at: https://doi.org/10.1111/1467-8268.12040.

McKinnon, R.I. (1973) *Money and capital in economic development*. Brookings Institution Press.

McShane, R.W. and Sharpe, I.G. (1985) 'A time series/cross section analysis of the determinants of Australian trading bank loan/deposit interest margins: 1962-1981', *Journal of Banking and Finance*, 9(1), pp. 115–136. Available at: https://doi.org/10.1016/0378-4266(85)90065-2.

Merton, M.H. (1977) 'Debt and Taxes', *The Journal of Finance*, 32(2), pp. 261–275.

Miller, H. (2013) 'Interest rate caps and their impact on financial inclusion', *Nathan Associates Policy Note* [Preprint], (February). Available at: http://partnerplatform.org/eps-peaks.

Miller, M.H. (1995) 'Do the M & M propositions apply to banks?', *Journal of Banking and Finance*, 19(3–4), pp. 483–489. Available at: https://doi.org/10.1016/0378-4266(94)00134-O.

Minogue, M. and Molly, J. (1974) *African Aims and Attitudes: Selected Documents*. Cambridge University Press.

Miyauchi, Y. (2014) 'Interest Rate Cap, Relationship Lending and Bank Competition in Bangladesh'.

Mlachila, M. *et al*. (2013) 'Banking in Sub-Saharan Africa: Challenges and Opportunities', *European Investment Bank, Economics Department in Luxembourg*, pp. 1–118.

Modigliani, F. and Miller, M.H. (1958) 'The Cost of Capital, Corporation Finance and the Theory of Investment', *American Economic Review*, 48(3), pp. 261–297. Available at: https://doi.org/10.1257/aer.103.1.i.

Modigliani, F. and Miller, M.H. (1963) 'Corporate Income Taxes and the Cost of Capital: A Correction', *American Economic Review*, 53(3), pp. 433–443. Available at: https://doi.org/10.2307/1809167.

Monti, M. (1972) *Deposit, credit and interest rate determination under alternative bank objective function*. North-Holland/American Elsevier.

Moundigbaye, M., Rea, W. and Reed, R. (2017) 'Which panel data estimator should I use?: A corrigendum and extension', (2017). Available at: http://www.economics-ejournal.org/economics/journalarticles/2018-4.

Mpofu, T.R. and Nikolaidou, E. (2018) 'Determinants of credit risk in the banking system in Sub-Saharan Africa', *Review of Development Finance*, 8(2), pp. 141–153. Available at: https://doi.org/10.1016/j.rdf.2018.08.001.

Mummolo, J. and Peterson, E. (2018) 'Improving the Interpretation of Fixed Effects Regression Results', *Political Science Research and Methods*, 6(04), pp. 829–835. Available at: https://doi.org/10.1017/psrm.2017.44.

Mwenda, A. and Mutoti, N. (2011) 'Financial Sector Reforms, Bank Performance and Economic Growth: Evidence from Zambia', *African Development Review*, 23(1), pp. 60–74. Available at: https://doi.org/10.1111/j.1467-8268.2010.00272.x.

Nakashima, K. (2016) 'An econometric evaluation of bank recapitalization programs with bank- and loan-level data', *Journal of Banking and Finance*, 63, pp. 1–24. Available at: https://doi.org/10.1016/j.jbankfin.2015.11.002.

Norden, L., Udell, G.F. and Wang, T. (2020) 'Do bank bailouts affect the provision of trade credit? 1', *Journal of Corporate Finance*, 60(September 2019), p. 101522. Available at: https://doi.org/10.1016/j.jcorpfin.2019.101522.

Noss, J. and Toffano, P. (2016) 'Estimating the impact of changes in aggregate bank capital requirements on lending and growth during an upswing', *Journal of Banking and Finance*, 62, pp. 15–27. Available at: https://doi.org/10.1016/j.jbankfin.2015.09.020.

Ogowewo, T.I. and Uche, C. (2006) '(MIS)USING BANK SHARE CAPITAL AS A REGULATORY TOOL TO FORCE BANK CONSOLIDATIONS IN NIGERIA', *Journal of African Law*, 50(2), pp. 161–186. Available at: https://doi.org/10.1017/S0021855306000143.

Ostas, J.R. (1976) 'Effects of Usury Ceilings in the Mortgage Market', *The Journal of Finance*, 31(3), pp. 821–834. Available at: https://doi.org/10.1111/j.1540-6261.1976.tb01925.x.

Papaioannou, E. (2007) 'Finance and Growth: A Macroeconomic Assessment of the Evidence from a European Angle', *SSRN Electronic Journal* [Preprint], (787). Available at: https://doi.org/10.2139/ssrn.1001457.

Paravisini, D. *et al.* (2015) 'Dissecting the Effect of Credit Supply on Trade: Evidence from Matched Credit-Export Data', *The Review of Economic Studies*, 82(1), pp. 333–359. Available at: https://doi.org/10.1093/restud/rdu028.

Peek, B.J.O.E. and Rosengren, E.S. (1997) 'The International Transmission of Financial Shocks : The Case of Japan', *American Economic Review*, 87(4), pp. 495–505.

Peek, J. and Rosengren, E. (1995) 'The Capital Crunch: Neither a Borrower nor a Lender Be', *Journal of Money, Credit and Banking*, 27(3), pp. 625–638.

Peek, J. and Rosengren, E.S. (2000) 'Collateral damage: Effects of the Japanese bank crisis on real activity in the United States', *American Economic Review*, 90(1), pp. 30–45. Available at: https://doi.org/10.1257/aer.90.1.30.

Peterson, C.L. (2008) 'Usury law, payday loans, and statutory sleight of hand: Salience distortion in American credit pricing limits', *Minnesota Law Review*, 92(4), pp. 1110–1164.

Peterson, R. (1983) 'Usury Laws and Consumer Credit: A Note', *The Journal of Finance*, 28(4), pp. 933–955.

Peydró, J.L. (2010) 'Discussion of "the effects of bank capital on lending what do we know and what does it mean"', *International Journal of Central Banking*, 6(4), pp. 55–69.

Pfleiderer, P.C. (2012) 'On the Relevancy of Modigliani and Miller to Banking: A Parable and Some Observations', *SSRN Electronic Journal* [Preprint], (93). Available at: https://doi.org/10.2139/ssrn.1727744.

Phillips, P.C.B. and Perron, P. (1988) 'Testing for a unit root in time series regression', *Biometrika*, 75(2), pp. 335–346. Available at: https://doi.org/10.1093/biomet/75.2.335.

Phillips, P.C.B. and Sul, D. (2003) 'Dynamic panel estimation and homogeneity testing under cross section dependence', *The Econometrics Journal*, 6(1), pp. 217–259. Available at: https://doi.org/10.1111/1368-423x.00108.

Piskadlo, D. (2018) 'The Politics (and Varied Experiences) of Interest Rate Caps | Center for Financial Inclusion'. Available at: https://www.centerforfinancialinclusion.org/the-politics-and-varied-experiences-of-interest-rate-caps.

Poghosyan, T. (2013) 'Financial intermediation costs in low income countries: The role of regulatory, institutional, and macroeconomic factors', *Economic Systems*, 37(1), pp. 92–110. Available at: https://doi.org/10.1016/j.ecosys.2012.07.003.

Porteous, D., Collins, D. and Abrams, J. (2010) *Policy Framing Note 4: Interest Rate Policy*. Financial Access Initiative Policy Framing Note.

Puddu, S. and Wälchli, A. (2013) 'TARP Effect on Bank Lending Behaviour: Evidence from the last Financial Crisis.', *Working Paper* [Preprint].

Pyle, D. (1971) 'On the Theory of Financial Intermediation', *Journal of Finance*, 26(3), pp. 737–747.

Rajan, R.G. and Zingales, L. (1998) 'Financial Dependence and Growth', *The American Economic Review*, 88(3), pp. 559–586.

Rambachan, A. and Roth, J. (2022) 'A More Credible Approach to Parallel Trends', *Working Paper*, 1745303. Available at: https://jonathandroth.github.io/assets/files/HonestParallelTrends_Main.pdf.

Rice, T. and Rose, J. (2016) 'When good investments go bad: The contraction in community bank lending after the 2008 GSE takeover', *Journal of Financial Intermediation*, 27, pp. 68–88. Available at: https://doi.org/10.1016/j.jfi.2016.02.001.

Rigbi, O. (2010) 'The Effects of Usury Laws : Evidence from the Online Loan', (June).

Ritchie, J. and Siwale, J. (2011) 'Failure by Design: The Rise and Fall of a Microfinance Institution in Zambia – a case of Pride Zambia', (June 2011), pp. 18–23.

Roa, M.J., Villegas, A. and Garrón, I. (2021) 'Interest rate caps on microcredit: evidence from a natural experiment in Bolivia', *Journal of Development Effectiveness*, 00(00), pp. 1–18. Available at: https://doi.org/10.1080/19439342.2021.1968934.

Romero, L.C. (2021) 'The Effects of Usury Ceilings on Consumers Welfare : Evidence from the Microcredit Market in Colombia'.

Roth, J. (2022) 'Pretest with Caution: Event-Study Estimates after Testing for Parallel Trends', *American Economic Review: Insights*, 4(3), pp. 305–322. Available at: https://doi.org/10.1257/aeri.20210236.

Roth, J. *et al.* (2022) 'What's Trending in Difference-in-Differences? A Synthesis of the Recent Econometrics Literature'. Available at: http://arxiv.org/abs/2201.01194.

Safavian, M. and Zia, B. (2018) 'The Impact of Interest Rate Caps on the Financial Sector Evidence from Commercial Banks in Kenya', *World Bank Policy Research Working Paper*, 8393(April), pp. 1–37.

Sandow, J.N., Duodu, E. and Oteng-Abayie, E.F. (2021) 'Regulatory capital requirements and bank performance in Ghana: evidence from panel corrected standard error', *Cogent Economics & Finance*, 9(1). Available at: https://doi.org/10.1080/23322039.2021.2003503.

Santiago, V. and Fernandez, F. (2007) 'The determinants of bank margins in European banking', *Journal of Banking and Finance*, 31, pp. 2043–2063. Available at: https://doi.org/10.1016/j.jbankfin.2006.06.017.

Sharpe, S.A. (1995) 'Bank Capitalization, Regulation, and the Credit Crunch: A Critical Review of the Research Findings', *Board of Governors of the Federal Reserve System Finance and Economics Discussion Series*, (95–20), pp. 1–36.

Shaw, E.S. (1973) *Financial Deepening in Economic Development*. Oxford University Press,.

Silva, D.F. *et al.* (2015) 'IMF Country Report No. 15/153 ZAMBIA SELECTED ISSUES', (15).

Simpasa, A. (2010) 'Performance of Zambian Commercial Banks in the Post-Liberalisation Period - Evidence on Cost Efficiency, Competition and Market Power', *PhD Thesis* [Preprint].

Simpasa, A., Nandwa, B. and Nabassaga, T. (2015) 'Bank lending channel in Zambia: empirical evidence from bank level data', *Journal of Economic Studies*, 42(6), pp. 1159–1174. Available at: https://doi.org/10.1108/JES-10-2014-0172.

Simpasa, A., Pla, L. and John, C. (2016) 'Sectoral Credit Concentration and Bank Performance in Zambia', (245).

Simpasa, A.M. (2013) *Competition and Market Structure in the Zambian Banking Sector*. Available at: www.afdb.org/.

Siwale, J. (2016) 'Microfinance and Loan Officers' Work Experiences: Perspectives from Zambia', *Journal of Development Studies*, 52(9), pp. 1289–1305. Available at: https://doi.org/10.1080/00220388.2016.1139692.

Slovik, P. and Cournede, B. (2015) 'Macroeconomic Impact of Basel III', *SSRN Electronic Journal* [Preprint], (844). Available at: https://doi.org/10.2139/ssrn.2650033.

Soludo, C.C. (2004) 'Consolidating the Nigerian banking industry to meet the development challenges of the 21st century'. BIS.

Sousa, M.M. de (2015) *Financial Inclusion and Global Regulatory Standards: An Empirical Study across Developing Economies*. The Centre for International Governance Innovation, p. 24.

Stiglitz, J.E. and Weiss, A. (1981) 'Credit Rationing in Markets with Imperfect Information', *The American Economic Review*, 71(3), pp. 393–410.

Sun, L. and Abraham, S. (2021) 'Estimating dynamic treatment effects in event studies with heterogeneous treatment effects', *Journal of Econometrics*, 225(2), pp. 175–199. Available at: https://doi.org/10.1016/j.jeconom.2020.09.006.

Sylla, R. and Homer, S. (2005) *A History of Interest Rates - Fifth Edition*, *John Wiley and Sons*. Available at: https://doi.org/10.2307/2977504.

Tabart, C. (no date) 'Background Paper: Basel Implementation in Low and Lower-Middle Income Countries – Early Adopters of Basel III 1', pp. 1–25.

Thakor, A. (1996) 'Capital Requirements, Monetary Policy, and Aggregate Bank Lending', *The Journal Of Finance*, 51(1), pp. 279–324.

Thakor, A.V. (1996) 'Capital requirements, monetary policy, and aggregate bank lending: Theory and empirical evidence', *Journal of Finance*, 51(1), pp. 279–324. Available at: https://doi.org/10.1111/j.1540-6261.1996.tb05210.x.

Thakor, A.V. (2014) 'Bank Capital and Financial Stability: An Economic Trade-Off or a Faustian Bargain?', *Annual Review of Financial Economics*, 6(1), pp. 185–223. Available at: https://doi.org/10.1146/annurev-financial-110613-034531.

USAID (2009) The legal and microfinance in environment for regulatory Egypt diagnostic study with focus on NGO-MFI transformation issues.

Vietnam Microfinance Working Group. (2013) *BULLETIN - MICROFINANCE REGULATIONS IN VIETNAM*.

Watanabe, W. (2010) 'Does a large loss of bank capital cause Evergreening? Evidence from Japan', *Journal of the Japanese and International Economies*, 24(1), pp. 116–136. Available at: https://doi.org/10.1016/j.jjie.2010.01.001.

Whitworth, A. (2015) 'Explaining Zambian Poverty: A History of (Nonagriculture) Economic Policy Since Independence', *Journal of International Development*, 27(7), pp. 953–986. Available at: https://doi.org/10.1002/jid.3049.

Wolken, J. and Navratil, F. (1981) 'The Economic Impact of the Federal Credit Union Usury Ceiling', *The Journal of Finance*, 36(5), pp. 1157–1168.

Wong, K.P. (1997) 'On the determinants of bank interest margins under credit and interest rate risks', *Journal of Banking and Finance*, 21(2), pp. 251–271. Available at: https://doi.org/10.1016/S0378-4266(96)00037-4.

Woolford, I. (2019) *Capital Review Paper 4: How much capital is enough?* RBNZ.

World Bank (2014) *Zambia Economic Brief - Financial Services: Reaching Every Zambian*. World Bank.

World Bank. (2004) 'Zambia, Country Economic Memorandum: Policies for Growth and Diversification', *World Bank Report*, (28069).

World Bank. (2013) *Zambia Economic Brief - Zambia's Jobs Challenge*. World Bank.

World Bank. (2017) 'Zambia Economic Brief: How Zambia Can Borrow Without Sorrow', (10). Available at: http://documents.worldbank.org/curated/en/782221512459934813/pdf/WP-P157243-PUBLIC-WorldBankthZambiaEconomicBriefDecemberFinalWEB.pdf.

Yan, M., Hall, M.J.B. and Turner, P. (2012) 'A cost-benefit analysis of Basel III: Some evidence from the UK', *International Review of Financial Analysis*, 25, pp. 73–82. Available at: https://doi.org/10.1016/j.irfa.2012.06.009.

Zambia (APF), A.-P. (2014) *A Market Study on Microfinance Services in Zambia, Agri-ProFocus Zambia (APF)*. APF, pp. 1–79.

Zinman, J. (2010) 'Restricting consumer credit access: Household survey evidence on effects around the Oregon rate cap', *Journal of Banking and Finance*, 34(3), pp. 546–556. Available at: https://doi.org/10.1016/j.jbankfin.2009.08.024.

Zins, A. and Weill, L. (2018) ‘Do Pan-African banks have the best of both worlds?’, *Economic Systems*, 42(4), pp. 665–681. Available at: https://doi.org/10.1016/j.ecosys.2018.06.001.