Essays on Mobile Money Services, Microenterprises and Role Models in Developing Countries

Emma Riley

New College
University of Oxford

A thesis submitted for the degree of Doctor of Philosophy in Economics

Michaelmas 2018

Abstract

Mobile money services have rapidly expanded across East Africa, with 66% of people using an account actively. Understanding the impact of these services is therefore an important area of research, which the first two chapters of my thesis contribute to.

Chapter One of my thesis uses survey data from Tanzania to understand how mobile money services affect the risk sharing capabilities of households. I look at the aftermath of rainfall shocks at the village level, and examine how consumption has been affected for users of mobile money services, non-users who reside in villages with other users, and households that live in villages without access to mobile money services. I find that after a village-based rainfall shock, only users of mobile money services are able to smooth their consumption. Non-users, regardless of whether they have access to mobile money or not, experience a 6% fall in their consumption. This raises questions of who benefits from the introduction of new financial services.

Chapter Two looks at integrating mobile money services into a microfinance loan product and the impact this has on the businesses of women in Uganda. In a field experiment, I randomly assigned women applying for a microfinance loan to receive the loan on a mobile money account or as cash. I find that women who receive their loan on a mobile money account invest 11% more of the loan into their business and experience 15% higher profits. These impacts are greatest for women who experienced high pressure to share money with family and friends at baseline. This study suggests that the manner in which a loan is disbursed has important implications for whether women are able to control how the loan is used.

The third chapter of my thesis looks at a different topic: role models and how they impact behaviour. I study the impact of being randomly assigned to see a movie featuring a potential role model, *Queen of Katwe*, or a placebo movie, on the exam results of secondary school students in Uganda. Seeing *Queen of Katwe* results in lower secondary school students scoring 0.11 standard deviations higher in their maths exam, mainly because they are less likely to fail maths, and upper secondary school students scoring 0.13 standard deviations higher in their finishing exams. Impacts are largest for female students. Further research is needed to uncover the mechanisms by which a role model can raise exam performance, with both changing beliefs and raising aspirations potential channels.

Essays on Mobile Money Services, Microenterprises and Role Models in Developing Countries

Emma Riley

New College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy in Economics

Michaelmas 2018

Abstract

Mobile money services have rapidly expanded across East Africa, with 66% of people using an account actively. Understanding the impact of these services is therefore an important area of research, which the first two chapters of my thesis contribute to.

Chapter One of my thesis uses survey data from Tanzania to understand how mobile money services affect the risk sharing capabilities of households. I look at the aftermath of rainfall shocks at the village level, and examine how consumption has been affected for users of mobile money services, non-users who reside in villages with other users, and households that live in villages without access to mobile money services. I find that after a village-based rainfall shock, only users of mobile money services are able to smooth their consumption. Non-users, regardless of whether they have access to mobile money or not, experience a 6% fall in their consumption. This raises questions of who benefits from the introduction of new financial services.

Chapter Two looks at integrating mobile money services into a microfinance loan product and the impact this has on the businesses of women in Uganda. In a field experiment, I randomly assigned women applying for a microfinance loan to receive the loan on a mobile money account or as cash. I find that women who receive their loan on a mobile money account invest 11% more of the loan into their business and experience 15% higher profits. These impacts are greatest for women who experienced high pressure to share money with family and friends at baseline. This study suggests that the manner in which a loan is disbursed has important implications for whether women are able to control how the loan is used.

The third chapter of my thesis looks at a different topic: role models and how they impact behaviour. I study the impact of being randomly assigned to see a movie featuring a potential role model, *Queen of Katwe*, or a placebo movie, on the exam results of secondary school students in Uganda. Seeing *Queen of Katwe* results in lower secondary school students scoring 0.11 standard deviations higher in their maths exam, mainly because they are less likely to fail maths, and upper secondary school students scoring 0.13 standard deviations higher in their finishing exams. Impacts are largest for female students. Further research is needed to uncover the mechanisms by which a role model can raise exam performance, with both changing beliefs and raising aspirations potential channels.

Supervisors This thesis was supervised by Stefan Dercon (University of Oxford) and Climent Quintana-Domeque (University of Exeter).

Total words This thesis contains approximately 65,000 words.

Statement of authorship All three chapters of my thesis are single-authored, and are fully my own work.

Acknowledgement of funding I acknowledge funding from the Economic and Social Research Council (ESRC) and the Department of Economics. Chapters two and three have been funded by the generous support of anonymous donors. Furthermore, I acknowledge financial support for conference presentations and research expenses from the George Webb Medley Fund, New College and the Royal Economic Society. I am very grateful for the support I have received.

Ethical approval For Chapter two and three, ethical approval has been sought and obtained from the University of Oxford CUREC committee. The reference numbers are: chapter two: ECONCIA16-17-006, and chapter three: ECONCIA1617-002.

Declaration The work in this thesis is based on research carried out by me at the University of Oxford between August 2014 and December 2018. It has not been submitted, either partially or in full, either for a degree of this University (except where the Special Regulations for the subject permit this), or for a degree of any other institution.

Notice of publication The first chaper of this thesis has been published in the Journal of Development Economics. The published version is available here: `https://doi.org/10.1016/j.jdeveco.2018.06.015`

Emma Riley (*December 2018)*

Acknowledgements

I could not have completed this thesis without the support of many people; my supervisors, mentors, colleagues, friends and family.

Firstly, I would like to thank my two DPhil supervisors, Professor Stefan Dercon and Professor Climent Quintana-Domeque, for their advice, support and guidance. Stefan opened up many opportunities for me by enabling me to work as a research assistant on one of his projects, which crucially helped me developed the skills I needed for my own fieldwork. Climent has given me a great deal of his time throughout the many revisions of the papers that constitute this thesis. He has also taught me many of the skills an academic needs, from effective presentation skills to how to review for a journal. It has been a pleasure and a privilege to learn from them both and I look forward to working with them in the future.

Additionally, I would like to thank Dr Simon Quinn and Dr Karlijn Morsink for their excellent supervision during my MPhil at Oxford, the thesis from which is incorporated into my DPhil. They both instructed me during my first academic thesis, and hence helped me develop the skills and tools I would need throughout the DPhil.

Many other people — in the Oxford economics department and beyond — influenced my work and contributed through comments, questions and discussions: Johannes Abeler, Abigail Adams, Siwan Anderson, Anja Benshaul-Tolonen, Rachel Glennester, Anandi Mani, Rossa O'Keeffe-O'Donovan, Kate Orkin and Natalie Quinn. I would like to thank Richard Sedlmayr for his support during the initial development of my third chapter and for being available to bounce ideas off throughout my DPhil. I would like to thank Tavneet Suri for her discussions around mobile money services and advice on data sources when I was first starting research on this theme.

I had the opportunity to work as a research assistant of Kate Orkin throughout my DPhil, and learnt the majority of my fieldwork and project management skills from her. I could not have run my own fieldwork projects if I hadn't first been through the experience with her. I would also like to specially thank Mahreen Mahmud, who has supported me throughout my doctorate as a colleague, but also been there as a friend.

Special mention goes to the past and current members of the Oxford Firms, Labour and Organisations reading group, where Rachel Cassidy, Lukas Hensel, Clement Imbert, Michael Koelle, Pramilla Krishnan, Pavel Luengas Sierra, Muhammad Meki, Ashley Pople, Christopher Roth, Vanessa Schreiber, Daniela Scur, Katherine

Stapleton, Hannah Uckat, Verena Weidemann, Marc Witte and Chris Woodruff have provided thoughtful and critical feedback on my work.

I was lucky enough to work with dedicated NGOs in Uganda. I would like to thank BRAC Uganda, and all the staff I worked with there who made my research possible. Particular mention goes to my field manager, Marion Kiconco, for organising field logistics so smoothly. In relation to my third chapter, I would like to thank Edward Kisitu, Rose Kusingura and Rachel Proefke for their assistance with logistics in Uganda, particularly recruiting schools and transportation for the students. I would like to thank the staff of the Initiative for Social and Economic Rights (ISER) in Uganda for their support recruiting schools and assistance at the cinema screenings.

I was lucky enough to have the support of a dedicated group of friends in Oxford and elsewhere throughout my DPhil. I'd particularly like to thank Charlotte Baarda and Lauren Burton for being there as we all worked through the ups and downs of the DPhil. New College MCR provided a warm and friendly environment to socialise and relax in and I'm grateful to have been a member of such a great community.

I will always be grateful to my parents for encouraging me when I decided to return to studying after a couple of years of work. Their support and love helped me get through the most stressful times, and my writing trips to France were always my most productive. Thanks dad for proof reading the entire thesis!

Finally, I would like to thank Marcus, for supporting me throughout the final, crucial stages of the DPhil. An acknowledgement is not enough to thank you for all your encouragement, love and support.

Emma Riley
Oxford, December 2018

Contents

Introduction

Overview This thesis is composed of three chapters that contribute to important issues in development economics: how individuals share risk; how to improve the impact of microfinance loans; and how to raise students' educational attainment. Two of the chapters look at the impact of mobile money services. Mobile money services use basic mobile phones to cheaply and easily send and store money. They have expanded rapidly across developing countries, and particularly in East Africa, since their introduction in Kenya in 2007. The first chapter examines how mobile money services can be used by households to smooth consumption in the face of unanticipated village-level rainfall shocks. The second chapter examines how changing the way a microfinance loan is disbursed, from cash to a mobile money account, affects the way borrowers are able to utilise the loan, to the benefit of their microenterprise. The final chapter, on a slightly different theme, examines how a role model affects the exam performance of secondary school students in Uganda.

Chapter One The first chapter of my thesis examines how the introduction of mobile money services in Tanzania allow remittances to flow into a village after a village-based rainfall shock, and to what extent these remittances are shared throughout the village. Informal networks of family and friends who reciprocally share income are one method that households use to insure their risk against unanticipated shocks (Dercon, 2002). However, when a shock is geographically clustered, e.g. a rainfall shock such as a drought or flood, insurance can only be sought from those outside the area. Mobile money services allow small amounts of money to easily, quickly and cheaply be sent between user's mobile phones across the country.

This chapter makes four main contributions. First, I add to early work on mobile money and risk sharing from Kenya (Jack and Suri, 2014) and show that in Tanzania

household consumption is also significantly affected by rainfall shocks, falling 6% on average, but that mobile money services are being used to provide insurance against these rainfall shocks. I find that amongst households that use mobile money services, consumption no longer falls in response to a large, village-based rainfall shock.

My second contribution is to examine the potential mechanisms through which mobile money allows consumption sharing, since mobile money can facilitate both saving and remittance flows. I find strong evidence for the remittance channel: households that use mobile money are more likely to receive remittances, and after an aggregate shock receive remittances worth 4% of per capita income, accounting for two-thirds of the negative effect of the shock.

My third contribution is to extend these results to examine spillovers between users of mobile money services and non-users. Spillovers would occur if, after an unexpected shock, users of mobile money shared some of the remittances they received with their neighbours. In this case, non-users of mobile money who live in villages with users of mobile money would not see as large a fall in their consumption after a rainfall shock as households that live in villages with no mobile money users. I find that non-users of mobile money services in villages with users still experience as large a fall in their consumption after an aggregate shock as those who reside in villages with no mobile money users. This raises the question of why remittances received via mobile money aren't being shared with others in the village after a rainfall shock. I suggest that a potential explanation for this is that users of mobile money services share risk with others across different locations and do not participate in risk sharing networks with others in the village.

My final contribution is to examine heterogeneity in these results by the distance to the nearest mobile money agent[1], whether the village is in a rural or urban location and whether the rainfall shock was a flood or drought. My results are principally driven by shocks in the form of droughts, which I confirm have a significant negative effect on crop yields, the primary income source for two-thirds of the households in my sample.

[1] An agent is needed to deposit and withdraw money from the account

Chapter Two The second chapter of my thesis uses a Randomised Controlled Trial (RCT) to examine the impact on women's businesses of changing the way microfinance loans are disbursed from cash to a mobile money account. Mobile money accounts offer a secure and private vehicle to keep the loan on until it is needed for the business, potentially protecting the loan from both own temptation to spend money and pressure from others to share it. I find that disbursing a microfinance loan onto a mobile money account results in women's business profits increasing 15% compared to the profits of women's businesses where the woman received the loan as cash. This seems to be because women who receive a loan on a mobile money account invest in 11% higher levels of business capital. The primary mechanisms driving this result is that the mobile money form of the loan enabled women to keep their loan better hidden from their spouse and family, and thus enabled women to use more of it in their businesses. I confirm this using data on expenditure patterns, transfers, and data on how the loan was used. I argue that neither the woman's self-control difficulties nor saving constraints can explain my results.

I also compare disbursing the loan onto a mobile money account (Mobile Disbursement) to just providing a mobile money account designated for the woman's business but with the loan as cash (Mobile Account). In principle, the Mobile Account treatment group could replicate the Mobile Disbursement treatment by depositing part or all of the loan themselves onto the mobile money account given to them. However, I find large differences between these groups, with no significant effects on business outcomes from the Mobile Account treatment. I argue this is due to women sticking with the default option.

Using administrative data from the telecoms provider, I am able to examine use of the mobile money accounts and balances held on them. For both the treatment groups, I find limited deposits by the women onto the accounts, with only 13% of women making a deposit onto the mobile money account in either treatment. The primary difference between these two treatments is therefore that the Mobile Disbursement treatment makes the default situation the removal of money from the account, whereas the Mobile Account treatment makes the default situation

adding money onto the account. This difference leads to the Mobile Disbursement treated group saving part of the loan left on the mobile money account, and securely keeping this money separate from other household funds and household members. Women assigned to the Mobile Disbursement treatment hold as balance on the account 7% of the loan value, or 22% of their baseline savings amount, during the first 30 days after the account is given to them. The Mobile Account treatment never hold significant amounts of money on the account. I argue that this default, secure saving of any part of the loan that is not withdrawn immediately from the mobile money account is what allows the Mobile Disbursement treatment group to invest more of the loan in their businesses rather than have left-over cash whittled-down by requests from relatives.

This chapter makes four main contributions to the literature. Firstly it adds to the literature on the question of what conditions enable microfinance loans to increase the profitability of women's businesses. Microfinance loans were initially championed as an effective way to encourage entrepreneurship and raise incomes, particularly of women. However, a series of studies looking at many forms of microfinance in different contexts failed to find any benefit to the business or household from receiving a loan (Banerjee et al., 2015). Since then, studies have shown that both cash grants and loans can have benefits for women's businesses if changes are made to the default structure to facilitate investment, by, for example, giving the grant in-kind or offering a longer grace period before repayment begins (Fafchamps et al., 2014, Field et al., 2013). Other papers have also focused on the role of the family in hindering women's enterprise growth, finding that if women can hide money or live in households without other business owners they can successfully grow their enterprise in response to a grant or loan (Bernhardt et al., 2017, Fiala, 2017). My study adds to this literature by showing if a microfinance loan is given to a woman in a way that hides it from her household and facilitates investment, it can have a large return for her enterprise.

Secondly, it adds to the literature looking at the pressure women experience to share money with their husband and family, and the strategies women use to mitigate

this pressure (Almas et al., 2015, Baland et al., 2011, Boltz et al., 2017, Jakiela and Ozler, 2015, Schaner, 2015, 2017). It also adds to the literature showing that giving women more control over money results in more of it being spent on their preferences (Aker et al., 2016, Ashraf, 2009, Attanasio and Lechene, 2002, Field et al., 2016).

Thirdly, my paper adds to a growing literature on how mobile money accounts can facilitate savings, contrasting with other papers that find that labelling a mobile money account as being for a particular form of saving, and providing financial inducements to save, can increase those savings (Bastian et al., 2018, Batista and Vicente, 2017, Dizon et al., 2017, Habyarimana and Jack, 2018, Lipscomb and Schechter, 2018). My findings are more in line with De Mel et al. (2018) who find little impact on deposits and savings of linking bank accounts to a mobile money interface. While other papers have looked at paying salaries and cash transfers directly onto mobile money accounts (Aker et al., 2016, Blumenstock et al., 2015), to the best of my knowledge this is the first paper examining paying a microfinance loan onto a mobile money account.

Finally, my paper adds to the literature on the importance of default effects. Though women assigned to the Mobile Account treatment could imitate the Mobile Disbursement treatment by depositing all or part of their loan onto the mobile money account, they did not do this. This is a similar finding to others which have shown default contributions to be the main driver of whether money is saved on a saving device (Blumenstock et al., 2018, Brune et al., 2017, 2016, Field et al., 2016, Somville and Vandewalle, 2018).

Chapter Three The third chapter of my thesis studies how exposure to a role model affects secondary school students' exam performance. A role model may improve exam results if performance is hampered by low aspirations or a lack of belief in oneself (Beaman et al., 2012, Bernard et al., 2014, Nguyen, 2008).

This chapter is based on a randomised experiment where I assigned 1,500 students immediately before their national exams to watch one of two movies: a treatment movie, *Queen of Katwe*, which details the true story of a young girl's struggle to get

into a top school to play international chess; or a placebo movie, *Miss Peregrine's Home for Peculiar Children*, a children's adventure story about fighting monsters. I use administrative data from the participating schools to quantify what impact seeing the treatment movie had on students' overall, maths and English test scores.

I find that amongst students taking their lower secondary school exam, treatment with *Queen of Katwe* one week before their exams results in the students scoring 0.11 standard deviations higher on their maths paper. This effect comes entirely from an 11 percentage point decrease in the probability of failing maths. The impact is largest for female students and those below the median in their prior academic performance, as measured on a mock exam two months before the experiment. These two groups would be expected to benefit the most from the role model since the role model is most 'like them': a female student coming from a background of extremely low academic performance.

I find that amongst students taking their finishing exams from secondary school, exposure to *Queen of Katwe* one month prior to their exams results in an overall increase in the exam score of 0.13 standard deviations. Again, it is female students who benefited most from the treatment.

Overall, this chapter makes four contributions to the literature. First, it contributes to a growing literature on the impact of the media on economic behaviours (La Ferrara, 2016) and is the first to study the impact of a form of media not designed with a specific educational goal on an educational outcome.

Secondly, I show that even with a brief exposure of two hours to a role model, important economic outcomes can be influenced. This is in comparison to other forms of media, such as a soap opera or radio show, that people are exposed to the messages of over a significant period of time.

Thirdly, I show that even when shown just a short time period before students sat their exams, exposure to a role model has a significant effect. Since the time frame is so short, the impact is more likely working through a motivational channel rather than through increased learning, though I do not have the data to confirm this.

Lastly, while my study is not the first to look at role model impacts on student educational attainment (Beaman et al., 2012, Nguyen, 2008) it is the first to show that the role model does not have to be in a real-life form or from the local community of the person to have a positive effect. This has implications for the ease of scaling up role model interventions to reach more people.

Organisation of the thesis The next three chapters consist of the essays just described, which are presented as self-contained journal-style articles. They all include their own sections on literature review, methodology, data description, results, and discussion. My contributions to existing research are outlined within each chapter.

Chapter 1

Mobile Money and Risk Sharing Against Village Shocks*

Abstract Households in developing countries have gained increased access to remittances through the recent introduction of mobile money services. I examine the impact of these mobile money services on consumption after a rainfall shock, such as a flood or drought, for both users of mobile money and for household that don't use mobile money but who reside in villages with other users. This allows me to determine the extent that remittances received via mobile money are shared within villages, creating wider benefits to the community. Using a difference-in-difference fixed effects specification, I find that after a village-level rainfall shock it is only users of mobile money who are able to prevent a drop in their consumption. There are no spillover effects to other members of the village. This finding has implications for how new technologies might change traditional risk sharing arrangements, and who might benefit and lose out from their spread.

1.1 Introduction

In developing countries, households use informal risk sharing networks to smooth their consumptions in response to unanticipated idiosyncratic shocks such as illness or death. Households within a network can insure their idiosyncratic shocks through cross-sectional risk sharing which allows a household in a network who is affected by a shock to receive transfers from those who aren't affected, assuming not everyone

*This chapter was published in the Journal of Development Economics in November 2018. The published version is available here: https://doi.org/10.1016/j.jdeveco.2018.06.015

in the same network is subject to the same shock at once. This sort of risk sharing has been found to exist both within and across villages, though it is often found to be incomplete (De Weerdt and Dercon, 2006, Gertler and Gruber, 2002, Kazianga and Udry, 2006, Kinnan, 2017, Townsend, 1994, Udry, 1994)

However, if the network is clustered in one geographical location, such as within a village, aggregate shocks to that village, such as a flood or drought, could occur, affecting all of the village at once. Larger risk sharing networks of friends and families in other villages could be used to insure this risk by sending remittances, but in practice it is costly and time consuming to send money long distances due to high transaction costs. Without a fast, cheap and secure method of sending remittances long distances, cross-village insurance networks will not be able to respond to shocks amongst their members.

Mobile money services are a new tool allowing small amounts of money to cheaply, quickly and safely be sent around the country via a mobile phone, dramatically increasing access to a wider remittance network that households can draw from. By facilitating risk sharing outside the village with people in other communities, which will be less likely to have experienced the same shock, mobile money allows households to insure themselves against aggregate shocks to their village.

This paper examines how the introduction of mobile money services allow remittances to flow into a village after a rainfall shock occurs in that village, and to what extent these remittances are shared throughout the village, allowing all households within the village to smooth their consumption. By comparing households in village with and without mobile money, and within villages with mobile money, households that do and do not use mobile money services, I can quantify the benefits of mobile money to both the recipient and to the rest of the village. While previous work has looked at the impact of mobile money on the user, no one has yet looked at the potential benefits to other members of a village when a household uses mobile money.

Likewise, previous work has not separately examined geographically clustered shocks at the village level (which I refer to here as village-level aggregate shocks)

and idiosyncratic shocks. I argue a key contribution of mobile money services is enabling risk sharing when an entire village experiences a shock at once through the flow of remittances from outside to inside the village. Previous to the introduction of mobile money services, even if a household had a network outside the village, methods of sending remittances were slow and costly. This paper will build upon other work showing the benefit of mobile money use to the user after an idiosyncratic shock (Jack and Suri, 2014) and focus on the extent of sharing of the benefits of mobile money use within the village after a village-level rainfall shock.

I begin by looking at village-level aggregate shocks in the form of floods and drought, which are geographically concentrated, large and unexpected, and hence cannot be insured within the village. I find that household consumption is significantly negatively affected by these rainfall shocks, with household consumption falling approximately 6%.

Secondly, I show that mobile money provides insurance against these rainfall shocks, resulting in household consumption of users no longer being negatively impacted by an aggregate shock to the village. The mechanism proposed here is that mobile money allows the user access to remittances from households outside the village. Mobile money therefore allows households to share risk with others outside their village, resulting in aggregate shocks to village consumption not being aggregate shocks to the household's network. When a mobile-money-user experiences a shock at the village level, they can ask for help from family and friends in other locations which have not experienced a negative shock and with whom they can reciprocally insure. Remittances can then be sent easily and cheaply via mobile money. This means users of mobile money are able to smooth their consumption after a village-based shock in a way non users aren't able to.

I confirm that remittances are the mechanism through which mobile money allows the smoothing of rainfall shocks using a cross section of remittance data from the final wave of my data. I find that mobile money users are both more likely to receive remittances and, after an aggregate shock, receive an increase in remittances of 4% of per capita income, accounting for 2/3 the negative effect of

the shock. While I show that remittances received via mobile money explain the majority of the fall in consumption after a rainfall shock, I cannot rule out that mobile money also helps households save and self-insure against shocks.

Thirdly, I examine the wider impact of mobile money transfers within a village, something that has not been looked at before in previous work. If insurance networks cover both users and non-users of mobile money within a village, then if a household uses mobile money and receives remittances after an aggregate shock the remittance will be shared with other non-user members of the insurance network within the village. Hence consumption of non-mobile-money users in villages with other mobile money users will also not decline as much after a village-level rainfall shock as that of households in villages without any mobile money users.

I find that while a user of mobile money is able to perfectly smooth the impact of a rainfall shock, non-users in villages with other mobile money users still experience the same fall in consumption as non-users in villages with no users. Users of mobile money are not sharing their remittances with other members of the community after a rainfall shock. A possible explanation for this is that recipients of mobile money are choosing not to participate in a risk sharing network with others in the village and instead are relying on networks outside the village and on the stream of remittances for insurance. Potential changes in the composition and extent of risk sharing relationships due to the introduction of mobile money are an exciting area for future research.

I examine heterogeneity in my results by distance to the nearest mobile money agent (essential for making transactions with mobile money), whether the village is in a rural or urban location and whether the rainfall shock is a flood or drought. I find suggestive evidence that droughts are the main driver of my results, and confirm this by showing droughts a have significant negative effect on crops yields, the main source of income for 2/3 of these households.

I confirm the robustness of my findings using a placebo test using two periods of data from before the introduction of mobile money and an instrumental variable regression using distance to the nearest mobile money agent as an instrument for

adoption. These allow me to test if mobile money use is correlated with ability to smooth risk. Neither test confirms selection, but both have large standard errors and so should be interpreted with caution.

This paper builds on the literature of the methods households in developing countries use to smooth consumption (Dercon, 2002) but shows use of these methods to be insufficient to fully smooth village-level rainfall shocks. I find that households in Tanzania without access to mobile money services suffer consumption shortfalls after aggregate shocks, here village-clustered rainfall shocks. They are unable to fully insure these shocks using other methods of smoothing consumption, such as savings or transfers from wider networks in the absence of mobile money.

However, I find that households that use mobile money are able to insure aggregate shocks, with the proposed mechanism being flows of remittances from networks of family and friends outside the village. This builds on a number of papers that have examined links to others outside the village and how households share risk across a larger network of family and friends (De Weerdt and Dercon, 2006, Fafchamps and Lund, 2003, Rosenzweig, 1988).

The mechanism proposed here through which mobile money allows risk sharing is remittances. Remittances have been shown to be an important channel through which households with family members outside the village can insure their consumption. Yang and Choi (2007) look at remittance patterns in the Philippines, finding that remittances move in opposite directions to income with 60% of the decline in income after a rainfall shock compensated for by increased remittances. Households without migrant members experience a fall in consumption. However, sending money across long distances by traditional channels such as through friends or via Western Union can be very costly, slow and unsafe, limiting the effectiveness of this channel. Mobile phone money transfer technology is able to overcome these barriers to sending remittances quickly and lower costs, allowing users access to their wider risk-sharing networks and assisting households in smoothing village-level shocks (Batista and Vicente, 2014, Blumenstock et al., 2016, Jack and Suri, 2014).

Jack and Suri (2014) use panel data to analyse how mobile money facilitates consumption smoothing in response to negative idiosyncratic income shocks in Kenya. They find that while the consumption of non-user households falls by 7%-10% after a shock, there is no corresponding fall for user households. They find that this effect is due to the improved ability to smooth risk via remittances; in the face of a negative shock, user households are 13% more likely to receive any remittances, receive more remittances and receive a larger total value amounting to 6-10% of annual consumption. Munyegera and Matsumoto (2016) also shows that remittances are the channel through which mobile money affects consumption in Uganda, with households with mobile money 20 percentage points more likely to receive remittances, receive remittances more frequently and the total value of the remittances received are 33% higher than for non-user households. However, none of the existing literature examine spillovers in mobile money use in the form of sharing of remittances within villages, something I can shed light on. Whether mobile money services only benefit users or also have spillover benefits to others in the community is an importance factor in assessing who benefits and loses from new technologies.

The remainder of this paper is organised as follows: I first provide some context on mobile money services. I summarise the data used in this paper in section 3, outline the empirical specifications and makes predictions to be tested in the data in section 4. Section 5 covers the main results, robustness checks and mechanisms. Finally, I conclude.

1.2 Mobile money services in Tanzania

Mobile money services allow users to transfer, deposit, withdraw and save money on a mobile phone. They work on basic mobile phones[2] using PIN secured SMS text messages and do not require the user to have a formal bank account. Mobile money services use agent networks that include airtime sellers and retail networks where customers can deposit and withdraw money. Users are charged a small fee for sending and withdrawing money using mobile money services and transaction sizes

[2]non-smart phones

are normally capped at a few thousand US dollars. To send money to someone, only their mobile phone number is required. Services have also expanded recently to include payment of bills and taxes and to allow the combining of a mobile money account with a formal bank account. Mobile money services have spread rapidly in the developing world since the launch of th first such service, MPesa in Kenya in 2007, by “leapfrogging" traditional banking services (Aron, 2018).

Mbiti and Weil (2011) describe the impact of M-Pesa in Kenya, finding that M-Pesa changes the pattern of remittance by increasing the frequency and volume of urban-rural transfers while lowering the price of competing remittance services such as Western Union. They find 25% of people report using M-Pesa for savings, and that it lowers the probability of people using informal saving mechanisms, such as ROSCAS, while raising the probability of them being banked. Jack and Suri (2011) also look descriptively at the use of M-Pesa for sending remittances in Kenya and find that remittances sent via M-Pesa are less likely to go to parents and more likely to go to friends and other relatives than other forms of remittance. This could signal that M-Pesa users have/take advantage of a broader network than non-users. For those who don’t use M-Pesa, the most commonly given reason is not owning a mobile phone followed by not needing the service. Less than 1% report not having access to an agent.

There are 4 mobile money providers in Tanzania; Vodacom’s M-Pesa, Zantel’s Z-Pesa and Zain’s Zap (now Airtel Money), all of which launched in 2008/9, and Tigo’s Tigo Pesa which launched in 2010. M-pesa is by far the largest of these with 72% of the market. Take-up of mobile money took off slowly, with only 0.5% of households having ever used mobile money in 2009 (Finscope, 2013), but after Vodacom initiated some changes at the end of 2009 the service took off, reaching a quarter of the population by the end of 2011 and a third by the end of 2013. From only 900 agents in September 2009, the service had 17,000 by December 2013.

Mobile money requires the user to have a mobile phone and sim card from the mobile money provider. The user must register for a mobile money account and can then deposit money through that mobile money providers’ agents, which are usually

located in shops. The cash is then electronically deposited in the customer's account. Customers can transfer money via SMS to other people even on different networks, and make withdrawals at their network's agents anywhere in the country. Users are charged a step-tariff rate for sending money and for withdrawing money from agents, with fees for M-Pesa of around 10% for withdrawing and 3% for sending $5 and falling with the amount. Depositing money on the account is free.

1.3 Data and summary statistics

1.3.1 Household panel

The data used comes from the Tanzania National Panel survey (NPS) 2008-9, 2010-11 and 2012-13, implemented by the Tanzania National Bureau of Statistics and downloaded from the World Bank LSMS microdata catalogue. The survey covers 3,265 households in 26 districts containing 409 Enumeration Areas (EAs), and is designed to be representative of Tanzania as a whole. Within each EA (village) an average of 8 households were randomly selected. The survey made particular effort to track respondents, with all adult former households members tracked to new location, resulting in over 97% of the round 1 households being re-interviewed in round 2 and a total panel attrition rate of 4.8%. The data includes weightings of the probability that an observation was included in the survey to take into account the fact that some areas were over surveyed to reflect the higher variance of the variables of interest (for example in cities).

The survey included questions on consumption, assets, finance, shocks, household characteristics and village characteristics. I combined the data by household since mobile money use is only recorded at the household level. Looking at the characteristics of the household head in table 1.1, the average household has 5 people, average years of education of the household head is just under 5 years, increasing slightly during the survey. 60% of household heads worked in agriculture, 10% in the private sector as paid workers and 15% were self employed. In 2008-09 annual real per person consumption was 742,386 TZ Shillings ($450), rising to 1,011,279 ($568) in 2012-13.

Table 1.1: HH summary stats by wave

	Wave 1		Wave 2		Wave 3	
	Mean	SD	Mean	SD	Mean	SD
Per capita consumption	743,386	725,334	862,266	782,264	1,011,279	1,090,465
Rainfall shock	0.21	0.41	0.21	0.41	0.29	0.46
Mobile money use	0.00	0.00	0.13	0.33	0.38	0.49
Rural	0.69	0.48	0.69	0.46	0.69	0.46
Education of head (yrs)	4.76	3.57	4.84	3.65	4.96	3.70
Female head	0.25	0.43	0.26	0.43	0.28	0.446
Age of head	45.53	15.05	45.85	15.53	46.71	16.04
Household size	5.08	2.86	5.28	3.13	5.02	3.05
Own mobile	0.45	0.50	0.62	0.48	0.71	0.45
Financial access						
Number of loans	0.07	0.30	0.10	0.35	0.12	0.39
Bank account	0.0	0.0	0.20	0.40	0.20	0.40
ROSCA	0.04	0.20	0.05	0.22	0.04	0.19
Wealthscore	-0.11	3.00	0.14	2.85	-0.05	2.56
Occupational dummies						
Agriculture/ Livestock	0.60	0.49	0.56	0.50	0.55	0.50
Fishing	0.02	0.14	0.01	0.12	0.01	0.12
Mining	0.00	0.05	0.00	0.06	0.00	0.06
Tourism	0.00	0.02	0.00	0.02	0.00	0.02
Employed: Government	0.06	0.23	0.06	0.24	0.06	0.23
Parastatal	0.01	0.08	0.01	0.08	0.01	0.07
Private sector	0.09	0.28	0.11	0.31	0.12	0.33
NGO/religious	0.01	0.09	0.01	0.08	0.01	0.09
Self-employed (non-agri) w employees	0.02	0.16	0.03	0.18	0.03	0.16
Self-employed (non-agri) w/o employees	0.15	0.35	0.14	0.35	0.15	0.36
Unpaid family work	0.01	0.05	0.01	0.11	0.01	0.11
Job seeker	0.00	0.05	0.00	0.07	0.00	0.04

I generated a wealth index of assets using principal component analysis (PCA) since the value of assets owned was not asked all the waves of the survey. Different components of wealth, such as the number of chickens owned or bicycle ownership, cannot easily be added up. PCA determines the relative importance of variables when seeking to summarize a set of variables. The first principal component accounts for the largest variance across the variables. In a wealth index, the first principal component is assumed to represent relative wealth. Based on this, each factor is given a factor weight representing its relative importance in constructing the principal component. I generated a wealth index score based on these factor weights.

Looking at mobile phone ownership and mobile money use, in 2008-9, 45% of households owned at least one mobile phone, increasing to 62% in 2010-11 and 71% in 2012-13. 13% of households had used a mobile money service in 2010-11 and 38% had by 2012-13. I am interested in both users and non users in villages with mobile money and non-users in villages without mobile money. Therefore I break down the number of villages and households by these categories in Figure 1.1. This figure shows the large increase in both villages with any mobile money users and the number of mobile money users and non-users within these villages. By the second round of the survey 47% of the communities had at least one person using mobile money. By 2012-13, 83% of the communities had at least one person using mobile money. The agent network also expanded rapidly during this period. In the second wave of the data 20% of villages had an agent in the village but this increased to 50% by the third wave.

Sending and receiving money are by far the most popular uses of mobile money, with 67% of users saying they send money and 82% of users saying they receive money. These two uses are also given as the most important use of mobile money services by 80% of respondents. 20% of people report using mobile money services to save up for emergencies and 12% have used it to pay for a good or service. 40% of people use the service at least monthly. The most common reason for not using mobile money was no mobile phone, given by just over 60% of respondents. Lack of

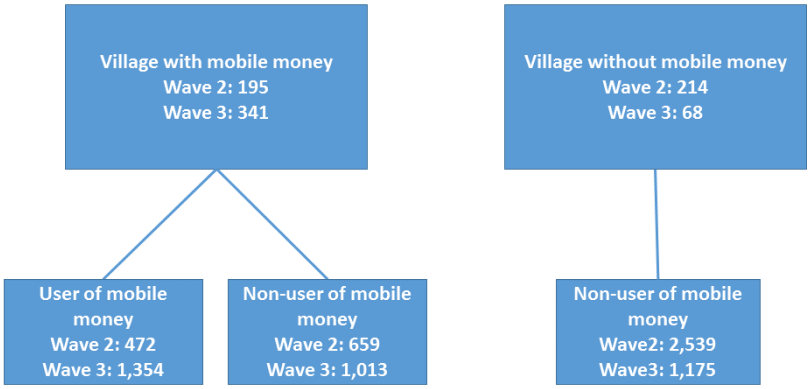


Figure 1.1: Break down of villages and households by mobile money use

proximity to an agent was only given as the reason for not using mobile money by 8% of respondents, also equal to those citing they don't understand the service.

In the third round of the survey there was detailed data on who sent the remittances to whom using what channel, where from and what their relationship was to the sender. 40% of remittances were sent physically via friends and family and 35% by mobile money. Only 2% was sent using a bank account, 1% using Weston Union and 0.4% using the Post Office. In the past it's probable that the majority of remittances were sent via friends and family with very little sent using any more formal channels. 40% of remittances were sent by a son or daughter with only 3.5% sent via a spouse, 7% by a parent and 17% by a sibling. 30% of remittances are sent from Dar Es Salaam and less than 3% are from abroad. This is consistent with a pattern of a family member migrating to another location such as the city within Tanzania and then sending remittances back to their family.

1.3.2 Rainfall measure

The panel data contains information on self reported shocks, including whether a household has experienced a drought or flood. This is a dummy equal to one if the household reported that they experienced a drought or flood in the year proceeding the survey wave. To the extent that households misreport or subjectively interpret a rainfall shock, for example saying they experienced a rainfall shock to justify a poor crop yield or exaggerating the importance of a rainfall shock in a year when

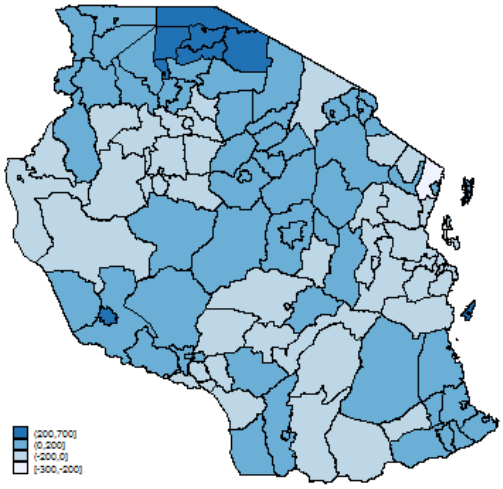
they have no other shocks, this measure of rainfall shocks will be subject to bias and measurement error. I therefore also calculate a rainfall shock measure per village using data from the NOAA's climate prediction centre FEWS (Famine Early Warning System). This is available at 0.1 degree resolution by latitude and longitude across Africa and was included in the Tanzania NPS summarized at the EA level.

I define a rainfall shock as more than a 1 standard deviation in absolute values from the 15 year mean by the nearest rainfall station to the village, as used in Jensen (2000). Deviations from the historic mean capture the extent that rainfall is different from what is expected, and 1 standard deviation is a large difference from normal (on average 200mm difference from an average annual rainfall of 800mm across the entire country). The absolute value is used because either too much or two little rainfall can be harmful. Only deviations greater than 1 standard deviation in absolute value are examined since a little bit too much or too little rain is unlikely to have a big effect, and initially more rain can have a positive effect on crop yields (Paxson, 1992, Suri, 2005). I am only interested in extreme, abnormal, rainfall deviations which could be classified as a drought or flood. The rainfall deviations in millimetres by year are shown in figures 1.2a to 1.2d.

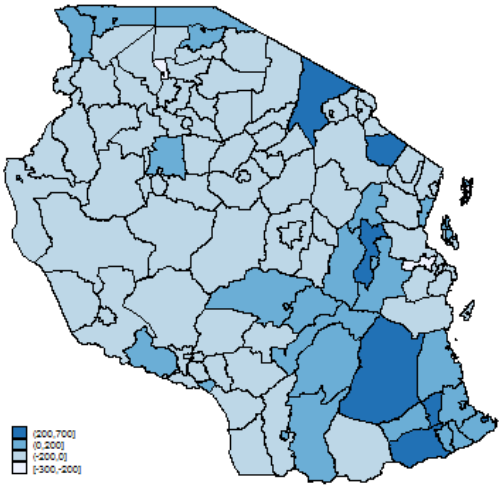
The mechanism through which rainfall shocks negatively affect income can be varied as I look at both rural and urban households. In rural households, droughts and floods will destroy crops leading to loss of income. In urban areas, flooding is likely to be the main mechanism though which rainfall shocks affect income by preventing people from working and by destroying property. For example, in Dar es Salaam in December 2011 there were severe floods resulting in over 6,000 people being displaced and left homeless. I examine the size of these different mechanisms by examining separately droughts and floods for both urban and rural households and by examining the impact of the shock on crop yields for agricultural households. In both these examples remittances can be sent to alleviate the loss of consumption, from urban to rural households after crop losses and from rural to urban households in the case of severe floods.

Figure 1.2: Deviations from mean rainfall, mm

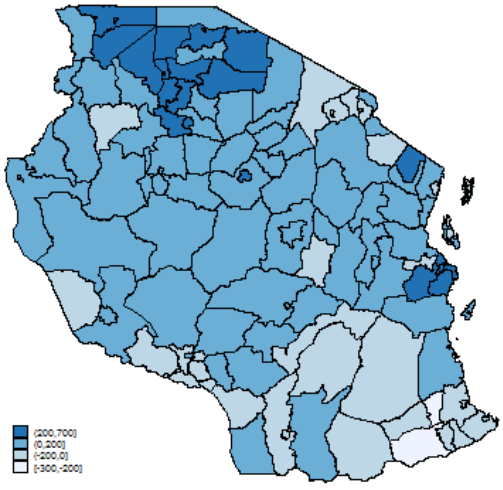
(a) July 2009-10



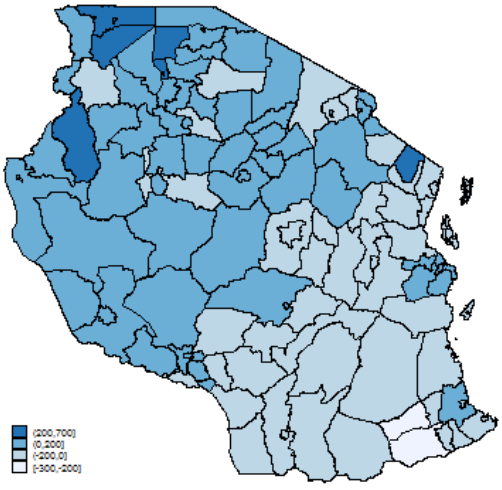
(b) July 2010-11



(c) July 2011-12



(d) July 2012-13



1.4 Empirical framework

1.4.1 Empirical Specification

I empirically test the impact of village-level rainfall shocks for households and villages with and without mobile money. If rainfall shocks are unable to be fully insured, they will lead to a drop in consumption. If mobile money allows for transfers to be made in response to an village-level rainfall shock, then consumption will no longer respond to aggregate shocks. If these transfers are shared with others in the village then there will be a positive spillover to non-users from other members of the community using mobile money. If transfers are kept by the user of mobile money then only the user will be able to smooth consumption after an aggregate shock.

I follow the literature[3] by writing an empirical specification of the impact of an aggregate rainfall shock on consumption for household with and without mobile money services:

$$\begin{aligned} C_{jvt} =& \gamma_a Shock_{jvt} + \mu MM_{jvt} + \lambda MMSp_{jvt} \\ &+ \beta_m MM_{jvt} \cdot Shock_{jvt} + \beta_{sp} MMSp_{jvt} \cdot Shock_{jvt} \\ &+ \boldsymbol{\theta X_{jvt}} + \boldsymbol{\psi X_{jvt}} \cdot Shock_{jvt} + \alpha_j + \delta_t + \varepsilon_{jvt} \end{aligned} \tag{1.1}$$

where C_{jvt} is household j's per capita log consumption in village v, $Shock_{jvt}$ is a rainfall shock in village v, MM_{jvt} is mobile money use by household j in village v, $MMSp_{jvt}$ is an indicator if household j doesn't use mobile money themselves but resides in a village v with at least one other mobile money user (a mobile money spillover household), $\boldsymbol{X_{jvt}}$ is a vector of controls consisting of household demographics, financial service use and occupation dummies to control for any other variables which might enable households to better smooth consumption, α_j is a household fixed effect, δ_t is a time trend and ε_{jvt} is a time varying error.

The parameters of interest are β_m, which allows for use of mobile money to affect the household's ability to smooth shocks and β_{sp}, which allows for being a mobile money spillover household to impact the household's ability to smooth shocks.

[3]This is the specification form used by Jack and Suri (2014) which also follows Gertler and Gruber (2002)

This gives the following predictions for the empirical estimation:

Prediction 1 For households in villages without mobile money (when $MM_{jvt} = 0$ and $MMSp_{vt}$=0), $\gamma_a < 0$ so that rainfall shocks have negative effects on consumption

Prediction 2 If users of mobile money receive remittances after an aggregate shock then $\beta_m > 0$.

Prediction 3 For households in villages with other mobile money user that don't use mobile money themselves, if there is some sharing of remittances within the village after an aggregate shock then $\beta_{sp} > 0$.

1.4.2 Identification strategy

I estimate (1.1) using a household fixed effects difference-in-difference specification. In the fixed effects case, difference-in-difference subtracts the average change in the control group (households in villages without mobile money) from the average change in the treatment group (users of mobile money or non-users in villages with mobile money), therefore removing biases from permanent differences between the two groups (captured in the fixed effects α_j) and changes due to a time trend (δ_t).

There are two main self-selection effects with regards to mobile money which could bias my results. The first is self-selection by households into using mobile money. If household selection into mobile money use is correlated with unobservables that affect their ability to deal with shocks that will bias my results, creating a spurious positive association between mobile money use and shock smoothing. Secondly there may be selection by mobile money agents into which villages they choose to locate in. Again, if this selection into villages is correlated with unobserved household ability to smooth shocks this will bias my results, again creating a spurious positive association between mobile money use and shock smoothing.

Starting with self-selection by a household into using mobile money: Self-selection effects into using mobile money are absorbed into the coefficient μ on MM_{jvt}, which is not the focus of my analysis, as I am only interested in the effects of using

mobile money after an exogenous aggregate shock. Self-selection into using mobile money is very likely, and in Table 1.2 column (1) I show that a variety of different observable characteristics predict mobile money adoption[4]. Despite adoption of mobile money not being the focus of my analysis, I still control for these observable characteristics which are correlated with mobile money use in all the regressions. Any unobservable fixed characteristics that may be correlated with mobile money use are controlled for using household fixed effects.

My analysis is focused on the ability of mobile money to help insure rainfall shocks. I am therefore concerned about selection into mobile money use being positively correlated with a household's ability to smooth shocks ($MM_{jvt} \cdot Shock_{jvt}$ being positively correlated with the error term ϵ_{jvt} conditional on the household fixed effects and other covariates). Likewise, for equation (1.1) to identify the impact of being a spillover household, $Shock_{jvt} \cdot MMSp_{vt}$ must be uncorrelated with the error term conditional on the household fixed effects and other covariates.

For this identifying assumption to be satisfied, the rainfall shock must be exogenous. I confirm this by running separate fixed effect regressions for different household characteristics for each of the rainfall shocks, the results of which are shown in Table 1.2 columns (3) and (4). For the self reported rainfall shocks in column (3), households that experience a 1 standard deviation rainfall shock are more likely to report experiencing a shock. Larger households and households with bank accounts are also more likely to report a shock. The F test of the joint significance of all the variables in Table 1.2 shows that they are jointly significant at the 10% level, suggesting the self-reported shock may be a biased measure.

In column (4), the 1 standard deviation rainfall shock, places with larger households and less ROSCA membership are correlated with more rainfall shocks, though this is no more than would be expected by chance. Jointly all the coefficients are insignificant. This suggest my 1 standard deviation rainfall shock is

[4]the table shows that being wealthier, having an older and better educated household head, owning a mobile phone and a bank account all increase the probability of a household using mobile money.

Table 1.2: Correlations of mobile money and rainfall shocks

	Mobile money		Rainfall	
	Use	Agent point	Self-reported	1 sd
	(1)	(2)	(3)	(4)
MM agent point	0.285		-0.045	0.056
	(0.028)***		(0.031)	(0.052)
1 sd rainfall shock	-0.007	-0.011	0.102	
	(0.029)	(0.067)	(0.023)***	
Self-reported rainfall shock	0.112	-0.039		
	(0.045)*	(0.212)		
Log per capita consumption	0.000	0.000	0.000	-0.001
	(0.000)	(0.000)	(0.024)	(0.006)
Wealth	0.019	-0.007	0.027	0.025
	(0.005)***	(0.027)	(0.011)*	(0.024)
Head education	0.000	-0.027	0.005	0.01
	(0.006)	(0.024)	(0.009)	(0.023)
Head age	0.001	0.000	0.002	-0.005
	(0.000)***	(0.000)	(0.001)*	(0.005)
Household size	0.027	0.023	0.022	0.026
	(0.003)***	(0.009)*	(0.004)***	(0.012)*
Mobile phone	0.175	0.219	0.048	0.077
	(0.004)***	(0.018)***	(0.03)	(0.126)
Number of loans	0.056	0.259	0.017	-0.164
	(0.027)*	(0.18)	(0.031)	(0.161)
Bank account	0.413	1.115	0.134	0.103
	(0.029)***	(0.164)***	(0.050)**	(0.125)
ROSCA	0.108	0.37	0.086	-1.071
	(0.047)*	(0.344)	(0.065)	(0.308)***
Occupation dummies				
Agriculture	-0.005	-0.291	0.067	-0.041
	(0.047)	(0.222)	(0.035)	(0.197)
Fishing	-0.069	0.329	0.038	-0.155
	(0.159)	(0.685)	(0.128)	(0.45)
Public sector	0.179	0.042	-0.056	0.328
	(0.073)*	(0.444)	(0.111)	(0.326)
Private sector	-0.012	0.000	0.021	0.087
	(0.042)	(0.225)	(0.07)	(0.193)
Self employed	0.006	-0.123	-0.086	0.014
	(0.039)	(0.211)	(0.054)	(0.172)
F test p value	0.002	0.147	0.071	0.249
Obs	9281	1228	9281	1228

Average marginal effects from fixed effects logit regressions. Each row is a separate regression. Columns (1) and (3) are at the household level and errors are clustered at the village level. Columns (2) and (4) are at the village level. *** p<0.01, ** p<0.05, * p<0.1

not correlated with characteristics of the households, confirming its use as an exogenous aggregate shock.

Since I have shown the rainfall shock to be exogenous, I am only concerned with factors which cause a household to use mobile money or reside in a village with other users also helping them smooth consumption following a shock. Bias such as this would inflate the size of the coefficients on these interaction terms and create the false conclusion that mobile money facilitates risk sharing rather than something else correlated with its uptake facilitating risk sharing. To control for observed factors which could be both correlated with mobile money use and help a household smooth consumption after an aggregate shock, I include the interaction of household characteristics with the rainfall shock in some of the specifications ($\boldsymbol{\psi X_{jvt}} \cdot Shock_{jvt}$).

To examine whether unobserved factors that influence mobile money update also affect shock smoothing, I use a placebo specification to test for common trends in previous shock smoothing before the introduction of mobile money, and an instrumental variable regression to instrument for mobile money use with something that can only influence consumption smoothing through the decision to use mobile money services.

The use of a difference-in-difference specification requires the common trends assumption. Common trends assumes that there are no differences in the consumption changes of users and non-users, had the users not actually used mobile money i.e. there are no time varying variables that differentially affect the mobile money using and non-using households consumption smoothing except adoption of mobile money. A example of such a violation of common trends is a time varying unobservable such as risk preferences which is correlated with both mobile money adoption and consumption smoothing ability. Such an unobservable would bias upwards my results. Under common trends, the counterfactual levels for the two groups can be different but the time trends must be the same so that in the absence of the use of mobile money the change in per capita consumption would have been the same for the two groups. I test this assumption by running a placebo test (see

Robustness section), examining if people who went on to adopt mobile money were already better able to smooth consumption in the past.

Secondly I test for endogeneity in mobile money use and consumption smoothing after a shock using an instrumental variables regression (see the Robustness Section). I use distance to the nearest mobile money agent as a variable which should be both correlated with mobile money use and uncorrelated with households' risk sharing ability. This relies on the assumption that mobile money agent location choice is also not correlated with consumption smoothing ability, bringing me onto the second potential source of bias.

The second potential source of bias is if there is self-selection by mobile money agents into villages. If mobile money agents are more likely to select into villages with certain citizen characteristics, such as wealthier inhabitants, whose income is also less sensitive to shocks, this could confound my results creating a spurious positive effect of mobile money. The roll-out of the agent network can shed some light on this. The majority of mobile money agents, especially early on in the launch of mobile money services, were existing sellers of airtime and sim cards. These small businesses already had links with the mobile operators and were spread throughout the country where mobile phone ownership was an already high 45% and cellular coverage was 75% of the population (Shkaratan, 2012).

Vodacom, the first and by far the largest mobile money operator in Tanzania, used aggregators to sign-up their existing airtime sellers as agents extremely quickly rather than dealing directly with thousands of outlets spread across the country (GSMA, 2010, International financial Coporation, 2010, USAID, 2013). This also allowed Vodacome to launch its mobile money services simultaneously nationwide instead of a regional roll-out. These aggregators provide liquidity to agents, allowing agents to be located in areas without bank access, and provide their initial training. Agents take a commission on the transactions and pay no fixed costs for being an agent, meaning that agents do not need a minimum number of mobile money users in their area to make the business viable. Since most agents operate out of an existing business selling airtime there is little movement of agents to, for

example, wealthier locations, though there is a higher density of agents in wealthier and more populated locations such as cities.

According to the Finscope (2013), in 2013 74% of households were within 1 hour of a mobile money agent, varying between 94% in urban areas and 64% in rural areas. This shows just how quickly the mobile money network was set up and how good the coverage is, especially compared to an alternative financial services such as Microfinance institution, which only 22% of the population were within 1 hours journey of in 2013. In the survey data used here only 8% of respondents reported lack of access to an agent as their reason for not using mobile money.

To examine whether mobile money agent presence is correlated with village characteristics, I run a logistic regression of the presence of a mobile money agent within the village on average observed characteristics of the village inhabitants and the rainfall shock indicators (with each covariate a separate regression) and village fixed effects to control for non-time varying characteristics of the village. Since the presence of a mobile money agent is determined at the village level, the level of observation is the village rather than the household and variables are village level means. In Table 1.2 column (2), I show the average marginal effect of each covariate.

Of these, mobile phone ownership and bank account ownership are significant at the 1% level, suggesting agents are locating in villages which show a tendency to adopt technology. Since mobile money use requires a mobile phone, it would be surprising if agent presence in a village was not correlated with mobile phone ownership. However the full set of covariates are jointly insignificant at the 15% level and other factors potentially correlated with the ability to smooth shocks such as wealth, credit access (loans) and education are not significant. Importantly, the rain shock is insignificant suggesting agents are not locating in areas which experience more or less shocks.

These correlations give an indication of whether contemporaneous characteristics of the village are correlated with agent presence but says nothing about the direction of the relationship or causation. It could equally be that mobile phone ownership increases if there is an agent in the village! I therefore also examine in the Robustness

section whether the introduction of an agent in a village was predicted by the change of average household characteristics and village services the previous year. This provides causal evidence using the time series nature of my data to see whether agents are responding to changing household and village characteristics.

As a final note, my analysis focuses on the impact of shocks which affect an entire village at once, since these are the types of shock for which mobile money accounts allow consumption smoothing, when previously this was difficult and costly. I therefore also confirm that the rainfall shocks are affecting most people in the village at once. To do this I look at the intra-class correlation, which measures the proportion of overall variance explained by within group variance. An intra-class correlation of 1 means the variable is the same for everyone in the class, here enumeration areas.

The intra-class correlation for the one standard deviation rainfall shock is 0.849 which is not surprising given rainfall was defined for an enumeration area. The fact that it is not one likely results from the clustering of city based enumeration areas together. The self-reported rainfall shock has an intra-class correlation of 0.13, showing some correlation within a village of households reporting a rainfall shock. The fact this is not higher could be because rainfall shocks have different effects on households depending on their household and plot characteristics, resulting in a house reporting a flood or drought when others in the village don't or vice versa. This also suggests the self-reported rainfall shock is less good as a measure of village shocks, and so my main results will be based on the 1 standard deviation rainfall shock definition, with the self-reported shock included for comparability with other research[5].

1.5 Results

This section begins by examining the impact of aggregate shocks on users of mobile money, non-users and non-users within villages with other mobile money users. I then look at heterogeneous effects by distance to the nearest agent, urban and

[5]such as Jack and Suri (2014)

rural households and droughts and floods. I finish by looking at remittances as the proposed mechanism through which mobile money allows shocks to be insured and examine crop yields as one mechanism through which the rainfall shock negatively affects consumption.

1.5.1 Main result

Table 1.3 shows the primary results of this paper: regression results of the impact of aggregate shocks on consumption for mobile money users and non-users in villages with and without other mobile money user, as in equation (1.1). All regressions include the full sets of household covariates from Table 1.1. All regressions also control for village characteristics which could affect the ease of sending remittances. These are the distance to the nearest main road, distance to nearest population centre and distance to nearest market. The data is also weighted in all regressions by the inverse of the probability that the observation is included in the survey. Standard errors are clustered at the village level since mobile money agents are located by village and so the decision to use mobile money will be correlated within villages but not across villages.

The first two columns of Table 1.3 show results using self reported droughts or floods, whereas the final two columns show results using the measure of a rainfall shock as a greater or less than 1 standard deviation difference from the mean. Mobile money (MM) use refers to households than use mobile money themselves. MM spillover refers to a household residing in a village with mobile money users but who doesn't use mobile money itself.

Columns (1) and (3) are difference-in-difference regressions with household fixed effect and columns (2) and (4) also include interaction terms of all the control variables with the rainfall shock. The interaction term with the shock variable, $\boldsymbol{X_{jvt}} \cdot Shock_{jvt}$, controls for any changes in observable household characteristics which might impact the household's ability to smooth shocks. It can be seen from Table 1.1 that many of the demographic variables changed over time including education, mobile phone ownership and loans which all increased across the three waves. These

Table 1.3: Impact of rainfall shocks on consumption for mobile money users and non-users

Dependent variable: Log consumption per capita				
	Self-reported shock		1 sd rainfall shock	
	(1)	(2)	(3)	(4)
Rain shock	-0.064**	-0.205	-0.068***	-0.040
	(0.029)	(0.164)	(0.020)	(0.144)
Shock*MM use	0.088*	0.042	0.121***	0.139***
	(0.051)	(0.058)	(0.041)	(0.045)
Mobile money use	0.005	0.005	-0.003	-0.007
	(0.026)	(0.026)	(0.026)	(0.027)
Shock*MM spillover	-0.057	-0.072	0.005	-0.012
	(0.049)	(0.052)	(0.040)	(0.040)
Mobile money spillover	0.002	0.003	0.005	-0.012
	(0.027)	(0.027)	(0.025)	(0.025)
Observations	9,281	9,281	9,281	9,281
Number of households	3,807	3,807	3,807	3,807
R-squared	0.194	0.198	0.196	0.202
Interactions with shock		YES		YES
Negative shock	-0.057**	-0.068***	-0.045***	-0.024
	(0.022)	(0.026)	(0.015)	(0.017)
(A) Shock MM user	0.024	0.007	0.053	0.059*
	(0.044)	(0.045)	(0.036)	(0.033)
(B) Shock Spillover	-0.121***	-0.126***	-0.063*	-0.043
	(0.041)	(0.040)	(0.033)	(0.034)
(C) Shock Non MM village	-0.064**	-0.074**	-0.068***	-0.042*
	(0.029)	(0.032)	(0.020)	(0.022)
(D) Shock Spillover $\|userX's$		-0.107**		-0.092***
		(0.032)		(0.048)
(E) Shock Non MM village $\|userX's$		-0.035		-0.081***
		(0.053)		(0.029)
F stat (A)=(C)	3.00*	2.71*	8.80***	6.31***
F stat (A)=(B)	6.31**	5.88**	5.55**	4.63***
F stat (B)=(C)	1.37	1.14	0.01	0.00
F stat (A)=(D)		3.95**		9.36***
F stat (A)=(E)		0.53		9.77***

Regressions include full set of household control variables from Table 1.1, errors clustered at the village level and control for village characteristics which could affect the ease of sending remittances (distance to the nearest main road, distance to nearest population centre and distance to nearest market). When there are interactions with the shock this set of control variables is interacted with the shock but not shown here for brevity. Mobile money use is a dummy variable equal to one if that household uses mobile money. Mobile money spillover refers to a household in a village where others use mobile money but who doesn't use themselves.
Village clustered standard errors in brackets, *** p<0.01, ** p<0.05, * p<0.1

could help a household smooth shocks, for example by mobile phone ownership providing access to information about shocks which makes it easier for households to smooth shocks. Including a set of covariates and interactions of these covariates with the shock controls for any effects of these variables on consumption smoothing.

Beginning with columns (1) and (3) (no shock interactions), the total effect of the rainfall shock is given below the coefficients and shows the average effect of the shock for the entire sample. The total effect of the rainfall shock for mobile money using households, spillover households and households in villages without mobile money are reported underneath the regression results for easy comparability across regression specifications (rows (A) to (C)).

When there are shock interactions (columns (2) and (4)), each control variable is interacted with the rainfall shock and included in the regression, though the coefficients are not reported here for brevity. To get the total effect of the rainfall shock overall and for each household type, I take the mean value of each control variable for that household type and multiply it by the coefficient on the interaction term[6]. These are summed to give the total effect of the rainfall shock for households with the mean characteristics of the entire sample, mobile money users, spillover household and non-spillover households respectively. I also show the overall effect of the shock for spillover households and households in villages without mobile money using the mean values of the control variables for households that use mobile money (rows (D) and (E)). This method accounts for how other differences in characteristics of mobile-money-using and spillover households might facilitate smoothing of the rainfall shock compared to non-users.

Looking at the regression results, the rainfall shock causes a drop in consumption of approximately 6% and is significant at at least the 10% level for households in

[6]For example, for household that use mobile money the average household size is 4.9, for spillover households its 4.8 and for non-spillover household 5.2. This mean value of the variable is multiplied by the coefficient on the shock interaction with household size for each type of household. This is repeated for all the interactions terms with the rainfall shock and these are added together to give the total effect of the rainfall shock for a household with the average set of control variables for mobile money users, spillover households and non-spillover households and for spillover and non-spillover households if they had the same mean value of the control variables as mobile money using households.

villages without mobile money (line C) in all specifications. Hence large rainfall shocks have a strongly negative effect on consumption per capita, confirming prediction 1 that in the absence of any mobile money use by the household or village, rainfall shocks have a negative effect on consumption.

Turning to the interactions with the shock dummy; the household mobile money use interaction with the aggregate shock is 9%-14% and significant in three out of four regressions. When a rainfall shock occurs the household using mobile money no longer experiences a drop in consumption and may even get a slight increase in consumption, as seen from the coefficients at the bottom of the table in row A being positive in all cases for both shock definitions. The F-test in the panel below comparing A to B and C confirms in all cases that the shock impact for mobile money users is significantly different from non-users, both in their own village and other villages. Additionally, in 3 out of 4 cases when the effect of the shock is evaluated for the spillover and non-spillover households using the mean values of the control variables for mobile-money-using households (comparing (A) to (D) and (E)) I can still reject the equality of the shocks. This confirms prediction 2; mobile money users are able to smooth consumption when a rainfall shock occurs.

In contrast, when a rainfall shock occurs, being a spillover household has a negative or zero and insignificant effect on ability to smooth a shock. Row B below the main results shows that spillover households experience a large fall in consumption of as much as 12.6% after a rainfall shock. However, the F-statistics comparing row B to C show that I cannot reject that the impact of the shock is the same (negative impact) as that of households in a village without mobile money. Hence households do not benefit from having other people using mobile money in the village when a rainfall shock occurs. This invalidates predicted 3.

The coefficient on mobile money use is small and insignificant in all the regressions. Therefore there is no increase in consumption for a household simply from using mobile money services, which could have been the case if for example using mobile money services allowed a new form of payments for a business or when selling agricultural goods, increasing sales and incomes. This is an interesting

result of itself and conflicts with Munyegera and Matsumoto (2016) but agrees with Batista and Vicente (2014) who find no impact of mobile money services on consumption. The coefficient on being a spillover household is not significant and is very small and precisely estimated. It therefore seems that non-mobile money users do not gain from having mobile money users residing in the same village as them.

Overall this suggests that when a negative shock for the village occurs, users keep any increase in remittances for themselves and do not give any detectable amount to the rest of the village to help others smooth the shock. There are many potential explanations for this result, such as changing risk sharing networks within the village or hidden income. For example, if mobile money users are choosing to insure themselves with a migrant in another location, due to a lower covariance of shocks, they may no longer participate in risk sharing within the village. I explore these in detail in the Discussion section, but understanding how mobile money services are affecting traditional risk sharing relationships is vital to understand if there are winners and losers from this new technology.

A comparison of the self-reported shock definition and actual rainfall shock definition shows that both result in an approximate 6% fall in consumption (row C). The coefficient on the shock interaction with mobile money use is smaller though, and only significant in one of the specifications. Since the self-reported shock is not highly covariate at the village level (the intra-cluster correlation is only 0.13) this could be due to it reflecting a shock that does not have a sufficient aggregate village-level impact and that can be partially smoothed within the village. This is in contrast to the 1 sd rainfall shock which impacts consumption of the entire village at once. This suggests that the 1 sd rainfall shock is a superior shock to use as an village-level shock and so only this rainfall shock definition will be used from here onwards where possible. Since the coefficients are very similar between specifications (3) and (4), only specification (3) will be used when reporting further results, though all regressions have also been run using specification (4) without any differences in results.

Heterogeneous effects

I examine a number of different effects to understand which factors might be driving the main result. These are reported in Table 1.4. I examine whether the results depend on the distance to the nearest mobile money agent, whether the effects differ by rural and urban households and lastly whether there are differential effects of rainfall shocks defined as a drought or flood ($<$ or $>$ 1sd). For all these regressions I only consider the 1 standard deviation rainfall shock measure.

The distance to the nearest mobile money agent could impact how easy it is for someone to send and receive remittances via mobile money and hence the benefit they receive from using this service. I therefore run specification (1.1) with interactions with dummy variables for whether an agent is within 1km of the village[7]. This will also indicate whether the distance to the nearest mobile money agent changes the pattern of sharing remittances within the village. For example, it might be easier to hide remittances the further away the mobile money agent is from the village.

The results for interactions with whether an agent is within or outside 1km of the village are reported in columns (1) and (2) of table 1.4. The coefficients are similar across the two regressions, and F-tests of the equality of the coefficients across the two columns fail to reject equally at the 10% level in all cases (not reported here). I therefore determine that having an agent in near proximity to the village does not affect the effectiveness of mobile money account use for smoothing rainfall shocks.

[7]I also look at agents within 5km and 10km and find similar results to 1km

Table 1.4: Heterogeneous effects of the impact of rainfall shocks on consumption for users and non-users of mobile money

Dependent variable: Log consumption per capita						
	agent<1km	agent>1km	rural	urban	drought	flood
	(1)	(2)	(3)	(4)	(5)	(6)
Rain shock	-0.051	-0.070***	-0.048**	-0.111***	-0.081***	-0.024
	(0.076)	(0.021)	(0.023)	(0.033)	(0.024)	(0.036)
Shock*MM use	0.113	0.106*	0.161**	0.120**	0.140*	0.071
	(0.085)	(0.060)	(0.066)	(0.052)	(0.081)	(0.051)
MM use	-0.005	-0.004	-0.017	-0.047	-0.011	0.025
	(0.031)	(0.031)	(0.039)	(0.041)	(0.029)	(0.028)
Shock*MM spillover	0.021	-0.039	-0.022	0.058	-0.003	0.015
	(0.095)	(0.051)	(0.056)	(0.047)	(0.065)	(0.052)
MM spillover	0.006	-0.024	0.028	-0.076**	-0.009	-0.028
	(0.031)	(0.030)	(0.037)	(0.033)	(0.026)	(0.026)
Obs.	9,281	9,281	6,518	2,763	9,281	9,281
Households	3,807	3,807	2,632	1,175	3,807	3,807
Shock	-0.036	-0.065***	0.110*	-0.081***	-0.056**	-0.009
	(0.061)	(0.019)	(0.062)	(0.024)	(0.025)	(0.026)
(A) Shock	0.061	0.036	0.113*	0.009	0.059	0.047
MM user	(0.042)	(0.059)	(0.062)	(0.042)	(0.080)	(0.037)
(B) Shock	-0.030	-0.108**	-0.070**	-0.052	-0.085	-0.009
spillover	(0.044)	(0.047)	(0.049)	(0.033)	(0.062)	(0.041)
(C) Shock	-0.051	-0.069***	-0.048**	-0.111***	-0.081***	-0.024
Non MM	(0.019)	(0.021)	(0.023)	(0.032)	(0.023)	(0.036)
F test (A)=(C)	1.75	3.10*	5.95**	5.26**	2.99*	1.95
F test (A)=(B)	2.20	3.75*	5.25**	1.29	2.02	1.03
F test (B)=(C)	0.05	0.59	0.16	1.54	0.00	0.08

Each column refers to a household fixed effects regression run when that condition is true. Rain shock is the 1 standard deviation rainfall shock dummy. Regressions include full set of household control variables from Table 1.1, errors clustered at the village level and control for village characteristics. MM spillover refers to a household in a village where others use mobile money but who doesn't use themselves. Mobile money use is a dummy variable equal to one if that household uses mobile money. The third panel reports the overall effect of the shock at the average proportion of each group in the sample and for each group separately. F tests compare whether the shock effect is equal for the three comparison groups. Village clustered standard errors in brackets, *** p<0.01, ** p<0.05, * p<0.1

Next I look at whether the results vary by whether the households resides in a rural or urban area by running the results separately for urban and rural households to see if there are differential effects for these groups. Mobile money services would be expected to benefit rural households more since they have less access to other ways to send remittances such as banks or designated money transfer shops (such as Weston Union) and are less likely to have friends or relative passing by regularly who could bring remittances. They are also more reliant on agriculture and so affected by rainfall shocks more directly in terms of crop losses than households in urban areas.

Separate urban and rural results are shown in columns (3) and (4) of Table 1.4. From the results it can be seen that rainfall shocks have negative effects in both rural and urban areas. You benefit from using mobile money when there is a rain shock in both rural and urban areas. In both these specifications I reject equality of the shock for mobile money users and non-users in villages with no mobile money users. These results suggest that both rural and urban households benefit from mobile money, possibly because they are engaging in a reciprocal relationship to share risk across urban and rural spaces, as the remittance data suggests, and because even in urban areas some households often undertake income generating activities reliant on rainfall[8]. F-tests of the coefficients across columns (3) and (4) fail to reject equality at the 10% level in all cases.

To see whether too much or too little rainfall have differential effects on consumption and the ability of mobile money to smooth these impacts, I separate out the effects of droughts compared to floods. A drought is defined as the difference in rainfall from the mean being more than one standard deviation below the mean and a flood as the difference in rainfall from the mean being more than one standard deviation above the mean. This is reported in Table 1.4, where droughts are reported in column (5) and floods in column (6).

It can be seen that it is only droughts which have a significant negative effect of 8% of per capita consumption. Floods have no significantly different effect from zero. Mobile money use is significant at the 10% level when there is a drought and the F

[8]11% of households in urban areas report the main income of the household head to come from agriculture or livestock compared to 89% in rural areas

test shows that mobile money using households are protected against the effects of a drought compared to households without mobile money users in their village. However, testing for equality of the coefficients across the columns I fail to reject that the coefficients on floods and drought interactions are equal at the 10% level.

1.5.2 Mechanisms

Remittances

The proposed mechanism in this paper is that mobile money allows remittances to be sent by friends and family in other locations in response to a rainfall shock at the village level and that this allows consumption smoothing. However, it is possible that mobile money affects consumption smoothing in other ways. One possible alternative is that mobile money allows funds to be safely stored on a mobile phone as savings which can be run down in response to a shock. A second is that households might be considered more creditworthy if they use mobile money and are able to borrow more when an adverse event happens. I am unable to explicitly reject either of these two alternative mechanisms due to limitations of the data, but the magnitude of remittance flows in response to a rainfall shock leads me to believe this is the dominant mechanism though which mobile money accounts facilitate consumption smoothing.

In the survey data used here, over 80% of respondents said they send and receive money as the main reason for using mobile money, with an equal spilt between sending and receiving as the main reason. 20% said they had ever used mobile money for savings but only 5% said this was the main reason they use mobile money. In the third round of the survey, questions were asked on who sent remittances, by what channel, from where and what their relation was to the recipient. 40% of remittances were sent by a son or daughter, 35% were sent via mobile money and 30% came form Dar es Salaam, the capital city. This suggests a story of families sharing risk with a migrant in the city via mobile money is a plausible one.

In order to test whether remittances are driving the way that mobile money protects against adverse shocks I use the data available on remittances in the third

round of the survey to run the following specification:

$$\begin{aligned} r_{jv} =& I_{jv} + \gamma_a Shock_{jv} + \mu MM_{jv} \\ &+ \beta_m MM_{jv} \cdot Shock_{jv} + \boldsymbol{\theta X_{jv}} + \varepsilon_{jv} \end{aligned} \tag{1.2}$$

where r_{jv} is whether a household received any remittances, and if they did receive any remittances the amount received by a household, and the other variables are as defined previously. Log income, I_{jv}, is included to control for income effects. $\boldsymbol{\theta X_{jv}}$ is the same vector of controls. Unfortunately, data on remittance amounts is only available for the final wave of the panel and so this specification can only be run as an OLS regression for one period. It still gives an indication though whether remittances are responding to negative shocks. If remittances are the channel through which mobile money smooths aggregate shocks then the following prediction will hold:

Prediction 4 $\beta_m > 0$

so that remittances increase for mobile money users in response to an aggregate shock.

Table 1.5 shows the OLS regressions of a dummy variable equal to one if any remittances were received (using any method of sending remittances) and, if any remittances were received, the amount received in Tanzanian Shillings, in wave 3. Mobile money use results in the households being 15% more likely to receive remittances, as can be seen in column (1). This shows how mobile money has increased the probability of receiving remittances compared to other forms of receiving remittance such as using friends and neighbours to transport money physically [9]. A rain shock does not increase the probability of receiving remittances alone or interacted with mobile money use. Rainfall shocks are unlikely to increase the probability of receiving remittances if there is no method of sending remittances. Why the rainfall shock interacted with mobile money use does not increase the probability of receiving remittances becomes more apparent when examining column (2).

[9] 35% of households send remittances via friends or family and 35% via mobile money

Table 1.5: OLS regression of remittances received after an aggregate shock

	(1)	(2)
	received remittances	value of remittances received
Rain shock	0.025	-1,589
	(0.024)	(7,010)
Rain shock*MM use	-0.019	29,582**
	(0.030)	(14,412)
MM use	0.154***	24,840***
	(0.020)	(8,404)
Control mean	0.177	35,058
Observations	3,240	3,240
R-squared	0.131	0.109

Full set of control variables as in Table 1.1 and errors clustered at the village level. Log income per capita is also controlled for. All regressions also control for village characteristics which could affect the ease of sending remittances. These are the distance to the nearest main road, distance to nearest population centre and distance to nearest market. MM use is a dummy variable equal to one if the household used mobile money in a given year. The values for amount of remittances received have been winsorized at 99%. Control mean refers to the mean value amongst those who don't use mobile money. Village clustered standard errors in brackets, *** p<0.01, ** p<0.05, * p<0.1

Looking at column (2), when there is a rainfall shock the value of remittances received increases but only for users of mobile money. When there is a rainfall shock, mobile money using households receive 29,600 shillings more ($10), approximately 4% of the median household's per capita income in wave 3 and of a similar magnitude to the fall in per capita consumption found in Table 1.3. This strongly supports the mechanism that mobile money users are able to smooth rainfall shocks since they receive more remittances, and confirms prediction 4. The point estimates for the rain shock is negative but insignificant and for mobile money users positive and significant. This suggests that mobile money is firstly increasing the probability a household receives remittances, as seen in column (1), and, when there is a rainfall shock, increases the value of remittances received. Households using mobile money receive larger values of remittances after an aggregate shock, but are not more likely to receive remittances in the first place, because households are not selecting into mobile money use based on experiencing a shock. This result therefore further

supports my use of the rainfall shock as an exogenous shock.

Overall, the results on remittances are supportive of the argument that mobile money allows the smoothing of rainfall shocks through remittance flows. However, it is still possible that mobile money accounts facilitate saving and it is through this channel that they enable consumption smoothing after a rainfall shock. The NPS data unfortunately does not include data on the value of savings in any form nor of the value of assets, meaning that I cannot directly examine saving changes by mobile-money-using and non-mobile-money-using households in response to a rainfall shock. I do however check whether household assets are declining in response to the rainfall shocks, and find no evidence of this, suggesting durable assets are not being used by households as a way to smooth consumption after the rainfall shock (results available on request).

Effect of rainfall shocks

I show that rainfall shocks have a negative effect on consumption for households without mobile money services. To examine why rainfall shocks have a negative effect on consumption I look at crop yields as the potential channel. For agricultural households, rainfall shocks would be expected to reduce crop yields, particularly droughts. To test this I create a measure of crop yields using the value of crops per acre (Gollin and Udry, 2017) to account for the large variety of different crops grown in Tanzania. I regress this on the rainfall shock dummy, controlling for plot characteristics and crop dummies .

The resulting regressions of rainfall shocks on crop yields are shown in Table 1.6. I define the rainfall shock in 3 ways; firstly as the 1 standard deviation rainfall shock, secondly only as negative deviations (droughts) and lastly as only positive deviations (floods). Overall there is no effect of rainfall shocks on crop yields but a strong negative effect of droughts and a strong positive effect of floods. When a drought occurs crop yields fall 13% and when more rainfall occurs crop yields increase 13%. This fits very much with the literature showing that crops benefit from too much rainfall (up to a point) but strongly decline from too little rainfall

Table 1.6: Effect of rainfall shocks on crop yields

Dependent variable: Log crop value per acre			
	(1)	(2)	(3)
	Rainfall shock	Drought	Flood
Shock	0.031	-0.134**	0.137***
	(0.043)	(0.062)	(0.051)
Observations	5,698	5,698	5,698
Number of households	2,646	2,646	2,646
R-squared	0.126	0.127	0.128

Crop yields defined as the log of value of crops per acre. Control variables include crop and district dummies to account for fixed characteristics of crops response to rainfall and of the value of crops in different areas. Rainfall shock refers to a 1 standard deviation in either direction deviation from mean. Drought is a negative deviation, flood a positive. Village clustered standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

(Paxson, 1992, Suri, 2005). It also fits with my heterogeneity results in Table 1.4 columns (5) and (6) showing that shocks negatively affect consumption and mobile money mitigates this is entirely driven by droughts, not floods[10]. These results therefore support my main conclusions.

1.5.3 Robustness

I run a number of difference tests to test the validity of my findings to selection into mobile money using being correlated with consumption smoothing ability. First I run a placebo regression using two periods of data before the introduction of mobile money to check if people who went on to use mobile money, and areas which gained agents, were already better able to smooth consumption compared to places with no users and agents before mobile money was introduced. Then I run an IV regression to further test for any potential self-selection into mobile

[10]These results are not directly comparable since the sample for crop yields only includes agricultural households whereas the main sample includes households not working in agriculture, for whom floods might have other negative repercussions on consumption and for whom droughts won't directly negatively affect. My results hold for urban as well as rural households suggesting other effects than just rainfall directly impacting crop yields are also causing rainfall to impact consumption

money use. These tests cannot reject my earlier results but should be interpreted with caution due to imprecise estimates.

Placebo test

The placebo test allows me to test if future mobile money exposure predicts past changes in consumption and the ability of households to smooth shocks, thus testing if my results are consistent with the common trends assumption required for a difference-in-difference specification. The placebo test can confirm whether there is something different about households and villages which will use mobile money in the future and if this difference allows households to better smooth shocks. For example, a location is more industrialised, mobile money agents are more likely to locate in industrial areas, households are more likely to use mobile money and households working in industry are better able to smooth rain shocks. This could create a spurious correlation positive relationship between mobile money use and smoothing shocks, whereas actually the industrialisation is allowing shocks to have less impact on consumption. If this case were true then the placebo test would show a positive effect of future mobile money use for smoothing rainfall shocks in the past.

I run a placebo test using the 2007 Tanzanian Household Budget survey (HBS) combined with the NPS 2008-9 wave 1 to construct two rounds of data prior to the introduction of mobile money services. A subsample of the 2007 HBS was re-sampled in the NPS, so by combining this sample with the first wave of the NPS I created a panel of 1,200 households in 191 villages covering 2 periods before the introduction of mobile money, which I call wave 0 (2007) and wave 1 (2008-9). I created a dummy variable for whether the household ever uses mobile money after it's introduction in 2009 and whether the household is ever a spillover household (lives in a village with other users but doesn't use themselves) and I use this to estimate the following equation:

$$\begin{aligned} C_{jit} =& \gamma_a Shock_{jvt} + \mu MMF_{jvt} + \beta_m MMF_{jvt} \cdot Shock_{jvt} \\ &+ \mu MMSF_{jvt} + \beta_m MMSF_{jvt} \cdot Shock_{jvt} \\ &+ \boldsymbol{\theta X_{jvt}} + \boldsymbol{\psi X_{jvt}} \cdot Shock_{jvt} + \alpha_j + \varepsilon_{jvt} \end{aligned} \tag{1.3}$$

where MMF_{jvt} is a dummy variable equal to zero in wave 0 and one in wave 1 if the household ever uses mobile money in the future and $MMSF_{jvt}$ is a dummy variable equal to one in wave 1 if the household ever lives in a village with other users but doesn't use mobile money itself in the future.

If the common trends assumption holds then the dummy for future mobile money use interacted with the shock will not be significant for the past data, confirming that it is not unobservable characteristics of households which choose to use mobile money that is driving my results. Likewise, the dummy for being a spillover household interacted with the shock will not be significant for the past data. A finding of no significance on these coefficients suggests that the households and villages where mobile money are used were similar in their ability to smooth shocks to the household and villages where no-one uses mobile money before the introduction of mobile money services, which is crucial for the validity of my results.

I constructed the same measures of control variables to match those used for in the main analysis such as household head education, household size, financial use and occupation dummies. The rainfall shock variable used here is a 1 standard deviation rainfall shock.

The results in Table 1.7 show the placebo test of future mobile money use or presence on consumption per capita[11].

I find no significantly different effect for households which use mobile money or reside in village with other users in the future on their consumption or ability to respond to a shock in the past. These are confirmed in the F statistics reported at the bottom of the table; the shock has no differential effect on households which go on to use mobile money or that become spillover households in the future. However, the rainfall shock is not leading to a significant fall in consumption in the placebo sample, and the coefficients on the shock interactions with future mobile money use are positive, and the standard errors are large, meaning I may be underpowered to detect effects. If anything though, the point estimate on the

[11]Note, the HBS and NPS used slightly different ways of recording consumption, principally in pricing of home produce, the number of food items, and recall versus diary methods (Tanzania National Bureau of Statistics 2009).

Table 1.7: Placebo test of future mobile money use 2007-2008/9

Dependent variable: Log consumption per capita	
	(1)
	FE
Rain shock	-0.023
	(0.053)
Shock*MM ever	0.024
	(0.079)
MM ever	-0.007
	(0.055)
Shock*MM spillover ever	0.057
	(0.088)
MM spillover ever	-0.030
	(0.058)
Observations	2,173
Number of households	1,183
R-squared	0.557
Negative shock	-0.008
	0.044
(A) Shock MM user	0.00
	0.073
(B) Shock spillover	0.034
	0.080
(C) Shock Non-MM village	-0.026
	0.054
F stat (A)=(C)	0.09
F stat (A)=(B)	0.09
F stat (B)=(C)	0.42

All regression include a full set of control variable from Table 1.1, time dummies and clustered errors at the village level. The shock variable here is a 1sd rainfall shock. This regression uses 2007 and 2008-9 data with a dummy variable MM ever if a household ever used mobile money after 2009. MM spillover ever refers to a household who will live in a village with at least one mobile money user in the future but never uses mobile money itself. Village clustered standard errors in brackets, *** p<0.01, ** p<0.05, * p<0.1

shock interacted with being a spillover household is over twice the size as the coefficient on the shock interacted with mobile money use, suggesting if anything spillover households were better able to insure shocks in the past and yet I find no positive spillover effect of mobile money.

These results are consistent with the common trends assumption needed for my use of a difference-in-difference specification to be valid. However the large standard errors suggest I cannot confirm that households which went on to use mobile money weren't already better able to smooth shocks in the past, which may create positive bias in my results.

I also test whether agents choose to locate in areas with certain characteristics. To do this I use the wave 0 data to examine whether changes in the average characteristics of households in a village or village services changes predict mobile money agent presence in a future period. The time series nature of the data therefore allows me to see whether the roll-out of agents was effectively random or could be causing spurious results by being correlated with village characteristics which facilitate risk sharing. In particular, this allows me to test if other aggregate shocks, which appear as changes in total village consumption, influence agent location choice. These results are shown in Table 1.8.

I find no effect of changes in any of the average characteristics of the households or the village services on either agent presence in a village or within 5km. Only one service is significant, with places that gained a primary school less likely to get a mobile money agent, but this could be due to chance considering the number of variables tested. Jointly, all the variables are not significant at predicting whether an agent decides to set-up in that villages, as shown by the very small F statistic. However, the standard errors on all the coefficients are very large, hence I should interpret this as a lack of evidence against my main results rather than support for them.

Table 1.8: Agent location choice

	(1) Agent in village	(2) Agent within 5km
Consumption	0.045	-0.030
	(0.102)	(0.100)
Rainfall shock	0.045	-0.011
	(0.077)	(0.075)
Distance to road	0.001	0.001
	(0.001)	(0.001)
Head age	-0.004	-0.010
	(0.010)	(0.010)
Head Education (yrs)	0.033	-0.001
	(0.035)	(0.034)
Household size	0.040	0.037
	(0.033)	(0.032)
Mobile phone	-0.013	0.029
	(0.182)	(0.178)
Loan	-0.017	-0.064
	(0.221)	(0.216)
ROSCA	-0.215	-0.260
	(0.220)	(0.215)
Occupations		
Agricultural	-0.234	-0.262
	(0.186)	(0.182)
Fishing	-0.272	-0.325
	(0.715)	(0.700)
Public sector	-0.130	0.057
	(0.774)	(0.757)
Private sector	-0.248	-0.073
	(0.284)	(0.279)
NGO/religious	0.422	0.721
	(1.014)	(0.993)
Self-employed	0.296	0.224
	(0.183)	(0.180)

In columns (1) and (2) each coefficient is run as a separate fixed effect regression at the village level with village fixed effects to control for non-varying characteristics of the villages. Mean dependent refers to the mean value of the dependent variable across the villages. Standard errors in brackets, *** p<0.01, ** p<0.05, * p<0.1

Table 1.8: continued...

Village services		
Bank	0.085	0.105
	(0.131)	(0.128)
Health centre	0.075	0.023
	(0.063)	(0.062)
Hospital	-0.052	-0.081
	(0.097)	(0.094)
Primary school	-0.207**	-0.278***
	(0.095)	(0.093)
Secondary school	0.041	0.028
	(0.068)	(0.069)
Police station	0.109	0.045
	(0.087)	(0.085)
Post office	-0.179	-0.149
	(0.126)	(0.123)
Mean dependent	0.26	0.30
Observations	382	382
F statistic	0.75	0.71

In columns (1) and (2) each coefficient is run as a separate fixed effect regression at the village level with village fixed effects to control for non-varying characteristics of the villages. Mean dependent refers to the mean value of the dependent variable across the villages. Standard errors in brackets, *** p<0.01, ** p<0.05, * p<0.1

IV regressions

As discussed in the Identification Section, if mobile money use is endogenous to household's ability to smooth consumption after a shock then my results will be biased. I can use instrumental variable regression to control for this. Given there are two potentially endogenous variables, mobile money use and the shock interaction with mobile money use, I need two instruments. An instrument would need to be correlated with mobile money use (relevant) but not affect consumption smoothing independently (exogenous).

An instrument that has been used in the literature is distance to the nearest mobile money agent. Distance to the nearest agent is correlated with mobile money use since it is easier to use mobile money if an agent is nearby. And it shouldn't be correlated with unobserved determinants of the ability of households in a village to

smooth risk. In addition to the distance to the nearest agent, which is only present in the case there is no agent within the village, I also have a dummy indicating whether there is a mobile money agent in the village or not. Together these two variables define access to a mobile money agent, since it is only when there is no agent in the village that the distance to the nearest agent matters. I therefore instrument with the distance to and presence of a mobile money agent and the interactions of distance to and presence of a mobile money agent with the rainfall shock.

Table 1.9: IV results

Dependent variable: Log consumption per capita

	(1) IV FE
Rain shock	-0.120**
	(0.058)
Rain shock*MM use	0.297
	(0.198)
Mobile money use	-0.184
	(0.295)
Rain shock*MM spillover	0.057
	(0.072)
MM spillover	-0.037
	(0.111)
Observations	4,448
Number of households	2,224
F-stat on excluded instruments for MM use	19.7
F-stat on excluded instruments for Rain shock*MM use	27.9
Cragg-Donald Wald F statistic	25.1
Underidentification test χ^2 p-value	0.00
Sargan-Hansen test χ^2 p-value	0.43

Instruments were distance to nearest mobile money agent and a dummy indicating a mobile money agent within the village, and their interactions with the shock variable. All regressions include full set of household control variables from Table 1.1, household fixed effects and errors clustered at the village level. All regressions also control for village characteristics which could affect the ease of sending remittances. These are the distance to the nearest main road, distance to nearest population centre and distance to nearest market. MM spillover refers to a household in a village where others use mobile money but who doesn't use themselves. Mobile money use is a dummy variable equal to one if that household uses mobile money. Rainfall shock is the 1 standard deviation rainfall shock.
Village clustered standard errors in brackets, *** p<0.01, ** p<0.05, * p<0.1

The instrumental variable results in Table 1.9 confirm that the rainfall shock has a negative effect on consumption per capita of 12%. None of the other coefficients are significant, though the point estimate on the mobile money shock interaction is positive and large, but with large standard errors. If anything the coefficient on the interaction of mobile money use with the shock seems of an overly large magnitude of over 2.5 times the size of the negative shock. The standard errors on all the coefficients are large, suggesting the strength of my instruments should be examined. Additionally, I am trying to instrument for two endogenous variables which also results in a loss of precision and hence high standard errors.

I report a number of different tests beneath the regression results to confirm the validity of my instruments. Firstly, the F-statistics on the first stage regressions for mobile money use and the interaction with the shock are moderately large, showing that the instruments do explain variation in these variables and hence supporting their use as valid instruments. The Cragg-Donald Wald F statistic is used to test the strength of more than one excluded instruments. For the IV regressions to have less than 5% of the bias of OLS the critical value for the F-statistic is 13.43. My Cragg-Donald Wald F statistic exceeds this critical value, again supporting the validity of my instruments.

The underidentification test tests the null hypothesis that the equation is underidentified. The null is rejected at the 1% level. The Sargan-Hansen test is an over-identification test which determines if the instruments are exogenous. The null hypothesis is that the instruments are uncorrelated with the error term. Since my errors are clustered, the Hansen J statistic is reported. I cannot reject the null hypothesis that all my instruments are valid.

Overall, my IV results cannot reject my main findings, but are not precise enough to definitely rule out endogeneity.

1.5.4 Discussion

Exploring why only mobile money users can smooth aggregate shocks would advance our understanding of how informal risk sharing arrangements operate in developing countries and how they are changing in an increasingly connected world. For example, are mobile money using households choosing to hide remittances (Kinnan, 2017) or are they opting out entirely from the risk sharing network in the village to share with a wider network of family anf friends (Fafchamps and Lund, 2003)? Looking at transfers within the village and from a migrant in response to shocks would allow these effects to be separated. My data was not detailed enough on specific network partners and remittance flows to answer this question, but the answer is important for understanding how new technologies change traditional risk sharing patterns and for a deeper understanding of how risk sharing networks are sustained.

There are many potential explanations for why remittances might not be shared within villages after an aggregate shock. The main one discussed here is mobile-money-using households are not participating in risk sharing networks within the village but instead insuring within wider networks of family and friends outside the village.

Mobile-money-using households have a cheap and easy way to share remittances across wide ranging geographical locations. They could insure both village-level rainfall shocks and idiosyncratic shocks through family migrants in other locations. The benefits of this risk sharing relationship would exceed those of the village risk sharing relationship as a cross-locational network allows for a lower correlation in income shocks and for the insurance of village-clustered shocks such as the rainfall shocks examined here. In this case, mobile money using households might no longer participate at all in village risk sharing.

It would be interesting future work to examine how money money is changing risk sharing relationships, potentially allowing them to cover wider geographical areas and so insure more different types of shocks. It is also important to understand under what circumstances and with who they still share risk within the village. This raises the question of if there is a shrinking pool of non-mobile-money-using

households in the village without a potential migrant or too old or poor to make use of mobile money struggling to share risk with each other. While mobile money benefits users, it could increasingly penalise those unable or not willing to use it if users share less and less within the village.

1.6 Conclusion

Household in developing countries are subject to large changes in consumption due to aggregate shocks, such as rainfall shocks, which negatively affect the consumption of the entire village at one. Droughts and floods are a major source of risk to developing households, and measures which help protect against these, ranging from social protection to micro-insurance, are key new areas of research. Mobile money services are a new and fast growing technology which can help households insure their consumption against aggregate shocks by providing access to remittances from other locations not affected by the shock.

In this paper I show that large rainfall shocks negatively affect household consumption but that mobile money use can mitigate this impact by allowing the easy sending and receiving of remittances. This effect is found regardless of whether a mobile money agent is nearby the village or not, and I cannot reject equal results regardless of whether a household is urban or rural or the shock is a drought or flood. I confirm the mechanism through which mobile money accounts enable risk sharing by looking at flows of remittances after a rainfall shock and find the value of these flows can cover the majority of a consumption drop. I also test the mechanism by which rainfall shocks lead to consumption shortfalls by examining the effect of rainfall shocks on crop yields for agricultural households, finding that droughts lead to a 13% reduction in crop yields. I confirm the robustness of my findings using a placebo test from before the introduction of mobile money to test that households which did and didn't go on to use mobile money didn't already differ in their ability to smooth risk. I also use instrumental variable techniques to address potential self-selection effects into mobile money use. Both of these are unable to reject my main results.

Mobile money only helps smooth rainfall shocks if your household uses mobile money, it is not enough for neighbours in your village to be users. This raises the question of why remittances aren't being shared with others in the village after an aggregate shock. One potential explanation discussed is that mobile money allows easy sending of remittances across networks with members in different locations, such as between family migrants to cities and the remaining family in a village, therefore decreasing reliance on traditional village risk sharing networks or networks confined to a small geographical area. This explanation would be interesting to understand in more detail and in particular to understand if this additional option for mobile money using households comes at the expense of non-mobile-money-using households in their villages who are more reliant on risk sharing with members of their village.

Research on the impact of mobile money services is still in its infancy and further work could explore in more detail how access and use of mobile money services changes traditional risk sharing arrangements and who the winners and losers are from this.

Chapter 2

Hiding loans in the household using mobile money

Experimental evidence on microenterprise investment in Uganda*

Abstract I examine whether changing the way microfinance loans are disbursed to utilise widespread mobile money services impacts the businesses of female microfinance borrowers. Using a field experiment of 3,000 borrowers of BRAC Uganda, I compare disbursement of a loan as cash to disbursement of a loan onto a mobile money account. After 9 months, women who received their microfinance loan on the mobile money account had 15% higher business profits and 11% higher levels of business capital. Impacts were greatest for women who experienced pressure to share money with others in the household at baseline, suggesting that providing the loan in a private account gives women more control over how the loan is used.

2.1 Introduction

Microfinance loans are extremely popular in developing countries, with an estimated 140 million clients worldwide, two-thirds of whom are women, and had client growth of 6% a year in 2017. This strong growth in borrowers is despite recent evidence showing that microfinance loans have not led to improvements in business

*The trial is registered at https://www.socialscienceregistry.org/trials/1836 and the pre-analysis plan was uploaded there on 11th December 2017 before endline data collection had finished and analysis begun. An amendment to the pre-analysis plan documenting further intermediate outcomes based on admin data was lodged on the 31st July 2018. All analysis in this paper follows these pre-analysis plans unless clearly stated otherwise.

profits, or wider improvements in household outcomes, particularly for women's businesses (Banerjee et al., 2015, De Mel et al., 2008). However, under certain conditions women's businesses can benefit from microfinance loans and grants (Bernhardt et al., 2017, Blattman et al., 2014, Fafchamps et al., 2014, Fiala, 2017, Field et al., 2013). This suggests that finding ways to help female entrepreneurs overcome key constraints to investing the loan in their business could improve business performance.

In this paper I examine how providing microfinance loans using a widespread financial service, mobile money accounts, impact women's businesses. I use a Randomised Controlled Trial of 3,000 female microfinance clients in Kampala, Uganda. Existing and new clients of BRAC Uganda who applied for a new loan for their business were individually randomised into two treatments.

Treatment One - Mobile Account: A mobile money account explicitly designated for the woman's business was provided to the woman, but the default method of disbursing the microfinance loan as cash was retained.

Treatment Two - Mobile Disbursement: The same mobile money account as Treatment One was provided to the woman, but the microfinance loan was disbursed onto the mobile money account rather than disbursed as cash.

A control group continued to receive their microfinance loan as cash and were not given a mobile money account. This sample already had very high (97%) rates of mobile money account usage, so these treatments principally look at the impact of designating a mobile money account for business use and the payment of a microfinance loan onto this account, rather than studying any impacts of the initial take-up of mobile money.

This study takes advantage of several features of mobile money accounts which facilitate secure saving, earmarking, and keep money safe from others: mobile money accounts enable separate accounts to be opened for different uses, are provided in an individual's name, can only be accessed or the balance checked by the individual and require the small barrier of going to an agent to withdraw money.

My research design allows me to examine two research questions. Firstly, how does the business-designated mobile money account, and obtaining the loan directly onto the mobile money account, impact women's savings and businesses? Secondly, do women with certain characteristics, such as low self-control or higher pressure to share money with spouse and family, benefit more from the mobile money account and disbursement of the loan onto the account?

I find that 8 months after providing the microfinance loan on a business-designated mobile money account (the Mobile Disbursement treatment) there is a 15% increase in business profits and an 11% increase in the value of business capital compared to providing the loan as cash. These findings are robust to multiple testing corrections and alternative specifications. I do not find that the Mobile Disbursement treatment has any impact on savings by the time of the endline survey, a result that is corroborated by transaction records, which show the balances on the accounts had fallen to near zero by 6 months after the loan disbursement.

I examine the potential mechanisms by which the Mobile Disbursement treatment had an impact on the women's businesses by looking at whether the treatment targeted primarily self-control difficulties, helped women to resist family pressure or just provided a safe place to store money. I find that those who experienced most pressure at baseline to share money with family experience the largest treatment impacts on their businesses from having their loan disbursed on a mobile money account: this group see a 25% increase in business profits from receiving their loan on a mobile money account and a 17% increase in business capital, compared to the control. I validate this by examining expenditure patterns, and see less of the loan going to the family, and less transfers to the spouse, of women assigned to the Mobile Disbursement treatment compared to those who got their loan as cash. I do not see heterogeneous impacts of either treatment by an index of self-control difficulty or evidence that the women were saving constrained. This suggests the Mobile Disbursement treatment worked primarily by providing a way to keep the loan hidden from family in a safe format.

I examine a number of different alternative hypothesis to explain the impact of the Mobile Disbursement treatment. Firstly, I look at whether the increase in profits is just a redistribution of income within the household, with other household businesses losing out. Secondly, I look at whether the mobile money account, which is designed for sending money, changes remittances flows. Thirdly, I examine experimenter demand effects and whether the Mobile Disbursement treatment led to misreporting of business outcomes. Fourthly, I look at measurement error and whether the Mobile Disbursement treatment increased the accuracy of business accounts. Lastly, I look at social network changes and whether the women saw a reduction in risk sharing as a result of the treatments. I do not find compelling evidence for any of these potential explanations.

Using transaction records provided by the telecoms operator, I see that the Mobile Disbursement treatment group did not seem to use the accounts for regular deposits of their own money: only 13% ever make a single deposit of their own onto the account, and these are for very small amounts (median 20,000 USH ($5.33)). This suggests that receiving the loan on a mobile money account did not cause women to learn about the benefits of putting money onto the account themselves. This fits with other research studies that have found that just depositing money into an account does not necessarily cause people to use it more (Field et al., 2016, Somville and Vandewalle, 2018).

Instead, I see those assigned to the Mobile Disbursement treatment hold significant balances on the account: on average, those who received their loan on the mobile money account hold 100,000 USH ($27), equal to 7% of the loan value, on the account during the first 30 days of account ownership[2]. The Mobile Disbursement group appear to retain some of the loan on the mobile money account and draw this down over a 6 month period by making multiple withdrawals. By the end of 6 months, balances are, on average, a very low 190 USH ($0.05). This suggest, since deposits are so small, that those assigned to the Mobile Disbursement

[2]I exclude the day of loan disbursement from this. Therefore if someone assigned to the Mobile Disbursement treatment withdrew the entire loan on the first day, their average balance over 30 days would be zero.

treatment retained some of the loan on the mobile money account after it was first disbursed and used the mobile money account as a way to safety and privately store the loan and draw on it as needed.

I do not find any effects just from getting a business-designated mobile money account, the Mobile Account treatment. Only 13% of the Mobile Account group also ever deposited money onto the account, and the amounts deposited are equally small as for the Mobile Disbursement group (median 27,000 USH ($7,20)). The lack of use may be because 97% of the sample had already used a mobile money account before at baseline and so simply designating an account for their businesses did not have a material impact. This is despite other studies finding large impacts on saving from similar treatments through mental accounting effects, although these studies offered additional incentives to save such as removing fees or paying interest (Dizon et al., 2017, Habyarimana and Jack, 2018). However, my results fit with another study that found that facilitating mobile money transfers into bank accounts did not led to more deposits for most people (De Mel et al., 2018).

This paper adds to the literature in four broad areas: returns to women from investment in microenterprises; social pressure for women to share money; the uses and benefits of mobile money services; and default effects in payment and saving mechanisms.

While a large body of the microfinance literature has found low returns to business investment by female entrepreneurs (Banerjee et al., 2015, De Mel et al., 2008), a subset has found that the returns can be high (Blattman et al., 2014, Field et al., 2013). If women with already profitable businesses receive in-kind grants instead of cash they see a rise in their business profits (Fafchamps et al., 2014). Likewise, if women are able to hide money from their spouse, or live in households with no other members who have businesses, they also see profit gains from business loans and grants (Bernhardt et al., 2017, Fiala, 2017). These papers suggest[3] providing loans or grants to women in a manner that's harder for other household members to dip into allows the money to be used for the woman's

[3]Note that Fafchamps et al. (2014) found evidence it was actually self-control not family pressure that was the mechanism by which the in-kind grant resulted in higher business investment.

business and hence leads to investment and profit growth. This paper adds to that subset by showing that if female entrepreneurs are given their loan in a manner that is easy to conceal, secure and private, on a mobile money account, they can experience high returns on their business investment.

This paper also adds to the literature looking at family pressure to share money by highlighting that it is through reducing this pressure that the mobile money accounts enable microfinance loans to be invested in the women's businesses. Women have frequently demonstrated a willingness to hide money from their spouse and family even if costly to do so (Almas et al., 2015, Boltz et al., 2017, Castilla, 2018, Jakiela and Ozler, 2015). Women may also prefer to take loans even when they have savings, in order to make their family think they do not have much money and so reduce sharing pressure (Baland et al., 2011). Women also use strategies to try to retain control over their money and reduce spousal access to it. When given the opportunity, women will choose to have an individual saving account over a joint account if they are not well matched on saving preference with their spouse, even if they give up interest as a result (Schaner, 2015). Women may also stop using bank accounts if access to the account by their spouse is made easier (Schaner, 2017).

These strategies benefit women by moving outcomes towards their preferences: evidence shows that there are improvements in female-decision-making power in the household when women are given access to their own saving accounts (Ashraf, 2009) and that providing money in a way women can control gives them more say over how the money is used (Aker et al., 2016, Attanasio and Lechene, 2002, Field et al., 2016). This study adds to this literature by showing that if microfinance loans are given in a way that takes into account the social constraints women face and facilitates hiding of money from the spouse, they gain more control over the use of these funds.

This research builds upon the early work on mobile money services (Jack and Suri, 2014) and examines how their integration into financial products can make these products better meet the needs of the poor. While mobile money accounts were discussed as a storage device from the earliest research studies, until recently there was limited evidence on their potential to act as saving devices (Jack and

Suri, 2011, Mbiti and Weil, 2011, Morawczynski, 2010). A series of recent RCTs have changed this by exogenously providing mobile money accounts labelled for specific uses along with varying interest rate incentives or automatic payment mechanisms (Bastian et al., 2018, Batista and Vicente, 2017, Blumenstock et al., 2018, Dizon et al., 2017, Habyarimana and Jack, 2018, Lipscomb and Schechter, 2018). This literature has found mobile money accounts to be an effective way to save for business expenditures, school fees, health expenses, agricultural inputs and unexpected shocks. However, my paper conflicts with these results by finding no impact of a labelled mobile money saving account.

Studies have also looked at whether integrating mobile money accounts into cash transfer and wage payment mechanisms changes how these income sources are used (Aker et al., 2016, Blumenstock et al., 2015). My study extends this literature by examining whether changing the payment mechanism of microfinance loans to take advantage of the saving features of mobile money accounts changes how the loans are used. To my knowledge, this is the first experiment looking at the impact of integrating a formal financial instrument of any kind into a microfinance loan product.

Lastly, my paper shows that the default choice matters: even small costs of switching prevent the Mobile Account treatment group from imitating the Mobile Disbursement group by depositing the loan onto the mobile money account provided to them. The default for the Mobile Disbursement group of keeping the loan on the mobile money account has large follow-on impacts by ensuring left-over funds remain on the account as savings, and also reducing the trickle of money from the account into other uses and other people's hands. A growing literature in developing countries has shown that defaulting savings into saving accounts or similar formal financial devices results in higher savings (Blumenstock et al., 2018, Brune et al., 2017, 2016, Somville and Vandewalle, 2018) and higher control of the money for women (Field et al., 2016). My research suggests that formal bank accounts or saving devices with restrictive commitment features aren't needed to help women save and

invest their microfinance loan in a way that's aligned with their needs. Instead the default just needs to be that the loan is kept as savings unless proactively spent.

The rest of this paper is organised as follows: Section 2 discusses the interventions and experiment design. Section 3 goes over the data used in this study. Section 4 contains the empirical specification and section 5 the results. Section 6 discusses mechanisms and section 7 alternative explanations for my results. Section 8 concludes.

2.2 Background and experiment Design

2.2.1 Background

Mobile money services began in Kenya in 2007, and rapidly spread in East Africa. 51% of the population used mobile money services in Uganda in 2017 (Demirguckunt et al., 2017) and over 40% of users are women. Mobile money services operate via a simple SMS-message interface on a sim card to allow the transfer and storage of up to $1000. The account is PIN protected and so can only be accessed by the owner provided this PIN number is kept private and the sim card secure. Withdrawal and deposit of money take place using widespread networks of mobile money agents, who are found throughout a city like Kampala. Mobile money services are increasingly being integrated in bank account offerings and the mobile money operators themselves are increasingly offering services ranging from bill payment to providing short term loans.

2.2.2 Setting

The study location is Kampala, Uganda, chosen as it has both a high prevalence of microfinance borrowing and high mobile money penetration, with 83% of the population owning a mobile mobile account (Mayanja Lwanga, Annet Adong, 2016). The study took place in 6 of the 14 microfinance branches of BRAC Uganda, chosen as they had a existing bank account with Stanbic bank, whose online banking platform had pre-existing mobile money transfer integration which was utilised to make the mobile money disbursements directly from the branch bank account.

BRAC Uganda is one of the largest providers of financial services to the poor in Uganda. It offer microfinance loans to women only of between 250,000 USH and 4mn USH ($70 - $1000) for expanding a small enterprise. Owning an existing enterprise is a pre-requisite for obtaining a microfinance loan, and a check of the business is carried out by credit officers before a loan is given. Loan durations vary between 20 and 40 weeks depending on the needs of the woman, with the interest rate set at 13% for the 20 week loan and 25% for the 40 week loan. Women apply for loans in groups of between 8 and 30 women, and each woman meets weekly with the other members of her group to repay their loans. Groups are not formally liable for repayment of their members' loans, and women each have a guarantor from outside the group who is meant to repay the loan if a woman defaults.

The study population was composed of any microfinance client applying for a new loan (whether as a first time borrower or a repeat loan) who owns a mobile phone of her own. The mobile phone requirement was not binding in this urban sample, and only 6 women were excluded from taking part in this study because they did not have their own mobile phone. This sample of women is therefore highly representative of female microfinance clients throughout Kampala, and likely similar to other urban populations of microentrepreneurs.

2.2.3 The intervention

The study involved two interventions:

Intervention One - Mobile Account

Women seeking a loan from BRAC were randomly offered a mobile money account designated for their business. Women were provided with a new sim card, helped in setting up their mobile money account and trained how to use it. The account was described as specifically for their business but no formal restrictions were placed on how they use the account nor money paid into the account. Women in this group receive their microfinance loan as cash.

Intervention Two - Mobile Disbursement

Women seeking a loan from BRAC were offered the same business mobile money account as in Intervention One but, additionally, their microfinance loan was paid directly into this account through a mobile money provider. An additional amount was included to cover the fee of approximately 1% of the loan amount for withdrawing the money from an agent so as not to disadvantage women receiving the loan this way[4]. This was fully explained so as to maximize take-up.

Features of the treatments

The treatments could have a number of different impacts both on how the loan is utilised and on savings. I classify these effects broadly as flexible saving device, mental accounting, commitment, and default effects.

Flexible saving device The mobile money account may provide a flexible and safe storage device for savings and so negate the need to hold savings as cash. This may decrease unplanned expenditures on personal items or pressure to give money to others. The latter has been shown to be particularly important for women, who are prepared to give up significant amounts of resources to keep money hidden rather than have to give it to family and friends (Boltz et al., 2017, Castilla, 2018, Goldberg, 2017, Jakiela and Ozler, 2015).

At baseline, 20% of the sample reported carrying some savings as cash, despite also using more structured saving devices like bank accounts and ROSCA. Prior research has shown that people are willing to pay to use mobile money accounts to avoid carrying cash (Economides and Jeziorski, 2017). The mobile money account may represent an in-between point of flexibility compared to the ways women currently save: it is more accessible than a bank account or ROSCA but less accessible than cash. It therefore may function more like a debit card does in developed countries, keeping money out of view but providing easy access when needed.

[4]This amount would cover 5 withdrawals of approximately one-fifth of the loan.

Mental accounting The mobile money account may increase savings through mental accounting effects. Evidence suggests that simply labelling something as a saving account can increase savings (Thaler, 1985, 1999). Previous studies have found that a separate, labelled mobile money account can increase saving for the labelled purpose (Dizon et al., 2017, Habyarimana and Jack, 2018, Lipscomb and Schechter, 2018). Money in this saving account is viewed as being unavailable for day-to-day spending. This therefore helps people to resist the temptation to spend the money on other things or to resist pressure to give money to other people. During focus group discussions, some of the women discussed using the fact that the loan was disbursed into a mobile money account explicitly for their business as a way to deter requests for money. They found it easier to argue that the loan can only be used for their business when it was so obviously in an account assigned for that purpose.

Soft Commitment device Providing the microfinance loan on a mobile money account may act as a soft commitment device compared to giving the loan as cash as it requires a trip to a mobile money agent to actively withdraw money before spending it. This contrasts with cash, which is easy to spend instantly. This would not necessarily be the case if paying for goods with mobile money was common, but less than 1% of mobile money users have used it to pay for goods at a store or shop (Intermedia, 2016).

The commitment features of the mobile money account may help to resist the pressure to give money to others. While sending money to others is a feature of mobile money accounts, it still requires more steps than to simply hand them some cash. It also requires the receiver to withdraw the money from an agent the other end and to pay a fee. This may therefore be enough to dissuade others that it is worth asking for money from the women.

The evidence on whether commitment is needed over just a safe place to store money is mixed. Individuals have been shown to have a demand for commitment, although strong take up alone doesn't mean savings will be large (Ashraf et al., 2006). Many papers are now showing that simply providing a safe storage device

is enough to increase savings and have benefical effects on downstream outcomes, including microenterprise growth, and commitment does not increase savings further (Blumenstock et al., 2018, Brune et al., 2016, Dupas and Robinson, 2013a,b, Habyarimana and Jack, 2018, Kast and Pomeranz, 2014, Lipscomb and Schechter, 2018, Prina, 2015, Schaner, 2018).

Default effects A common theme across these mechanisms is the default difference between the treatments. The Mobile Account treatment requires active deposit of funds onto the account for any of its saving, mental accounting or commitment features to be relevant. The Mobile Disbursement treatment however, automatically provides a safe place labelled for the business to store the loans until the money is actively withdrawn. Prior literature has shown default effects about whether money is given as cash or into a saving account to be an important predictor of savings (Blumenstock et al., 2018, Brune et al., 2017, Somville and Vandewalle, 2018).

2.2.4 Experiment design

The study involved 3,000 female micro-entrepreneurs, split as follows. 1,000 acted as controls receiving the microfinance loan in the usual way as cash and nothing else. 1,000 were signed up for a business designated mobile money account but still received their loan as cash. 1,000 were signed up for the business designated mobile money account and received their loan on that account.

All other aspects of the BRAC microfinance loan product remained the same, including the requirement to be physically present at the branch for the disbursement of the loan and signing of final agreements, and the repayment of the loans via weekly group collection meetings within the borrower's community.

Randomisation took place weekly in blocks of 150-200 women determined by the timing of requesting a new loan. All women who were both accepted for a loan with BRAC and who had a mobile phone were individually randomised into the treatment or control groups. This continued for approximately 5 months until the sample size of 3,000 was achieved.

The randomisation was done in Stata. It was stratified by 5 variables: present bias, behaviour in a willingness-to-pay-to-hide-money game (see Section 2.3.1), first time borrower with BRAC, microfinance branch and also by business profits at baseline. The first two variables were chosen based on the idea that women who are present bias or show a desire to hide money from their spouse might benefit more from having their loan disbursed on a mobile money account instead of as cash. I stratified by first time borrower and branch in case there were systematic differences between new and established entrepreneurs and to ensure an even amount of mobile money disbursement by branch. I stratified by profit since Fafchamps et al. (2014) showed heterogeneous effects of loans for women based on their profitability.

For those assigned a treatment, the treatment was offered when the woman went to have her loan disbursed. At this point, if she was assigned to the Mobile Account treatment she was offered a mobile money account and trained in how to use it. The account was framed as for her business, but without any constraints on how it was actually used. Women were free to refuse the account if they wanted.

If she was assigned to the Mobile Disbursement treatment, she was offered both the mobile money account and to have her loan disbursed on this account. She could refuse either the disbursement and/or the sim card, permitting partial compliance if she wanted the sim card but not the disbursement. The additional amount to cover fees was explained to the woman and the same training and framing as in as for the Mobile Account treatment given.

Although the treatments were offered to individuals as they applied for a new loan, it remained the case that women met in groups to repay their loans. Therefore, within any group, there would be a mix of women over time who were recruited into the study and assigned to the treatment and control groups, as well as some women who were still paying back a previous loan and were not in the study.

2.3 Data

I have four sources of data for the analysis. Firstly, a baseline survey was conducted on all women applying for a new loan at the 6 BRAC microfinance branches. Baseline

surveys were conducted between January and June 2017 before randomisation and assignment to treatment group occurred. Approximately 1 week after the baseline survey, randomisation took place and the woman's loan was disbursed by BRAC in the assigned manner. Lists of treatment assignment were sent to the BRAC branches weekly, and only women who had been baselined and assigned a treatment could have a loan disbursed to them. This ensured that all women applying for loans during this 5 month period were part of the study.

Secondly, an endline survey was completed. The endline survey began in October 2017 and ran until January 2018. This is approximately 8 months after the loan disbursement, and was chosen so that those women who had 40 week loans[5] were still repaying them when the endline survey took place, helping to reduce attrition.

Thirdly, focus groups were conducted with a sample of 16 women from 3 different microfinance groups during September 2018. There were 8 women from the Mobile Disbursement treatment, 5 from the Mobile Account treatment and 3 from the control group. The purpose of these focus groups was to obtain qualitative, descriptive information on how women used the mobile money accounts and how they felt they affected their businesses, along with a comparison to the control group. Though this is a small sample, the focus groups give richness and a deeper understanding into the mechanisms by which the treatments had an impact.

Finally, I obtained transaction records obtained from MTN Uganda of all the mobile money transactions between January 2017 and January 2018 made using the mobile money accounts provided to clients as part of the study. All respondents gave their consent for the transaction records from these accounts to be used for the study and this data includes the type of transaction (including transfer, payment, cash-in, cash-out), account numbers for whom the transaction was from and to, date and time, amount, fee and balance on the account. The transaction records are available for both treatment groups but not the control group.

[5]BRAC began offering a new 30 week loan just before the start of the study. 40 week loans were therefore a lower proportion than expected, but still the majority (51%). 25% had 30 week loans and 25% 20 week loans

2.3.1 Behavioural games

In order to test whether the women who benefit most from receiving the loan on a mobile money account are those who are most likely to give in to temptation goods or most subject to pressure to transfer money to others, incentivised games were played at baseline to elicit time preferences and willingness to pay to hide money from the spouse.

The time preference games used were standard multiple price lists (Andersen et al., 2008), which have been used frequently in a developing country context (Ashraf et al., 2006). Individuals were asked to choose between a fixed monetary reward in one period and various larger rewards in a later period. The periods were either today and 2 weeks or 2 weeks and 4 weeks time. The near payment was fixed at \$2 and the far payment varies between \$1.8 and \$8. One in five respondents was randomly chosen to be paid one of her choices from this game at the specified time period.

The propensity to pay to hide money from the spouse game has been used as a measure of women's empowerment in the literature (Almas et al., 2015, Fiala, 2017, Mani, 2011). Here I expand upon the version used in Fiala (2017) by conducting a variant of the (Almas et al., 2015) game with multiple choices between whether the woman or her spouse receives set amounts of money the next day. Women had to make a series of 8 choices between receiving a fixed amount of money themselves (\$2) or having their spouse receive varying amount of money between \$1.8 and \$8.

One in five respondents was randomly chosen to be paid one of her choices from this game to either herself or her spouse tomorrow. Tomorrow was chosen to be the payment date to remove effects of strong present bias and to allow the enumeration team time to contact and find the spouse if necessary.

2.3.2 Balance test and baseline characteristics

I confirm the validity of my randomisation by performing a balance test, results of which are shown in Table 2.1. I perform an F-test of equality of the means across the three groups for each characteristic, as shown in the final column. None of the

characteristics are significantly different across the 3 groups at the 10% level. I also perform a joint orthogonality test for each treatment separately. This regresses all the characteristics on each treatment indicator and tests if all the characteristics are jointly zero. This has a p-value of 0.63 for the Mobile Account treatment and 0.84 for the Mobile Disbursement treatment. Thus I cannot reject overall balance.

A few characteristics of the sample are worth highlighting: Looking at the game behaviour; 20% of the women displayed hyperbolic preferences, which is similar to the level found in other studies (Ashraf et al., 2006). 60% of them switched above the median in the hiding money game, meaning they are willing to give up $6 in order retain control over $2 rather than their spouse be given it. Again this large amount of hiding is similar to that found in other studies (Almas et al., 2015, Fiala, 2017)

Moving onto demographics; the sample was well educated with 80% of women completing primary school and 15% completing secondary school. On average, they were 35 years old with 3 other household members. Two-thirds of them were married and 20% had a job in addition to their business.

The average loan was 1.4mn USH ($370) and half the loans were for 40 weeks. Women reported making 440,000 USH ($120) a month in their businesses. The households earned on average 1mn USH ($274) a month, so the woman's business brought in just under half the household income, and spent 900,000 USH ($245) a month. Their business capital was predominantly in inventory, which made up 80% of the total. Married women lived in a household where their spouse had a business 57% of the time, and all women in the sample lived in a household with another business 43% of the time. Nearly 90% had savings, and these averaged 430,000 USH ($100). 97% of women reported already having used mobile money before and the nearest mobile money agent was less than 5 minutes from their home. They owned nearly 3.4mn USH ($1000) in household assets on average.

Table 2.1: Summary statistics and balance test

	Mobile disburse			Mobile account			Control			
	mean	sd	obs	mean	sd	obs	mean	sd	obs	p
branch1	0.23	0.42	984	0.23	0.42	993	0.24	0.42	982	0.98
branch2	0.24	0.43	984	0.24	0.43	993	0.26	0.44	982	0.53
branch3	0.12	0.33	984	0.15	0.36	993	0.13	0.33	982	0.19
branch4	0.12	0.32	984	0.11	0.31	993	0.13	0.33	982	0.52
branch5	0.11	0.31	984	0.11	0.31	993	0.10	0.30	982	0.68
branch6	0.18	0.38	984	0.16	0.37	993	0.16	0.36	982	0.49
high profits	0.47	0.50	984	0.48	0.50	993	0.48	0.50	982	0.91
hide money	0.65	0.48	641	0.63	0.48	647	0.62	0.49	659	0.64
repeat borrower	0.82	0.38	984	0.82	0.38	993	0.81	0.39	982	0.83
hyperbolic	0.21	0.40	984	0.22	0.41	993	0.18	0.39	982	0.13
respondent age	35.78	8.70	984	36.01	9.06	993	35.99	8.95	981	0.82
married	0.65	0.48	984	0.66	0.48	993	0.67	0.47	982	0.60
hh size	4.22	1.70	984	4.27	1.55	993	4.30	1.65	982	0.54
completed primary	0.81	0.39	984	0.81	0.40	993	0.79	0.41	982	0.70
completed secondary	0.14	0.35	984	0.12	0.32	993	0.14	0.35	982	0.11
job	0.21	0.41	984	0.19	0.39	993	0.19	0.39	982	0.47
loan amount	1382	749	967	1430	774	985	1372	767	977	0.20
loan 40	0.52	0.50	984	0.52	0.50	993	0.50	0.50	982	0.46
monthly profit	634	729	980	633	750	992	612	644	980	0.73
monthly profit (self-report)	436	406	980	443	425	992	422	378	980	0.49
business asset value	550	890	984	577	890	993	568	880	982	0.80
inventory value	1971	2098	980	2002	2178	992	1978	2007	980	0.94
weekly hours business	96.21	46.89	984	98.82	47.28	993	99.53	47.77	982	0.26
spouse business	0.57	0.50	575	0.58	0.49	592	0.58	0.49	587	0.91
household business	0.43	0.50	874	0.45	0.50	885	0.45	0.50	883	0.66
have saving	0.88	0.33	984	0.87	0.34	993	0.86	0.35	982	0.35
amount saved	434	703	984	462	761	993	465	818	982	0.60
mobile account	0.97	0.18	984	0.96	0.19	993	0.97	0.18	982	0.68
agent distance (min)	4.46	5.68	984	4.43	6.06	993	4.70	5.60	982	0.56
household income	1042	887	984	1040	804	993	1037	829	982	0.99
household asset value	3371	2631	984	3440	2786	993	3351	2476	982	0.73
household consumption	917	510	984	906	495	993	922	492	982	0.75

All monetary amounts in '000 Ugandan Shilling and winsorised at the 99% level

2.3.3 Take-up

Since women were free to accept or reject the assigned treatment, take-up rates were a concern. However, the interventions had high take-up rates. 94% of the individuals assigned to Mobile Account (Treatment One) received a mobile money account. 71% of those assigned to Mobile Disbursement received this in full.

Additionally, 14% of those assigned Mobile Disbursement received only a mobile money account and their loan as cash (they were assigned to receive Mobile Disbursement and got Mobile Account). The reasons for those assigned to Mobile Disbursement getting Mobile Account was both refusal of the disbursement of the loan onto the mobile money account (5%), but also external problems completing mobile disbursement, such as power cuts or networks outages (10%). Lastly 15% of women assigned to Mobile Disbursement refused the entire treatment (sim card and mobile disbursement). This is summarized in Table 2.2 below.

Table 2.2: Treatment compliance

	Mobile Account	Mobile Disbursement
Received mobile money account	-	700
and loan as mobile money	-	(71%)
Received mobile money account	931	
and loan as cash	(94%)	
Refused mobile disbursement		51
		(5%)
Technical problem for		88
mobile disbursement		(9%)
Received no mobile money	62	145
account (refused)	(6%)	(15%)
Total	993	984
	(100%)	(100%)

I look at correlates with treatment take-up, and find only one variable predicts take-up. Appendix Table 2.19 shows OLS regression results from regressing baseline variables one-by-one on take-up dummy variables for each of the two treatments. For the Mobile Account treatment, the only variable that predicts take-up is the hiding game: women who always hid money from their spouse are 5 percentage

points less likely to accept the sim card. The reason for this is unclear, but could be random given the number of variables I look at. The only factor that predicts take-up of the Mobile Disbursement treatment, and only at the 10% significance level, is completing secondary school, which could indicate more educated women are more comfortable with technology. Below each table, I also include a p-value from an F-test of regressing all the characteristics on the take-up dummies. I cannot reject that all the characteristics are jointly zero.

2.3.4 Attrition

The survey team made a great effort to follow up with this highly mobile population of women. Even though the endline survey was on average only 8 months after the baseline, half the sample had taken loans of a shorter duration than this and so were not necessarily still attending their microfinance groups. Despite this 90% of the sample were found and re-surveyed for endline. Of the 10% who were not resurveyed, 25 refused to be surveyed and 292 couldn't be found. Attrition rates of approximately 10% are common in mobile populations such as this urban sample.

Table 2.3: Attrition

	(1) attrition
Mobile account	0.008
	(0.014)
Mobile disbursement	0.011
	(0.014)
Constant	0.101***
	(0.010)
Observations	2,959
R-squared	0.000
p-value T1=T2	0.83

Linear regression of treatment indicators on a variable equal to one if the woman was not surveyed at endline. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

However, of concern is whether treatment was correlated with attrition. I test for this in Table 3.3 by regressing a dummy variable indicating if the woman was

not found at endline on treatment indicators. I find no significant differences in attrition rates across treatment arms. Correlates of attrition are shown in Appendix Table 3.4. Three variables are significant at the 5% level: older women, those in larger households and those with larger loans are less likely to be surveyed at follow-up. The size of the coefficients are very small, and less than 2% of attrition is explained by the baseline characteristics I examine.

2.4 Empirical strategy

McKenzie (2012) showed that in the case of a single baseline and follow-up with an autocorrelation less than 0.5 (as is the case for business profits, saving and spending), power is highest when regressing an outcome measure at endline on baseline covariates, the treatment measure and the baseline value of the outcome measure. I therefore estimate intent-to-treat (ITT) effects using an ANCOVA specification of the form:

$$Y_{i1} = \alpha_0 + \alpha_1 T_{1i} + \alpha_2 T_{2i} + \alpha_X X_{i0} + Y_{i0} + \epsilon_{i1} \tag{2.1}$$

Where Y_1 is the outcome of interest, T_1 the Mobile Account treatment dummy, T_2 the Mobile Disbursement treatment dummy, X a set of randomization strata dummies (Bruhn and McKenzie, 2009), Y_0 the baseline value of the outcome (if measured at baseline, otherwise excluded) and ϵ random error for individual i.

For every outcome, I test whether each treatment had significant effect ($\alpha_1 = 0$, $\alpha_2 = 0$), as well as whether the treatments differ from each other ($\alpha_1 = \alpha_2$).

As I am considering three primary outcome measures (profit, saving and business capital), I adjust the p-values of the coefficients of interest for multiple statistical inference by calculating sharpened q-values that control for the false discovery rate (FDR). These q-values correct for the fact that I conduct 3 tests across the 3 primary outcomes. Rather than pre-specifying a single q, I report the minimum q-value at which each hypothesis is rejected, following Anderson (2008) and Benjamini et al. (2006).

For some summary measures of outcome families, I group several related variables into index variables following Anderson (2008). I construct the indices in three steps. First, I re-code all contributing outcomes so that higher values correspond to treatment effects in the same direction ("better" outcomes). Second, I standardize the individual outcomes using the baseline mean and standard deviation of the control group for that outcome. Third, I calculate the average of the standardized constituent outcomes, weighted by the inverse covariance matrix. Where a specific outcome value is missing for a respondent, I calculate the value of the index for that respondent using the remaining outcomes.

When looking at secondary and intermediate outcomes I do not correct for multiple testing as this analysis is informative for exploratory analysis of additional impacts, robustness checks and mechanisms analysis, not the main impact.

2.4.1 Administrative data

The administrative data is only available for the two treatment groups that I gave mobile money accounts to, not the control group. Analysis will therefore give the additional impact of disbursing the loan on the mobile money account on how it is used.

I estimate ITT effects for the administrative data using an OLS regression of the form:

$$Y_i = \alpha_0 + \alpha_2 T_{2i} + \alpha_X X_i + \epsilon_i \tag{2.2}$$

Where Y is the outcome of interest, T_2 the Mobile Disbursement treatment dummy, X a set of randomization strata dummies and ϵ random error, for individual i.

For the administrative data, I test whether disbursement of the loan onto the mobile money account had a significant effect ($\alpha_2 = 0$) as compared to just being given the mobile money account.

2.5 Results

2.5.1 Impacts on mobile money transactions and balances

I look at mobile money account usage outcomes based on administrative data collected from the mobile telecoms operator, MTN. This data gives an indication of how the accounts were used, allowing me to understand if the accounts were primarily used to facilitate business transactions or for the saving and safe storage of the loan and other funds. This data also allows me to verify that indeed the loan was successfully disbursed onto the mobile money account for the 690 of the 982 women assigned to Mobile Disbursement, matching the take-up numbers recorded in the survey data.

A summary of some of the mobile money account usage outcome statistics is shown in Table 2.4. The first thing to note is the ever deposit variable. This captures if the woman ever deposited money onto the mobile money account, for example, by topping up the account herself, receiving money from someone else or by being paid for goods or services on the account. It excludes the loan disbursement for the Mobile Disbursement group. As seen in the table, both groups are similarly likely to deposit money onto the account, with 13% ever depositing. This means that for the Mobile Account group, only 13% ever used the account (since they could not withdraw or save money without first depositing some). Both groups make similar low numbers of deposits (0.6-0.8 of a deposit at the mean, though some make as many as 50), and the deposit amount conditional on making a deposit is similar for both treatments at around 50,000 USH ($13). While the maximum deposits made onto the accounts are relatively large, 600,000 ($160) and 1mn USH ($270) for the Mobile Account and Mobile Disbursement treatments respectively, the most common outcome for both groups is that they don't deposit anything.

Larger differences appear between the treatments when looking at withdrawals. The Mobile Disbursement treatment group make a withdrawal 83% of

the time[6]. For the Mobile Account group withdrawals are similar to deposits at 12% ever making one.

Additionally the number of withdrawals is much higher for the Mobile Disbursement group. This is interesting, as in principal the Mobile Disbursement group could just withdraw all the loan the day they got it and so only needed to make 1 withdrawal. However, on average, women in the Mobile Disbursement treatment makes nearly 4 withdrawals. Likewise, the average withdrawal amount was less than the average loan - 600,000 USH ($160) compared to 1.4mn USH ($370) for the Mobile Disbursement group. Qualitative questions and survey responses suggest this was not because mobile money agents didn't have enough float to withdraw all the loan at once, but because the women were choosing to retain some money on the accounts.

I examine the outcomes summarised in Table 2.4 as well as the balances held on the accounts over time, using regression analysis in Figure 2.1 and Table 2.5 below.

Table 2.4: Summary statistics of mobile money account usage

	Mobile account						Mobile disburse					
	obs	mean	sd	max	min	median	obs	mean	sd	max	min	median
ever deposit	894	0.13	0.33	1.00	0.00	0.00	828	0.14	0.35	1.00	0.00	0.00
number deposit	894	0.61	2.86	47.00	0.00	0.00	828	0.78	3.88	63.00	0.00	0.00
deposit amount (USH)	112	48.15	84.03	635.00	1.00	26.90	119	53.78	120.01	1002.75	0.30	20.20
total deposits (USH)	894	26.44	136.37	1687.00	0.00	0.00	828	32.37	204.80	4011.00	0.00	0.00
ever withdrawal	894	0.12	0.33	1.00	0.00	0.00	828	0.83	0.38	1.00	0.00	1.00
number withdrawal	894	1.09	6.11	103.00	0.00	0.00	828	3.82	7.45	101.00	0.00	2.00
withdrawal amount (USH)	108	43.25	128.09	1250	0.50	16.78	686	647.99	600.64	3484.80	1.00	502.68
total withdrawals (USH)	894	29.29	172.98	3326.00	0.00	0.00	828	1107.04	894.00	7631.00	0.00	966.01

Monetary outcomes are in '000 Ugandan Shillings. All variables are defined over the first 180 days after the account was provided. I cap transactions at 180 since the last mobile money accounts were given out in June 2017 and the administrative data ends in January 2018. Deposits always excludes the loan disbursement for the mobile disbursement treatment group. Deposit amount and withdrawal amount summarises cumulative deposits/withdrawals to the account. Ever deposit and withdraw are dummy variables if at least one transaction of that type occurred. Number of deposits and withdrawals is the count of each transaction for an account. Total deposits and withdrawals are cumulative transactions on an account.

In Figure 2.1, I show the average balance on the mobile money account for various time periods along with the 95% confidence interval. The periods are the first 15 days, 15-30days, 30-45days, 45-60days, 60-90days and 90-180days. I also show the final balance from the last transaction before the 180 day cut-off.

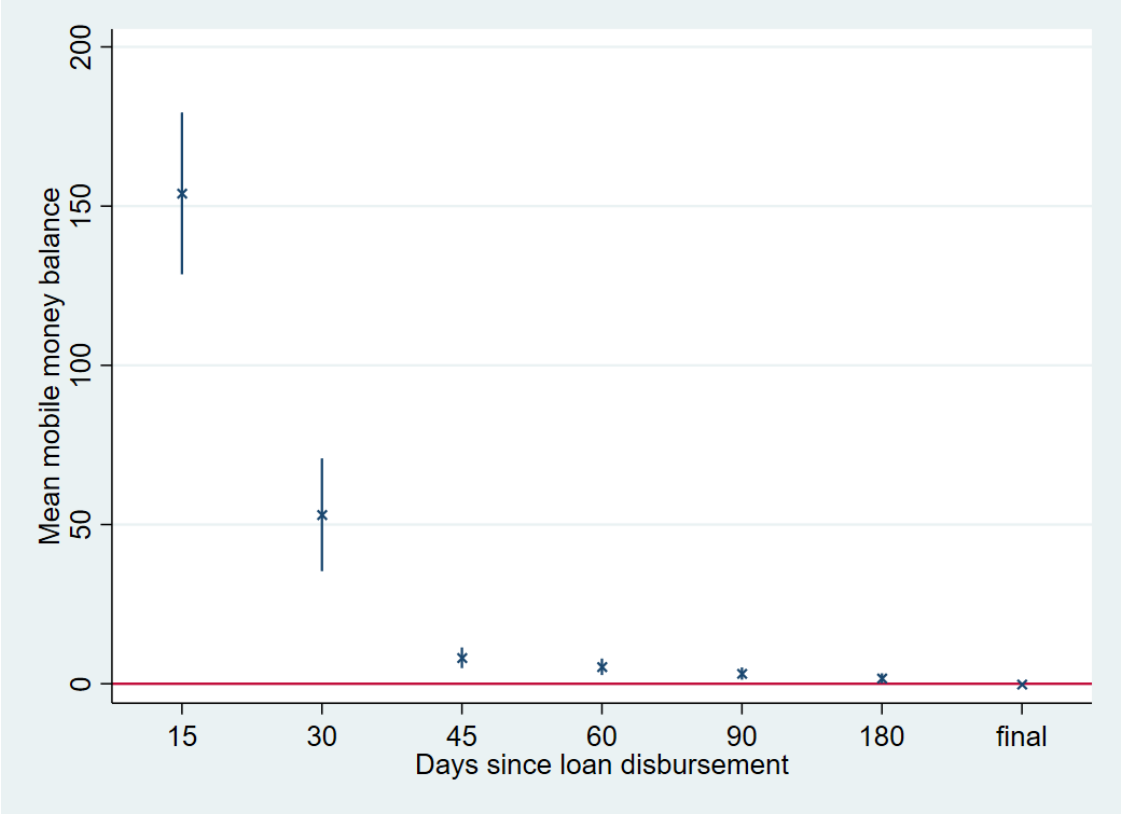
[6]This is not 100% as some of the Mobile Disbursement group did not receive their loan on the mobile money account, but were still given the mobile money account (see section 2.3.3)

Figure 2.1 clearly shows that the average balance on the mobile money account for the Mobile Disbursement treatment is large and statistically significant compared to the Mobile Account treatment. The Mobile Account group hold almost zero average balances throughout the period. The average balance for the Mobile Disbursement treatment also declines over time, though remains significantly different than the Mobile Account treatment until the final balance. During the first 15 days after loan disbursement, women in the Mobile Account group are typically holding 150,000 USH ($40) on the account, approximately 10% of the loan value or 34% of the mean household saving at baseline. Between 15 and 30 days this falls to 50,000 USH ($14). This indicates that microfinance clients treated with Mobile Disbursement are choosing to hold some of the loan as a balance on their accounts, which they are slowly dipping into and running down over time. While some clients in the Mobile Account treatment do deposit into the mobile money account, they are few and their balances are tiny.

Turning to Table 2.5, I show a series of variables capturing whether the mobile money account was used and how intensely, split into transactions relating to depositing and withdrawing money. Table 2.5 shows that it is only for withdrawals that there are significant differences between the Mobile Account and Mobile Disbursement treatments. Mobile disbursement treated women are 70 percentage points more likely to make a withdrawal than the Mobile Account treatment. This seems reasonable considering they needed to withdraw the loan and 70% of them took-up the treatment according to the survey data. On average, Mobile Disbursement treated clients make 4 withdrawals, significantly different from the Mobile Account mean of 1 withdrawal. The fact that withdrawals for the Mobile Disbursement group are greater than 1 corroborate the finding that clients are leaving a balance on the accounts which they are slowly drawing down over time.

Finally, looking at the value of withdrawals. Average withdrawal is the value of the average withdrawal made to the account. This is shown only for the sample of women who made any withdrawals. The total withdrawals are the total value of

Figure 2.1: Treatment effects of Mobile Disbursement on average balances in mobile money account



all withdrawals to the accounts over the first 180 days of account ownership[7]. The typical withdrawal is 600,000 USH ($160) for the Mobile Disbursement treatment compared to only 40,000 USH ($11) for the Mobile Account treatment group who made any withdrawals. Total withdrawals for the Mobile Disbursement group are 1.1mn USH ($290), 300,000 USH ($80) less than the average loan size[8].

This is also backed-up by examining the average percentage of the loan withdrawn on the same day as it was deposited. 71% of the Mobile Disbursement group withdrew some of the loan on the day it was disbursed and the average withdrawal amongst this group was 54% of the loan value. Focus groups validated that this was not due to liquidity constraints among agents: women who did encounter an agent with insufficient float could easily go to one of the many other agents concentrated

[7]I look only at the first 180 days of account ownership since the last disbursements of the loans for study women were in June 2017 and the administrative data only goes until January 2018

[8]This will not add up to the balance on the accounts due to fees paid on transactions

Table 2.5: Treatment effects on intermediate usage outcomes

	(1) Ever deposit	(2) Number deposit	(3) Average deposit	(4) Total deposit	(5) Ever withdraw	(6) Number withdrawals	(7) Average withdrawal	(8) Total withdrawals
MD	0.02	0.21	-13.54	6.46	0.70***	2.86***	598.57***	1,074***
	(0.02)	(0.18)	(17.20)	(8.36)	(0.02)	(0.33)	(69.93)	(31.92)
Constant	0.12***	0.59***	58.02***	26.12***	0.12***	1.02***	48.71	30.44
	(0.01)	(0.12)	(10.44)	(5.64)	(0.01)	(0.22)	(63.59)	(21.52)
Obs.	1,722	1,722	231	1,722	1,722	1,722	794	1,722
R-squared	0.24	0.22	0.74	0.31	0.63	0.35	0.47	0.57
Control mean	0.13	0.61	48.15	26.47	0.12	1.09	43.25	29.32

Impacts amongst those who received sim cards. All regressions include strata dummies. Monetary outcomes in '000 Ugandan Shillings. All variables are defined over the first 180days after the account was provided. MD (Mobile Disburse) is the treatment where a mobile money account was provided and the loan also disbursed onto this account. Control mean refers to the mean in the mobile account group. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

around Kampala. Most women are therefore leaving some balance on the accounts beyond the disbursement day and making multiple withdrawals over time.

Overall, the summary of transaction records suggests that for both treatments the mobile money accounts were not used for frequent deposit and withdrawal of money. This means the accounts were not used by the majority of women for either business transactions or to frequently save either business or other income. This differs to the findings of Dizon et al. (2017) and Habyarimana and Jack (2018) who find that labelling a mobile money account for a saving goal increases savings, even if those people already had another mobile money account, though they provided additional monetary incentives to save. It also conflicts with Bastian et al. (2018) who find providing information about a mobile saving account increases saving, though partly through crowding out other forms, and Batista and Vicente (2017) who find a mobile money linked saving account increased savings in Mozambique[9]. This could suggest that actually people will not use mobile money for saving unless induced by other incentives, such as offering interest on balances, at least in an urban context with access to alternative forms of saving. I discuss this further in section 2.6.4.

However, my findings fit with evidence from mobile linked saving accounts in Sri Lanka, which had relatively low levels of use and did not led to higher overall savings (De Mel et al., 2018). My study may be most similar to De Mel et al.

[9]Again, bonus interest rates were offered to induce savings in this study

(2018) in that women already had access to other forms of saving such as bank accounts at relatively high levels (38% already used a bank account at baseline). Also being in an urban setting means women are extremely close to other methods of saving such as a bank, and so any reduction in transaction costs from using mobile money is likely to be small.

Instead, it appears as though the accounts were predominantly used by the Mobile Disbursement group to save some of the loan and withdraw it down over time. This is similar to the findings of Somville and Vandewalle (2018) and Field et al. (2016) who also both compare in different contexts paying money as cash versus into a saving account. They both find that the saving account payment results in higher levels of savings from retaining some of the money paid into the account, but no increases in own payments into the account.

2.5.2 Impact on primary business outcomes

As outlined in my pre-analysis plan, the primary outcomes of this study are profits, savings and the value of enterprise capital (defined as the value of business assets and inventory). The results for intent-to-treat estimate on those three outcomes are shown in Table 2.6. I find a positive and significant effect on both profits and business capital for the Mobile Disbursement treatment. Both of these results also remain after a multiple testing correction is applied. Those in the Mobile Disbursement treatment experience a 15% increase in their profits and a 11% increase in the value of their business capital compared to the control group. These results are consistent with the hypothesis that disbursing the loan on a mobile money account increased the amount of the loan used to invest in the business and that this increased businesses investment led to gains in profit.

There are no effects of the Mobile Disbursement treatment on the amount of saving and I find no significant or large coefficients from the Mobile Account treatment on any of the three outcomes. I am able to reject equality of the treatment effects for the Mobile Account and Mobile Disbursement treatments for both business profits and business capital, but not savings. These results are

Table 2.6: Treatment effects on primary outcomes

	(1) profit	(2) savings	(3) capital
Mobile account	10.41	3.33	-30.40
	(13.01)	(33.98)	(86.65)
	[0.99]	[0.99]	[0.99]
Mobile disburse	63.72***	30.44	213.08***
	(12.73)	(36.82)	(82.92)
	[0.00]	[0.74]	[0.03]
Observations	2,639	2,639	2,639
R-squared	0.44	0.41	0.60
Control mean endline	395.3	559.2	2473
Control mean baseline	419.8	483.6	2488
p-value T1=T2	0.00	0.50	0.00

Intent-to-treat estimates. All outcomes are winsorized at the 99% level. '000 Ugandan Shillings. All regressions include strata dummies and include the baseline value of the outcome. Mobile Account is the treatment where only a mobile money account was provided and the loan was disbursed as cash. Mobile Disburse is the treatment where a mobile money account was provided and the loan also disbursed onto this account. Profits refers to the self-reported monthly business profit. Savings is individual savings held by the woman. Capital is the value of all assets the woman uses in her business plus the value of inventory held for her business. Control mean endline is the mean value of the outcome in the control group at endline. Control mean baseline is the mean value of the outcome in the control group at baseline. False discovery rate (FDR) adjusted p-values, also known as q-values, were used to correct for multiple hypothesis testing. They are shown in square brackets. These were calculated following the method of Benjamini et al. (2006)

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

consistent with the fact that by 6 months after the loan disbursement, neither treatment group was holding significant balances on the mobile money account, and in the case of the Mobile Account treatment they never deposited money onto the account. Hence it is unsurprising that by 8 months there is no saving impact for the Mobile Disbursement group, and, since they didn't use the accounts, I find no impact for the Mobile Account group.

Also of note from this table is the difference for the control group between baseline and endline. In the control groups, profits actually decline by 25,000 USH ($5), 6%, between baseline and endline despite the control group obtaining a loan. This result matches that of other studies which have found no overall impact of

getting a microfinance loan on a woman's business (Banerjee et al., 2015). Across all treatment groups, savings increase by 100,000 USH ($25), 21%, and in the control group there is no change in business capital. Later results will show that the control group appear to use the loan mainly to buy household assets and pay for school fees, hence why no overall business impacts are seen.

2.5.3 Impacts on secondary outcomes

I pre-defined additional outcomes for each of my three primary outcome families. These additional and secondary outcomes shine light on why the primary outcomes are affected by the treatments. I do not multiple-hypothesis correct the secondary outcomes.

Business outcomes

I examine 4 additional business outcomes in Table 2.7: monthly and weekly sales and calculated monthly and weekly profits. These are alternative outcomes of business performance to supplement the self-reported profit measure used as the primary outcome. I see large significant effects of the Mobile Disbursement treatment on all these outcomes. Sales are approximately 15% higher for the Mobile Disbursement group both weekly and monthly. Similarly profits in the Mobile Disbursement group are 10% higher than the control group, a similar increase as the self-reported profit measure.

I see no significant impacts from the Mobile Account treatment, but I cannot reject that the treatments had equal effects for the two alternative measures of profits.

Savings

Secondary savings outcomes are reported in Table 2.8. I look at saving specifically with mobile money, to see if the treatment caused a shift in savings from other forms to saving on mobile money account. I look at whether the woman saves at all with mobile money and, if so, the amount she saves with mobile money. Since the mobile money account was framed as an account for the business I also look at whether when are more likely to report that they are saving for their business.

Table 2.7: Treatment effects on secondary business outcomes

	(1) monthly sales	(2) weekly sales	(3) monthly profit	(4) weekly profit
Mobile account	66.59	20.07	19.98	12.37
	(66.15)	(18.48)	(25.10)	(10.39)
Mobile disburse	211.07***	52.18***	61.83**	26.06**
	(67.80)	(18.52)	(24.10)	(10.72)
Observations	2,606	2,606	2,606	2,606
R-squared	0.34	0.28	0.29	0.17
Control mean endline	1356	351.4	564.5	132.6
Control mean baseline	1399	353.7	607.9	151.4
p-value T1=T2	0.03	0.09	0.13	0.23

Intent-to-treat estimates. All outcomes are winsorized at the 99% level. '000 Ugandan Shillings. All regressions include strata dummies and include the baseline value of the outcome. Mobile Account is the treatment where only a mobile money account was provided and the loan was disbursed as cash. Mobile Disburse is the treatment where a mobile money account was provided and the loan also disbursed onto this account. Monthly and weekly profit are calculated by subtracting the corresponding expenditures from sales. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

As robustness checks, I also look at the calculated value of savings in each form and net savings over the last 30 days.

I see no effects on calculated savings or net savings in the last 30 days, in the same way as I didn't on self-reported savings. The calculated saving variable is very similar to self-reported total savings and they in fact have a high correlation (0.75), showing that women do have a good idea of their savings. On reporting that their main saving goal is for the business, I see an effect from the Mobile Account treatment but not the Mobile Disbursement treatment, but this is only significant at the 10% level and I can't reject equality of the treatment effects.

I see effects for both treatments on whether a woman reports saving with mobile money and the amount of savings held on the mobile money account. Those given the Mobile Account treatment are 4 percentage points more likely to report using mobile money to save, while those given the mobile money disbursement treatment are 9 percentage points more likely. This is from a control mean of only 12%, meaning the Mobile Disbursement treatment almost doubled savings on mobile money and the

Table 2.8: Treatment effects on secondary saving outcomes

	(1) calculated savings	(2) net savings	(3) saves mobile money	(4) amount mobile money	(5) saving goal business
Mobile account	-23.18	-8.48	0.04**	5.89*	0.04*
	(44.34)	(12.57)	(0.02)	(3.08)	(0.02)
Mobile disburse	21.36	-8.48	0.09***	12.08***	0.01
	(47.19)	(8.39)	(0.02)	(3.17)	(0.02)
Observations	2,642	2,642	2,642	2,642	2,642
R-squared	0.16	0.12	0.19	0.25	0.19
Control mean endline	581.15	72.91	0.12	13.34	0.24
p-value T1=T2	0.31	1.00	0.01	0.09	0.11

Intent-to-treat estimates. All outcomes are winsorized at the 99% level. '000 Ugandan Shillings. All regressions include strata dummies. Mobile Account is the treatment where only a mobile money account was provided and the loan was disbursed as cash. Mobile Disburse is the treatment where a mobile money account was provided and the loan also disbursed onto this account. All outcomes reported here were only collected at endline. Calculated savings is the sum of savings in each form of saving. Net savings is additions-withdrawals from savings in the last month. Saves mobile money is a dummy equal to one if the the respondent reported saving on a mobile money account. Amount mobile money is the value of savings on a mobile money account. Saving goal business is a dummy if the reported goal of saving is to use it for the business. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Mobile Account treatment increased them by one-third. This is from an extremely low value though: saving on a mobile money account were only 13,000 USH ($3.5) in the control group, less than 2% of all savings by value. The two treatments only increase this to 18,000 ($4.8) and 25,000 USH ($6.7)[10] in the Mobile Account group and Mobile Disbursement group respectively, still less than 5% of total savings. The impacts are therefore significant statistically but of low economic significance.

These impacts suggest both treatments induced a shift towards use of mobile money for savings. The fact that there are no overall effects on savings suggests

[10]The balances saved in mobile money are larger when self-reported compared to the administrative data. This could be because the majority of the women already had mobile money accounts but I only have records for those accounts given out as part of the study. These records will therefore always undervalue total mobile money balances, assuming the women continue to also use their private accounts.

that this is a reallocation of savings between forms rather than higher savings. However the coefficients for both calculated and self-reported savings are of a positive magnitude for the Mobile Disbursement treatment and so there could be a small increase in savings I am unable to detect[11].

Business assets

I examine an index of business assets formed by taking the first principal component of a series of dummy variables for whether or not an asset is used in the business. This measure enables me to capture changes in the number of different assets used in the business, rather than just changes in the value of assets used.

Looking at Table 2.9, I find a significant positive effect of the Mobile Disbursement treatment on the asset index, implying that it is not simply that those who receive their loan on a mobile money account are purchasing higher value assets, or more of the same assets. They also seem to be increasing the diversity of assets used in the business. This could reflect the idea that getting the loan on the mobile money account makes it easier to purchase a number of different, moderate valued assets, rather than trying to tie-up as much of the cash loan as possible into an asset as soon as possible. I find no significant impact of the Mobile Account treatment on the business asset index.

I also examine the value of a business assets, which was a component of the primary outcome capital, along with the value of inventory. Inventory was by far the largest component of capital (80%), but even looking just at business assets I still see a significant impact of the Mobile Disbursement treatment of 130,000 USH ($35). I also find a significant effect of the Mobile Disbursement treatment on inventory value alone, of 120,000 USH ($32). This shows that women treated with Mobile Disbursement invest in more business assets and higher value assets, as well as greater inventory.

[11]I am only powered to detect 0.1 standard deviations. Since the variance of savings is very high this is 80,000 USH

Table 2.9: Treatment effects on secondary business asset outcomes

	(1) PCA of business assets	(2) value of business assets	(3) inventory value
Mobile account	0.10	49.75	-82.79
	(0.07)	(44.92)	(69.96)
Mobile disburse	0.38***	132.73***	122.41*
	(0.07)	(43.49)	(62.75)
Observations	2,642	2,610	2,606
R-squared	0.32	0.42	0.62
Control mean endline	-0.109	643.7	1887
Control mean baseline	0.0541	577.4	1968
p-value T1=T2	0.00	0.03	0.00

Intent-to-treat estimates. All outcomes are winsorized at the 99% level. '000 Ugandan Shillings. All regressions include strata dummies and include the baseline value of the outcome. Mobile Account is the treatment where only a mobile money account was provided and the loan was disbursed as cash. Mobile Disburse is the treatment where a mobile money account was provided and the loan also disbursed onto this account. Principal component analysis of assets used in the business. Higher values mean a larger number of different assets are used in the business. Control mean endline is the mean value of the outcome in the control group at endline. Control mean baseline is the mean value of the outcome in the control group at baseline. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

2.5.4 Robustness

I perform a permutation test to compute exact test statistics which do not depend on asymptotic theorems. To do this I use Stata's permute function which randomly assigns women to the two treatments and control group and calculates the probability of observing the treatment effect I did under the null hypothesis that there is no treatment effect. I use 1000 permutations within strata.

These are reported in Appendix Tables 2.21 underneath the robust p-values and q-values. The permutation p-values reject the null hypotheses at the same levels as the robust p-values.

My results are robust to alternative specifications and the treatment of outliers. I include a time trend of the number of days between disbursement and endline, both linearly and as a quadratic. This will control for seasonality effects, which

could be important as the endline finished just before Christmas. Including a time trend does not affect my results[12], as seen in appendix Table 2.22.

I also examine alternative treatment of outliers by winsorizing at the 0.5 and 2% levels. This makes no difference to my results, as seen in Tables 2.23 and 2.24.

I show average treatment on the treated effects from instrumenting actual take-up of the treatments with random treatment assignment in the Appendix in Table 2.25. Since my take-up was relatively high at 71%, these are approximately one-quarter larger than the estimates in Table 2.6.

2.6 Mechanisms: Internal or external constraints to microenterprise investment?

There are three main channels through which mobile money accounts, and disbursement of loans onto those accounts could impact women's businesses:

Firstly, the Mobile Disbursement treatment in particular, may have facilitated both learning and credibility about saving in a mobile money account and so relaxed saving constraints. Secondly, disbursement of the loan onto the mobile money account may have helped women to exercise self-control, both through mental accounting effects of having an earmarked account for the business and through the soft commitment of having to withdraw money from the account rather than have it as cash in hand. Finally, the mobile money accounts, and the disbursement of loans onto these accounts, may have hidden money from family and so given the woman more control over the loan.

2.6.1 Saving constraints

One reason the mobile money accounts could have an effect is if the women were saving constrained. The mobile money accounts may then have presented the women with a new avenue to save with. In this case, getting the loan on the mobile money account may have had a larger impact due to learning effects: the women

[12]The mean (and median) number of days between loan disbursement and the endline survey was 200 days, or 7 months

might have not thought to save on a mobile money account before, or at least to save large amounts. The disbursement of the loan onto the mobile money account may therefore have taught the women that its possible to save so much on a mobile money account. They may also have implicitly assumed BRAC was validating that keeping so much money on a mobile account is safe and a good idea, helping them to overcome any reservations about doing this.

At first glance it seem unlikely that women who already have mobile money accounts (as 97% of them do) would not think to use them to save. However, according to survey data collected by the Financial Inclusion Initiative (2013) only 3% of households that use mobile money have used it to 'Save money for a future purchase or payment'. A further 5% use mobile money to 'Set money aside just in case/for an undetermined purpose'. Similarly in my data I find only 12% of the control group reported saving on a mobile money account. This suggests very low use of mobile money services for saving. A reason for this could be that people must learn about saving on a mobile money account, and build trust that money would be as safe in the mobile money account as in, say, a bank.

The Mobile Disbursement treatment may have provided a shock that forced women to at least temporarily hold a lot more money on the mobile money account than they were used to. BRAC also was implicitly providing information that this was a safe thing to do. The women were also told that they could use the mobile money account to safety store business funds.

However, there are potential problems with this explanation: if the Mobile Disbursement treatment group had learnt that mobile money accounts were a good place to save money I'd expect to see more deposits onto the accounts as women shift to putting more of their savings there. Instead I see no differences between the two treatment groups in terms of deposits into the accounts. Self-reported savings with mobile money, while significantly different for both treatments from the control group, are of economically tiny magnitudes (See Table 2.8 - the treatments increase mobile money savings from 2% of all savings to 3% and 5% in the Mobile Account and Mobile Disbursement treatments respectively). The women also already had

access to many other forms of saving, including over one-third who save in a bank account. If the women did learn that mobile money accounts are a good way to save, it seems difficult to reconcile this with the data on how they actually use the accounts. This makes me doubtful that saving constraints can explain my effects.

2.6.2 Self-control

To examine if self-control difficulties are a key channel trough which the accounts had an impact, I look at heterogeneity by an index of self-control difficulties at baseline. I construct this index in a standard fashion[13] using the method of Anderson (2008). The index is composed of whether a woman had hyperbolic time preferences (stratified) at baseline, whether she was impatient at baseline, where impatience was defined as always preferring money now over the future in the near-far time frame, and whether she didn't report saving for her business. It's important to note that while a component of the self-control index was used to stratify the original randomisation, the other variables could be picking up a correlation with another variable.

I show these results in Table 2.10. I see large heterogeneous effects by the index of prior self-control difficulties for the Mobile Disbursement treatment on profits. However, this does not survive a multiple testing correction (q value 0.103). The results for business capital are more noisy, and I only find a significant overall impact of the Mobile disbursement treatment for those who had self-control difficulties at baseline. Overall, while there seems to be some evidence that those women with self-control difficulties at baseline benefited more from the Mobile Disbursement treatment, the effects are not strong enough to explain all my results. This contrasts with Somville and Vandewalle (2018) who argue self-control difficulties explain their findings well.

[13] e.g. see Fafchamps et al. (2014)

Table 2.10: Heterogeneous treatment effects by baseline self control index

	(1) business profit	(2) total savings	(3) business capital
MA*self control	-5.24	2.72	77.62
	(28.97)	(73.91)	(183.20)
	[0.99]	[0.99]	[0.99]
MD*self control	63.17**	26.04	180.92
	(27.65)	(76.73)	(176.28)
	[0.12]	[0.99]	[0.84]
Mobile account	12.88	4.44	-24.04
	(16.95)	(45.66)	(108.75)
	[0.99]	[0.99]	[0.99]
Mobile disburse	38.39**	22.86	144.67
	(16.19)	(50.36)	(106.59)
	[0.10]	[0.99]	[0.48]
Observations	2,639	2,639	2,639
R-squared	0.44	0.35	0.56
Control mean	387.2	510.0	2366
Control mean baseline	425.2	445.4	2499
Overall effects			
mobile account	7.64	7.16	53.59
	(22.27)	(55.31)	(143.4)
mobile disburse	101.6***	48.89	325.6**
	(21.69)	(55.48)	(136.9)
T1=T2	0.12	0.69	0.10
T1=T2 interaction	0.00	0.44	0.03

Intent-to-treat estimates. Monetary outcomes are winsorized at the 99% level and reported in '000 Ugandan Shilling. All regressions include strata dummies. Mobile Account (MA) is the treatment where only a mobile money account was provided and the loan was disbursed as cash. Mobile disburse (MD) is the treatment where a mobile money account was provided and the loan also disbursed onto this account. Heterogeneous indexes are defined in section 2.6. The interaction is for someone who is above the median in the index. Profit is self-reported monthly profit. Capital is composed of business assets and inventories. Control mean refers to the mean value of the outcome variable for the interaction condition being true in the control group, and control mean baseline the mean in the control group at baseline when the interaction condition is true. The panel labelled overall effects gives the total impact of each treatment for someone who is above the median in the index variable. False discovery rate (FDR) adjusted p-values, also known as q-values, were used to correct for multiple hypothesis testing. They are shown in square brackets. These were calculated following the method of Benjamini et al. (2006). Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

2.6.3 Family pressure

During focus groups prior to the research beginning, the women discussed the pressure they experience to share some of the loan with their family when they first get it. This is compounded by the difficulty of hiding a large amount of cash in small denomination bills. The women discussed the many strategies they employ to hide all or part of the loan when they first receive it.

Hiding the loan when its disbursed onto a mobile money account is likely to be considerably easier than hiding large amounts of cash in hand. Additionally, though mobile money accounts were designed to send money, they still involve multiple steps to making a transfer, which are considerably more of an obstacle compared to taking cash out of a pocket. The fact that the money was disbursed onto a mobile money account may also make it more credible for the woman to argue that this money was given to her by BRAC for her business, and that it would be known if she used it for other things. This may make it easier for her to argue that this money is earmarked only for her business. Both treatment groups could also use the account to hide business profits by making deposits to the account.

To examine whether facilitating hiding of money was a key channel by which the mobile money treatments affected women's businesses, I look at an index of family pressure at baseline and examine heterogeneous effects by this index. I construct this index in the same way as for self-control using the method of Anderson (2008). The index is composed of the following components at baseline[14]: whether they always preferred to receive the money themselves on the hiding game (stratified); whether they were married; whether they scored below the median in an index of decision making; whether they reported that when they have money on hand their spouse and family takes it; and whether their spouse or another household member had a business at baseline. Heterogeneous effects by this index are shown in Table 2.11.

I find strong heterogeneous effects for the Mobile Disbursement treatment by the index of family pressure at baseline for both profit and business capital. Those with high family pressure at baseline see a additional increase in their profits of 74,000

[14]A component of the family pressure index was used to stratify the original randomisation

Table 2.11: Heterogeneous treatment effects by baseline family pressure index

	(1) business profit	(2) total savings	(3) business capital
MA*family pressure	7.09	-33.41	100.60
	(28.09)	(77.41)	(186.16)
	[0.99]	[0.99]	[0.99]
MD*family pressure	74.53***	22.33	451.46**
	(27.17)	(82.57)	(178.34)
	[0.03]	[0.99]	[0.03]
Mobile account	6.77	19.32	-43.10
	(18.58)	(44.42)	(113.89)
	[0.99]	[0.99]	[0.99]
Mobile disburse	29.80*	20.34	7.93
	(17.49)	(44.57)	(111.58)
	[0.49]	[0.99]	[0.99]
Observations	2,639	2,639	2,639
R-squared	0.44	0.35	0.57
Control mean	386.7	659.5	2657
Control mean baseline	423.4	546.1	2639
Overall effects			
mobile account	13.85	-14.09	57.49
	(19.69)	(59.08)	(138.7)
mobile disburse	104.3***	42.67	459.4***
	(19.76)	(65.50)	(131.5)
T1=T2	0.19	0.98	0.61
T1=T2 interaction	0.00	0.34	0.00

Intent-to-treat estimates. Monetary outcomes are winsorized at the 99% level and in '000 Ugandan Shillings. All regressions include strata dummies. Mobile Account (MA) is the treatment where only a mobile money account was provided and the loan was disbursed as cash. Mobile Disburse (MD) is the treatment where a mobile money account was provided and the loan also disbursed onto this account. Heterogeneous indexes are defined in section 2.6. The interaction is for someone who is above the median in the index. Profit is self-reported monthly profit. Capital is composed of business assets and inventories. Control mean refers to the mean value of the outcome variable for the interaction condition being true in the control group, and control mean baseline the mean in the control group at baseline when the interaction condition is true. The panel labelled overall effects gives the total impact of each treatment for someone who is above the median in the index variable. False discovery rate (FDR) adjusted p-values, also known as q-values, were used to correct for multiple hypothesis testing. They are shown in square brackets. These were calculated following the method of Benjamini et al. (2006). Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

USH ($20) from getting the Mobile Disbursement treatment, or approximately 17% of profits in the control group. There is still a small impact of the Mobile Disbursement treatment for those who didn't experience above median pressure to share with family at baseline, but this is only significant at the 10% level and doesn't survive multiple test correcting.

Looking at the overall effect, those who at baseline experienced high family pressure to share money see a 100,000 USH ($27) increase in their business profits from treatment with Mobile Disbursement, or a 25% increase compared to those in the control group who experience high family pressure at baseline. The control mean actually shows that those who experienced strong family pressure to share money at baseline see a decline in their business profits between baseline and endline, from 423,000 ($110) to 386,000 USH ($103). Treatment with Mobile Disbursement is therefore not only mitigating an otherwise decline in profits but actually leading to an increase.

I likewise see similar heterogeneity for business capital by family pressure at baseline for the Mobile Disbursement treatment. Overall, those who experience family pressure to share at baseline see their business capital increase by 460,000 USH ($123), or 17%, from the Mobile Disbursement treatment. The heterogeneous effects by family pressure survive a multiple testing correct for profits and business capital, remaining significant at the 5% level. I see no heterogeneous effects from the Mobile Account treatment and no heterogeneous effects for the saving outcome.

Expenditure patterns

If the Mobile Disbursement treatment helped women to resist family pressure to share money then this should appear in the expenditure data[15]. I have measures of the amount of money the women reports giving to her spouse. I therefore examine examine whether the treatments changed the amount and whether the woman reports giving money to her spouse[16] This is shown in Table 2.12.

[15]these outcomes were not pre-specified and are exploratory only

[16]This was no pre-specified as an outcome

Table 2.12: Treatment effects on amount and whether the woman gave money to her spouse and amount received from her spouse

	(1) amount given spouse	(2) dummy gave money to spouse	(3) amount received spouse
Mobile account	-4.19	-0.03	4.15
	(3.83)	(0.03)	(9.63)
Mobile disburse	-10.78***	-0.09***	-1.82
	(3.54)	(0.03)	(9.85)
Observations	1,613	1,613	1,613
R-squared	0.24	0.29	0.27
Control mean endline	21.88	0.297	157.8
Control mean baseline	11.81	0.218	160.1
p-value T1=T2	0.0727	0.0974	0.538

Not in pre-analysis plan. Intent-to-treat estimates. All outcomes are winsorized at the 99% level. '000 Ugandan Shillings. All regressions include strata dummies and include the baseline value of the outcome. Mobile Account is the treatment where only a mobile money account was provided and the loan was disbursed as cash. Mobile Disburse is the treatment where a mobile money account was provided and the loan also disbursed onto this account.
Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

I find that women who received the Mobile Disbursement treatment give significantly less money to their spouse, 10,000 USH ($2.7) on a mean of 22,000 USH ($5.4), or nearly 50% less. They are also significantly less likely to give any money to their spouse, with the Mobile Disbursement treatment group being 9 percentage points less likely to give money to their spouse. This is on a mean of one-third of women giving any money to their spouse. What is interesting about these results is that between baseline and endline the control group go from giving 11,000 USH ($2.9) to 22,000 USH ($5.4) and from 22% of them giving money to 30% of them giving money. For the Mobile Disbursement group there is no change in the probability of giving or the amount given to spouse.

I find no significant impact of the Mobile Account treatment on money given to the spouse or the probability of giving money to the spouse, though the coefficients are negative and I cannot reject equality with the Mobile Disbursement treatment at the 10% level.

This suggests that following receipt of the loan, spouses are receiving higher amounts of money from their wives. The Mobile Disbursement treatment mitigates this impact, and allows the women who receive this treatment to continue giving to their spouse at the baseline level. This suggests that receiving their loan on a mobile money account assists women in resisting pressure to share with their spouses.

I also confirm that as a result of giving less money to her spouse, or because she has higher income from her business, those treated with Mobile Disbursement don't receive less money from their spouses. This is shown in column (3) of Table 2.12. While women only give 20,000 USH ($5) to their spouse, they receive on average 160,000 USH ($40) from their spouses. This is unchanged between baseline and endline and does not differ by treatment. The spouse is therefore not giving the woman less money in light of her higher income, suggesting that this increased income may be hidden from him.

I additionally collected data on how the loan was used immediately after disbursement[17]. It's important to note that these questions about use of the loan in the week following disbursement were asked on average 8 months later, and so may be subject to large measurement error and recall bias compared to other questions which ask about the current period. They may also be more sensitive for the women to answer, since the loan is meant to be explicitly for their business, and so show over reporting of business expenditures. This bias however, would not be expected to differ by treatment group. I also did not pre-specify this outcome in the pre-analysis plan. Despite this, finding out how the loan was used immediately after disbursement provides important information about how the Mobile Disbursement treatment had an impact on business outcomes.

Results for how the loan was used across 7 categories are shown in Table 2.13. Spending on the business was the largest use of the loan immediately after disbursement, with an average of 760,000 USH ($200) or 54% of the mean loan size of 1.4mn USH ($370). However, spending on other categories was also large, with 135,000 USH ($36) going to sharing with others (10%), 112,000 USH ($30) on school

[17]this outcome was not pre-specified in the pre-analysis plan

fees (8%) and 110,000 USH on the household assets (8%). On average only 150,000 USH ($40) of the loan is 'saved' after the first week, suggesting that the loan is put to use very quickly rather than held as savings or spent on the business over a longer time period. On average, women reported expenditures accounting for 1.27mn ($340) of the 1.4mn USH loan, suggesting some under reporting may be occurring.

I see significant differences for the Mobile Disbursement treatment in the composition of loan spending. The Mobile Disbursement treatment group spend 29,000 USH ($7.7) less giving money to their family, 29,000 USH less on their home and save 45,000 USH ($12) more beyond the first week. This suggests a general slow down in spending as well as less spending on non-business expenditures. Combined with the findings of largest effects from the Mobile Disbursement treatment on profits and business capital for women who felt pressure to share money with family, and the reduction in transfers to the spouse, this suggests the Mobile Disbursement treatment could be helping women to protect their loan from their family, and as a result they are able to both spend the loan more slowly and spend more of it on their business.

This evidence on heterogeneity, money given to the spouse and use of the loan is further supported by anecdotes from focus groups carried out with a small sample of women from the study. A common theme that ran through all the discussions was the control that the Mobile Disbursement treatment gave to women to use the loan in the way they intended rather than spending it on other things or giving it to other people. Women described the disbursement of the loan onto the mobile money accounts as helping them to refuse requests for money by arguing that 'BRAC gave me this money for my business and placed it in this account so that I would only use it for my business. If I give some to you they'll (BRAC) will know[18]'. Women may therefore have used the loan being on the mobile money account as a method of refusing to give money to others in a way that wouldn't be seen to be violating

[18]BRAC never had access to the account transaction data, only the researcher did, and the women were informed of this at the start of the study. The woman saying this in the focus group knew BRAC didn't actually have the ability to know what she used the loan for if it came out of the mobile money account, but seemed to be using the fact that other people didn't know this to refuse their requests for money

Table 2.13: Treatment effects on secondary loan use outcomes

	(1) business	(2) sharing	(3) school	(4) home	(5) expenditure	(6) saving	(7) loan
Mobile account	11.88	7.11	4.75	9.44	-0.21	-9.52	-0.00
	(23.76)	(5.12)	(6.67)	(10.60)	(0.32)	(11.58)	(0.32)
Mobile disburse	17.32	-28.76***	-4.67	-29.30***	0.15	44.71***	0.04
	(23.56)	(4.90)	(6.24)	(9.48)	(0.34)	(12.24)	(0.25)
Observations	2,642	2,642	2,642	2,642	2,642	2,642	2,642
R-squared	0.20	0.21	0.17	0.16	0.21	0.16	0.11
Control mean endline	764.39	135.11	111.98	110.89	0.88	153.85	0.43
p-value T1=T2	0.821	0.00	0.14	0.00	0.25	0.00	0.89

Not specified in pre-analysis plan. Intent-to-treat estimates. All outcomes are winsorized at the 99% level. '000 Ugandan Shillings. All regressions include strata dummies. Mobile Account is the treatment where only a mobile money account was provided and the loan was disbursed as cash. Mobile Disburse is the treatment where a mobile money account was provided and the loan also disbursed onto this account. Amount of loan spent on each category 1 week after receiving loan. Business is business inventory and assets, sharing is money given to the spouse, friends or other family members, both at home and elsewhere, school is money spent on school fees and related expenditures, home is money spent on items for the home or home improvements, expenditure is money spent on food, clothes, transport etc. and loan is money spent paying back other loans. Recall 8 months later. Robust standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

social norms. I discuss social norms in section 2.7, but I do not find any evidence that treatment affected women's place in or amount of support from social networks.

2.6.4 Why didn't the Mobile Account treatment imitate the Mobile Disbursement treatment?

One puzzle about the results found here is why the Mobile Account treatment, and even the control treatment, did not just imitate the treatment received by the Mobile Disbursement group. In other words, why didn't the Mobile Account group take their loan and deposit some of it directly onto the sim card I gave them?

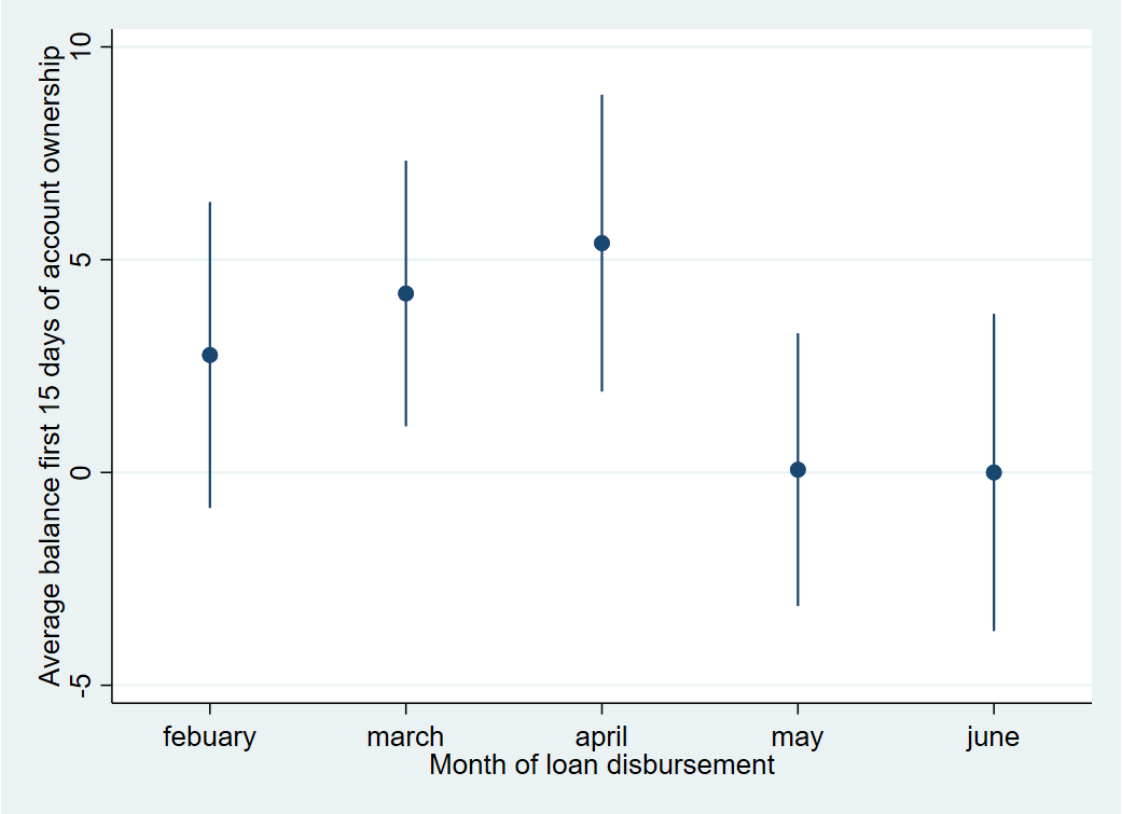
Equally, why didn't the control group, the majority of whom already did have a mobile money account, also deposit some of their loan onto the account?

There are a number of possible reasons why the control and Mobile Account groups might not have imitated the Mobile Disbursement. Firstly, as already discussed saving via mobile money was not very popular with less than 20% of the sample doing this. The amounts saved on a mobile money account were also relatively small, with a mean of 135,000 USH ($36) and a median of 100,000 USH ($27), compared to total savings of 800,000 USH ($210) (median 500,000 USH ($130)) for those that saved using mobile money. There may therefore have been learning effects around keeping money on a mobile money account and it being safe to store so much money on the account, since the average loan size was 10 times what the average women saved on mobile money. BRAC might also have legitimised that keeping so much money on a mobile money account is a safe and secure thing to do.

However, if this was true I'd expect to see the Mobile Account group becoming more likely to deposit the loan on their mobile money account over time, as they increasingly saw members of their group receive the loan on the mobile money account. I examine this by regressing the average balance on the mobile money account over the first 15 days since opening on month dummies for the month of the study, between February and June 2017. The coefficients on the month dummies are shown in figure 2.2. While at first it looks like the balances added to the account by the Mobile Account group are increasing over time, this trend breaks down in May and June. In fact, less than 10% of the Mobile Account group have a positive balance during the first 15 days after the loan is disbursed, and so a small number of people depositing relatively large amounts (300,000 USH ($80)) are skewing the numbers in the early part of the year. Overall this evidence suggests that there is no learning by the Mobile Account group to deposit their loan on the mobile money account, and so casts doubt that learning and validating by BRAC as a safe way to store money are responsible for my findings.

Secondly, a key benefit of receiving the loan on the mobile money account is the hiding of money. It is possible that going to an agent yourself and deposit

Figure 2.2: Average account balance Mobile Account group over time '000 USH



some of the loan still makes the loan more visible than BRAC depositing the money for you. While this could be a concern in village communities where the mobile money agent knows everyone, it is unlikely to be the case in Kampala where the woman could go to any number of agents and presumably could easily find one with no connection to her.

A third hypothesis relates to the time investment in depositing the loan into the mobile money account. Evidence has shown that even small costs can have large impacts on behaviour, particularly for those with hyperbolic preferences (O'Donoghue and Rabin, 1999). I confirm that distance to the nearest mobile money agent does not vary by treatment in the balance Table 2.1, and on average the women are less than 5 minutes from a mobile money agent. This suggests that transaction costs at least in terms of finding an agent are extremely low. However, even this cost combined with the costs of waiting in line and depositing the money with the agent may have been enough of a deterrent to the women to prevent them

depositing the loan themselves. Considering that 20% if the sample have hyperbolic preferences and 34% are defined as impatient, I cannot rule out that small time costs combined with some amount of procrastination could explain why the Mobile Account group does not imitate the Mobile Disbursement group.

A final explanation is default effects. Default effects have been shown to have large impacts on behaviour, including saving behaviour (Chetty et al., 2014, Choi et al., 2004). A number of studies have also looked at default effects as a driver of low savings in developing countries. Two studies have found that when people are given a bank account and then paid in either cash or directly onto that account, even when payment takes place at the bank itself those paid in cash do not deposit the money onto the accounts and as a result save less than those paid directly onto the account (Brune et al., 2017, Somville and Vandewalle, 2018). The reason for this is argued to be default effects, since the cost of transacting in this setting is so small. Additionally, when people are encouraged to save part of their salary, defaults were found to be equivalent to a 50% matching incentives in terms of the increase in savings they induced (Blumenstock et al., 2018). In my study, Mobile Account makes the default around adding savings onto the mobile money account. Mobile Disbursement makes the default removing money from the account. It is therefore very possible that the lack of imitation of Mobile Disbursement by those assigned to Mobile Account is entirely due to default effects and the inertia associated with them.

2.7 Alternative explanations

I examine a number of different potential reasons for the results I find. Firstly, the results may be simply a reallocation within the household that may actually leave the household worse off. Secondly, since the mobile money account facilitate remittances, any benefit to the household in terms of higher income may have been eroded by higher transfers to others. Thirdly, there may be experimenter demand effects combined with the salience of the loan being disbursed onto a business-designated mobile money account that made households report better business outcomes. Fourthly, there may be measurement error in business outcomes

and the mobile money disbursement of the loan may have helped households keep better track of their finances and so report better outcomes. Lastly, if women give less to their social networks, they may receive less in return, damaging their ability to withstand shocks.

2.7.1 Redistribution within the household

It is possible that if the mobile money disbursement helped women retain use of the loan for their own business over transferring it to other members of the household, that this could lead to a reduction in total household income and welfare if other household members have higher returns to capital in their businesses (Bernhardt et al., 2018). I therefore examine whether the income of other household members changed as a result of the treatments, as well as household consumption.

Looking at Table 2.14, I see an overall increase in household income of just under 90,000 USH ($24) for households in which the woman got her loan disbursed on the mobile money account. This is a similar figure to the increase in income I see for the woman's business (60,000 USH ($16)), with the difference seeming to be made up of (insignificant) increases in wage earnings for both the spouse and other household members. I see small and insignificant at the 5% level reductions in women's wage earnings from both treatments.

I see no differences in either the spouse or other household members business earnings as a result of giving women mobile money accounts. An important point to note here is that at baseline, business and wage incomes were not distinguished for the spouse and other household members but combined under primary and secondary sources of income in general. At endline, I explicitly distinguish between household business and wage income and collect more detailed information on other household business, including since when and for how long other household members have been running businesses. Even looking at total spouse and other household member earnings, I find no significant impacts of either treatment, and, if anything, the coefficients on the Mobile Disbursement treatment is positive.

Table 2.14: Treatment effects on secondary income outcomes

	(1) total household income	(2) woman's wage earnings	(3) spouse wage earnings	(4) spouse business earnings	(5) other hh wage earnings	(6) other hh business earnings	(7) spouse all earnings	(8) other hh all earnings
Mobile account	10.04	-7.86*	-11.52	10.30	11.83	-2.57	1.05	9.02
	(35.66)	(4.18)	(18.67)	(7.83)	(24.81)	(7.08)	(27.93)	(11.70)
Mobile disburse	87.14**	-2.39	11.03	12.99	-2.35	-3.95	18.67	10.31
	(36.48)	(4.49)	(19.16)	(8.07)	(24.78)	(7.04)	(28.83)	(11.72)
Observations	2,642	2,642	2,561	2,642	2,642	2,642	2,561	2,642
R-squared	0.33	0.18	0.25	0.24	0.16	0.15	0.33	0.27
Control mean endline	1010	25.42	187.1	56.56	281.09	38.34	477.55	99.31
Control mean baseline	1041	66.40	-	-	-	-	423.46	126.48
p-value T1=T2	0.03	0.19	0.22	0.75	0.57	0.84	0.53	0.91

Intent-to-treat estimates. All outcomes are winsorized at the 99% level. '000 Ugandan Shillings. All regressions include strata dummies and include the baseline value of the outcome. Mobile Account is the treatment where only a mobile money account was provided and the loan was disbursed as cash. Mobile Disburse is the treatment where a mobile money account was provided and the loan also disbursed onto this account. All incomes are monthly. Note at baseline spouse and household wage and business income was captured as a combined total. At endline they were captured separately. Difference between total household earnings and columns in this table is woman's business earnings. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

These findings suggests that in fact enabling the loan to be used by the woman for her business generates more income for the household. These results differ to the interpretation in Bernhardt et al. (2018), where women are investing the loan in whichever household businesses has the highest return, and on average women's businesses have lower returns in multi-business households. This could be because significant amounts of hiding are occurring in this sample[19], which may differ from other contexts, and so women may be engaging in costly hiding strategies to retain control over their loans. If the mobile money disbursement of the loan alleviates costly hiding by providing a more effective hiding device, then more profit and overall household income can be generated from the loan. Potential costly hiding strategies were discussed in Section 2.6 when I examined how the loan was spent immediately after disbursement, and found that significant amounts of the loan were used for household spending and that 90% of the loan was spent within the first week after disbursement. These findings are in line with Goldberg (2017) who finds households given a windfall income predict and actually spend more of it in the weeks immediately after getting it if the windfall is public.

I also validate that the increase in profits from the woman's business is feeding through into higher consumption[20]. Looking at consumption in Table 2.15, I see significant increases in overall consumption for the Mobile Disbursement treatment. This shows that the increase in profit from obtaining the Mobile Disbursement treatment is feeding through into higher household welfare overall. This increase in consumption is of a similar value to the increase in business profits seen (50,000 USH ($13) compare to a 60,000 USH ($16) profit increase), and so suggests the majority of the profit increase is actually being spent by the household. This could also explain why I find no impacts on savings from the treatment, as any additional income is being spent.

[19] 55% of the sample would be willing to give up $7 to retain control of money over giving it to their spouse

[20] since I find no increase in saving, this additional income must appear in consumption, remittances or as assets

Table 2.15: Treatment effects on secondary consumption outcomes

	(1) total	(2) food	(3) non-food exl school	(4) school
Mobile account	27.19	9.61	7.87	12.42
	(23.99)	(9.83)	(8.49)	(12.76)
Mobile disburse	50.66**	20.50**	4.87	22.06*
	(24.26)	(10.31)	(8.45)	(12.04)
Observations	2,642	2,642	2,642	2,642
R-squared	0.34	0.24	0.20	0.39
Control mean endline	973.6	406	252.5	300.6
Control mean baseline	886.6	398.3	224.3	252.7
p-value T1=T2	0.334	0.293	0.732	0.433

Intent-to-treat estimates. All outcomes are winsorized at the 99% level. '000 Ugandan Shillings. All regressions include strata dummies and include the baseline value of the outcome. Mobile Account is the treatment where only a mobile money account was provided and the loan was disbursed as cash. Mobile Disburse is the treatment where a mobile money account was provided and the loan also disbursed onto this account. All values are monthly for the entire household. Non-food consumption excludes temptation spending and transfers. Control mean endline is the mean value of the outcome in the control group at endline. Control mean baseline is the mean value of the outcome in the control group at baseline. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Anecdotally, women described school expenditures as one of the easiest ways to hide excess income, as they are a large and variable expenditure once costs of schools supplies are included and the spouse has little idea of the true cost. I do indeed see that women treated with the Mobile Disbursement spend more on their children's schooling in Table 2.15.

2.7.2 Remittances

Mobile money accounts make it easier to send remittances (Jack and Suri, 2011). Any benefits of the accounts in terms of ease of saving money may therefore be outweighed by the increased ease of sending money. I examine this by looking at remittance flows.

Looking at remittances in Table 2.16, which are defined as money sent/received from non-household members, I see relatively large coefficients on amount of

money sent for both the Mobile Account and Mobile Disbursement treatments of approximately 10,000 USH ($3). However, only the coefficient on the Mobile Account treatment is significant at the 10% level. I see no other large or significant effects of the treatments on amount received as remittances, the net amount received (amount received minus amount sent), whether the woman used a mobile money account to send the remittances or the probability that she received or sent remittances. Overall, this suggests there might be a small increase in the amount of remittances sent as a result of treatment, but no increase in use of mobile money or likelihood of sending remittances using other forms.

This mitigates concerns about the treatments that any beneficial effect of receiving the loan on a secure mobile money account might be outweighed by the fact that the mobile money account makes it easier to send money to others. The fact that I see little to no effects on remittances might be partly because the mobile money account provided in the study was a second mobile money account for most of the women. If the account had been the first and primary mobile money account for the women it is possible more leakages of the loan in the form of remittances might have occurred.

2.7.3 Experimenter demand effects

The salience of giving mobile money accounts designed for the business and of disbursing the loans specifically onto those accounts may have caused those women who received the Mobile Disbursement treatment to over report their business outcomes because they believed that is what the study intended to do. However, they are unlikely to also over report improvements in other household outcomes not linked to the business.

While it is possible that the Mobile Disbursement treatment made it more salient to households that they should be investing in their business and so caused them to report more of their household assets as used for the business, they are unlikely to also over inflate their household assets. In that case I would see a rise in business asset for the Mobile Disbursement group without any overall increase

Table 2.16: Treatment effects on secondary remittance outcomes

	(1) amount sent	(2) amount received	(3) net amount received	(4) used mobile money	(5) Recei-ved dummy	(6) Sent dummy
Mobile account	11.37*	-5.29	1.71	-0.01	-0.03	0.02
	(6.89)	(10.38)	(6.32)	(0.02)	(0.02)	(0.02)
Mobile disburse	10.37	-3.83	1.39	-0.01	-0.02	0.03
	(6.68)	(10.27)	(5.56)	(0.02)	(0.02)	(0.02)
Observations	2,642	2,642	2,642	2,642	2,639	2,639
R-squared	0.23	0.21	0.14	0.19	0.18	0.21
Control mean endline	58.03	85.86	6.83	0.37	0.34	0.34
p-value T1=T2	0.88	0.89	0.95	0.94	0.53	0.83

Intent-to-treat estimates. All outcomes are winsorized at the 99% level. '000 Ugandan Shillings. All regressions include strata dummies. Mobile Account is the treatment where only a mobile money account was provided and the loan was disbursed as cash. Mobile Disburse is the treatment where a mobile money account was provided and the loan also disbursed onto this account. All outcomes reported here were only collected at endline. Remittances defined as money given to a non-household member. Robust standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

in household assets. For the households in this sample the distinction between household and business assets is not clear, and often the same asset is used both by the household and by the woman's business. The survey therefore asked for all assets owned by (anyone in) the household, and of those, which were used for the woman's business. Total household assets then by definition captures all those used in the woman's business.

I test this by seeing if there is an overall increase in assets regardless of how they are used. In Table 2.17, I see that the Mobile Disbursement treatment led to a significant increase in overall asset levels of 340,000 USH ($90) compared to control. Since household assets by definition includes those used in the business, this measure confirms an increase in businesses assets. In Table 2.9, we saw in column (2) that the value of business assets is 132,000 USH ($35) higher for the Mobile Disbursement treatment. This implied that 200,000 USH ($55) was

additionally invested by the Mobile Disbursement group in household assets. In addition the control group increased by 1mn USH ($270) their household assets between baseline and endline in household assets. This means that actually one of the key uses of the loan for all the women in the study is increasing household assets, and the Mobile Disbursement treatment appears to have increased both business and household assets even further.

Table 2.17: Treatment effects on secondary wealth outcomes

	(1) Total asset value
Mobile account	137.97
	(151.95)
Mobile disburse	342.67**
	(154.94)
Observations	2,642
R-squared	0.30
Control mean endline	4398
Control mean baseline	3384
p-value T1=T2	0.18

Intent-to-treat estimates. All outcomes are winsorized at the 99% level. '000 Ugandan Shillings. All regressions include strata dummies and include the baseline value of the outcome. Mobile Account is the treatment where only a mobile money account was provided and the loan was disbursed as cash. Mobile Disburse is the treatment where a mobile money account was provided and the loan also disbursed onto this account. Total asset value includes the value of all household and business assets. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

As already noted, consumption in the Mobile Disbursement women's households increased by close to the amount that woman's business profits increased. Since it is less clear why the woman would inflate her consumption because she thinks we wanted her business to grow, this provides further evidence that the business improvement is not due to experimenter demand effects.

Additionally, it is not clear that just providing a business-designated mobile money account is significantly less salient as a treatment designed to improve their business than also providing the loan on the account. If experimenter demand effects were strong in this population, it would be strange to see no effect of this treatment too.

Experimenter demand effects have been found to be relatively small (de Quidt et al., 2018), and so combined with the fact I find impact across a range of household, not just business, outcomes, I do not believe experimenter demand effects could be driving my results.

2.7.4 Measurement error

The mobile money accounts may have made it easier to keep track of business outflows, sales and profits if the mobile money account was used for these activities. The disbursement of the loan onto the mobile money account may also have made it easier to keep track of what the loan was spent on. These are unlikely to be responsible for the impacts I see for the following reasons.

Firstly, the mobile money accounts given to either treatment group were not used by the majority of the women for frequent deposits and withdrawals of funds. They therefore are unlikely to have made it easier to keep track of regular business expenses and sales since these activities did not take place on the accounts. Additionally, I would only see impacts from use of mobile money accounts correcting measurement error if measurement error only downwardly biased estimates of profit and business capital. It is not clear why measurement error would only downward bias reported business outcomes.

Secondly, while the Mobile Disbursement treatment may have made it easier to track the use of the loan, this would only be expected to impact capital expenditures on inventory and assets. There should not be any additional effect on profits, or the downstream outcomes of household consumption.

Overall, this suggests that the idea that the mobile money accounts corrected measurement error in the tracking of business outcomes seems unlikely as an explanation for the impacts I see.

2.7.5 Social networks

I argue that the Mobile Disbursement treatment helped women resist pressure to give money to others. However, if women are giving less to their social network they may also receive less and be less able to withstand shocks. I did not collect survey data on social network links or experiences of negative shocks. However, I do have some data on money given to and received from others and on the number of people the woman can rely on when in need from her microfinance group. I can use these as proxies for social networks.

Firstly, I do not see any changes for either treatment group in the amount of remittances either given by or received from others, see Table 2.16, suggesting women are not contributing less or being cut off from wider remittance networks. Instead, I argue it is primarily the spouse and immediate household who receive less.

Secondly, I look at women's peers in the microfinance group. Many of the women described their friends in the microfinance group as those they rely on most when in need. I asked questions on the number of women in the microfinance group a woman talks to at least once a week outside the group, how many they could ask for financial help from and how many they'd offer financial help to. The results of treatment on each of these outcomes is shown in Table 2.18. On average, women talk to 7 other group member at least once a week outside the group but would ask for help from, and be happy to give help to just around 4 of these. This is from a mean group size of 21 women. I find no difference by treatment status, suggesting getting the loan on a mobile money account did not isolate women from other members of their microfinance group.

Table 2.18: Treatment effects on number of women in the microfinance group you'd interact with in each of the situations

	(1) talk to at least once a week outside the group	(2) ask for financial help from if you needed money	(3) give financial help to if she needed money
Mobile account	0.14	-0.09	-0.11
	(0.26)	(0.20)	(0.21)
Mobile disburse	0.05	0.09	0.08
	(0.26)	(0.20)	(0.22)
Observations	2,642	2,642	2,642
R-squared	0.18	0.20	0.19
Control mean endline	6.96	3.77	3.90
p-value T1=T2	0.74	0.37	0.39

Intent-to-treat estimates. All regressions include strata dummies. Mobile Account is the treatment where only a mobile money account was provided and the loan was disbursed as cash. Mobile Disburse is the treatment where a mobile money account was provided and the loan also disbursed onto this account. Outcomes only measured at endline. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

2.8 Conclusion

This paper shows that the manner in which loans are disbursed to microfinance clients leads to significant differences in how those loans are used. Women assigned to receive the loan on the mobile money account hold significant balances equal to 7% of the loan value or 22% of household savings on their account during the first 30 days after getting the loan. They draw down this balance over a 6 month period. Clients who receive their loan on a mobile money account invest in 11% more business capital and as a result have 15% higher profits. These impacts are largest for women who experiences family pressure to share money at baseline, and result in them giving less of their loan to their spouse and other household members. This suggests the benefits to women's business from the Mobile Disbursement treatment come from a safe and private way to store the loan, saving it to invest when needed.

My study suggests that microfinance loan providers should consider disbursing the loan onto a private account, as opposed to the current default in much of the

world of cash. This small change could have significant benefits to the profitability of female entrepreneurs. With the increasing spread of mobile money services, this intervention is a low cost way to raise the benefits of microenterprise loans to women and an easy policy recommendation for NGOs and other organisations disbursing microfinance loans to follow.

One limitation of this study is the short time horizon over which it took place: 8 months was chosen as the follow up period to allow the endline survey to be completed before most clients loan repayment period had ended, thus improving tracking. However, as a result of this design it is not clear whether the benefits to the women's profitability would persist going forward. This is especially true since BRAC Uganda reverted to disbursing loans using only cash after the study ended, despite many clients expressing their preference for mobile money[21]. Even if BRAC had continued with Mobile Disbursement, it is entirely possible that over time a woman's family would learn about the mobile money disbursement and find more effective means of gaining access to the funds there. Ideally, future work would both replicate my findings and also look at how the effects persisted over a longer period of time of making loan disbursements using mobile money.

A second limitation is that my study only took place in an urban sample amongst women familiar with mobile money services. Women in rural locations may stand to benefit more from disbursement of a loan onto a mobile money account if they also are saving constrained. However, they may struggle to use the service and require more training, and limitations in the amount of float that agents hold in rural areas may prevent them cashing out as much of the loan as they'd like. Further research is needed to understand how my results generalise to rural locations and other contexts where people are less familiar with mobile money.

[21]BRAC Uganda are currently transforming to a full banking license, and are planning to pilot mobile money loan disbursement again once they are able to do the disbursement themselves as opposed to through a partner

Appendices of Chapter 2

2.A Additional tables

Table 2.19: Correlates of treatment take up

	(1) Mobile Account	(2) Mobile Disburse
respondent age	-0.001	-0.002
	(0.001)	(0.001)
married	0.021	-0.046
	(0.015)	(0.026)
household size	-0.003	-0.007
	(0.005)	(0.007)
primary school	0.011	0.011
	(0.018)	(0.031)
secondary school	0.01	0.084*
	(0.023)	(0.035)
job	0.000	-0.059
	(0.018)	(0.03)
loan amount	0.000	0.000
	(0.000)	(0.000)
weekly profit	0.000	0.000
	(0.000)	(0.000)
high profits	-0.002	-0.035
	(0.014)	(0.025)
current client	-0.016	-0.045
	(0.019)	(0.032)
amount saved	0.000	0.000
	(0.000)	(0.000)
mobile money account	-0.020	0.017
	(0.031)	(0.070)
hyperbolic	-0.034	0.024
	(0.017)	(0.031)
impatient	-0.025	0.018
	(0.015)	(0.025)
woman's income share	-0.01	-0.032
	(0.023)	(0.039)
hides money	-0.047**	0.047
	(0.017)	(0.033)
family takes	0.022	-0.026
	(0.015)	(0.026)
spouse business	-0.026	-0.01
	(0.019)	(0.035)
household business	-0.004	-0.043
	(0.016)	(0.027)
Observations	984	956
R-squared	0.033	0.029
Mean control	0.946	0.823
F-test p-value	0.50	0.66

Each row represents a separate OLS regression of whether the individual accepted that treatment on the baseline characteristics specified. I also include a p-value from an F-test of regressing all the characteristics on the take-up dummies. I count as those who consented to the Mobile Disbursement treatment both those who got it (71%) and the 10% who consented but couldn't receive treatment due to random technical errors. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 2.20: Correlates of attrition

	(1) attrition
respondent age	-0.003***
	(0.001)
married	-0.011
	(0.012)
household size	-0.012***
	(0.003)
primary school	0.017
	(0.014)
secondary school	0.036*
	(0.017)
job	-0.008
	(0.014)
loan amount	-0.000**
	(0.000)
weekly profit	0.000
	(0.000)
high profits	-0.011
	(0.011)
current client	-0.003
	(0.015)
amount saved	-0.000
	(0.000)
mobile money account	0.018
	(0.028)
hyperbolic	-0.006
	(0.014)
impatient	-0.000
	(0.012)
woman's income share	-0.004
	(0.018)
hides money	0.001
	(0.014)
family takes	-0.031*
	(0.012)
Observations	2,959
R-squared	0.017
F-test p-value	0.000

Linear regression of baseline characteristics on a variable equal to one if the woman was not surveyed at endline. Each row represents a separate regression. Monetary amounts in '000 Ugandan Shilling and winsorized at the 99% level. The F-test p-value comes from regressing the attrition variable on all the characteristics and testing if they are jointly zero. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 2.21: Treatment effects on primary outcomes - permutation test

	(1) profit	(2) savings	(3) capital
Mobile account	10.41	3.33	-30.40
	(0.42)	(0.92)	(0.98)
	[0.99]	[0.99]	[0.99]
	{0.17}	{ 0.82}	{0.93}
Mobile disburse	63.72***	26.66	213.08***
	(0.00)	(0.47)	(0.01)
	[0.00]	[0.86]	[0.03]
	{0.00}	{0.95}	{0.03}
Observations	2,639	2,639	2,639
R-squared	0.44	0.41	0.60
Control mean endline	395.3	580.6	2473
Control mean baseline	419.8	483.6	2488
p-value T1=T2	0.00	0.50	0.00

Intent-to-treat estimates. All outcomes are winsorized at the 99% level. '000 Ugandan Shillings. All regressions include strata dummies and include the baseline value of the outcome. Mobile Account is the treatment where only a mobile money account was provided and the loan was disbursed as cash. Mobile Disburse is the treatment where a mobile money account was provided and the loan also disbursed onto this account. Profits refers to the self-reported monthly business profit. Savings is individual savings held by the woman. Capital is the value of all assets the woman uses in her business plus the value of inventory held for her business. Control mean endline is the mean value of the outcome in the control group at endline. Control mean baseline is the mean value of the outcome in the control group at baseline. False discovery rate (FDR) adjusted p-values, also known as q-values, were used to correct for multiple hypothesis testing. They are shown in square brackets. These were calculated following the method of Benjamini et al. (2006). Permutation p-values are shown in curly brackets. These used the permute command in Stata and 1000 repetitions. Robust p-values in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 2.22: Primary outcome results with linear and quadratic time trend of the number of days between loan disbursement and endline

	(1) profit	(2) savings	(3) capital
Mobile account	12.91	0.80	13.37
	(12.99)	(33.91)	(85.26)
	[0.99]	[0.99]	[0.99]
Mobile disburse	61.69***	28.99	202.34**
	(12.76)	(37.24)	(83.06)
	[0.00]	[0.80]	[0.04]
Observations	2,639	2,639	2,639
R-squared	0.44	0.41	0.57
Control mean endline	395.3	580.6	2473
Control mean baseline	419.8	483.6	2488
p-value T1=T2	0.00	0.42	0.02
p-value T1=T2=0	0.00	0.67	0.02

Intent-to-treat estimates. All outcomes are winsorized at the 99% level. '000 Ugandan Shillings. All regressions include strata dummies and include the baseline value of the outcome. Mobile Account is the treatment where only a mobile money account was provided and the loan was disbursed as cash. Mobile Disburse is the treatment where a mobile money account was provided and the loan also disbursed onto this account. Profits refers to the self-reported monthly business profit. Savings is individual savings held by the woman. Capital is the value of all assets the woman uses in her business plus the value of inventory held for her business. Control mean endline is the mean value of the outcome in the control group at endline. Control mean baseline is the mean value of the outcome in the control group at baseline. False discovery rate (FDR) adjusted p-values, also known as q-values, were used to correct for multiple hypothesis testing. They are shown in square brackets. These were calculated following the method of Benjamini et al. (2006)
Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 2.23: Primary outcome results with winsorizing the top 2%

	(1) profit	(2) savings	(3) capital
Mobile account	8.50	9.93	22.41
	(12.22)	(32.28)	(80.85)
	[0.99]	[0.99]	[0.99]
Mobile disburse	54.39***	25.26	223.22***
	(11.93)	(34.33)	(78.19)
	[0.00]	[0.85]	[0.01]
Observations	2,639	2,639	2,639
R-squared	0.41	0.38	0.55
Control mean endline	393.4	559.5	2420
Control mean baseline	415.4	441.8	2443
p-value T1=T2	0.00	0.64	0.01
p-value T1=T2=0	0.00	0.76	0.00

Intent-to-treat estimates. Mobile account is the treatment where only a mobile money account was provided and the loan was disbursed as cash. Mobile disburse is the treatment where a mobile money account was provided and the loan also disbursed onto this account. '000 Ugandan Shillings. Profits refers to the self-reported monthly business profit. Savings is individual savings held by the woman. Capital is the value of all assets the woman uses in her business plus the value of inventory held for her business. All outcomes are winsorized at the 98% level. All regressions include strata dummies and include the baseline value of the outcome. Control mean endline is the mean value of the outcome in the control group at endline. Control mean baseline is the mean value of the outcome in the control group at baseline. False discovery rate (FDR) adjusted p-values, also known as q-values, were used to correct for multiple hypothesis testing. They are shown in square brackets. These were calculated following the method of Benjamini et al. (2006). Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 2.24: Primary outcome results with winsorizing the top 0.5%

	(1) profit	(2) savings	(3) capital
Mobile account	9.47	-6.65	-17.91
	(13.93)	(36.66)	(86.19)
	[0.99]	[0.99]	[0.99]
Mobile disburse	75.13***	29.46	196.06**
	(13.84)	(41.43)	(84.50)
	[0.00]	[0.88]	[0.06]
Observations	2,639	2,639	2,639
R-squared	0.46	0.40	0.57
Control mean endline	396.5	597.2	2488
Control mean baseline	421.2	491.3	2522
p-value T1=T2	0.00	0.34	0.01
p-value T1=T2=0	0.00	0.63	0.01

Intent-to-treat estimates. All outcomes are winsorized at the 99.5% level. '000 Ugandan Shillings. All regressions include strata dummies and include the baseline value of the outcome. Mobile Account is the treatment where only a mobile money account was provided and the loan was disbursed as cash. Mobile Disburse is the treatment where a mobile money account was provided and the loan also disbursed onto this account. Profits refers to the self-reported monthly business profit. Savings is individual savings held by the woman. Capital is the value of all assets the woman uses in her business plus the value of inventory held for her business. Control mean endline is the mean value of the outcome in the control group at endline. Control mean baseline is the mean value of the outcome in the control group at baseline. False discovery rate (FDR) adjusted p-values, also known as q-values, were used to correct for multiple hypothesis testing. They are shown in square brackets. These were calculated following the method of Benjamini et al. (2006). Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 2.25: Treatment effects on primary outcomes - ATT

	(1) profit	(2) savings	(3) capital
Mobile account	11.07	3.54	2.29
	(13.70)	(37.21)	(85.89)
Mobile disburse	88.94***	42.47	297.6***
	(18.04)	(49.04)	(113.2)
Observations	2,610	2,610	2,610
R-squared	0.261	0.221	0.473
Control mean endline	395.3	559.2	2473
Control mean baseline	419.8	483.6	2488
p-value T1=T2	0.000	0.383	0.004

Average treatment on the treated estimates using treatment assignment as an instrument for actual take-up. All outcomes are winsorized at the 99% level. '000 Ugandan Shillings. All regressions include strata dummies and include the baseline value of the outcome.Mobile Account is the treatment where only a mobile money account was provided and the loan was disbursed as cash. Mobile Disburse is the treatment where a mobile money account was provided and the loan also disbursed onto this account. Profits refers to the self-reported monthly business profit. Savings is individual savings held by the woman. Capital is the value of all assets the woman uses in her business plus the value of inventory held for her business. Control mean endline is the mean value of the outcome in the control group at endline. Control mean baseline is the mean value of the outcome in the control group at baseline. False discovery rate (FDR) adjusted p-values, also known as q-values, were used to correct for multiple hypothesis testing. They are shown in square brackets. These were calculated following the method of Benjamini et al. (2006). Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 2.26: Heterogeneous treatment effects on business profit

	(1) high profits	(2) hide money	(3) current loan	(4) hyper - bolic	(5) impa- tient	(6) high saving	(7) high asset	(8) mar- ried	(9) high empower	(10) sent family	(11) family takes	(12) saving goal bus	(13) spouse bus	(14) hh bus
MA*interaction	-33.76	11.67	5.53	11.87	1.30	-14.21	-4.48	5.80	-18.84	29.09	-1.12	-45.11	26.77	25.69
	(27.65)	(38.40)	(29.83)	(32.16)	(29.90)	(28.13)	(27.84)	(30.39)	(28.23)	(28.96)	(30.15)	(36.22)	(41.02)	(28.61)
MD*interaction	109.73***	76.04**	21.74	36.14	55.39*	29.90	28.50	66.83**	4.68	-27.75	70.03**	-31.84	74.51**	56.31**
	(27.17)	(35.56)	(29.55)	(30.22)	(29.29)	(26.78)	(26.53)	(28.51)	(27.54)	(28.19)	(28.27)	(34.35)	(37.84)	(27.63)
Mobile account	18.70	-12.70	-2.80	0.49	1.79	9.16	4.61	-2.11	11.27	-16.59	0.32	12.02	-18.69	-9.97
	(15.00)	(32.89)	(26.23)	(14.68)	(16.32)	(17.75)	(18.15)	(24.56)	(18.56)	(22.43)	(16.24)	(14.78)	(31.12)	(18.46)
Mobile disburse	18.33	34.14	52.50**	63.13***	47.51***	55.32***	56.54***	25.73	67.37***	87.65***	45.07***	77.26***	28.69	45.00**
	(15.84)	(29.20)	(25.73)	(14.39)	(15.88)	(16.91)	(17.27)	(22.53)	(18.05)	(22.40)	(15.47)	(14.36)	(29.88)	(17.95)
Observations	2,606	1,726	2,606	2,606	2,606	2,606	2,606	2,606	2,606	2,606	2,606	2,606	1,738	2,606
R-squared	0.42	0.51	0.42	0.42	0.42	0.42	0.42	0.42	0.42	0.42	0.42	0.42	0.50	0.42
Control mean	469.88	380.94	401.98	362.34	399.47	434.58	409.64	387.26	406.18	397.79	396.38	419.80	368.71	367.20
Control mean base	657.95	412.95	437.77	381.10	438.72	480.86	430.95	417.21	436.01	410.68	467.98	389.12	397.75	392.29
Interaction mean	0.48	0.63	0.82	0.20	0.35	0.49	0.48	0.66	0.51	0.64	0.34	0.22	0.58	0.45
Overall effects														
mobile account	-15.06	-1.04	2.73	12.35	3.09	-5.05	0.14	3.68	-7.57	12.50	-0.80	-33.09	8.09	15.72
	(22.35)	(19.87)	(14.65)	(28.42)	(23.89)	(20.48)	(19.95)	(16.12)	(19.82)	(16.82)	(24.19)	(31.85)	(23.28)	(20.24)
mobile disburse	128.06***	110.18***	74.24***	99.27***	102.90***	85.22***	85.04***	92.56***	72.05***	59.90***	115.10***	45.42	103.20***	101.30***
	(20.89)	(20.00)	(14.38)	(26.44)	(23.33)	(19.76)	(19.34)	(16.02)	(19.30)	(15.89)	(22.95)	(30.15)	(21.65)	(19.50)
p-value for testing														
T1=T2	0.98	0.15	0.02	0.00	0.00	0.01	0.00	0.22	0.00	0.00	0.00	0.00	0.08	0.00
T1=T2 interaction	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00

Intent-to-treat estimates. Self-reported business profits in '000 Ugandan Shillings. Monetary outcomes are winsorized at the 99% level. All regressions include strata dummies. Mobile Account (MA) is the treatment where only a mobile money account was provided and the loan was disbursed as cash. Mobile Disburse (MD) is the treatment where a mobile money account was provided and the loan also disbursed onto this account. Control mean refers to the mean value of the outcome variable for the interaction condition being true in the control group, and control mean base the mean in the control group at baseline when the interaction condition is true. Heterogeneous variables are defined in section 2.6. Note that hide money and spouse bus are only reported for married women who have a spouse. The panel labelled overall effects gives the total impact of each treatment for someone who is the interaction term is true for. The panel labelled overall effects gives the total impact of each treatment for someone who is the interaction term is true for. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 2.27: Heterogeneous treatment effects on business capital

	(1) high profits	(2) hide money	(3) current loan	(4) hyper - bolic	(5) impa- tient	(6) high saving	(7) high asset	(8) mar- ried	(9) high empower	(10) sent family	(11) family takes	(12) saving goal bus	(13) spouse bus	(14) hh bus
MA*interaction	26.17	-158.87	-81.66	102.55	-171.49	-125.30	-142.47	130.03	21.76	36.80	43.81	-13.66	88.81	26.16
	(184.37)	(240.43)	(213.46)	(217.42)	(196.33)	(183.54)	(185.91)	(192.13)	(185.18)	(180.63)	(196.29)	(230.42)	(249.81)	(186.71)
MD*interaction	143.50	131.33	192.57	6.45	14.19	171.84	-34.40	403.09**	148.09	-291.01*	400.97**	-8.05	395.11*	367.86**
	(176.97)	(206.91)	(209.72)	(203.59)	(185.98)	(174.32)	(176.47)	(179.57)	(174.84)	(167.52)	(176.69)	(213.78)	(223.48)	(174.48)
Mobile account	-41.47	129.35	37.07	-51.46	27.49	32.75	40.22	-115.35	-35.40	-53.54	-50.84	-24.54	-14.39	-43.57
	(114.17)	(189.68)	(191.70)	(96.79)	(106.94)	(110.65)	(101.38)	(148.02)	(137.09)	(132.39)	(106.89)	(98.23)	(180.20)	(116.24)
Mobile disburse	191.51*	308.67**	104.97	260.25***	241.06**	175.89*	277.93***	-5.33	191.35	445.00***	116.48	261.95***	182.23	102.99
	(108.22)	(147.33)	(191.46)	(90.37)	(101.15)	(106.77)	(99.23)	(142.62)	(126.66)	(120.45)	(100.62)	(90.40)	(164.58)	(111.62)
Observations	2,606	1,726	2,606	2,606	2,606	2,606	2,606	2,606	2,606	2,606	2,606	2,606	1,738	2,606
R-squared	0.61	0.64	0.61	0.61	0.61	0.61	0.61	0.61	0.61	0.61	0.61	0.61	0.67	0.61
Control mean	2766	2711	2649	2520	2589	3004	3225	2622	2311	2660	2496	2593	2671	2661
Control mean base	3020	2717	2702	2610	2476	3175	3446	2689	2437	2607	2625	2471	2768	2653
Interaction mean	0.48	0.63	0.82	0.20	0.35	0.49	0.48	0.66	0.51	0.64	0.34	0.22	0.58	0.45
Overall effects														
mobile account	-15.30	-29.52	-44.59	51.09	-144.00	-92.55	-102.20	14.68	-13.64	-16.73	-7.03	-38.20	74.42	-17.41
	(138.4)	(153.3)	(96.53)	(195.7)	(159.6)	(140.6)	(150.3)	(112.6)	(115.7)	(117)	(159.8)	(203.3)	(162.1)	(139)
mobile disburse	335**	440**	297.5***	266.7	255.2*	347.7***	243.5*	397.8***	339.4***	154	517.4***	253.9	577.3***	470.9***
	(131.3)	(146.3)	(88.38)	(182.6)	(149.1)	(129.8)	(138.3)	(101.6)	(111)	(110.4)	(141.7)	(190.3)	(144.8)	(126)
P-value for testing														
T1=T2	0.02	0.30	0.70	0.00	0.03	0.18	0.01	0.38	0.07	0.00	0.07	0.00	0.22	0.17
T1=T2 interaction	0.00	0.00	0.00	0.15	0.00	0.00	0.01	0.00	0.00	0.08	0.00	0.09	0.00	0.00

Intent-to-treat estimates. Monetary outcomes are winsorized at the 99% level and in '000 USH. All regressions include strata dummies. Mobile Account (MA) refers to the treatment where women got a mobile money account and their loan as cash. Mobile Disburse (MD) refers to the treatment where women got a mobile money account and the loan disbursed onto the account. Business capital is composed of business assets and inventories. Heterogeneous variables are defined in section 2.6. Note that hide money and spouse bus are only reported for married women who have a spouse. Control mean refers to the mean value of the outcome variable for the interaction condition being true in the control group, and control mean base the mean in the control group at baseline when the interaction condition is true. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 2.28: Heterogeneous treatment effects on saving

	(1) high profits	(2) hide money	(3) current loan	(4) hyper - bolic	(5) impa- tient	(6) high saving	(7) high asset	(8) mar- ried	(9) high empower	(10) sent family	(11) family takes	(12) saving goal bus	(13) spouse bus	(14) hh bus
MA*interaction	24.05	22.20	78.99	-23.92	40.37	58.49	15.45	-13.82	41.84	72.71	-71.81	-54.14	-30.60	25.30
	(73.62)	(93.71)	(70.99)	(96.05)	(70.83)	(73.55)	(74.76)	(78.56)	(74.25)	(70.89)	(86.48)	(93.95)	(97.50)	(74.53)
MD*interaction	-84.38	32.13	-12.76	-68.02	55.31	-72.43	-70.40	-160.41*	-51.04	-95.67	-4.55	-113.17	-175.68*	-159.84**
	(80.49)	(101.01)	(76.77)	(95.37)	(74.94)	(78.04)	(82.21)	(87.26)	(78.95)	(82.25)	(89.96)	(97.77)	(103.37)	(81.11)
Mobile account	-8.83	-26.61	-62.83	5.43	-9.62	-25.38	-2.46	14.50	-16.55	-43.08	28.37	16.10	16.62	-7.94
	(45.48)	(73.04)	(58.67)	(36.21)	(46.46)	(31.98)	(38.53)	(63.69)	(54.90)	(51.46)	(38.10)	(38.38)	(65.82)	(47.46)
Mobile disburse	66.33	-48.34	36.62	38.80	10.81	62.49	60.87	134.36*	54.26	87.92	26.93	52.85	92.58	97.51*
	(50.97)	(79.95)	(62.78)	(41.82)	(52.42)	(38.76)	(45.98)	(70.16)	(62.09)	(62.64)	(41.37)	(42.48)	(72.62)	(52.92)
Observations	2,639	1,744	2,639	2,639	2,639	2,639	2,639	2,639	2,639	2,639	2,639	2,639	1,752	2,639
R-squared	0.41	0.48	0.41	0.41	0.41	0.41	0.41	0.41	0.41	0.41	0.41	0.41	0.51	0.41
Control mean	691.37	627.95	617.38	680.81	460.23	829.50	726.99	619.70	483.33	591.45	699.47	622.30	635.16	622.38
Control mean base	629.78	522.50	529.01	500.33	390.03	896.39	631.98	497.71	424.08	504.08	606.39	482.22	485.06	480.19
Interaction mean	0.48	0.63	0.82	0.20	0.35	0.49	0.48	0.66	0.51	0.64	0.34	0.22	0.58	0.45
Overall effects														
mobile account	15.22	-4.42	16.16	-18.48	30.75	33.11	12.98	0.68	25.29	29.63	-43.43	-38.05	-13.98	17.36
	(54.46)	(56.42)	(39.41)	(89.00)	(49.79)	(63.69)	(60.87)	(41.78)	(45.23)	(46.21)	(74.32)	(83.15)	(63.19)	(53.48)
mobile disburse	-18.06	-16.21	23.87	-29.22	66.12	-9.94	-9.53	-26.05	3.22	-7.74	22.39	-60.32	-83.10	-62.33
	(58.47)	(61.51)	(43.27)	(84.96)	(50.05)	(65.61)	(63.77)	(45.99)	(44.74)	(48.77)	(77.07)	(85.25)	(67.88)	(56.75)
P-values for testing														
T1=T2	0.09	0.77	0.07	0.39	0.67	0.02	0.16	0.09	0.20	0.03	0.97	0.36	0.23	0.04
T1=T2 interaction	0.56	0.83	0.85	0.89	0.50	0.49	0.70	0.52	0.63	0.40	0.37	0.79	0.28	0.14

Intent-to-treat estimates. Amount saved in '000 Ugandan Shillings. Monetary outcomes are winsorized at the 99% level. All regressions include strata dummies. Mobile account is the treatment where only a mobile money account was provided and the loan was disbursed as cash. Mobile disburse is the treatment where a mobile money account was provided and the loan also disbursed onto this account. MA refers to the treatment where women got a mobile money account and their loan as cash. MD refers to the treatment where women got a mobile money account and the loan disbursed onto the account. Control mean refers to the mean value of the outcome variable for the interaction condition being true in the control group, and control mean base the mean in the control group at baseline when the interaction condition is true. Heterogeneous variables are defined in section 2.6. Note that hide money and spouse bus are only reported for married women who have a spouse. The panel labelled overall effects gives the total impact of each treatment for someone who is the interaction term is true for. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Chapter 3

Role Models in Movies

The Impact of *Queen of Katwe* on Students' Educational Attainment*

Abstract This paper presents experimental evidence on the impact of a role model on secondary school student exam performance in Uganda. Students preparing to take their national exams were individually randomised to see either a movie featuring a potential role model, *Queen of Katwe*, or to see a placebo movie. I find that treatment with the role-model movie leads to lower secondary school students being less likely to fail their maths exam a week later: 84% of those who watched *Queen of Katwe* passed the exam, whereas only 73% of those who didn't passed. This effect is strongest for female and lower ability students. For upper secondary school students, treatment with *Queen of Katwe* 1 month before their exams results in an increase in their total exam score of 0.13 standard deviations. This study highlights the power of a movie role model as an alternative way to improve secondary school students' educational attainment, particularly of the worst performing students.

*The trial is registered at https://www.socialscienceregistry.org/trials/1832/history/14361 and the pre-analysis plan was uploaded there on 23rd February 2017 before endline data collection had finished and analysis begun. An amendment to the pre-analysis plan was uploaded there on the 18th July 2017 before additional data collection was finished and additional analysis begun. All analysis in this paper follows these pre-analysis plans unless clearly stated otherwise.

3.1 Introduction

There are high returns to investment in education in Uganda, with those finishing secondary education with passing grades set to earn 78% more than those with just primary education (Kavuma, 2014). Yet only 35% of those who enrol in secondary school remain in school 4 years later (MoES Uganda, 2015). Achievement is also low, with only half of male students who've completed a year of secondary school rated as proficient in mathematics and only one-third of girls. One reason why continuation and achievement in school might be so low is due to low beliefs about the personal returns to education.

Social learning theory argues that a person's beliefs are shaped by those they encounter around them (Bandura, 1977). These beliefs in turn affect their investment decision, including in the formation of human capital (Genicot and Ray, 2017, Jensen, 2010, Lybbert and Wydick, 2018). The poor, who are more likely to lack references of other people making successful investments in their future, may therefore become trapped in communities characterised by a low beliefs, low investment and low aspirations poverty trap (Genicot and Ray, 2017).

Role models can act as a powerful way to update beliefs about the returns to investments (Beaman et al., 2012, Bernard et al., 2014, Nguyen, 2008). However for a role model to be meaningful they must be from a similar background to the audience (Ray, 2006). Wilson (1987) argued that individuals form their beliefs about returns to education from individuals 'like them'. This is especially important where heterogeneity in returns may be present and so the example of an 'average' individual may not provide appropriate information (Nguyen, 2008). If an individual lacks role models in their immediate environment, a likely scenario if the successful individuals move away, a role model in the form of media can be used to exposure the individual instead (La Ferrara, 2016).

In this paper, I examine whether a movie featuring a potential role model can improve exam performance. I do this through the randomised exposure of 1,500 secondary school students in Kampala, Uganda to a treatment movie, *Queen of Katwe*, featuring a potential role model, versus a placebo movie. Students preparing

to take their national exams at the end of lower and upper secondary school were individually randomised to see the treatment or placebo movie between 1 week and 1 month before their exams. This design allows me to test the impact of the role model in the movie on academic performance in the short run.

A number of recent studies have shown that role models can affect economic behaviours (Beaman et al., 2012, Bernard et al., 2014, Chong and Ferrara, 2009, Jensen and Oster, 2009, La Ferrara et al., 2012, Nguyen, 2008). Role models can present information in the form of a story that is more salient and meaningful than information provision in the form of facts (Green et al., 2004). A role model may therefore lead to an updating of beliefs about what can be achieved by similar people (Nguyen, 2008), resulting in changes in behaviour. Exposure to role models has also been shown to lead to higher aspirations for the future (Beaman et al., 2012, Bernard et al., 2014). A role model might cause people to reassess and raise their goals, ambitions and effort level, breaking them out of an aspirations induced poverty trap (Dalton et al., 2016).

The treatment examined here, the movie *Queen of Katwe*, is based on the true story of a teenage girl from the slums of Kampala, Uganda striving to become a chess master through hard work and perseverance. Along the way she must overcome many obstacles to achieving her dream, including learning to read and write and getting into the top school in Uganda in order to play chess. She may therefore act as a role model to teenage students in Uganda, particularly female students and students who have done less well academically. Through watching her story, student may change their beliefs about the importance of education, having bold dreams and working hard to achieve your dreams in the fact of obstacles, leading to behavioural changes towards increased study effort. The use of a placebo movie, here *Miss Peregrine's Home for Peculiar Children*, allowed me to exclude any beneficial effect to exam performance from the novelty of going to the cinema and media exposure in general (Bernard et al., 2015).

The form of exposure in an entertaining movie could also magnify any impact of the role model. A movie allows the narrative of the role model to be presented in an

engaging and immersive way, causing the viewer to experience vicarious cognitive and emotional responses to the story as it unfolds (Green and Brock, 2000). Involvement with the characters and the storyline to allow the individual to feel 'transported' into the plot have been shown to be key determinants of the persuasive effects of edutainment programmes (La Ferrara, 2016, Moyer-Gusé, 2008). Video-based media has been shown to be particularly effective at tailoring information to individuals in a way that individuals relate to (Bernard et al., 2015). Exposure to a successful role model through a movie also allows a wider group of people to be exposed to a role model who may lack one in their immediate environment.

I find that amongst students taking qualifying exams for lower secondary school, seeing the treatment movie results in a 0.11 standard deviation improvement in maths scores. This is similar in magnitude to another study that looked at the impact of a role model on exam performance in a developing country (Nguyen, 2008) and equal in size to other types of interventions such as teacher incentives, school management practices or textbooks (Crawfurd, 2017, Duflo et al., 2012, Glewwe et al., 2004). Decomposing this improvement in maths score into the effect on the probability of achieving each grade (A-F), I find the entire effect is coming from a 30% decrease in the probability a student fails maths. I find no effect of the treatment movie on the probability of achieving any particular grade in other compulsory subjects than maths or on a student's aggregate exam score.

When examining these findings by pre-defined subgroups, I find that it's female students and students performing the worst prior to the exam who benefit most from treatment. Female students go from failing their maths exam 32% of the time to 18% of the time after viewing *Queen of Katwe*, a 44% decrease in the probability of failing. When looking at prior ability as measured by a mock exam taken in the summer before the study began, the entire benefit from seeing the treatment movie is experienced by low ability students. Students whose scores in the mock exam were below the mean increase their maths scores by 0.28 standard deviations when exposed to the treatment movie and decrease their probability of failing maths by 50%, from 54% to 27%. Students who scored above the mean in their mock

exam experience no effects on their maths scores or probability of failing maths from seeing the treatment movie. These results suggest role models could be a particularly effective intervention for the worst performing students.

Amongst students taking their finishing exams from upper secondary school, I find an overall improvement in their performance of 0.13 standard deviations. This effect is coming from improvements in their chosen subject papers. Again, it is women who benefit from seeing the treatment movie and see the largest improvement in their overall exam scores of 0.20 standard deviations. Students are also 6 percentage points more likely to get a place at public university, suggesting the treatment could have longer term beneficial effects on human capital acquisition beyond performance in a single, though highly important, exam.

I also perform a number of pieces of exploratory analysis that were not pre-specified in my pre-analysis plan but that could help indicate who benefits most from treatment and why. Firstly, I examine heterogeneity by school characteristics. I find it is lower ranked schools and schools charging lower fees where students taking the lower secondary school exam benefit the most from watching the *Queen of Katwe*. This suggests it is the worst students at poorly performing schools who benefit most from treatment. At the higher level, it is students at the best performing schools charging higher fees who get the most benefit from seeing the treatment movie. The opposing effects on different subgroups between lower and upper secondary could be due to selection of student's into upper secondary and the difference in time span between when students were treated and started their exams, factors which I explore in more detail in the Discussion section of this paper.

The second piece of exploratory analysis looked at whether students who saw the treatment movie in lower secondary were more likely to continue onto upper secondary school. I find indicative evidence that treatment led to students being nine percentage points more likely to enrol in upper secondary, with the effect particularly strong for female students. Again this may indicate that treatment not only improved an exam score but has longer term benefits for students' educational attainment.

Lastly, I examine persistent of the effects, by using natural variation in the number of days between exposure to the treatment movie and the student's exams. I can't reject that there is no difference in treatment effect regardless of whether a student had relatively early or late exams. This provides suggestive evidence that at least over the 1 month period during which upper secondary school students took exams the treatment effect persisted.

This paper contributes to a growing literature on the impact of media on economic behaviours (La Ferrara, 2016). The intervention used in this paper is closest to "edutainment" interventions, where information is presented through a narrative with an explicit policy change goal. Banerjee et al. (2018) use randomised screenings of a TV show called *Shuga*, produced by MTV with the aim of reducing risky sexual behaviour, encouraging testing of HIV and reducing stigma against those who are HIV positive, to assess the effectiveness and mechanisms behind edutainment. They find striking changes in behaviour and knowledge about HIV, and present evidence that the entertainment component of the intervention was a key reason for its impact. Likewise Paluck and Green (2009) randomize exposure to a soap opera in Rwanda and find effects on behaviours and social norms. This paper is unique, however, in using as the intervention a narrative with the aim of purely being entertaining to bring about a behavioural change, rather than one explicitly designed for a social purpose.

This paper shows that behavioural change is possible after a brief (2 hour) exposure to a role model, and impacts on exam outcomes are seen even as soon as 1 week after exposure. This complements work which has looked at the impact of brief media exposure to role models and found large behavioural change over time. Bernard et al. (2014), in Ethiopia, invited people to watch 15 minute documentaries about how people from similar backgrounds to them had improved their socio-economic position. Six months later, the treated group had higher aspirations and displayed behavioural changes: they saved more, took out more loans, and increased school enrolment of their children. My study takes this type of intervention into

a new setting, student educational attainment, and shows there are likewise real economic effects from role model exposure.

There is also non-experimental evidence from developing countries that exposure to the lives of alternative role models through TV, who rural individuals might not have encountered in their ordinary experience, can result in major shifts in behaviour. La Ferrara et al. (2012) and Chong and Ferrara (2009) show that exposure to soap operas, which include themes of women's empowerment and criticism of traditional family values, led to a reduction in fertility and an increase in divorce in Brazil. Likewise, Jensen and Oster (2009) show that the introduction of cable TV in India changed norms around the acceptability of domestic violence, increased women's autonomy and reduced fertility. In a developed country context, exposure to a TV show depicting teen pregnancy led to a reduction in the teen birth rate, with changes in attitudes leading to increased commitment to avoiding pregnancy (Kearney and Levine, 2015). My study adds to these by showing that the media exposure to the role model can even be as brief as a 2 hour movie and still lead to significant economic effects.

There have also been studies looking at the impact of exposure to local role models and the effects this has on education. Beaman et al. (2012) look at the effect of random exposure to female role models on village councils in India, finding that exposure closed the gender gap in aspirations, particularly for education and occupation-related aspirations. In Madagascar, Nguyen (2008) used a randomised experiment to compare giving information about schooling returns to exposure to a role model in the form of a former student from either a rich or poor background. She finds 0.17 standard deviation impacts on test scores from being exposed to a role model but only if the role model is from a similar poor background to the student. The effect is even larger for the poorest students, improving test scores by 0.27 standard deviations. This suggests role models can be a powerful tool, particularly for the poorest, by changing beliefs about both the returns to education and the probability of success. My study complements these by showing that

the role model does not have to be available in real life or from the viewers local community to have a positive effect on students.

There is extensive evidence from developed countries that exposure to role models change beliefs and improves aspirations, particularly among young adults. Stout et al. (2011) find that contact with same-sex role models in the form of advanced peers, professors and professionals in STEM (science, technology, engineering, and mathematics) subjects enhanced self-efficacy, attitudes towards and motivation to pursue STEM subjects and helped overcome negative stereotypes for women. Dasgupta and Asgari (2004) show the power of role models in overcoming stereotypes relating to academic achievement, and that exposure to role models can change beliefs about what is possible. Dennehy and Dasgupta (2017) show that female mentors increase female students' feelings of belonging in engineering, their retention and their aspirations for pursuing postgraduate engineering study. Male mentors didn't have these effects. Research has even shown the power of female role models to increase enrolment and majoring in Economics amongst undergraduates in the USA (Porter and Serra, 2017). My study adds to this large literature but in a developing country context, with a large sample size and on an important educational outcome.

This intervention also shows that substantial impacts can be had on exam performance even when the intervention is as short as 1 week before the exam. Over such a short time span, there is limited opportunity for increased study effort to affect exam performance and so effects relating to motivation during the exam are likely to dominate. The size of effect seen in this paper is of a similar magnitude to that seen in experiments which offer to pay students for performance immediately before an exam, thus removing all effects from increased studying and enabling only motivational effects. Levitt et al. (2016) found 0.12-0.22 standard deviation effects from paying students for performance. This effect is especially strong for maths, which has generally been found to be more elastic than other subjects, where students can improve their scores simply by trying harder and more persistently on a problem (Bettinger, 2012).

In terms of policies to improve performance in school in developing countries, this intervention was extremely costs effective, costing only $5 per student for a cinema screening and transport and so could easily be scaled up through screenings in schools. My findings therefore demonstrate that a low cost, one-off and brief exposure to a role model can have as powerful effects on education outcomes as larger and more complex programmes, such as teacher incentives, instructional materials or reducing class sizes (Evans and Popova, 2015). It also shows that costly materials designed to specifically affect certain behaviours are not needed to achieve the desired effect, suggesting wide potential to repackage existing materials for new aims.

The rest of this paper is organised as follows: Section 2 discusses the interventions and study design, section 3 the data used. Section 4 contains the empirical specification and results. Section 5 discusses the cost effectiveness and policy implications of the findings and section 6 concludes.

3.2 Intervention and Study Design

The study involved randomised exposure to either a treatment or a placebo intervention:

The treatment intervention involved a cinema screening of *Queen of Katwe*, the inspirational story of a young girl, Phiona Mutesi, from the slums of Kampala's rise out of poverty to become a world chess champion. The film is based on a true story.

The placebo intervention involved a cinema screening of *Miss Peregrine's Home for Peculiar Children*, a fantasy story about children with paranormal abilities.

3.2.1 Treatment movie

The (true) story of Phiona Mutesi is an inspiring rags-to-riches tale; Phiona goes from nothing, living in the slums and selling corn to passing drivers, to getting into the top school in Kampala, playing international-level chess and achieving her dreams.

The real life Phiona becomes one of the first two women in Uganda to become a titled chess player. The movie version of Phiona's story, *Queen of Katwe*, was produced by Disney and ESPN and directed by Mira Nair. It received widespread acclaim from critics[2], being both nominated for and winning multiple awards.

The movie begins with a quote form Ellen Johnson Sirleaf "The size of your dreams must always exceed your current capacity to achieve them". This idea of having bold goals and fighting to achieve them in the face of obstacles is the central theme of the story.

The movie sets up the story by showing the poverty and daily struggles of Phiona and her family to survive. Phiona only discovers chess after she approached a children's chess club because they were offering free food. When Phiona is concerned she does not belong at the club, after the other children make fun of her smell and tatty clothes, the club's coach tells her "Sometimes the place you're used to is not the place you belong. You belong where you believe you belong." Phiona returns to the chess club the next day.

One of the first things Phiona learns about chess is the idea that you can become bigger than you are "In chess, the small one can become the big one", meaning that even the lowest piece, a pawn, can become the most powerful, a Queen. The story then charts Phiona's own metaphorical rise from pawn to Queen.

The film uses chess as a metaphor for life: it doesn't matter how strong, intelligent or wealthy you are, you can learn to strategize your way to a better life. As their coach tells his class "Use your minds. Make a plan. There you will find safety." The concepts of sacrifice and winning and losing are repeated frequently throughout the film. Failing is shown as a key part of life, with their chess coach telling his class "Losses happen to everyone. But then you reset the pieces and play again". However the students are strongly encouraged to never give up, being told "Do not be quick to tip your king. You must never surrender." and "This is a place for fighters."

A key narrative of the film is the fact that Phiona can only play in chess tournaments if she can get into a top school with a chess programme. Phiona

[2]The movie was scored 73/100 by metacritic and 7.4/10 by IMBD in their aggregates of critics scores. The New York times scored the movie 90/100

never learnt to read and write, so she first has to learn how to read and write in order to go to school, spending hours a day studying with her mentor, and taking time away from playing chess. Central messages of the story are therefore that education can be a means of achieving other goals and that intelligence is not fixed but can be gained by learning.

In *Queen of Katwe* Phiona displays a number of different positive psychological behaviours throughout the movie. These include: perseverance and hard work; over-coming hardship; shaping her own life (Rotter, 1966); a growth mindset (Dweck, 2000) and belief that her lack of knowledge is from lack of education not lack of intelligence; goal setting; achieving long term goals through small incremental steps (Locke and Latham, 2002); and reaching out to others for help. If the students relate to Phiona strongly as a role model they may change their behaviour to imitate her and increase their academic success as a result (Lockwood and Kunda, 1997).

Phiona, has many characteristics which have been shown in a large psychology literature to make her a meaningful role model with whom secondary school students in Kampala could identify. Phiona is similar in multiple dimensions to many of the students in my sample, and hence relevant to them and easy for them to relate to (Lockwood and Kunda, 1997). However, certain characteristics are more of less meaningful to different groups of students, making her a more relevant role model to some students compared to others. This informs me about which subgroup of students and in which subjects I am likely to see the largest effects on attainment.

Phiona is the same age (a teenager) and from the same country and even city as the students in this study (Kampala, Uganda). The fact she is a woman is also important because research has shown that women require same-sex role models in order to identify with them whereas men have been shown to identify equally well with role models of either gender (Lockwood, 2006). This means that Phiona will act as a potential role model to both male and female students.

Additionally, Phiona is a counter-stereotype in that she is a woman doing well at what is traditionally a male dominated game, chess (Dasgupta and Asgari, 2004). Exposure to a counter-stereotype has been shown to change attitudes and

“inoculate" those exposed against applying stereotypes to themselves (Stout et al., 2011). Effects of exposure to Phiona's story might therefore be expected to be largest for female students in subjects they experience negative stereotypes about, the STEM subjects (science, technology, engineering and maths).

Lastly, Phiona begins from a position of extremely low educational attainment. She has dropped out of school and only qualified to sell vegetables on the side of the road. She rises from this low position to get into a top private school in Uganda so that she can pursue her love of chess. Phiona might therefore be a particularly relevant role model for student at the bottom of the ability distribution and show that academic ability need not be innate (Claro et al., 2016, Dweck, 2000).

3.2.2 Placebo movie

Going to the cinema is an affluent activity in Uganda, reserved for the middle classes for a special occasion. Most of the students in the study would have never been to the cinema before, or been very few times. The placebo movie was therefore important to remove any potential effects simply from going to the cinema. For example, the very act of going to the cinema may have made students want to do well academically so they could get good jobs and afford to go to the cinema! The placebo movie allows me to remove any effect from simply the activity of attending the cinema and instead ascribe any effects to seeing the treatment movie in particular.

The placebo movie was chosen carefully to be appealing to this age group. It was important the movie was entertaining and suitable for the students, containing characters of a similar age but without a Ugandan background. The content was purely an adventure story focused on overcoming monsters threatening the characters. There was no educational content.

3.2.3 Sample

Secondary schools were approached during September 2016 in the urban Kampala area. The outreach to schools was done by an NGO, the Initiative for Social and Economic Rights (ISER), that was connected to the study via the funder.

ISER approached 22 schools whom they had previously worked with and asked if their students sitting national exams in 2016 (the S4 and S6 classes) would like to participate in the study. There were no criteria for a school being recruited into the study except for being known to ISER, being within 1 hours drive of the cinema (in normal traffic) and consenting to provide student records and later exam data. 13 schools agreed to participate in the study.

The study was pitched to schools as looking at the impact of film on exam performance. Schools were not told that the study was looking at the *Queen of Katwe* movie in particular. Schools were given a list of 4 possible movies, including the treatment and placebo movie, so they could assess their suitability for their students to see, but not told which of them their students would be seeing. The students were unaware of which movies they would be seeing until they arrived at the cinema. Schools signed consent forms for the students to be transported to and attend the cinema, and agreed to provide student lists and exam results once they became available.

Schools were recruited until a sample size of approximately 1500 students was reached. The students were equally split between male and female and the S4 and S6 classes taking national exams. The schools provided their entire cohort of S4 or S6 students, such that the only untreated students in the year group were students who were absent from school on the day of the screening. Using the registrars of students enrolled for the national exams at each school, I confirmed that schools did indeed provide their entire S4 and S6 classes for the screenings and that at most 2-3 students were missing from a given class.

Consenting schools were allocated to one of five consecutive screening days in the second week of October, and either a morning or afternoon session. This was based on their geographical proximity to each other, the number of students at the school and the capacity of the cinema screens. Schools with less than 100 students were combined into a screening session with another school nearby. The cinema had 3 screens which could be use for screening the movies, two screens of 100 person capacity and one screen of 300 person capacity. If there were less than 200 student

attending the screening the two small screens would be used, if between 200 and 300 students one screen of 100 and one of 300 would be used and for more than 300 students both screens of 100 and the 300 person screen would be used.

3.2.4 Randomisation

The movie screenings began on the day that both *Queen of Katwe* and *Miss Peregrine's Home for Peculiar Children* were released in Uganda, Friday 7th October 2016. Two sessions, each screening both movies, were run per day, one at 11am and one at 2pm, for 5 days, finishing on Tuesday 11th October. The chosen cinema was one of two multi-screen cinemas in Kampala which allowed us space to conduct a randomisation and complete control over the movie schedule and times.

The students were collected by mini vans hired for the study, which arrived at the cinema 1 hour before the screening to allow time for the randomisation. Students were individually randomised into the treatment or placebo movie upon arrival at the cinema for a screening. This was done by students lining up outside the cinema and one by one entering, upon which an assistant picked a ticket out of a bag without looking and handed it to the student. The bag was opaque and the tickets identical except for the name of the movie printed in small print at the bottom of the ticket. An assistant was chosen to actually pick the ticket to further reduce any probability that a student might try and pick a particular ticket.

Immediately after getting a ticket, students were steered to the designated registration desk for that movie, where their ticket was checked and they registered their name, school, age and gender before proceeding into the theatre. These registration lists were later combined with lists from the schools of student index numbers, which uniquely identify student exam results. Once a ticket had been selected, students with tickets for different movies were kept separate the entire time, even using different bathrooms. I am therefore confident that all students saw the movie they were assigned to. The students also had between 2 and 5 teachers accompanying them depending on the class size. These teachers were

split between the theatres randomly such that half the teachers attended each movie and could supervise their students.

Due to the difference in the sizes of the cinema screens, students within individual schools did not have an equal probability of seeing the treatment and placebo movie. For example, if a school had 250 students then 150 would have to see one movie and 100 the other. This was randomised and balanced over different sessions so that overall we issued 794 treatment movie tickets and 706 placebo movie tickets to students in classes taking national exams. School fixed effects will be used to control for this difference in treatment probability within a school.

Tables 3.1 and 3.2 show balance tests by class for the individual and exam choice characteristics collected during the intervention and from the schools. No significant differences are found between the samples. Looking at Table 3.1, students in the S4 class were on average just over 17 years old, half of them were female and most were taking 10 subjects in the exams. The standardised mock score was approximately zero in both the treatment and control groups, as would be expected from a standardised score, and not different between them. At S6 level, Table 3.2 shows that students are now two years older, at 19 years old on average, half are female and one third are taking maths or science as an optional paper. Again, the standardised mock scores were approximately zero and not significantly different between the treatment and control groups.

Attrition occurred in the form of students not taking the national exam. Since I had the students' exam index numbers I could always obtain exam results if they existed. Missing results meant either that the index number obtained for that student was incorrect or that the student didn't take the exam. All cases of no results for an index number were double checked with the school, with remaining cases due to students not taking the exam. Attrition was balanced across the treatment and control groups, as shown in Table 3.3 below. 21 students in the placebo and 33 in the treatment group did not take their national exams, 3.6% of the sample.

Attrition varied greatly by school, with some of the schools in particular having very few candidates at S6 level taking the exams and many of these students deciding

to not actually take the exam. I examined whether student or school characteristics were correlated with attrition in Table 3.4. Students at Christian schools are more likely to take the exam, as are older students and students in the S4 class.

Table 3.1: Balance test S4 class

	Placebo		Treatment			
	mean	sd	mean	sd	difference	p-value
Age	17.28	1.25	17.25	1.23	0.03	(0.76)
Female	0.51	0.50	0.51	0.50	0.00	(0.61)
Number of subjects	9.73	0.62	9.68	0.60	0.04	(0.34)
Mock total score	0.01	0.98	-0.01	1.01	0.02	(0.74)
Observations	344		391		735	

Age refers to age in years, Number of subjects is the number of subjects the student had been entered for exams in. Mock total score is the standardised score achieved in the mock exam taken prior treatment.

Table 3.2: Balance test S6 class

	Placebo		Treatment			
	mean	sd	mean	sd	difference	p-value
Age	19.09	1.24	19.00	1.13	0.09	(0.31)
Female	0.47	0.50	0.50	0.50	-0.03	(0.40)
STEM	0.33	0.47	0.30	0.46	0.02	(0.53)
Mock total score	-0.02	0.97	0.04	1.01	-0.06	(0.45)
Observations	341		370		711	

Age refers to age in years, STEM is a dummy if the student is taking maths, biology, chemistry or physics as one of their subject choices. Mock total score refers to the standardised test score in the mock exam taken prior to treatment.

Table 3.3: Attrition Balance Test

	Placebo		Treatment			
	mean	sd	mean	sd	difference	p-value
Attrition rate	0.03	0.17	0.04	0.20	-0.01	(0.22)
Observations	706		794		1500	

Differences in mean attrition between placebo and treatment. Attrition means the students didn't take their national exams.

Table 3.4: Individual and school characteristics correlated with attrition

	(1) Attrition
Boarding	0.01
	(0.01)
High fees	-0.02*
	(0.01)
Christian	-0.05***
	(0.01)
Age	-0.02***
	(0.01)
Female	-0.01
	(0.01)
S4	-0.02*
	(0.01)
Observations	1,498
R-squared	0.05

This table shows attrition correlates with 3 school characteristics and 3 student characteristics. Boarding refers to whether the school only has boarding pupils, high fees if the fees charged are above the median in this sample, Christian is the schools religious affiliation. Age is the age in years, S4 is a dummy if that student is in the S4 class.

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

3.3 Data

3.3.1 Student data

Limited information about the students was collected upon registration at the cinema. This was their name, age, gender, class (S4 or S6) and school, along with which movie they saw. This data was combined with lists of exam enrolment provided by each school. The exam enrolment information listed the student's name, age, index number and subjects entered for exams. The index number was particularly important as it is a unique identifier of a student's exam results. The registration and school exam enrolment data were combined using double data

entry with any discrepancies checked. This resulted in a data set of 1,500 students who saw a movie and were due to take a national exam.

Mock exam results from a practice for the national exam administered by the schools in the summer before the study began were also obtained for all students in the study. This data is described in the test score data section.

3.3.2 School data

Information was collected about the schools at the same time as exam results data was collected. This involved asking the schools about their religious affiliation, fees and whether they offered boarding. Publicly available rankings of the schools were also collected. This information is shown in the Appendix in Table 3.21.

Schools all had a strong religious affiliation as either Christian or Islamic. This determined the kind of prayers they took part in during the school day and was a strong part of the school's character. Nearly all the schools had some boarding element, with only one, Kulumba, having none. This is very common in secondary schools in Uganda, where students might come from very far away and transport is difficult. Note, all the schools were private schools, as two-third of secondary schools in Uganda are (MoES Uganda, 2015).

National rankings of the schools were obtained from the Ministry of Education. At the advanced examination taken by S6 students, UACE, there are 1,900 schools ranked, of which the schools in my sample varied markedly, from near the top to one of the very worst. Likewise, at the UCE level taken by S4 students, the schools are also spread out in the ranking out of 3,300, though not so dispersed as for UACE. The fees the schools charge for boarding and day students in the S4 and S6 classes were also collected from schools and display a wide spread, with the most expensive boarding school 900,000 USH a year, or $257 at current exchange rates, while the least expensive is half that at only 440,000 USH or $125 a year. The fees can act as a very rough proxy for the income level of the students' parents.

The schools also differed in which classes they provided to attend the cinema. I offered to take both the S4 and S6 classes if the school wanted. For mainly

timetabling reasons and staff constraints, some schools only offered one class. The schools are also dramatically different sizes, with the largest having nearly 200 S6 students and the smallest only five.

3.3.3 Test score data

The primary outcomes in this paper are standardised exam scores on national exams. Secondary school students sit their national exams in October and November and the results are released in January and February of the following year. Ordinary exams are taken after 4 years of secondary school by the S4 class, the Uganda Certificate of Education (UCE), and began on the 19th October 2016, 1 week after the last movie screening. Advanced level exams are taken after a further 2 years of study by the S6 class, the Uganda Advanced Certificate of Education (UACE), and began on the 14th November 2016, 1 month after the last screening.

The exams sat by the students had already been chosen and registered for well before the intervention occurred and so neither the subject choices nor the number of subjects could be changed as a result of the intervention. They are pre-determined with respect to treatment.

Data on national exam results was collected in February 2017 once the exam data of both the S4 and S6 classes had been released. The data was collected in two ways. Initially, results were collected directly from schools in the form of printouts of all the student's results provided by the exam board. These were double entered into Excel. In the case of a few schools not wanting to provide us with the exam results of their students[3], an SMS exam results collection system was used. The Ugandan National Exam Board allows you to text in an index number to obtain results for that student. Results obtained in this way provide an equal amount of information as those provided by the schools. Results were collected via the text interface for all remaining students that results were missing for and

[3]All the schools signed consent forms agreeing to provide exam results as part of being in the research study. Some schools, particularly those with poor results, later changed their minds about providing us with copies of results. However they were all aware and informed that since we had the index numbers of the students we could obtain the results directly from the exam board.

entered into Excel. A random sample of results obtained via text-messaging were audited to ensure they had been entered correctly.

Mock exam results data was collected for all the students in the study. This data was provided by the schools. All students sit a mock exam during the summer before their national exam. This corresponded to August 2016 for the study sample, two months before treatment took place and one month before schools were approached about taking part in the study. This mock exam is administered by the schools and based on previous exams. Students in the S4 class sit mock exams in English and Maths only. Students in the S6 class sit the mock exam in the principal and subsidiary subjects they are registered for in the national exam. Schools were requested to provide the complete subject-by-subject mock results. However, some schools only provided the aggregate score across all subjects. The mock exam results will be used here as a baseline test score.

Standard 4 Exam

After 4 years of secondary education candidates take the UCE exam. The UCE comprises six mandatory subjects administered in English; these are mathematics, English language, biology, chemistry, physics, and a choice of either geography, or history. Two other optional subjects are also chosen from subjects such as music and business. Candidates must register for a minimum of 8 and a maximum of 10 subjects. The exams are graded with a score from 1-9 with 1 being the best score and 9 the worst. Passing grades are considered to be an 8 or lower. For a candidate sitting 10 exams, the best score is therefore 10 and the worst 90.

For this analysis scores have been inverted so that a 9 becomes 0 and a 1 becomes 8. This is so that a higher score can be interpreted as a better performance, while a higher aggregate score can indicate better performance per paper or more papers taken.

Standardized test scores have been created for each subject by subtracting the mean and dividing by the standard deviation of the control group. An overall aggregate of exam performance was calculated by summing standardised test scores

across all subjects and renormalising. A core index of exam performance was calculated by summing test scores across the six core subjects and renormalising.

For students taking UCE exams the following outcomes are examined:

1. **Exam score aggregate**: aggregate score composed of exam scores across all eight-ten subjects taken by a student

2. **Core exam score**: composed of exam score in the six mandatory subjects taken by all students

3. **Individual subject grade**: Standardised score achieved in maths and English subjects

Effects of treatment are more likely to be expected on subjects related to chess, of which maths has the clearest link. I therefore examine the particular effect of the treatment on the maths exam outcome, and also look at English individually since it is a common outcome examined in the education literature.

Standard 6 Exam

In their final year of secondary school, students sit the UACE exam. The UACE is taken in five subjects, three of which are from a list of principle subjects, one in a subsidiary subject out of mathematics or computer and one in a general paper. The subsidiary subjects and general paper are graded on a 1-9 scale, with 1 being the best and 9 the worst grade. Grades 7 and above are fails. Any student achieving a 6 or below on a subsidiary paper or the general paper gets one point. The principal papers are marked on a A, B, C scale, with an A earning 6 points, a B 5 points etc. The maximum of 2 points earned on the subsidiary and general paper are added to the points earned on the principal papers. This means the highest total score a subject could earn is three As and passes on the subsidiary and general paper, giving 20 points.

Standardised test scores were constructed for each subject by subtracting the mean and dividing by the standard deviation of the control group. An overall

index of exam performance was calculated by summing test scores across all subjects and renormalising.

For students taking UACE exams, the following outcomes were examined:

1. **Total exam score**: aggregate exam score composed of exam scores across all principal and subsidiary subjects taken by a student, with subsidiary subjects scoring a maximum of 1 point.

2. **Principal score**: aggregate score in the principal papers only.

3. **General paper and subsidiary paper score**: standardised score on the general paper and subsidiary paper in maths or computer taken by all students. This will be an inverted scale of the 1-9 score on these papers.

An additional outcome examined is a dummy variable for whether a student achieves the grades to get into public university. Public University in Uganda is the best type of tertiary education and the grades required are set nationally. The requirement is passing grades in two principal subjects, where a pass is any score greater than 0. I therefore construct a dummy variable equal to one if a student got at least two passes in their principal subjects and zero otherwise. Note that this outcome was not pre-specified in the analysis plan as I was not aware of the common grade requirement for university entrance at this time.

3.3.4 University place

In an amendment to the original pre-analysis plan two further outcomes were specified before data was collected on them:

1. An indicator for whether the student obtained a government scholarship

2. An indicator for whether the student gained entry to Public University

These outcomes were obtained from the Ugandan National Council of Higher Education who hold records on all public University entry and determine scholarship awards. These records are publicly available and include identifiable information

for the students, such as name, id number and school, which I used to match this data to my study sample.

3.4 Empirical Strategy and Results

3.4.1 Empirical strategy

To examine the effect of the treatment on exam outcomes, I run the following regression:

$$y_{is1} = \beta_0 + \beta_1 \text{QofK} + y_{is0} + \boldsymbol{x_i'} \cdot \gamma + \theta_s + \epsilon_{is}, \tag{3.1}$$

where i indexes student at school s, y_{is1} denotes the exam outcome of interest, QofK is an indicator variable equal to one for if the student saw the movie *Queen of Katwe*, $\boldsymbol{x_i'}$ is a vector of individual characteristics, θ_s is a vector of school fixed effects and ϵ_{is} is a random error.

y_{is0} is the standardised mock exam result from before treatment. If available, the mock result in the specific subject outcome will be controlled for. If the equivalent mock result is not available for an outcome, the aggregate result constructed from the available mock papers will be controlled for instead.

Specification 3.1 is the basic specification used here, as set out in the pre-analysis plan. Any departures from the contents of that plan will be clearly stated.

The parameter of interest is β_1, the average treatment effects of the *Queen of Katwe* movie on an exam outcome. The school fixed effects, θ_s, are included to account for differential treatment probability depending on which movie was played on the larger cinema screen. They also control for substantial school heterogeneity (and so improve precision), as seen in Table 3.21 in the large dispersion of rankings of the schools. Robust standard errors are calculated to allow for heteroskedasticity.

Individual characteristics, $\boldsymbol{x_i'}$, are included to improve precision. These are:

1. dummy for whether the student is female

2. the age of the student in years

3. the number of subjects taken (for S4 students)

4. whether the student choose to take any subjects in maths or science (STEM subjects) at S6 level

3.4.2 Main Results

S4 Class

Table 3.5 shows the impact of assignment to see the treatment movie on the S4 exam outcomes defined in section 3.3.3. I show results both with and without individual control variables. Treatment assignment has no effect on the total score, core score or English standardised scores. However, treatment does result in an increase of 0.11 standard deviations in maths score, significant at the 5% level when controls are included, and 0.14 standard deviations still significant at the 5% level without any controls. This is a large positive effect on the maths exam outcome, and is examined in more detail below.

Mock exam performance is a strong predictor of national exam score, with each additional standard deviation scored in the mock associated with a 0.99 standard deviation increase in total score. I will examine heterogeneity by mock exam performance later.

To examine the effect of treatment on the maths exam performance in more detail, I break down the maths exam into dummies by grade achieved. As mentioned, the exam is graded from 0, fail, to 8, the maximum result. I look at the impact of treatment on a dummy for obtaining each grade in Table 3.6. In column (1) it can clearly be seen that seeing the treatment movie reduces the probability that a student obtains the bottom, failing, grade in maths by 11 percentage points from a control group mean of 27%. This is a 40% decrease in the probability of failing maths. Seeing the treatment movie increases the probability by 5 percentage points a student scores 2 or 3 on the maths test, suggesting that seeing the treatment movie might be pushing students to the next couple of grades above what they would have achieved, though this is only significant for grade 3 at the 10% level. No impact is seen for higher scores, and in fact less than 1% of students achieve the

Table 3.5: Impact of treatment assignment on S4 standardized test scores

	(1) Total score	(2) Total score	(3) Core score	(4) Core score	(5) Maths	(6) Maths	(7) Eng.	(8) Eng.
Treatment	0.01	-0.02	0.00	-0.03	0.14**	0.11**	-0.04	-0.06
	(0.07)	(0.03)	(0.07)	(0.03)	(0.07)	(0.05)	(0.07)	(0.05)
Age		0.00		-0.00		-0.03		-0.03*
		(0.01)		(0.01)		(0.02)		(0.02)
Female		-0.01		-0.01		-0.03		0.10**
		(0.03)		(0.03)		(0.05)		(0.05)
No. subjects		-0.06**		-0.08**		-0.16***		-0.05
		(0.03)		(0.03)		(0.05)		(0.05)
Mock score		0.99***		0.95***		0.80***		0.77***
		(0.02)		(0.02)		(0.03)		(0.03)
Constant	-0.67***	0.32	-0.66***	0.61	-0.35***	1.70***	-0.53***	0.60
	(0.09)	(0.36)	(0.12)	(0.43)	(0.13)	(0.65)	(0.12)	(0.60)
Observations	735	729	735	729	735	729	735	729
R-squared	0.31	0.88	0.29	0.82	0.17	0.57	0.25	0.62

Total score refers to standardised aggregate score across all subjects taken in the exam. Core score refers to standardised aggregate score in the 6 mandatory subjects at S4 level. Standardized test scores composed of subject standardized scores and renormalised. All regressions include school fixed effects.
Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

top grade at all in this sample. I find no effect of treatment on the probability of failing any other core subject (see Robustness section, Table 3.20).

I also look at the effect of treatment on the probability of getting each maths grade using an ordered logit regression to improve power. Since the largest number of students getting a particular grade is students failing (27%), I have most power to detect an effect here. An ordered logit will allow me more power to see effects elsewhere in the grade distribution. The marginal effect of treatment on each grade are shown in Table 3.7. Here the coefficients shown give the probability of achieving each grade level as the treatment indicator goes from 0 to 1. Seeing the treatment movie reduces the probability of achieving the lowest 3 scores, particularly the probability of getting the failing score declines by 5% percentage points. The

Table 3.6: Impact of treatment on probability of getting each maths grade

	(1) Fail	(2) 1	(3) 2	(4) 3	(5) 4	(6) 5	(7) 6	(8) 7	(9) Top
Treatment	-0.11***	0.05	0.05*	-0.01	0.01	0.01	0.00	-0.00	0.01
	(0.03)	(0.03)	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)
Age	0.01	0.01	-0.02**	0.00	-0.00	0.00	-0.01	-0.00	0.00
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)
Female	0.03	-0.05*	0.02	0.02	0.02	-0.04*	-0.00	0.01	-0.00
	(0.03)	(0.03)	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)
No. Subjects	0.01	0.07**	0.01	-0.02	-0.04	-0.01	-0.02	-0.02	0.01*
	(0.03)	(0.03)	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)	(0.01)	(0.00)
Mock score	-0.17***	-0.11***	-0.04***	0.02	0.05***	0.09***	0.09***	0.06***	0.03***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)
Constant	-0.04	-0.67*	0.50	0.33	0.46*	0.05	0.29	0.16	-0.07
	(0.35)	(0.38)	(0.33)	(0.33)	(0.27)	(0.27)	(0.21)	(0.15)	(0.05)
Mean in control	0.27	0.17	0.13	0.17	0.10	0.08	0.05	0.03	0.01
Obs	729	729	729	729	729	729	729	729	729
R-squared	0.31	0.09	0.02	0.06	0.06	0.10	0.10	0.08	0.05

Maths papers are graded from fail (0) to highest grade (8). All regressions include school fixed effects. Mean in control is the mean proportion of the control group obtaining that grade. Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

treatment movie also increases the probability of obtaining higher grades, with the effect significant at at least the 10% level in all cases, though the magnitudes are small. The largest effect is seen on grades 6 and 7 where seeing the treatment movie increases the probability of obtaining that grade by 1.4 percentage points. I also do the same ordered logit for the other subjects in the core exams and find no effect of seeing the treatment movie on the probability of getting any particular grade.

To further understand where on the grade distribution the treatment effect is I plotted histograms by subject. Plots of the entire distribution of results for both treated and placebo students are shown in Figure 3.1. The histograms of total score, core score and English show no statistically significant impact of treatment in the distribution. To formally test this I perform a Kolmogorav test. For total score, core score and English the p-value on the test of equality of the distributions are 0.25, 0.25 and 0.28 respectively. Hence I cannot reject equality of the distributions. However, in the histogram of maths score it can be seen that the histogram is shifted to the right, particularly at the lower end to just above the mean. Now the p-value for the Kolmogorav test is 0.008, so I can reject equality of the distributions at the

Table 3.7: Ordered logit regression of the impact of treatment on maths grade at S4

Grade	treatment
Fail	-0.048***
	(0.015)
1	-0.013***
	(0.004)
2	-0.000
	(0.001)
3	0.009***
	(0.003)
4	0.010***
	(0.003)
5	0.014***
	(0.005)
6	0.014***
	(0.005)
7	0.009***
	(0.003)
8	0.004***
	(0.002)
Observations	729

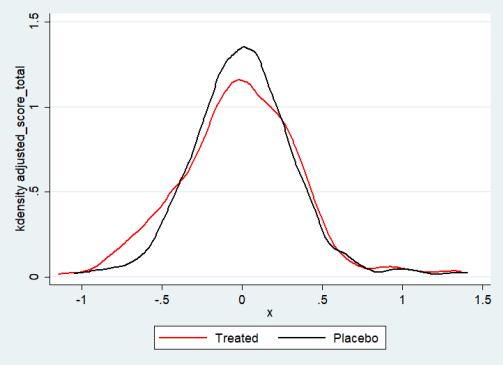
Maths papers are graded from fail (0) to highest grade (8). Regressions include school fixed effects and individual control variables (age, gender, number of subjects taken and standardised mock score). Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

1% significance level and confirm that treated students achieve higher maths scores.

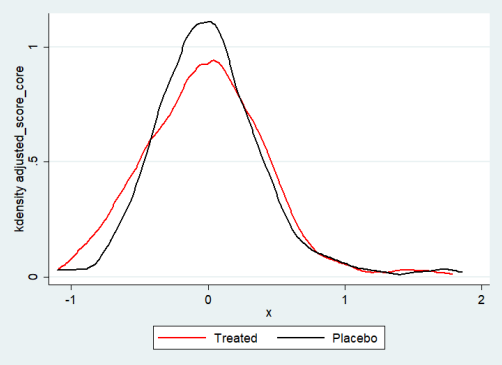
It is also possible that any improvement in maths scores and reduction in failing maths came at the expense of other subjects. This could be the case if effort was directed away from other subjects and towards maths rather than increased overall. While I see no significant negative on English or total or core scores, it is possible there are small decreases in other subjects across the board that can't be seen when aggregated with the maths improvement. I test this by looking at the impact of treatment on the aggregate score excluding maths. Results for this are shown in the Appendix in Table 3.23. I find no significant effect of treatment on the total exam score excluding maths. Though the treatment coefficient is negative it is

Figure 3.1: Histograms of S4 student results by treatment assignment

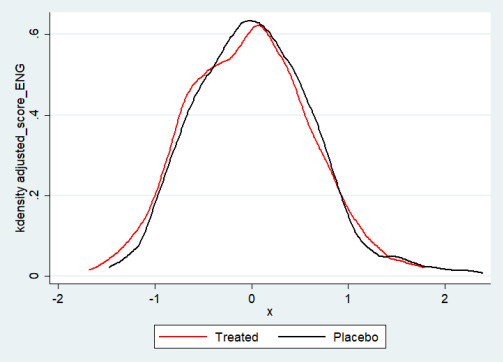
(a) Total score



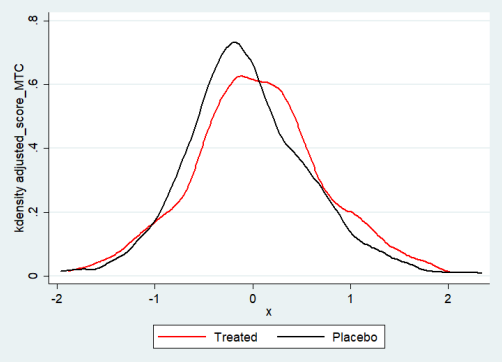
(b) Core score



(c) English



(d) Maths



small, at negative 3 percentage points with controls. This is suggestive that any improvement in maths is not coming from a reduction in effort in other subjects.

S6 class

Table 3.8 shows the impact of assignment to see the treatment movie on the S6 exam outcomes defined in section 3.3.3. Results are shown both without and with individual control variables, but I will discuss only the results with control variables for brevity. Seeing the treatment movie results in the overall exam score being 0.13 standard deviations higher, a large effect on an education outcome. This is significant at the 1% level. The score on the principal exam papers increases by 0.13 standard deviations, significant at the 5% level. There is no effect on the score achieved on the general and subsidiary papers.

I include the pre-specified control variables; age in years, a female dummy and a dummy for if at least 1 subject out of maths, biology, physics or chemistry were taken (STEM) and the baseline mock score. Students taking STEM subjects do significantly better,particularly on the subsidiary paper which is more science focused. A 1 standard deviation better performance on the mock is associated with a 0.76 standard deviation better performance on the overall and principal papers, but only a 0.28 standard deviation improvement on the subsidiary subjects.

I also look at the effect of treatment assignment on different parts of the results distribution. I do this by looking at the impact of treatment on a dummy for each decile of overall score and by examining histograms of adjusted scores by treatment assignment. Firstly, the impact on decline of overall score, shown in Table 3.9. I include control variables but the results do not change without them. Treatment has no effect on the probability that a student's final score is in a particular decile except for the very top decile shown in column (9). Treatment results in an increase of 3 percentage points in the probability the student scores in the top decile, though this is only significant at the 10% level.

Again, I plot histograms by these three outcomes to see where the treatment effect is shifting the distribution. Plots of the entire distribution of results for both treated and placebo students are shown in Figure 3.2. The histograms all show a shift to the right for students assigned to treatment around the middle of the distribution. There is also an effect at the top of the distribution for principal

Table 3.8: Impact of treatment on S6 standardized test scores

	(1) Overall score	(2) Overall score	(3) Principal subjects	(4) Principal subjects	(5) Subsidiary subjects	(6) Subsidiary subjects
Treatment	0.17**	0.13***	0.17**	0.13**	0.08	0.07
	(0.07)	(0.05)	(0.08)	(0.05)	(0.06)	(0.06)
Age		-0.03		-0.02		-0.04
		(0.02)		(0.02)		(0.03)
Female		0.05		0.10*		-0.09
		(0.05)		(0.05)		(0.06)
STEM		0.55***		0.40***		0.79***
		(0.06)		(0.06)		(0.07)
Mock score		0.76***		0.77***		0.28***
		(0.03)		(0.03)		(0.03)
Constant	-1.50***	-0.13	-1.48***	-0.25	-0.36	0.69
	(0.42)	(0.43)	(0.41)	(0.43)	(0.35)	(0.57)
Observations	711	708	711	708	711	708
R-squared	0.20	0.62	0.15	0.59	0.29	0.44

Overall score refers to the aggregate score in the principal and subsidiary papers. Principal subjects refers to the standardised score on the 3 chosen subject papers. Subsidiary subjects refers to the standardised score on the two mandatory subsidiary papers. Standardized test scores composed of subject standardized scores and renormalised. All regressions include school fixed effects.
Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

subjects, with the top tail of high scores extending further for treated students. This supports the results found in the decile analysis that its the top students who benefit most from seeing the treatment movie and perhaps indicates there are also some positive effects around the mean of the distribution that I am not powered to detect.

I again look at the Kolmogorov test for equality of the distributions. For the overall score, the p-value is 0.046 so I reject equality. The treatment shifts the distribution to the right. For principal papers, the Kolmogorov p-value is 0.073, so I can just reject equality at the 10% level. For the subsidiary paper I get a p-value of only 0.041, so I can also reject equality of the distributions here. This suggests that while I cannot detect differences in the mean subsidiary outcome by treatment status there may be differences elsewhere in the distribution, suggesting heterogeneity is important here.

Table 3.9: Deciles of overall score at S6

	(1) decile 1	(2) decile 2	(3) decile 3	(4) decile 4	(5) decile 6	(6) decile 7	(7) decile 8	(8) decile 9	(9) decile 10
Treatment	-0.02	-0.01	-0.02	0.03	-0.00	-0.00	0.01	-0.01	0.03*
	(0.02)	(0.02)	(0.02)	(0.03)	(0.02)	(0.03)	(0.02)	(0.02)	(0.02)
Age	0.01	-0.01	0.00	0.01	0.01	-0.01	0.01	-0.01	-0.00
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Female	-0.01	0.00	-0.04	0.04	-0.01	-0.00	-0.02	0.03	0.01
	(0.02)	(0.02)	(0.02)	(0.03)	(0.02)	(0.03)	(0.02)	(0.02)	(0.02)
STEM	0.11***	-0.01	-0.00	-0.03	-0.04	0.05	-0.05**	-0.03	-0.01
	(0.03)	(0.03)	(0.03)	(0.04)	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)
Mock score	-0.13***	-0.08***	-0.05***	-0.05***	0.01	0.08***	0.04***	0.07***	0.11***
	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)
Constant	0.35	0.11	-0.07	-0.16	0.13	0.34	-0.17	0.30	0.16
	(0.23)	(0.22)	(0.22)	(0.31)	(0.24)	(0.28)	(0.19)	(0.21)	(0.18)
Control mean	0.15	0.10	0.10	0.18	0.10	0.16	0.07	0.09	0.06
Observations	708	708	708	708	708	708	708	708	708
R-squared	0.25	0.07	0.04	0.04	0.02	0.06	0.05	0.13	0.24

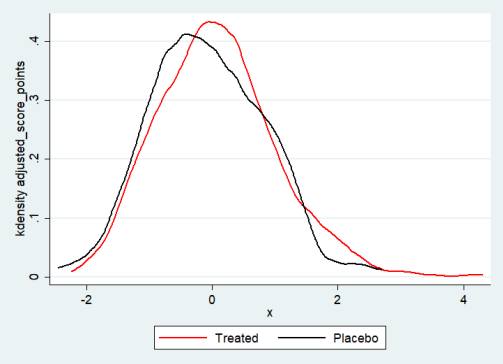
Decile 1 is worst and decile 10 best. Overall score refers to the aggregate score in the principal and subsidiary papers. STEM refers to taking a principal paper in maths or science. Robust standard errors in parentheses. Regressions include school fixed effects.
*** p<0.01, ** p<0.05, * p<0.1

In Table 3.10 I report whether the student achieved the required 2 passes in principal papers to get into public university and whether they obtained a place at public university or not [4]. Students who saw the treatment movie were 4 percentage points more likely to get the necessary grades and 6 percentage points more likely to get a place at University. However both these results are only just significant at the 10% level and so should only be taken as indicative. This is from a mean of 79% getting the required grades to get into university and 31% being offered a place at University in the control group. Seeing the treatment movie therefore increases by 20% the odds that a student will get a place at University. This result shows that not only has seeing the treatment movie improved exam outcomes but that there will be long term effects from students seeing the treatment movie being more likely to get a place at university.

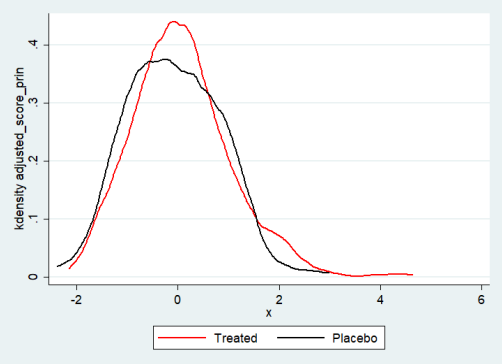
[4]Whether a student obtained the grades to get into University was not included in the original pre-analysis plan. Whether a student obtained a place at University was included in the pre-analysis plan amendment. I pre-specified whether seeing the treatment movie increase the odds that a student obtained a government scholarship. However I find no effects here on the tiny sample of 16 students who obtained scholarships from my study sample, and so I do not include the results.

Figure 3.2: Histograms of S6 student results by treatment assignment

(a) Total score



(b) Principal papers



(c) Subsidiary/generate papers

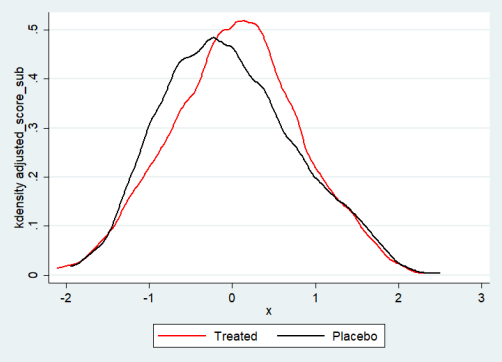


Table 3.10: Impact of treatment on probability obtain scores to get into public university

	(1) University passing grade	(2) University passing grade	(3) Place at University	(4) Place at University
Treatment	0.06*	0.04*	0.06*	0.06*
	(0.03)	(0.02)	(0.03)	(0.03)
Age		-0.01		0.01
		(0.01)		(0.02)
Female		-0.01		-0.01
		(0.03)		(0.04)
STEM		-0.24***		-0.08*
		(0.03)		(0.04)
Mock score		0.15***		0.10***
		(0.01)		(0.02)
Constant	0.37	0.82***	0.16	0.18
	(0.22)	(0.29)	(0.20)	(0.34)
Mean in control	0.79		0.31	
Observations	711	708	711	708
R-squared	0.05	0.32	0.10	0.15

Regressions include school fixed effects. University passing grade refers to the minimum two principal passes to get into public University - it was not pre-specified in a pre-analysis plan. Place at University refers to obtaining a space at a public University - it was pre-specified in a pre-analysis plan. Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

3.4.3 Heterogeneity

Heterogeneous treatment effects across variables collected at treatment assignment are tested by augmenting equation 3.1 to include the variable and the interaction between treatment and that variable. This gives the following specification:

$$\begin{aligned} y_{is1} =& \beta_0 + \beta_1 \text{QofK}_i + \boldsymbol{x_i'} \cdot \gamma + y_{is0} \\ & (\boldsymbol{x_i'} \cdot \text{QofK}) \cdot \lambda + \theta_s + \epsilon_{is}, \end{aligned} \tag{3.2}$$

where i indexes student in school s, y_{is1} denotes the exam outcome of interest, y_{is0} is the standardised mock value of the same outcome variable[5], QofK is an indicator variable equal to one for if the student saw the movie *Queen of Katwe*, $\boldsymbol{x_i'}$ is a vector of student characteristics, θ_s is a vector of school fixed effects and ϵ_{is} is a random error. The parameter of interest here is λ, the heterogeneous treatment effect of seeing *Queen of Katwe*.

Individual characteristics

The individual student variables I examine are:

1. An indicator equal to one if the respondent is female
2. An indicator variable equal to one if respondent's age is above the sample median for that grade.
3. An indicator variable equal to one if the respondent is taking fewer subjects than the median for that grade (at S4 level)
4. An indicator variable equal to one if the student is taking at least one principal science subject (maths, physics, biology and chemistry) at S6 level
5. An indicator equal to one if the student was below the median exam performance in their mock exams.
6. Indicators of decile of exam performance in the mock exam

[5]if provided by the school, if not available the standardised mock total score will be used instead.

The heterogeneity by gender would reveal whether the treatment movie, featuring a female protagonist, appealed more to women or equally to each gender, as research from psychology suggests it might (Lockwood, 2006). The heterogeneity by age would pick up whether older students, who might have been held back years or had to postpone their studies for a while, perform less well as a result but benefit more from seeing the treatment movie featuring a girl who had stopped school but returns to it. Reports from those familiar with the Ugandan education system suggests students who are struggling often take fewer subjects at UCE in order to trade off quality versus quantity. Looking at heterogeneity by students taking fewer subjects than the median would reveal if the weakest students benefited most from the movie.

At UACE students are encouraged to commit to a sciences or humanities subject track. Since the treatment movie was most concerned with what is regarded as a scientific game, looking at heterogeneity would reveal if the treatment movie had greater effects on students taking more scientific subjects who might more closely relate to the protagonist. Lastly heterogeneity by mock exam performance will reveal whether students at the bottom or top of the ability distribution benefited more from treatment, with the expectation being that those most similar to Phiona, those at the bottom of the distribution, would benefit the most from seeing the treatment movie.

Heterogeneous treatment effects on maths exam performance for the S4 class are shown in Table 3.11. I only show here heterogeneity by maths score and failing maths as there are no heterogeneous effects for the total score, core score or English score. At S4 level, 50% of students are female, 35% are above the median age of 17, 28% are taking fewer than the median number of 10 subjects and 49% score below the median score on the mock exam.

Looking first at gender in columns (1) and (2), though the point estimate for the interaction of female and treatment for maths score is positive, and for failing maths negative, neither are significant. However the overall effect of treatment for women is a positive and significant 0.17 standard deviations for maths score and a 14 percentage point decrease in the probability of failing maths. The simple treatment effect for maths score is no longer significant, and for failing maths is

only a significant 9 percentage point decrease. This suggests that girls benefit more from seeing the treatment movie than boys.

Amongst older students (columns (3) and (4)), the point estimate on the interaction term for maths score is actually negative and for failing maths positive, implying treatment could worsen maths performance for older student. However, again these are not significant. Overall, students older than the median have no benefit from seeing the treatment movie on their maths score but still experience a 9 percentage points decrease in the probability of failing maths.

There is a large point estimate for the interaction between taking fewer subjects than the median and treatment on the maths score, but it is not significant. The overall effect for those taking fewer subjects is a 0.25 standard deviation improvement in maths score, though this is only significant at the 10% level. The point estimate on the interaction for failing maths in column (6) is significant and negative, resulting in those students who have chosen to take fewer subjects being 19 percentage points less likely to fail maths after treatment compared to 9 percentage points less likely for those taking more subjects than the median. It could be the case that those taking less subjects find it easier to shift effort from one subject to another, resulting in treatment having larger effects.

Looking at students who scored below the median in their mock exam, there is a large and significant heterogeneous effect of treatment. Maths scores increase by 0.26 standard deviations amongst this group from treatment, and the probability of failing maths falls an additional 31 percentage points. The coefficient on the simple treatment effect on the maths score result becomes only 0.02 standard deviations and is no longer significant, likewise for failing maths. The entire improvement in maths from watching the treatment movie is seen from the group who were in the bottom half of performance on the mock exam. Overall, students who performed below the median in the mock exam are 27 percentage points less likely to fail maths. Compared to a mean of 54% of students with below median mock scores failing maths, this means treatment reduced the probability a previously poorly performing

Table 3.11: Heterogeneity in treatment effect for S4 by gender, age and number of subjects

	(1) maths	(2) fail maths	(3) maths	(4) fail maths	(5) maths	(6) fail maths	(7) maths	(8) fail maths
Treatment	0.12	-0.09**	0.17**	-0.13***	0.11	-0.09***	0.02	0.04
	(0.08)	(0.04)	(0.07)	(0.03)	(0.07)	(0.03)	(0.08)	(0.04)
Female	-0.10	0.06	-0.07	0.04	-0.07	0.04	-0.07	0.03
	(0.08)	(0.04)	(0.06)	(0.03)	(0.06)	(0.03)	(0.06)	(0.03)
Age	-0.09	0.02	-0.06	0.00	-0.09	0.02	-0.10	0.02
	(0.07)	(0.03)	(0.09)	(0.04)	(0.07)	(0.03)	(0.07)	(0.03)
Subjects	0.10	0.02	0.10	0.02	0.02	0.07	0.11	0.01
	(0.07)	(0.03)	(0.07)	(0.03)	(0.10)	(0.05)	(0.07)	(0.03)
Mock	-1.15***	0.25***	-1.15***	0.25***	-1.14***	0.25***	-1.28***	0.41***
	(0.07)	(0.03)	(0.07)	(0.03)	(0.07)	(0.03)	(0.09)	(0.04)
Treat * female	0.05	-0.05						
	(0.11)	(0.05)						
Treat * age			-0.07	0.03				
			(0.12)	(0.06)				
Treat * subjects					0.14	-0.10*		
					(0.13)	(0.06)		
Treat * mock							0.26**	-0.31***
							(0.12)	(0.05)
Constant	0.66***	0.06	0.63***	0.08	0.66***	0.06	0.69***	0.02
	(0.14)	(0.06)	(0.13)	(0.06)	(0.13)	(0.06)	(0.13)	(0.06)
Overall treatment effect								
Female	0.17**	-0.14***						
	(0.08)	(0.04)						
Age			0.11	-0.09**				
			(0.10)	(0.05)				
Subjects					0.25*	-0.19***		
					(0.11)	(0.05)		
Mock							0.28***	-0.27***
							(0.08)	(0.04)
Mean in control		0.32		0.29		0.42		0.54
Observations	730	730	730	730	730	730	730	730
R-squared	0.42	0.28	0.42	0.28	0.42	0.28	0.42	0.31

Maths is a standardized maths score. Fail maths is a dummy for whether a student got a fail in the maths exam. Age refers to a dummy if the student is above the median age for students in S4. Subjects is a dummy variable if the student is taking less subjects than the median for the UCE exams. Mock is a dummy variable for if the student performed below the median in a mock exam. The middle panel shows the overall treatment effect for each group. The mean in control shows the control mean for that sub-group. Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

student failed maths by 50%. These are very large effects, and suggest that those students struggling are more able to improve their maths scores from treatment.

I breakdown the treatment effect by mock exam performance further by interacting the treatment with each decile of mock score. These are shown in Appendix Table 3.24, again only for the maths score and dummy variable for failing maths since there are no significant effects for total score, core score or English outcomes. The bottom decile, 1, is the excluded group. The results for both the maths score and the probability of failing maths suggest it is those in the bottom of the distribution who are benefiting from the treatment and not those at the top.

Moving onto the S6 class, heterogeneous treatment effects on exam performance are shown in Table 3.12 for the total score outcome variable only. Results for the principal subjects score and subsidiary paper score are similar. At S6 level, 49% of the students are women, 30% are above the median age of 19 years, 31% are taking a STEM subject and 43% scored below the median mock score.

Looking first at gender, the total effect of treatment is significant for women $(0.12 + 0.08)$ at the 5% level and the coefficient on the simple treatment effect is no longer significant. This suggests that the beneficial effects from treatment are going to female students. For students above the median age, taking stem subjects and below the median in the mock, the point estimates on the interaction with treatment are negative, but the standard errors are very large. There is no overall effect of treatment for these groups, suggesting it may be younger students, those not taking stem subjects and those who did better in the mock exam than the median who benefit from treatment.

Once again, I breakdown the treatment effect by mock exam performance further by interacting the treatment with each decile of mock score. These are shown in Appendix Table 3.25 for the total score only (there are no differences for principal score or subsidiary score). I find no differential effects of treatment by decile.

Table 3.12: Heterogeneity in treatment effect for S6 by gender, age and taking STEM subjects

	(1)	(2)	(3)	(4)
	Dependent variable: total score			
Treatment	0.08	0.17**	0.17**	0.16**
	(0.08)	(0.07)	(0.07)	(0.08)
Female	-0.03	0.03	0.03	0.03
	(0.08)	(0.06)	(0.06)	(0.06)
Age	-0.15**	-0.09	-0.15**	-0.15**
	(0.07)	(0.09)	(0.07)	(0.07)
STEM	0.45***	0.45***	0.50***	0.45***
	(0.07)	(0.07)	(0.09)	(0.07)
Mock	-1.20***	-1.20***	-1.20***	-1.16***
	(0.06)	(0.06)	(0.06)	(0.08)
Treatment * female	0.12			
	(0.11)			
Treatment * age		-0.11		
		(0.12)		
Treatment * STEM			-0.10	
			(0.12)	
Treatment * mock				-0.07
				(0.11)
Constant	-0.51	-0.56	-0.56	-0.56
	(0.37)	(0.37)	(0.37)	(0.36)
Overall treatment effect				
Female	0.20**			
	(0.08)			
Age		0.06		
		(0.10)		
STEM			0.07	
			(0.10)	
Mock				0.10
				(0.08)
Observations	711	711	711	711
R-squared	0.50	0.50	0.50	0.50

Total is the aggregate score achieved, prin the score in 3 principal papers, sub the score on the subsidiary papers. All scores are standardized. Age refers to being above the median age for students in S6. Mock refers to scoring below the median score on a mock exam. STEM refers to choosing maths, biology, chemistry or physics as a principal subject. The bottom panel shows the overall effect for each group.
Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

3.4.4 Exploratory analysis

School Characteristics

Since the Pre-Analysis Plan was lodged and exam results data analysed, additional data on the schools was collected. This data covered:

- The national ranking of the school
- The fees of the school

and is summarised in Appendix Table 3.21.

Heterogeneous treatment effects are also analysed by these school characteristics since information about the schools could provide important information about which types of students benefit most from seeing the treatment movie. For example, both the national ranking of schools and the fees charged by the school give an indication of how good the school is. If students at worse schools benefit more from seeing the treatment movie than those at better schools, then treatment might be able to partially mitigate factors related to poor schooling, such as teacher quality.

Heterogeneous treatment effects on maths exam performance for the S4 class are shown in Table 3.13. I only show here heterogeneity by maths score and failing maths as there are no effects for the total score, core score or English score. First looking at whether the school is a top 500; schools in Uganda are nationally ranked and this is freely available from the Ministry of Education. Out of the 3300 schools included in the ranking I consider the top 500 as the leading schools and look at treatment heterogeneity by this variable. The top 500 defines in the ranking what are considered good schools. 46% of the schools in my sample are ranked in the top 500.

The treatment interactions with being at a top 500 school for the maths score is small and insignificant. The interaction coefficient on failing maths is positive 10 percentage points and significant at the 10% level. Overall, treatment results in students at school in the top 500 scoring 0.16 standard deviations higher on their maths exam, though this is only significant at the 10% level. There is no benefit from treatment to students at top 500 school in terms of failing maths,

perhaps because students at top 500 schools already rarely failed maths (only 13% fail it in the control group). This result indicates that it's schools not in the top 500, so schools performing less well nationally, which benefit most from seeing the treatment movie in terms of the probability of failing maths, with treatment resulting in students at lower ranked schools being 16 percentage points less likely to fail maths. Seeing the treatment movie may therefore help to mitigate some aspects of being at a poorly performing school.

Table 3.13: Heterogeneity in treatment effect for S4 by school characteristics

	(1) maths	(2) fail maths	(3) maths	(4) fail maths
Treatment	0.14*	-0.16***	0.20***	-0.17***
	(0.08)	(0.04)	(0.07)	(0.03)
Treatment * top 500	0.03	0.10*		
	(0.12)	(0.05)		
Top 500	0.54***	-0.45***		
	(0.20)	(0.09)		
Treatment * high fees			-0.17	0.18***
			(0.13)	(0.06)
High fees			-0.09	-0.09
			(0.16)	(0.07)
Constant	-0.09	0.51***	0.61***	0.11*
	(0.19)	(0.08)	(0.14)	(0.06)
Overall treatment effect				
Top 500	0.16*	-0.06		
	(0.09)	(0.04)		
High Fees			0.03	0.01
			(0.10)	(0.05)
Mean in control		0.13		0.05
Observations	730	730	730	730
R-squared	0.42	0.28	0.42	0.29

Maths is a standardized maths score. Fail maths is a dummy for whether a student got a fail in the maths exam. Top 500 refers to if the school is within the top 500 out of 3300 nationally ranked schools. High fees refers to if a school charges above the median of school fees in the sample. All regressions include school fixed effects and student individual characteristics (age, gender, mock score and number of subjects taken). Mean in control refers to the control mean of that sub-group.
Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

I also look at whether a school charges above the median fees in my sample, of which 31% do. High fees may be a proxy for the income of students' parents, useful since I do not have data on household income. For schools charging the highest fees, the interaction with the treatment has large and negative but not significant effect on the maths score. This would balance out against the pure effect from treatment of positive 0.20 standard deviations, resulting in students at high fees schools getting no overall benefit to their maths score from seeing the treatment movie. Likewise, the coefficient is large, positive and significant for the interaction term for the likelihood of failing maths outcome which balances out against the large negative simple treatment effect giving no overall impact for students at high fees schools from seeing the treatment movie on failing maths.

This indicates it is students at lower fee charging schools which see improvements in their maths scores and reduction in the probability of failing maths from watching the treatment movie. Again, this is likely because students at high fee schools already do much better, with only 5% of students at high fee schools failing maths.

This could be interpreted as because high charging schools are already doing so much better than low charging schools, that incremental improvements are much harder. If a student is at a poor school and almost failing maths, then small increases in effort or motivation could push that student over the threshold to no longer fail. If a student is at a high fee paying school, where presumably more and better quality resources are already being invested in their education, then if they are one of the few students failing maths, seeing the movie is not enough to improve their performance compared to everything already being done. Improvements in overall maths score are also much harder for students already doing very well at schools invested in their education.

Overall these heterogeneous effects highlight that it is students at worse performing schools, whether by ranking or fees who benefit most from seeing the treatment movie.

Turning to the S6 class, heterogeneous effects by school characteristics are shown in Table 3.14. Here I show the effect on the total score, principal paper score and

subsidiary paper score. I examine the impact of a school being in the top 200. I use the top 200 to make it comparable to the top 500 out of 3300 schools at the S4 level, since at S6 only 1800 schools provide teaching at this level. I find some large but insignificant effects for a school being in the top 200 interacted with treatment on all the outcomes. These result in overall positive and significant effects from being at a top 200 school on total and principal paper scores. There is no significant effect on the subsidiary paper score. However the simple treatment effect is actually negative for the subsidiary score, suggesting if there is a positive effect of treatment on this score it is all coming through top 200 schools.

Secondly I examine whether there are heterogeneous effects by whether the school fees charged are higher than the median. Higher fee charging schools have students which perform better on the exams, especially for the subsidiary paper. The coefficients on the interaction of treatment with being at a high fee school are significant for both the overall score and subsidiary paper score, resulting in overall positive effects from treatment on all the outcomes for high fee schools, ranging from 0.16 standard deviations to 0.23 standard deviations. This suggests that the only group experiencing positive effects on the subsidiary paper is the high fees schools.

Note, there is no heterogeneity by school fixed effects, and so results are not shown here for brevity.

Continuation in school

Since writing the pre-analysis plan and the amendment, while collected results data from the schools, the opportunity came up to collect data on whether students from S4 chose to continue onto upper secondary school in S5 class. This data could give an indication if not only are effort levels in the exams increasing, but if also student's ambitions for continuing in schooling are changing.

It is important to highlight a number of limitations to this data though. Firstly, it was not included in the pre-analysis plan so should be taken as exploratory at best. Secondly, only 3 schools provided this data. Whether a school provided this data depending on two things: firstly whether they actually offered upper

Table 3.14: Heterogeneity in treatment effect for S6 by school characteristics

	(1) total	(2) prin	(3) sub	(4) total	(5) prin	(6) sub
Treatment	-0.01	-0.01	-0.11	0.12*	0.11*	0.00
	(0.11)	(0.12)	(0.14)	(0.07)	(0.07)	(0.08)
Treatment * top 200	0.18	0.18	0.22			
	(0.13)	(0.13)	(0.16)			
Top 200	0.14	0.17*	0.11			
	(0.10)	(0.10)	(0.12)			
Treatment high fees				0.06	0.05	0.23*
				(0.10)	(0.11)	(0.13)
High fees				0.19**	0.10	0.33***
				(0.07)	(0.08)	(0.09)
Constant	0.15	0.07	-0.08	0.48	0.40	0.37
	(0.44)	(0.46)	(0.55)	(0.44)	(0.45)	(0.53)
Overall treatment effect						
Top 200	0.17***	0.17***	0.12			
High fees				0.18**	0.16**	0.23**
Observations	708	708	708	708	708	708
R-squared	0.58	0.56	0.30	0.58	0.55	0.34

Total is the aggregate score achieved, prin the score in 3 principal papers, sub the score on the subsidiary papers. All scores are standardized. Top 200 refers to if the school is within the top 200 out of 1800 nationally ranked schools.High fees refers to if a school charges above the median of school fees in the sample. Regressions include school fixed effects and individual controls (age, gender, mock score and number of subjects taken). Mean in control refers to the control mean of that sub-group.
Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

secondary school, which typically only the larger schools did[6], secondly whether they agreed to provide this information to us. This data is therefore only available for half (331) the S4 students at 3 of the largest schools. This reduced sample also substantially limits my power, particularly for heterogeneity analysis. Additionally, just because a student enrols in S5 does not mean they complete S5, continue to S6 or take the final exams at the end of S6.

Taking these limitations into account, results for the impact of treatment on whether students from S4 continued onto S5 are presented in Table 3.15 and

[6]At the smaller secondary schools students moved to a different school if they wanted to continue to upper secondary

heterogeneity by individual characteristics is shown in Table 3.16. I find that treated students are 9 percentage points more likely to continue to upper secondary school, significant at the 10% level. Since 1/5 of the control group continue to secondary school, this is a 50% increase in the probability of continuing to upper secondary school. Looking at heterogeneity in this effect in Table 3.16, I see that this effect is primarily coming from female students who are 13 percentage points more likely to go to upper secondary school, significant at the 10% level. Since only 18% of girls in the control group continue to upper secondary school, this is an increase of 72%. Taking the limitations of this data into consideration, this could be indicative that the treatment encourages girls in particular to continue in school.

Table 3.15: Impact of treatment assignment on continuation to upper secondary school

Dependent variable: continuation to S5		
	(1)	(2)
Treatment	0.09*	0.09*
	(0.05)	(0.05)
Age		-0.02
		(0.02)
Female		-0.03
		(0.05)
No. subjects		0.05
		(0.05)
Mock score		-0.06*
		(0.03)
Constant	0.22***	0.16
	(0.06)	(0.67)
Mean control	0.20	0.20
Observations	331	330
R-squared	0.01	0.03

Continuation to S5 is a dummy variable equal to one if that student continued to the first grade of upper secondary school, S5. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 3.16: Impact of treatment assignment on continuation to upper secondary school

Dependent variable: continuation to S5				
	(1)	(2)	(3)	(4)
Treatment	0.05	0.08	0.07	0.09
	(0.07)	(0.06)	(0.06)	(0.06)
Female	-0.06	-0.02	-0.02	-0.02
	(0.07)	(0.05)	(0.05)	(0.05)
Above median age	-0.07	-0.08	-0.07	-0.07
	(0.06)	(0.08)	(0.06)	(0.06)
Below median subject	-0.04	-0.04	-0.08	-0.04
	(0.05)	(0.05)	(0.07)	(0.05)
Below median mock	0.11**	0.11**	0.11**	0.11
	(0.05)	(0.05)	(0.05)	(0.07)
Treatment*female	0.08			
	(0.09)			
Treatment*above median age		0.02		
		(0.10)		
Treatment*below median subject			0.06	
			(0.10)	
Treatment*below median mock				-0.01
				(0.10)
Constant	0.28***	0.26***	0.27***	0.26***
	(0.08)	(0.08)	(0.08)	(0.08)
Overall treatment effect				
Female	0.13			
	(0.07)*			
Above median age		0.10		
		(0.08)		
Below median subject			0.13	
			(0.09)	
Below median mock				0.08
				(0.09)
Mean control	0.18	0.18	0.15	0.27
Observations	331	331	331	331
R-squared	0.03	0.03	0.03	0.03

Continuation to S5 is a dummy variable equal to one if that student continued to the first grade of upper secondary school, S5. Above median age refers to a dummy if the student is above the median age for students in S4. Less median subject is a dummy variable if the student is taking less subjects than the median for the UCE exams. Below median mock is a dummy variable if the student scored below the median mark on the mock exam. The middle panel shows the overall treatment effect for each group. The mean in control shows the control mean for that sub-group. Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Persistence of effects

To try and ascertain if effects on exam performance tail off with time or have a more persistent effect I used the natural variation in exam date depending on the subjects chosen at S6 to see if the treatment effects differ for students taking subjects closer or further away from the intervention[7].

Exams for different subjects at S6 level took place between 14th November 2016 and 29th November 2016. Depending on which combinations of subjects student's chose, some students had their exams closer to the intervention than others, presenting natural variation I can exploit to see if treatment effect vary over time. Note the subject choice and exam timing are exogenous with respect to treatment as there were all determined before the movies were shown to students. To examine heterogeneity of the treatment over time I construct two dummy variables capturing whether students took exams closer or further to the intervention.

The first dummy variable captures whether the first exam the student took was below the median date of the first exam in the sample. The median first exam for the S6 class was 36 days after the intervention. The student's first exam date varied between 34 and 44 days after the intervention. I classify those whose first exam was less than 36 days after treatment as early first exam students and the rest as late first exam student.

The second dummy looks at the average days since intervention of all the students' exams, not just their first exam. I take the mean number of days since the intervention of all their exams and compare this to the median of the sample, 40 days. The mean date of a student's exams varied from a minimum of 35 days after the intervention to a maximum of 48 days after the intervention. I define students as having early exams if the mean days since intervention of all their exams is less than 40 days, and as having late exams otherwise.

Results for these two outcomes are shown in Table 3.17, with the early first exam indicator heterogeneity in columns (1)-(4) and the early average exams indicator in

[7]I also try and do a similar exercise for the S4 class but since the screenings took place over only 5 days and all students had the same first exam there is not enough variation and I find no differential effects by treatment day

Table 3.17: Impact of treatment assignment on standardized test scored by students taking early exams

	(1) Over-all score	(2) Prin. sub-jects	(3) Sub. sub-jects	(4) Uni pass grade	(5) Over-all score	(6) Prin. sub-jects	(7) Sub. sub-jects	(8) Uni pass grade
Treatment	0.12*	0.09	0.07	0.08**	0.11	0.09	0.06	0.07**
	(0.07)	(0.07)	(0.08)	(0.03)	(0.07)	(0.07)	(0.08)	(0.03)
Treat*early	0.04	0.09	-0.01	-0.06				
first exam	(0.10)	(0.11)	(0.12)	(0.05)				
Early first	-0.01	-0.11	0.15	-0.02				
exam	(0.11)	(0.11)	(0.12)	(0.05)				
Treat*early					0.05	0.08	0.04	-0.06
average exams					(0.10)	(0.11)	(0.12)	(0.05)
Early					0.03	0.04	0.07	0.02
average exams					(0.08)	(0.08)	(0.09)	(0.04)
Constant	-0.12	-0.22	0.69	0.81***	-0.13	-0.24	0.68	0.80***
	(0.50)	(0.53)	(0.59)	(0.25)	(0.51)	(0.53)	(0.59)	(0.25)
Overall treatment effect								
Early first	0.15**	0.18**	0.07	0.01				
exam	(0.07)	(0.07)	(0.08)	(0.04)				
Early average					0.16**	0.17**	0.09	0.01
exam					(0.07)	(0.08)	(0.09)	(0.04)
Observations	708	708	708	708	708	708	708	708
R-squared	0.62	0.59	0.44	0.33	0.62	0.59	0.44	0.32

Total is the aggregate score achieved, prin the score in 3 principal papers, sub the score on the subsidiary papers, uni pass grade that they got 2 passes on the principal papers. All scores are standardized. Early first exam means the first exam that student took was before the median first exam for all students. Early average exam means the average days since the intervention of all that student's exams was below the median for all students. Regressions include school fixed effects and individual controls (age, gender, mock score and number of subjects taken).
Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

columns (5)-(8). I find similar results regardless of which indicator I use: Treatment interacted with either early exam indicator is insignificantly different from zero at at least the 10% level in all cases. However, the overall treatment effect is a larger magnitude and of higher significance for those taking early exams (by either measure) on their overall and principal subject scores, while there is no overall effect for those taking early exams on their likelihood of getting the required two principal passes

to get into university. Overall, I cannot reject that treatment effects are the same for those taking early and later exams by either method of indicator construction.

3.4.5 Robustness

Randomisation test

I use permutation tests to compute exact test statistics which do not depend on asymptotic theorems. To do this I use Stata's permute function to randomly assign students to the treatment and control group and calculates the probability of observing the treatment effect I did under the null hypothesis that there is no treatment effect. I do this for the S4 and S6 outcomes defined in section 3.3.3 using 10000 permutations and without individual control variables, only school fixed effects.

Table 3.18: S4 main results robustness tests

	(1) Total score	(2) Core score	(3) Maths	(4) English
Treatment	0.006	-0.003	0.140	-0.044
Robust p-value	(0.930)	(0.966)	(0.045)**	(0.517)
Permutation p-value	(0.917)	(0.962)	(0.038)**	(0.493)
Observations	735	735	735	735
R-squared	0.310	0.293	0.172	0.249

Total score refers to standardised aggregate score across all subjects taken in the exam. Core score refers to standardised aggregate score in the 6 mandatory subjects at S4 level. Standardized test scores composed of subject standardized scores and renormalised. Regressions include school fixed effects. Permutation p-value calculated using 10000 permutations. *** p<0.01, ** p<0.05, * p<0.1

These are reported in Tables 3.18 and 3.19 underneath the robust p-values. At the S4 level, treatment still only has a positive impact on the maths exam and is still significant at the 5% level. At the S6 level, for both the overall score and principal subjects score, though the permutation p-values are higher than the robust p-values, treatment still has a positive effect, significant at the 5% level.

Table 3.19: S6 main results robustness tests

	(1) Overall score	(2) Principal subjects	(3) Subsidiary subjects
Treatment	0.169	0.165	0.079
Robust p-value	(0.024)**	(0.032)**	(0.222)
Permutation p-value	(0.041)**	(0.046)**	(0.409)
Observations	710	710	710
R-squared	0.196	0.150	0.288

Overall score refers to the aggregate score in the principal and subsidiary papers. Principal subjects refers to the standardised score on the 3 chosen subject papers. Subsidiary subjects refers to the standardised score on the two mandatory subsidiary papers. Standardized test scores composed of subject standardized scores and renormalised. Regressions include school fixed effects. Permutation p-value calculated using 10000 permutations.
*** p<0.01, ** p<0.05, * p<0.1

Multiple hypothesis testing

The outcomes examined in the main results section were pre-specified in the pre-analysis plan as well as conforming to the standard when examining educational outcomes of focusing on overall scores and maths and English. As a robustness check, I illustrate the treatment effect on failing any core subject to highlight the stability of the maths result to multiple testing. To do this, I compare the result of treatment for failing maths to the impact of treatment on the probability of failing each of the core exams listed in Section 3.3.3. I perform this only for the core subjects that all students took.

To correct for multiple hypotheses, I compute sharpened q-values. Q-values adjust p-values to control for the false discovery rate. The false discovery rate is an approach which controls for the expected proportion of rejected hull hypotheses that are false (incorrectly rejected). It therefore controls for the rate of type I errors when testing many hypotheses. This is a less stringent approach than those controlling for the probability of any type I error, such as the Bonferroni correction, and it therefore allows more power with a trade off of a higher rate of type I error. The method used here is Benjamini, Krieger and Yekutieli (2006) sharpened q-values as described in Anderson (2008) and using the code provided by Anderson online.

This is one of the least conservative methods to control for false discovery rates. However the findings are unchanged using conservative methods such as Bonferroni.

In Table 3.20 I show the impact of treatment assignment on the probability of failing each of the core exam subjects and display both robust p-values and sharpened q-values below each coefficient. It can be seen that the only subject for which the significant result is robust to multiple hypothesis testing is the maths result, where treatment results in a reduced probability of failure and is significant at the 1% level even using sharpened q-values. Of the other core subjects, none is significant using either conventional p-values or sharpened q-values.

Table 3.20: Multiple hypothesis test for failing core subjects

	(1) Maths	(2) Eng	(3) Chemistry	(4) Biology	(5) Physics	(6) History	(7) Geography
Treatment	-0.11	-0.01	0.04	-0.00	0.01	0.00	0.01
p-value	(0.000)***	(0.703)	(0.123)	(0.959)	(0.654)	(0.867)	(0.76)
q-value	(0.001)***	(1.00)	(0.740)	(1.00)	(1.00)	(1.00)	(1.00)
control mean	0.27	0.11	0.38	0.35	0.49	0.16	0.12
Observations	729	729	728	729	728	727	706
R-squared	0.31	0.28	0.51	0.57	0.51	0.36	0.35

Core subjects are taken by all students at S4 level. All regressions include school fixed effects and individual controls (age, gender, number of subjects taken and mock exam score). Q-values calculated using the 2 step procedure of Benjamini et al. (2006)
Robust p-value in parentheses. *** p<0.01, ** p<0.05, * p<0.1

3.5 Cost effectiveness and discussion

3.5.1 Cost effectiveness

The study was extremely costs effective, with the total cost of the intervention only $3 per student for the cinema screening and $2 per student for transport to the cinema. This means there was only a cost of $5 per student to improve education outcomes by 0.11 standard deviations in maths at S4 and 0.13 standard deviations

overall at S6. To compare this to some other education interventions, I use the method in Kremer et al. (2013) of comparing the standard deviation of impact that could be had for $100 of spending. In my study, you could improve test scores by 2.2 to 2.6 standard deviations (by raising the scores of 20 students by 0.11-0.13 standard deviations) for $100 of spending. This is comparable to a remedial education programme in India which generated a 3 standard deviation test score gain per $100 spent (Banerjee et al., 2007) or to teacher incentives in Kenya (Glewwe et al., 2010). Baird et al. (2016) find similar effects of 0.15 standard deviation on maths scores from giving a $5 conditional cash transfer a month to girls if they stay in school. Equally importantly is that my study had an effect over a period of 1-4 weeks whereas many studies are finding similar impacts after years of an intervention.

Additionally, this study showed the *Queen of Katwe* movie at a cinema for three main reasons: one, so that students could see the movie immediately upon its release before sitting their exams; two, because it was logistically simpler and faster then arranging screenings at schools; and three, to allow an individual randomisation at the cinema. If the study was scaled up though, the movie could be shown in schools to entire classes, perhaps through a projector or a specially arranged screening for many schools, and this might lower the cost further.

3.5.2 Impact on earnings

Any effects on earnings of those who saw the treatment movie cannot be known as this data was not collected as part of this study. However, it is possible to calculate an estimate of the income gains associated with the gain in educational attainment based on the returns to education in Uganda. Treatment led to a 6 percentage point increase in the likelihood of getting a place at university, increasing enrolment from 30% to 36%. According to government statistics, 94% of students admitted to government universities complete their programmes (Basheka, 2013). Kavuma (2014) found that those finishing university earn 120% more than those finishing just secondary school, so there are sizeable potential income gains from helping students get into university.

Treatment also led to a 9% percentage point increase in the likelihood of continuing to S5, increasing the enrolment from 20% to 29%. Of those who start S5, 94% finish the year and go onto S6 (MoES Uganda, 2015). The return to an additional year of schooling is an 11% increase in wages (Kavuma, 2014), so again these students could see a large gain in their wages.

3.5.3 Interpretation of effects

The results for S4 and S6 students appear to come from different parts of the ability distribution. For the S4 class, it is lowest ability students as measured on a mock exam who benefit most from the treatment. This suggests the treatment is helping to compensate for being a poor student. Likewise effects are concentrated amongst students at lower ranked schools charging lower fees. This suggests effects are greatest at lower performing schools. At the S6 level I find it is students at the highest ranked and highest fee schools who gain the most from treatment, and if anything, the heterogeneity analysis suggests the effects are most pronounced at the top of the distribution.

There are a number of possible reasons for these differences in effect. Firstly, the class profiles are very different. All schools have minimum requirements to go from S4 to S6, and on average nationally only one-third of students continue to S5 (MoES Uganda, 2015). The students for which *Queen of Katwe* had an effect on at S4 are therefore unlikely to continue onto S6.

Secondly, the profile of subjects also differs greatly between S4 and S6. At S4, most subjects are compulsory and students take many subjects, whereas at S6, the principal papers, for which I see an improvement from seeing the treatment movie, are all optional and students take only three subjects.

Thirdly, there were very different time lags between seeing the treatment and the exam for S4 (one week) and S6 (one month). Having only one week between seeing the treatment and the exam means that there is only time to make a limited amount of improvement in the S4 class. If we assume it is easier to improve an exam score from fail to one above fail than from a B to an A, then it will be the students

who would have failed if they hadn't seen the treatment movie whose increased effort would most easily be seen in an improvement in exam grade. For them, a very small amount of extra effort could translate into a higher grade, whereas at the top grades more effort is needed to reach a higher grade. Hence the time constraint might have meant I could only detect effects at the bottom of the distribution.

The fact that effects were only seen for the maths exam is a common result for this type of short term intervention and maths exams are considered more elastic then English or reading exams (Bettinger, 2012). At S6 they had at least a full month between seeing the treatment movie and the exams, over which time frame longer term motivation and perseverance comes into play and there may be larger complementaries with being in a good school. This might explain why I only see the strongest effects for above average students at S6, at top quality and more expensive schools.

Lastly, schools are primarily judged in Ugandan at S4 level by how many students get the highest score, rather than how many get low scores, and so generally invest less in low performing S4 students and focus instead on a few best students. This could leave a large cohort of low performing students for which a small investment can have large payoffs in terms of exam performance. At S6, this effect is less pronounced and the focus is more on helping students achieve the grades to get into public university (2 passes in principal papers), rather than just those at the top. This could explain the more across the board effects seen in S6.

3.6 Conclusion

I find that exposing secondary school students to a movie featuring a potential role model improves national exam performance. Amongst S4 students completing lower secondary school, seeing the treatment movie increases maths scores by 0.11 standard deviations, with the effect coming from lower ability students at worse schools being 30% less likely to fail maths. Effects are also strongest for female students. At S6 level, amongst students trying to achieve the grades to get into university, I find seeing the treatment movie improves overall exam performance

by 0.13 standard deviations and increases the probability by 6 percentage points that they get the necessary grades for university. Again, I see heterogeneity, with women gaining the most from seeing the treatment movie. This time I see the largest effects for top students.

An implication of these findings is that schools should place more emphasis on having appropriate role models in schools, whether through showing a movie or through having former students come in to tell their stories. It is also important that schools do not just focus on the best performing students and leave the weakest behind. The fact that the *Queen of Katwe* movie had such a big effect on S4 students failing maths, especially at the worst schools, suggests that small changes at those schools could also have a big effect. One way to do this is to place more emphasis on motivation and inspiration through example, to give more meaning to the students of how education can help them to achieve their life goals.

However, this work had a number of limitations which would benefit from further research. Firstly, there was no pure control group, so it is possible that both the treatment and placebo movie actually reduced exam performance, just *Queen of Katwe* less so, or that *Queen of Katwe* was neutral and *Miss Peregrine's Home for Peculiar Children* reduced exam performance. This seems unlikely given the fact that a lot of thought went into finding and reviewing movies that would be appropriate for the age group and not have any potential negative effects by being frightening, for example. It therefore seems unlikely to me that *Miss Peregrine's Home for Peculiar Children* could actively have reduced exam performance.

Secondly, due to limitations of time and money, no individual surveys were done with the students, preventing a deeper understanding of the mechanisms by which the treatment movie had an effect. Due to this, this study is best viewed at providing evidence on whether a role model in a movie can affect economic behaviours, and if so who experiences the largest effect. It presents a starting point for further work on the importance of role models for education and that this role model can be as popular media.

Further work would hope to understand potential mechanisms for how seeing the treatment movie led to a change in behaviour. Previous papers have highlighted channels such as information presented in a more salient way (Nguyen, 2008), norms (Banerjee et al., 2018, Jensen and Oster, 2009, Paluck and Green, 2009) and aspirations (Beaman et al., 2012, Bernard et al., 2014) as potential mechanisms through which role models work. Whether the movie only impacted exam effort, or also led to changes in other areas of the students' lives is also important to understand. Understanding which if any of these mechanisms the *Queen of Katwe* movie worked through would increase our understanding of what limits educational achievement. Additional work would also seek to understand the persistence of these effects.

Appendices of Chapter 3

3.A Additional tables

Table 3.21: School Characteristics

School	Religion	Boarding	Ranking		Fees				Class size	
			UACE Rank /1882	UCE Rank /3294	S4 board	S4 day	S6 board	S6 day	S4	S6
Hope	Islamic	Boarding only	7	94	650,000		650,000		93	65
Paul Musaka	Christian	Mixed day and boarding	220	199	680,000	340,000	680,000	360,000	136	80
Kyandondo	Islamic	Mixed day and boarding	271	537	730,000	530,000	730,000	530,000		187
Makerere	Christian	Mixed day and boarding	342	464	450,000	250,000	500,000	300,000	85	47
Royal	Christian	Boarding only	461	32	600,000		650,000		110	93
Kinaawa	Islamic	Boarding only	492	430	900,000		900,000			94
Jakayza	Islamic	Mixed day and boarding	525	1047	460,000	230,000	480,000	245,000		25
Mukono	Christian	Mixed day and boarding	527	472	600,000	450,000	600,000	450,000	82	57
Atlas	Christian	Mixed day and boarding	529	170	900,000	450,000	920,000	470,000		40
Gayaza	Islamic	Mixed day and boarding	931	2020	470,000	208,000	500,000	220,000		12
Dynamic	Christian	Mixed day and boarding	1423	2036	550,000	180,000	400,000	180,000	141	
Kulumba	Islamic	Mixed day school	1782	1205		170,000		220,000	21	5
Devine	Islamic	Mixed day and boarding	1799	2007	440,000	210,000	500,000	250,000	53	5

Religion is the religious affiliate reported by the school. Students are taken to pray on religious days and 5 times a day at Islamic schools. Fees are in Ugandan Shillings per year. Class size refers to the size of the class if it participated in the study. Schools either gave the entire class or not at all, never part of a class.

Table 3.22: Ordered logit regression of the impact of treatment on core subject grade for S4

	(1) Eng	(2) Chem	(3) Phy	(4) Bio	(5) Hist	(6) Geog
Fail	0.011	0.021	0.016	0.014	0.002	-0.005
	(0.010)	(0.017)	(0.019)	(0.015)	(0.012)	(0.011)
2	0.005	-0.000	-0.003	0.001	0.000	-0.001
	(0.005)	(0.001)	(0.003)	(0.001)	(0.001)	(0.002)
3	0.003	-0.004	-0.004	-0.001	0.000	-0.001
	(0.002)	(0.003)	(0.004)	(0.001)	(0.001)	(0.001)
4	0.002	-0.003	-0.003	-0.004	0.000	-0.000
	(0.002)	(0.003)	(0.003)	(0.004)	(0.001)	(0.001)
5	-0.005	-0.004	-0.002	-0.004	0.000	0.000
	(0.005)	(0.004)	(0.002)	(0.004)	(0.000)	(0.001)
6	-0.008	-0.004	-0.002	-0.003	-0.000	0.002
	(0.007)	(0.003)	(0.002)	(0.003)	(0.001)	(0.004)
7	-0.004	-0.004	-0.002	-0.002	-0.001	0.002
	(0.004)	(0.003)	(0.002)	(0.002)	(0.003)	(0.005)
8	-0.002	-0.001	-0.002	-0.001	-0.001	0.002
	(0.002)	(0.001)	(0.002)	(0.001)	(0.004)	(0.003)
9	-0.002	-0.001	-0.000	-0.000	-0.001	0.001
	(0.002)	(0.001)	(0.000)	(0.000)	(0.007)	(0.002)
Observations	729	728	728	729	727	706

Core subjects are taken by all students at S4 level. All regressions include school fixed effects and individual characteristics (age, gender, number of subjects taken and mock score). Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 3.23: Impact of treatment assignment on S4 standardised tests scores

	(1) Total score excl. maths	(2) Total score excl. maths
Treatment	-0.01	-0.03
	(0.05)	(0.02)
Age		0.01
		(0.01)
Female		-0.01
		(0.02)
No. subjects		-0.04
		(0.02)
Mock score		0.81***
		(0.01)
Constant	-0.58***	0.06
	(0.07)	(0.29)
Observations	735	729
R-squared	0.32	0.88

Total score excl. maths refers to the total score excluding the maths score. Standardized test scores composed of subject standardized scores and renormalised. All regressions include school fixed effects.
Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Table 3.24: Heterogeneity in treatment effects for S4 by mock decile

	(1) maths	(2) fail maths
treatment	0.14	-0.24***
	(0.14)	(0.07)
treatment * decile 2	-0.05	-0.00
	(0.23)	(0.11)
treatment * decile 3	0.14	0.02
	(0.22)	(0.10)
treatment * decile 4	0.26	-0.13
	(0.22)	(0.10)
treatment * decile 5	0.24	0.14
	(0.21)	(0.10)
treatment * decile 6	-0.16	0.37***
	(0.22)	(0.11)
treatment * decile 7	0.03	0.28***
	(0.21)	(0.10)
treatment * decile 8	-0.30	0.31***
	(0.24)	(0.11)
treatment * decile 9	-0.22	0.19*
	(0.21)	(0.10)
treatment * decile 10	-0.21	0.27**
	(0.22)	(0.11)
Overall treatment effect		
decile 1	0.14	-0.24***
	(0.14)	(0.07)
decile 2	0.10	-0.25*
	(0.18)	(0.08)
decile 3	0.28*	-0.22***
	(0.16)	(0.09)
decile 4	0.41**	-0.37***
	(0.17)	(0.08)
decile 5	0.39***	-0.10
	(0.15)	(0.07)
decile 6	-0.02	0.12
	(0.17)	(0.08)
decile 7	0.17	0.03
	(0.15)	(0.07)
decile 8	-0.16	0.07
	(0.19)	(0.09)
decile 9	-0.08	-0.05
	(0.15)	(0.07)
decile 10	-0.07	0.03
	(0.17)	(0.08)
Observations	730	730
R-squared	0.56	0.40

Decile refers to decile of mock exam score. Regressions include school fixed effects and individual controls of age, gender, number of subjects taken and mock decile. The second panel shows the overall effect of treatment for each decile. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 3.25: Heterogeneity in treatment effects for S6 by mock decile

	(1) total score
treatment	0.13
	(0.12)
treatment * decile mock 2	0.12
	(0.20)
treatment * decile mock 3	-0.19
	(0.19)
treatment * decile mock 4	0.01
	(0.20)
treatment * decile mock 5	0.03
	(0.21)
treatment * decile mock 6	-0.18
	(0.21)
treatment * decile mock 7	0.07
	(0.23)
treatment * decile mock 8	0.07
	(0.21)
treatment * decile mock 9	0.07
	(0.21)
Overall treatment effect	
decile 2	0.26
	0.16)
decile 3	-0.06
	(0.14)
decile 4	0.14
	(0.16)
decile 5	0.16
	(0.17)
decile 6	-0.05
	(0.16)
decile 7	0.20
	(0.19)
decile 8	0.20
	(0.17)
decile 9	0.20
	(0.17)
Observations	708
R-squared	0.58

Decile refers to decile of mock exam score. Regressions include school fixed effects and individual controls of age, gender, if taking a stem (maths or science) paper and the mock decile. The second panel shows the overall effect of treatment for each decile. There are 9 deciles since two deciles had equal scores associated with them. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Conclusion

This thesis studies three themes in development economics. Two of the chapters look at a relatively new type of financial service, mobile money services, and how they can contribute to improving the lives of the poor. The final chapter looks at what is also a new area of focus for economists: role models and how these can change behaviours. All three papers centred on East Africa, two in Uganda and one in Tanzania, where GDP per capita is $1700 and $2700 respectively at purchasing power parity (PPP). This puts these countries just below and right at the mean of GDP per capita PPP in developing countries. More work therefore remains to be done to improve the lives of people living in these countries.

In Chapter One, I show how mobile money services can contribute to risk sharing after village-level rainfall shocks. Households in developing countries are still reliant on primarily informal methods of risk sharing which leave them exposed to drops in consumption after a rainfall shock that affects the entire village (Dercon, 2002, Udry, 1994). Mobile money services allow the user of the service to smooth their consumption and no longer experience a fall after the shock. I show that mobile money services primarily benefit users of the services through the receipt of remittances, but that these remittances are not shared with other non-users in the same village.

It is surprising that, after a shock that reduces consumption, others in the village don't also ask for remittances. This suggests that either these households do not have access to someone outside the village they can request remittances from, or that they face a barrier to adoption of mobile money services that would allow them to receive remittances. In either case, more work is needed to understand how the arrival of new financial services affects traditional risk sharing relationships, who benefits and loses out from their introduction and what could be done to help those who lose out.

In Chapter Two, I examine how mobile money accounts can be used to allow female entrepreneurs to invest more of a microfinance loan in their business. I find that disbursing a loan onto a mobile money account provided to the woman for her business results in higher business investment and higher profits compared to giving the loan as cash. The mechanism through which this works is that mobile money disbursement results in less of the loan being given to the woman's spouse and family. The benefits of mobile money loan disbursement are therefore highest for women who most suffered from pressure to share money with others at baseline. I do not find compelling evidence that the mobile money account satisfied saving constraints or assisted the woman with self-control issues. I find no impact of just providing a mobile money account for the woman's business to the woman, the loan had to also be disbursed onto the account for there to be any impacts on the woman's business. This suggests a strong role for default effects, as in principal those getting just the mobile money account could have deposited the loan onto it themselves.

This study contributes to a literature on how women in developing countries benefit from greater control over their money. It raises questions about whether the effects would persist once it became common knowledge that women received their loans on a mobile money account. In principal, it is not hard for family to take the sim card and demand the PIN number. Whether this would happen depends on if the prior situation involved the family controlling the majority of the loan or if the family instead requested small amounts of the loan, which whittles away its value to the business. If the latter then mobile money accounts may permanently protect the loan from requests. If the former then the family may just switch to controlling the mobile money account. It is therefore important for future work to understand the different dynamics by which family members support each other and have control over their funds.

Likewise, I did not find any negative effects of the treatment on the amount of money the woman received from her spouse, which could happen if by refusing requests for money the woman ostracised herself from her spouse or if he believed she could sufficiently support herself without the need for him to give her funds. I

also did not detect any changes in support from her peers within the microfinance group. Other work though has found that mobile money accounts can change risk sharing from based on social networks to based on own saving (Dizon et al., 2017) and it is therefore important to examine who, if anyone, becomes worse off from this.

This study also contributes to the discussion around under what conditions microfinance loans can benefit households. I found that keeping the loan separate from other household cash, and away from requests from other household members allowed greater investment of the loan. There is a small but consistent literature which has found, under certain conditions, that microfinance loans can have high returns for female entrepreneurs (Bernhardt et al., 2017, Blattman et al., 2014, Fafchamps et al., 2014, Fiala, 2017, Field et al., 2013). These conditions generally focus on allowing lumpy investments of the loan or grant. Understanding in more detail what these conditions are, and how can they be replicated, is an important area of further research.

A key limitation of my study is that it is not clear whether women's ability to hide the loan with the mobile money would remain once it became common knowledge that BRAC was disbursing loans this way. More research would be needed to understand the dynamics of this treatment. Additionally, this experiment took place in an urban setting amongst women who were already literate with mobile money services. The impacts could be larger for rural microfinance borrowers if they are saving constrained and also gain access to a safe saving device. However, the impacts could be lower if they struggle to use the account and find it hard to cash out the loan. More research is needed to understand if my results would generalise outside an urban sample.

The final chapter of my thesis looks at the impact of a role model on the exam performance of secondary school students in Uganda. This adds to a relatively new area of economics looking at role models and how these impact beliefs and aspirations. It also speaks to the impact of media on behaviour (La Ferrara, 2016). My research is the first to look at the direct impact of a role model for students on their educational attainment. I do this by randomly assigning secondary school

students preparing to take their national exams to see either a movie featuring a potential role model, *Queen of Katwe*, or a placebo movie, *Miss Peregrine's Home for Peculiar Children*. Those students who see the movie featuring the role model achieve 0.11 standard deviations higher scores on their lower secondary maths exam, and are 11 percentage points less likely to fail the exam. The impacts are even stronger for female students and those who were already performing poorly, as measured by a previous mock exam. Students taking their upper secondary leaving exams see a 0.13 standard deviation increase in their test scores, again stronger in female students.

A limitation of this study is its lack of mechanisms, so it is not possible to explain why the role model improved students' test scores. Potential explanations from other research include that the role model changed beliefs about what can be achieved by someone 'like them' (Bandura, 1977, Nguyen, 2008, Wilson, 1987) or that the role model raised aspirations (Beaman et al., 2012, Bernard et al., 2014). Future research would attempt to understand these channels. Additionally, my study was over a short time frame. While other studies have found evidence of persistent effects from role models (Bernard et al., 2017), long term impacts in my case depend on whether the improvement in test scores reflected a real increase in learning or perhaps just a temporary improvement in self belief which improved scores in one set of exams. Again, future research would examine how these impacts evolve over time.

* * *

Mobile money services have rapidly expanded since their launch in Kenya in 2007, both within Africa and worldwide. 66% of adults in East Africa now use mobile money actively (GSMA, 2017). Understanding the impact of new financial services like mobile money is important for both researchers and policy makers. Additionally, thinking through how to use these new services to improve the impact of existing services for the poor, such as microfinance loans, can bring large benefits to those populations.

Media-based interventions have the potential to reach more people at a large scale and low cost. The media is therefore increasingly being examined as a method to influence behaviours in a wide population (La Ferrara, 2016).

In this thesis I have touched upon both these themes: showing the benefits of mobile money services to users in terms of smoothing consumption: the benefits of mobile money accounts to female microentrepreneurs in enabling them to retain control over their loans; and the benefit of a movie-based role model to secondary school students who need to improve their exam results. These are exciting new areas which could benefit from a deeper understanding, and are important topics both for researchers, but also policy makers, going forwards.

References

Aker, J. C., R. Boumnijel, A. Mcclelland, and N. Tierney (2016). Payment Mechanisms and Anti-Poverty Programs : Evidence from a Mobile Money Cash Transfer Experiment in Niger. *Economic Development and Cultural Change 65*(1), 1–37.

Almas, I., P. Carneiro, O. Attanasio, A. Armand, O. Attanasio, P. Carneiro, and O. Attanasio (2015). Measuring and Changing Control : Women's Empowerment and Targeted Transfers. *Working paper*.

Andersen, S., G. W. Harrison, M. I. Lau, and E. E. Rutström (2008). Eliciting risk and time preferences. *Econometrica 76*(3), 583–618.

Anderson, M. L. (2008). Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects. *Journal of the American Statistical Association 103*(484), 1481–1495.

Aron, J. (2018). Mobile Money and the Economy: A Review of the Evidence. *TheWorld Bank Research Observer 33*(2), 135–188.

Ashraf, N. (2009). Spousal Control and Intra-Household Decision Making: An Experimental Study in the Philippines. *American Economic Review 99*(4), 1245–1277.

Ashraf, N., D. Karlan, and W. Yin (2006). Tying odysseus to the mast: Evidence from a commitment saving product in the Philippines. *Quarterly Journal of Economics 121*(2), 635–672.

Attanasio, O. and V. Lechene (2002). Tests of Income Pooling in Household Decisions. *Review of economic dynamics 5*(4), 720–748.

Baird, S., C. Mcintosh, and B. Özler (2016). When the Money Runs Out Do Cash Transfers Have Sustained Effects on Human Capital Accumulation? *World Bank Policy Research Working Paper*.

Baland, J.-M., C. Guirkinger, and C. Mali (2011). Pretending to Be Poor : Borrowing to Escape Forced Solidarity in Cameroon. *Economic Development and Cultural Change 60*(1), 1–17.

Bandura, A. (1977). *Social Learning Theory*. Englewood Cliffs, NJ: General Learning Press.

Banerjee, A., D. Karlan, and J. Zinman (2015). Six Randomized Evaluations of Microcredit: Introduction and Further Steps. *American Economic Journal: Applied Economics 7*(1), 1–21.

Banerjee, A., E. La Ferrara, and V. Orozco (2018). The Entertaining Way to Behavioral Change: Fighting HIV with MTV. *Working paper*.

Banerjee, A. V., S. Cole, E. Duflo, and L. Linden (2007). Remedying Education: Evidence from Two Randomized Experiments in India. *Quarterly Journal of Economics 122*(3), 1235–1264.

Basheka, B. C. (2013). The State of Higher Education and Training in 2013/14. Technical report.

Bastian, G., I. Bianchi, M. Goldstein, and J. Montalvao (2018). Short-Term Impacts of Improved Access to Mobile Savings, With and Without Business Training: Experimental Evidence from Tanzania. *Center for global development working paper 476*(March).

Batista, C. and P. C. Vicente (2014). Introducing Mobile Money in Rural Mozambique: Evidence from a Field Experiment. *NovaAfrica working paper series*.

Batista, C. and P. C. Vicente (2017). Improving Access to Savings through Mobile Money: Experimental Evidence from Smallholder Farmers in Mozambique.

Beaman, L., E. Duflo, R. Pande, and P. Topalova (2012). Female Leadership Raises Aspirations and Educational Attainment for Girls. *Science 335*(6068), 582–586.

Benjamini, Y., A. M. Krieger, and D. Yekutieli (2006). Adaptive Linear Step-Up Procedures that Control the False Discovery Rate. *Biometrika 93*(3), 491–507.

Bernard, T., S. Dercon, K. Orkin, and A. S. Taffesse (2014). The Future in Mind: Aspirations and Forward-Looking Behaviour in Rural Ethiopia. *CSAE Working Paper Series 44*(April).

Bernard, T., S. Dercon, K. Orkin, and A. S. Taffesse (2015). Will Video Kill the Radio Star? Assessing the Potential of Targeted Exposure to Role Models through Video. *World Bank Economic Review 29*, S226–S237.

Bernard, T., S. Dercon, K. Orkin, and A. S. Taffesse (2017). The Future in Mind : Long-Run Impact of an Aspirations Intervention in Rural Ethiopia. *Mimeo*.

Bernhardt, A., E. Field, R. Pande, N. Nigol, S. Schaner, and C. Troyer-moore (2018). Male Social Status and Women's Work. *AEA Papers and Proceedings 108*, 363–367.

Bernhardt, A., E. Field, R. Pande, and N. Rigol (2017). Household Matters: Revisiting the Return to Capital among Female Micro-Entrepreneurs. *NBER Working Paper No. 23358*.

Bettinger, E. P. (2012). Paying to Learn: The Effect of Financial Incentives on Elementary School Test Scores. *The Review of Economics and Statistics 94*(3), 686–698.

Blattman, C., N. Fiala, and S. Martinez (2014). Generating Skilled Self-Employment in Developing Countries: Experimental Evidence from Uganda. *The Quarterly Journal of Economics 129*(2), 697–752.

Blumenstock, J., M. Callen, and T. Ghani (2018). Why Do Defaults Affect Behavior? Experimental Evidence from Afghanistan. *American Economic Review 108*(10), 2868–2901.

Blumenstock, J. E., M. Callen, T. Ghani, and L. Koepke (2015). Promises and Pitfalls of Mobile Money in Afghanistan: Evidence from a Randomized Control Trial. *Proceedings of the Seventh International Conference on Information and Communication Technologies and Development. ACM.*.

Blumenstock, J. E., N. Eagle, and M. Fafchamps (2016). Airtime Transfers and Mobile Communications: Evidence in the Aftermath of Natural Disasters. *Journal of Development Economics 120*, 157–181.

Boltz, M., K. Marazyan, and P. Villar (2017). Income Hiding and Informal Redistribution: A Lab-in-the-Field Experiment in Senegal. *UMR working paper 4*.

Bruhn, M. and D. McKenzie (2009). In Pursuit of Balance: Randomization in Practice in Development Field Experiments. *American Economic Journal: Applied Economics 1*(4), 200–232.

Brune, L., X. Giné, J. Goldberg, and D. Yang (2017). Savings Defaults and Payment Delays for Cash Transfers: Field Experimental Evidence from Malawi. *Journal of Development Economics 129*, 1–13.

Brune, L., X. Giné World Bank, J. Goldberg, and D. Yang (2016). Facilitating Savings for Agriculture: Field Experimental Evidence from Malawi. *Economic Development and Cultural Change 64*(2), 197–220.

Castilla, C. (2018). What's Yours is Mine, and What's Mine is Mine: Field Experiment on Income Concealing Between Spouses in India. *Journal of Development Economics*, 1–38.

Chetty, R., J. Friedman, S. Leth-petersen, T. Heien nielsen, and T. Olsen (2014). Active vs. Passive Decions and Crowd-Out in Retirement Savings Accounts:

Evidence from Denmark. *The Quarterly Journal of Economics 129*(3), 1141–1219.

Choi, J. J., D. Laibson, B. C. Madrian, and A. Metrick (2004). For Better or for Worse: Default Effects and 401(k) Savings Behavior. In David Wise (Ed.), *Perspectives in the Economics of Aging*, pp. 81–121. University of Chicago Press.

Chong, A. and E. L. Ferrara (2009). Television and Divorce: Evidence From Brazilian Novelas. *Journal of the European Economic Association 7*(2-3), 458–468.

Claro, S., D. Paunesku, and C. S. Dweck (2016). Growth Mindset Tempers the Effects of Poverty on Academic Achievement. *Proceedings of the National Academy of Sciences 113*(31), 8664–8668.

Crawfurd, L. (2017). School Management in Uganda. *RISE working paper 13*.

Dalton, P. S., S. Ghosal, and A. Mani (2016). Poverty and Aspirations Failure. *Economic Journal 126*(590), 165–188.

Dasgupta, N. and S. Asgari (2004). Seeing is Believing: Exposure to Counterstereotypic Women Leaders and its Effect on the Malleability of Automatic Gender Stereotyping. *Journal of Experimental Social Psychology 40*(5), 642–658.

De Mel, S., C. Mcintosh, K. Sheth, and C. Woodruff (2018). Can Mobile-Linked Bank Accounts Bolster Savings? Evidence from a Randomized Controlled Trial in Sri Lanka. *NBER working paper series*.

De Mel, S., D. Mckenzie, and C. Woodruff (2008). Returns to Capital in Microenterprises: Evidence from a Field Experiment. *The Quarterly Journal of Economics 123*(4), 1329–1372.

de Quidt, J., J. Haushofer, and C. Roth (2018). Measuring and Bounding Experimenter Demand. *American Economic Review 108*(11), 3266–3302.

De Weerdt, J. and S. Dercon (2006). Risk-Sharing Networks and Insurance against Illness. *Journal of Development Economics 81*(2), 337–356.

Demirguc-kunt, A., L. Klapper, D. Singer, S. Ansar, and J. Hess (2017). *The Global Findex Database*.

Dennehy, T. C. and N. Dasgupta (2017). Female Peer Mentors Early in College Increase Women's Positive Academic Experiences and Retention in Engineering. *Proceedings of the National Academy of Sciences 114*(23), 5964–5969.

Dercon, S. (2002). Income Risk, Coping Strategies, and Safety Nets. *The World Bank Research Observer 17*(2), 141–166.

Dizon, F., E. Gong, and K. Jones (2017). The Effect of Savings on Informal Risk-Sharing : Experimental Evidence from Vulnerable Women in Kenya.

Duflo, E., R. Hanna, and S. P. Ryan (2012). Incentives Work: Getting Teachers to Come to School. *American Economic Review 102*(4), 1241–1278.

Dupas, P. and J. Robinson (2013a). Saving Constraints and Microenterprise Development: Evidence from a Field Experiment in Kenya. *American Economic Journal: Applied Economics 5*(1), 163–192.

Dupas, P. and J. Robinson (2013b). Why Don't the Poor Save More? Evidence from Heath Savings Experiments. *American Economic Review 103*(4), 1138–1171.

Dweck, C. S. (2000). *Self-Theories: Their Role in Motivation, Personality, and Development*. Pscyhology press.

Economides, N. and P. Jeziorski (2017). Mobile Money in Tanzania. *Marketing Science 36*(6).

Evans, D. K. and A. Popova (2015). What Really Works to Improve Learning in Developing Countries? An Analysis of Divergent Findings in Systematic Reviews. *World Bank Policy Research Working Paper*.

Fafchamps, M. and S. Lund (2003). Risk-Sharing Networks in Rural Philippines. *Journal of Development Economics 71*(2), 261–287.

Fafchamps, M., D. Mckenzie, S. Quinn, and C. Woodruff (2014). Microenterprise growth and the flypaper effect: Evidence from a randomized experiment in Ghana. *Journal of Development Economics 106*, 211–226.

Fiala, N. (2017). Business is tough, but family is worse: Household bargaining and investment in microenterprises in Uganda. *Working papers 2017-05, University of Connecticut, Department of Economics.*.

Field, E., R. Pande, J. Papp, and N. Rigol (2013). Does the Classic Microfinance Model Discourage Entrepreneurship among the Poor? Experimental Evidence from India. *American Economic Review 103*(6), 2196–2226.

Field, E., S. Schaner, and C. T. Moore (2016). An Account of One's Own: Can Targeting Benefits Payments Address Social Constraints to Female Labor Force Participation?

Finscope (2013). Finscope Tanzania 2013 report. Technical report.

Genicot, G. and D. Ray (2017). Aspirations and Inequality. *Econometrica 85*(2), 489–519.

Gertler, P. and J. Gruber (2002). Insuring Consumption against Illness. *American Economic Review 92*(1), 51–70.

Glewwe, P., M. Kremer, S. Moulin, and E. Zitzewitz (2004). Retrospective vs. Prospective Analyses of School Inputs: The Case of Flip Charts in Kenya. *Journal of Development Economics 74*(1), 251–268.

Goldberg, J. (2017). The Effect of Social Pressure on Expenditures in Malawi. *Journal of Economic Behavior and Organization 143*(C), 173–185.

Gollin, D. and C. Udry (2017). Heterogeneity, Measurement Error, and Misallocation: Evidence from African Agriculture. *Department of International Development, Oxford University, Mimeo*.

Green, M. C. and T. C. Brock (2000). The Role of Transportation in the Persuasiveness of Public Narratives. *Journal of Personality and Social Psychology 79*(5), 701–721.

Green, M. C., T. C. Brock, and G. F. Kaufman (2004). Understanding media enjoyment: The role of transportation into narrative worlds. *Communication Theory 14*(4), 311–327.

GSMA (2010). What makes a Successful Mobile Money Implementation? Learnings from M-PESA in Kenya and Tanzania. Technical report.

GSMA (2017). State of the Industry Report on Mobile Money: Decade Edition 2006-2016. Technical report.

Habyarimana, J. and W. Jack (2018). High Hopes: Experimental Evidence on Saving and the Transition to High School in Kenya.

Intermedia (2016). Uganda financial inclusion insights survey. Technical report.

International financial Coporation (2010). M-Money Channel Distribution Case – Tanzania Vodacom Tanzania M-PESA. Technical report.

Jack, W. and T. Suri (2011). Mobile Money: The Economics of M-PESA. *NBER Working Paper Series*.

Jack, W. and T. Suri (2014). Risk Sharing and Transactions Costs : Evidence from Kenya' s Mobile Money Revolution. *American Economic Review 104*(1), 183–223.

Jakiela, P. and O. Ozler (2015). Does Africa Need a Rotten Kin Theorem? Experimental Evidence from Village Economies. *Review of Economic Studies 83*(1), 231–268.

Jensen, R. (2000). Agricultural Volatility and Investments in Children. *American Economic Review 90*(2), 399–404.

Jensen, R. and E. Oster (2009). The Power of TV: Cable Television and Women's Status in India. *The Quarterly Journal of Economics 124*(3), 1057–1094.

Jensen, R. T. (2010). The (Perceived) Returns to Education and the Demand for Schooling. *The Quarterly Journal of Economics 125*(2), 515–548.

Kast, F. and D. Pomeranz (2014). Saving More to Borrow Less: Experimental Evidence from Access to Formal Saving Accounts in Chile. *NBER working paper series*.

Kavuma, S. N. (2014). Private Returns to Education for the Wage-employees and Self-employed in Uganda.

Kazianga, H. and C. Udry (2006). Consumption Smoothing? Livestock, Insurance and Drought in Rural Burkina Faso. *Journal of Development Economics 79*(2), 413–446.

Kearney, M. S. and P. B. Levine (2015). Media Influences on Social Outcomes: The Impact of MTV's 16 and Pregnant on Teen Childbearing. *American Economic Review 105*(12), 3597–3632.

Kinnan, C. (2017). Distinguishing Barriers to Insurance in Thai Villages.

Kremer, M., C. Brannen, and R. Glennerster (2013). The Challenge of Education and Learning in the Developing World. *Science 340*(6130), 297–300.

La Ferrara, E. (2016). Mass Media and Social Change: Can We Use Television to Fight Poverty. *Journal of the European Economic Association 14*(4), 791–827.

La Ferrara, E., A. Chong, and S. Duryea (2012). Soap Operas and Fertility: Evidence from Brazil. *American Economic Journal: Applied Economics 4*(4), 1–31.

Levitt, S. D., J. A. List, S. Neckermann, and S. Sadoff (2016). The Behavioralist Goes to School: Leveraging Behavioral Economics to Improve Educational Performance. *American Economic Journal: Economic Policy 8*(4), 183–219.

Lipscomb, M. and L. Schechter (2018, nov). Subsidies Versus Mental Accounting Nudges: Harnessing Mobile Payment Systems to Improve Sanitation. *Journal of Development Economics 135*, 235–254.

Locke, E. A. and G. P. Latham (2002). Building a Practically Useful Theory of Goal Setting and Task Motivation A 35-Year Odyssey. *American Psycologist 57*(9), 705.

Lockwood, P. (2006). "Someone Like Me Can Be Successfull": Do College Students Need Same-Gender Role Models? *Psychology of Women Quarterly 30*(1), 36–46.

Lockwood, P. and Z. Kunda (1997). Superstars and Me: Predicting the Impact of Role Models on the Self. *Journal of Personality and Social Psychology 73*(1), 91–103.

Lybbert, T. J. and B. Wydick (2018). Poverty, Aspirations, and the Economics of Hope. *Economic Development and Cultural Change 66*(4), 709–753.

Mani, A. (2011). Mine, Yours or Ours? The Efficiency of Household Investment Decisions: An Experimental Approach. *University of Warwick, Department of Economics CAGE Online Working Paper 64*.

Mayanja Lwanga, Annet Adong, M. (2016). A Pathway to Financial Inclusion: Mobile Money and Individual Savings in Uganda. Technical report.

Mbiti, I. M. and D. Weil (2011). Mobile banking: The impact of M-Pesa in Kenya. *National Bureau of Economic Research No. w17129*.

McKenzie, D. (2012). Beyond Baseline and Follow-up: The case for More T in Experiments. *Journal of development Economics 99*(2), 210–221.

MoES Uganda (2015). Education Statistical Abstract. Technical report.

Morawczynski, O. (2010). *Examining the Adoption, Usage and Outcomes of Mobile Money Services - The Case of M-PESA in Kenya*. The University of Edinburgh.

Moyer-Gusé, E. (2008). Toward a Theory of Entertainment Persuasion: Explaining the Persuasive Effects of Entertainment-Education Messages. *Communication Theory 18*(3), 407–425.

Munyegera, G. K. and T. Matsumoto (2016). Mobile Money, Remittances, and Household Welfare: Panel Evidence from Rural Uganda. *World Development 79*, 127–137.

Nguyen, T. (2008). Information, Role Models and Perceived Returns to Education: Experimental Evidence from Madagascar. *MIT Job Market Paper*.

O'Donoghue, T. and M. Rabin (1999). Doing It Now or Later. *American Economic Review 89*(1), 103–124.

Paluck, E. L. and D. Green (2009). Deference, Dissent, and Dispute Resolution: An Experimental Intervention Using Mass Media to Change Norms and Behavior in Rwanda. *American Political Science Review 103*(4), 622–644.

Paxson, C. H. (1992). Using Weather Variability to Estimate the Response of Savings to Transitory Income in Thailand. *American Economic Review 82*(1), 15–33.

Porter, C. and D. Serra (2017). Gender Differences in the Choice of Major: The Importance of Female Role Models. *Southern Methodist University, Department of Economics Departmental Working Papers 1705*.

Prina, S. (2015). Banking the Poor via Saving Accounts: Evidence from a Field Experiment. *Journal of Development Economics 115*, 16–31.

Ray, D. (2006). Aspirations, Poverty and Economic Change. *In Understanding Poverty. Oxford: Oxford University Press.*, 409–421.

Rosenzweig, M. R. (1988). Risk, Implicit Contracts and the Family in Rural Areas of Low-Income Countries. *The Economic Journal 98*(393), 1148–1170.

Rotter, J. B. (1966). Generalized Expectancies for Internal versus External Control of Reinforcement. *Psychological monographs: general and applied 80*(1), 1–28.

Schaner, S. (2015). Do Opposites Detract? Intrahousehold Preference Heterogeneity and Inefficient Strategic Savings. *American Economic Journal: Applied Economics 7*(2), 135–174.

Schaner, S. (2017). The Cost of Convenience? *Journal of Human Resources 52*(4), 919–945.

Schaner, S. (2018). The Persistent Power of Behavioral Change : Long-Run Impacts of Temporary Savings Subsidies for the Poor. *American Economic Journal: Applied Economics 10*(3), 1–66.

Shkaratan, M. (2012). Tanzania's Infrastructure A Continental Perspective. Technical report.

Somville, V. and L. Vandewalle (2018). Saving by Default: Evidence from a Field Experiment in India. *American Economic Journal: Applied Economics 10*(3), 39–66.

Stout, J. G., N. Dasgupta, M. Hunsinger, and M. A. McManus (2011). STEMing the Tide: Using Ingroup Experts to Inoculate Women's Self-Concept in Science, Technology, Engineering, and Mathematics (STEM). *Journal of personality and social psychology 100*(2), 255–270.

Suri, T. (2005). Spillovers in Village Consumption: Testing the Extent of Partial Insurance. *New Haven, Conn.: Yale University. mimeo.*

Thaler, R. (1985). Mental Accounting and Consumer Choice. *Marketing Science 4*(3), 199–214.

Thaler, R. H. (1999). Mental accounting matters. *Journal of Behavioral Decision Making 12*(3), 183–206.

Townsend, R. M. (1994). Risk and Insurance In Village India. *Econometrica 62*(3), 539–591.

Udry, C. (1994). Risk and Insurance in a Rural Credit Market: An Empirical Investigation in Northern Nigeria. *Review of Economic Studies 61*(3), 495–526.

USAID (2013). Tanzania Mobile Money Assessment and Case Study. Technical report.

Wilson, W. (1987). *The Truly Disadvantaged*. University of Chicago Press.

Yang, D. and H. J. Choi (2007). Are Remittances Insurance? Evidence from Rainfall Shocks in the Philippines. *World Bank Economic Review 21*(2), 219–248.