Information Theory and Banking

Markus Parlasca

Merton College



Thesis submitted in fulfilment of the requirements
for the degree of
Doctor of Philosophy in Economics

Department of Economics
University of Oxford

Trinity Term 2020

Information Theory and Banking

Markus Parlasca

Abstract

This thesis uses information theory to study policy questions in banking, finance, and industrial organization. It consists of four independent papers.

The first paper, "Time Inconsistency in Stress Test Design", shows that central banks face a time inconsistency problem when publishing bank stress test results. Before a stress test, they want to *appear* tough as the threat of letting banks fail the stress test incentivizes prudent behaviour. After the stress test, they want to *act* soft by releasing only partial information in order to reassure financial markets about bank health. The paper characterises an institutional design solution to this commitment problem: a social planner sets the framework within which the central bank communicates. We find that a hurdle rate framework, where all banks are judged to pass or fail relative to a common threshold, is optimal in many settings as it generates intermediate levels of both incentives and reassurance. We show that, in such a hurdle rate framework, stress tests become an informational contagion channel.

The second paper proposes a model of imperfect competition in markets with selection and investigates whether imperfect competition mitigates the inefficiency which, in these markets, arises under perfect competition. Focusing on markets with advantageous selection, such as consumer credit markets, we find that no degree of competition can restore the first best efficient allocation. We discuss policy alternatives.

The third paper shows that, in markets with adverse selection, tying contracts can prevent unravelling and thus have a substantial positive effect on welfare. We illustrate this effect for two insurance markets. When one characteristic (such as life-expectancy) is hidden information and causes agents to have high expected cost for one product (life insurance) but low expected cost for another (pension annuity), then for tying contracts cost differences partially cancel out. This can prevent unravelling.

In the fourth paper, which is joined work, we build on a field experiment, find that some agents increase effort when it becomes more costly, and rationalize this in a signalling model. In the field setting, agents also choose whether to truthfully disclose their effort. The data on this form of strategic communication adds further support to the interpretation that effort choices are driven by signalling motives.

To my parents.

Declaration

I declare that this thesis is my own work except where explicitly stated otherwise.

Chapters 2-4 are single authored papers. Chapter 3 builds on parts of my MPhil thesis. While the topic of imperfect competition in markets with selection and the Hotelling model approach already formed part of my MPhil thesis, chapter 3 in its current form differs substantially from my MPhil thesis. New content includes, but is not limited to, studying price discrimination in addition to studying gate prices, investigating the role of alternative policy remedies, and an updated literature review. Moreover, the chapter has been rewritten relative to my MPhil thesis. Chapter 2 and 4 were not part of my MPhil thesis.

Chapter 5 is co-authored and was not part of my MPhil thesis. I was responsible for the theoretical idea and the formalisation of the theoretical argument. I wrote the theory section of the paper and contributed to the overall write up. I was not involved in the empirical work or in conducting the field experiment.

Markus Parlasca
June 2020

Word count: 93 472 (508 words on representative page 6, 184 pages)

Acknowledgements

I am indebted to John Vickers for invaluable guidance and support. His mentoring was extraordinary. Thank you!

I benefited tremendously from discussions with members of the Micro Theory and the Finance Group at Oxford. Listing risks omission, but a special thanks goes to Alan Morrison for welcoming me into the Finance Group, and to Margaret Meyer for excellent comments at confirmation of status and for providing ample opportunities for me to present my work.

I was fortunate to meet many great young economists at Oxford. It was a pleasure to discuss ideas with them and reflect on research. This includes, but is not limited to, Jan David Bakker, Maria Balgova, George Charlson, Andrea de Mauro, Zoe Fannon, Alexis Grigorieff, Lukas Hensel, Ines Lee, Christopher Roth, Erkki Vihriälä, and Ben Waltmann.

I gratefully acknowledge financial support from the Economic and Social Research Council and Merton College, Oxford.

Contents

1 Introduction **1**

2 Time Inconsistency in Stress Test Design **4**

2.1 Introduction 5

2.2 Literature Review 7

2.3 Model 11

2.4 Reassurance 15

2.5 Incentives 22

2.6 Central Bank Design 27

2.7 Informational Contagion 32

2.8 Conclusion 33

2.9 Appendix 36

3 Imperfect Competition in Markets with Selection **55**

3.1 Introduction 56

3.2 Literature 60

3.3 The Model 65

3.4 Efficient Allocation 70

3.5 Perfect Competition and Monopoly 72

3.6 Imperfect Competition 75

3.7 Policy Remedies 83

3.8 Discussion 86

3.9 Conclusion 87

3.10 Appendix 88

4 Tying Contracts in Selection Markets: A Solution to the Unravelling Problem **104**

4.1 Introduction . . . 105

4.2 Literature Review . . . 107

4.3 A Model of Two Selection Markets . . . 113

4.4 Equilibrium Purchasing Behaviour . . . 116

4.5 Welfare . . . 124

4.6 Conclusion . . . 132

4.7 Appendix . . . 133

5 Signalling Through Costly Effort Provision: Theory and Evidence from the Field **155**

5.1 Introduction . . . 156

5.2 Signalling Model . . . 158

5.3 Field Evidence . . . 162

5.4 Conclusion . . . 165

5.5 Theoretical Appendix . . . 167

5.6 Main Figures . . . 168

5.7 Main Tables . . . 170

5.8 Appendix Tables . . . 175

6 Conclusion **184**

7 Bibliography **185**

Chapter 1

Introduction

Firms operating in the financial sector have adjusted to information asymmetries. Banks address moral hazard and adverse selection problems by screening and monitoring loan applicants, insurance providers use co-insurance and deductibles, and shareholders have a right to truthful and timely disclosure of financial information. However, in many settings, it is an open question how public policy should respond to information asymmetries and to firms' responses caused by these information asymmetries.

Central bank communication: While many public authorities are required to be transparent and reveal information to the public, it is not clear whether this should also apply to central banks. Should central banks fully disclose all information they have on bank health and financial stability even when this might trigger bank runs and could further undermine bank health?

Competition policy in markets with selection: While competition policy generally aims to achieve high levels of competition, it has been argued that in markets with advantageous selection, such as credit markets, competition policy can be harmful (Mahoney and Weyl, 2017). This argument builds on the result that perfectly competitive markets with advantageous selection can experience inefficient overprovision. Should competition policy therefore aim to mitigate overprovision by tolerating intermediate degrees of competition with the resulting quantity reduction? Can such a policy achieve the first best efficient allocation, or do alternative more targeted policies exist?

Tying contracts in markets with selection: A firm's decision to condition the sale of one product on the sale of another (tying, bundling) can indicate the presence of market power and the need for intervention by competition authorities. Is there an alternative rationale for why firms might use tying contracts when these contracts cover markets with adverse selection? Can tying contracts prevent unravelling and thus be socially valuable?

This thesis uses information theory to study these policy questions. The thesis consists of four independent papers. Chapters 2 - 4 are single authored theory papers which study central bank communication about financial stability, imperfect competition and competition policy in markets with selection, and the social benefit of tying contracts

when they span two markets with adverse selection. One example for tying contracts are insurance contracts which combine life-insurance features with a pension annuity. In addition to these three long single-authored theory chapters, chapter 5 is a short paper which uses information theory to rationalize behaviour of individuals which was observed in a field experiment. In this project, co-authored with empirical researchers (Lukas Hensel, Johannes Hermle, Christopher Roth, Anselm Hager), we show that some agents (political activists) increase their effort (door-to-door canvassing) when it becomes more costly (bad weather) and interpret this observation in a signalling model.

Chapter 2 studies central bank communication about the stability of banks. We show that central banks face a time inconsistency problem when publishing bank stress test results. Before a stress test, they want to *appear* tough as the threat of letting banks fail the stress test incentivizes prudent behaviour. After the stress test, they want to *act* soft by releasing only partial information in order to reassure financial markets about bank health. We characterize an institutional design solution to this commitment problem: a social planner sets the framework within which the central bank communicates. We find that a hurdle rate framework, where all banks are judged to pass or fail relative to a common threshold, is optimal in many settings as it generates intermediate levels of both incentives and reassurance. With a hurdle rate framework, stress tests become an informational contagion channel, as changes in one bank's fundamental health affect the perceived health of other banks. Thus, informational contagion can be a feature of a socially optimal institutional design in the presence of a time inconsistency problem.

Chapter 3 proposes a model of imperfect competition in markets with selection. While we know that in markets with advantageous selection, such as consumer credit or insurance markets, perfect competition can lead to inefficient overprovision and that monopoly results in underprovision, we do not know whether intermediate degrees of competition achieve the efficient allocation. We find that no degree of competition can restore the first best efficient allocation. Overprovision induced by advantageous selection and underprovision caused by market power do not cancel out perfectly. Instead, they co-exist for different segments of the consumer taste spectrum. However, we show that the efficient allocation can be reached through a combination of a corrective tax and tough competition policy. This suggests that if we view the build up of credit prior to the recent financial crisis as inefficient overprovision in a market with advantageous selection, then we should not conclude that competition in credit markets was "excessive", but rather conclude that we lacked the appropriate corrective taxation. Overall, our paper cautions against viewing imperfect competition as a solution to the inefficiency introduced by advantageous selection.

Chapter 4 shows that, in markets with adverse selection, tying contracts can prevent unravelling and thus have a substantial positive effect on welfare. We illustrate this effect for two insurance markets. When one characteristic (such as life-expectancy) is hidden information and causes agents to have high expected cost for one product (life insurance) but low expected cost for another (pension annuity), then for tying contracts

cost differences partially cancel out. Thus, cost heterogeneity is lower for the tying contract than for individual insurance contracts. This can prevent unravelling. The largest welfare benefit occurs if without tying contracts both individual markets unravel but with tying contracts all agents are insured against both risks.

Chapter 5 proposes a model of signalling which endogenizes agents' decisions on whether to communicate their costly effort to signal their unobserved type. The model predicts that types that have high reputational concerns and high motivation *increase* their effort and the likelihood of sharing their effort choices, when the action becomes *more* costly. We test the predictions of the model using novel field data from an online application which records canvassing effort and the communication of canvassing effort among a sample of political activists. As predicted by our model, activists who revealed to have reputational concerns by having previously communicated their effort level on social media, increase their canvassing effort if effort becomes more costly, while the opposite pattern holds for individuals without prior communication of effort choices. Furthermore, canvassers with prior communication of effort are more likely to communicate their effort if it was more costly. Our paper provides theoretical and empirical insights on agents' strategic decision on whether to communicate their costly effort choices in order to signal their unobserved types.

Appendices are given after each chapter. In order to avoid duplication of references, a bibliography covering all chapters is at the end of the thesis.

Chapter 2

Time Inconsistency in Stress Test Design

We show that central banks face a time inconsistency problem when publishing bank stress test results. Before a stress test, they want to *appear* tough as the threat of letting banks fail the stress test incentivizes prudent behaviour. After the stress test, they want to *act* soft by releasing only partial information in order to reassure financial markets about bank health. We characterize an institutional design solution to this commitment problem: a social planner sets the framework within which the central bank communicates. We find that a hurdle rate framework, where all banks are judged to pass or fail relative to a common threshold, is optimal in many settings as it generates intermediate levels of both incentives and reassurance. With a hurdle rate framework, stress tests become an informational contagion channel, as changes in one bank's fundamental health affect the perceived health of other banks. Thus, informational contagion can be a feature of a socially optimal institutional design in the presence of a time inconsistency problem.

Keywords: *Bank Stress Tests, Strategic Communication, Contagion, Central Bank Design, Financial Stability*

JEL Classification: D83, E58, G21

2.1 Introduction

"*Had we been fully open and fully transparent about what was going on during the financial crisis, it would, let me tell you, have been a lot, lot worse. That would have been shouting 'fire' in the theatre.*"

— Andy Haldane, Chief Economist, Bank of England[1]

"*The stress test has been a catalyst for pressure to raise capital.*"

— Andrea Enria, Chair of the ECB Supervisory Board[2]

Following the recent financial crisis, the frequent publication of bank stress test results has become the centrepiece of central bank communication about bank stability. While central banks view the communication of bank stress tests as serving two purposes – incentivize banks to take prudent actions and reassure financial markets about bank stability – the optimal stress test design literature has largely focused on the latter (Goldstein and Leitner, 2018; Inostroza and Pavan, 2018).

This paper studies stress tests as serving both purposes. We find that a trade-off between these two purposes exists and that they create a time inconsistency problem which results in suboptimal combinations of incentives and reassurance. The trade-off arises because incentives for banks to improve their health are created by the central bank's threat to reveal bank weakness, which to financial markets is the opposite of a reassuring message. Additionally, these two roles create a time-inconsistency problem: Before the stress test, when banks decide whether to take prudent actions, the central bank wants to *appear* tough, i.e. create a threat of revealing bank weakness. After the test, when banks cannot alter their health anymore, the central bank wants to *act* soft, i.e. not reveal bank weakness. Rational agents anticipate this. Banks know that the threat is not credible and therefore do not take prudent actions.

We propose an institutional design solution to this time inconsistency problem. A social planner chooses an institutional design before banks can take prudent actions. After banks have taken prudent actions and having observed bank health, the central bank (CB) communicates stress test results within the constraints set by the institutional design. A design consists of a communication framework which is either full disclosure, zero disclosure, or a hurdle rate framework where all banks are judged to pass or fail relative to a common level of stress.[3] A design also specifies the CB's objective function (henceforth: mandate). This approach echoes Rogoff's (1985) work on time inconsistency in monetary

[1]Quoted in *The Times*, 6 October 2017. Andy Haldane was the Bank of England's Executive Director for Financial Stability during the financial crisis.

[2]Quoted in *Financial Times*, 15 July 2011. Andrea Enria was Chairperson of the European Banking Authority from 2011 to 2018.

[3]A full disclosure framework is currently used in the Euro Area whereas a hurdle rate framework is used by the Bank of England.

policy.

We find that a hurdle rate framework is optimal for a large set of parameters as it provides an intermediate combination of both incentives and reassurance. A full disclosure framework generates strong incentives but no reassurance. The reverse is true for a zero disclosure framework.

We show that, in a hurdle rate framework, *strategic delegation* provides additional benefits. This means the planner maximises ex-ante welfare by setting the central bank's mandate to differ from ex-post welfare maximisation. The optimal mandate can differ in either direction. If taking prudent actions has high costs for banks, the planner optimally specifies the CB's mandate to be "tougher" than welfare maximisation, so that the CB is less concerned about supporting weak banks than society. This makes it credible that the CB will reveal bank weakness by letting weak banks fail the stress test and thus generates strong incentives. If the costs of prudent actions are low, the optimal mandate is "softer" than welfare maximisation, so that the CB is very concerned about weak banks. This achieves high reassurance while still generating sufficient incentives to induce the prudent action.

Stress tests become an informational contagion channel when a hurdle rate framework is used. This arises because in some cases the central bank's optimal response to a deterioration in the health of one bank is to lower the severity of stress so that all banks continue to pass. For all banks, the resulting pass results are interpreted less favourably than at the initial higher severity of stress. Thus, a change in one bank's health leads the CB to send a different message, which in turn affects beliefs financial markets hold about all banks. More surprisingly, there are also cases where a deterioration of one bank's health improves the beliefs about other banks. This is the case when the CB raises the level of stress. As a result, the weak bank fails the test and beliefs about its health deteriorate. However, beliefs about other banks improve as they pass a tougher test. In a hurdle rate framework, both of these contagion patterns arise for all sender mandates, including "tough" mandates which focus on the average health in the banking system and place no extra weight on supporting weak banks.

Informational contagion can thus be a feature of an optimal design in the presence of a time-inconsistency problem; contagion need not reduce welfare. The mechanism that makes the hurdle rate framework optimal for some parameters is the same mechanism that turns stress tests into an informational contagion channel. In a hurdle rate framework, the CB is constrained by having to judge all banks against one common level of stress. From the planner's perspective, this constraint is valuable, as it results in some but less than full revelation of bank health and thus achieves intermediate levels of both incentives and reassurance. Without the constraint, there would be no contagion but also no resolution to the time inconsistency problem.

Our examination of stress tests has broader implications for economic theory. We contribute to the strategic communication literature by characterizing the equilibrium in a

verifiable disclosure communication game with a cross-message constraint. In these games, the sender communicates about multiple banks, truthfully releases pass or fail results for each of them, but is constrained by having to judge all banks relative to a common level of stress, rather than using bank specific stress levels. The sender therefore faces a trade-off: While higher levels of stress result in more favourable messages for those banks that pass, they have the cost of more banks failing the test which is a very unfavourable message.

Our paper also contributes to the literature on incentives generate by communication, which has commonly been framed in the context of teachers grading exams (Boleslavsky and Cotton, 2015; Dubey and Geanakoplos, 2010). Our model can be reinterpreted in that setting: A class teacher (central bank) wants to incentivize students (banks) to study for an exam (take a prudent action) by threatening to let bad students fail the test. However, once the exam has been taken, the teacher wants to give out good marks to help students get good jobs. A head teacher (planner) tries to solve this time inconsistency problem by specifying the school's grading policy (stress test design).[4]

The remainder of this paper is structured as follows: Section 2.2 reviews the related literature. Section 2.3 outlines the model and discusses key assumptions. Section 2.4 focuses on reassurance. First, it takes an ex-post perspective and characterises how much reassurance a central bank achieves for given health outcomes in equilibrium in a given design. For a hurdle rate framework, this involves solving the verifiable disclosure game with a cross-message constraint. Then, it takes the planner's ex-ante perspective and characterises how much reassurance different designs achieve in expectation, i.e. across possible health outcomes. Section 2.5 compares designs in terms of incentives. Section 2.6 combines incentive and reassurance concerns to solve for the optimal institutional design. Section 2.7 shows that a hurdle rate framework implies that stress tests become an informational contagion channel. Section 2.8 concludes.

2.2 Literature Review

This paper is related to three different strands of the literature. Our idea of studying optimal institutional design solutions to a time inconsistency problem is related to the monetary economics and particularly the "central bank design" literature (Reis, 2013). Rather than focusing on a monetary policy maker who is tempted to boost output by creating surprise inflation we focus on a bank regulator tempted to shield weak banks from market pressure by not releasing information. Our modelling framework for this communication game builds on and contributes to the theoretical literature on strategic communication. Our topic of bank stress tests is rooted in the central bank communication literature.

[4]This paper differs from Boleslavsky and Cotton (2015) in not assuming commitment of the teacher (central bank) and in focusing on effort of a third party (students, banks) rather than the sender. It differs from Dubey and Geanakoplos (2010) in focusing on effort prior to and thus in anticipation of communication rather than afterwards.

Central Bank design and strategic delegation

In monetary economics, a time inconsistency problem exists when the policy maker would like to generate surprise inflation to increase output (Lucas, 1972, 1975; Kydland and Prescott, 1977; Barro and Gordon, 1983a,b). Agents with rational expectations anticipate this and therefore expect inflation to be so far above target that the monetary policy maker is indifferent between the cost of surprise inflation and the benefit of output gains. The result is "inflation bias": an equilibrium with inflation above target and no output gains.[5]

Institutional design solutions which reduce inflation bias range from strategic delegation to performance contracts. Rogoff (1985) argues that delegating policy to a "conservative central banker" who places more weight than society on inflation stabilisation relative to output stabilisation reduces the inflation bias and thus increases welfare. Walsh (1995) uses insights from principal agent theory to argue that inflation bias can be reduced by a performance contract which rewards the central banker for low inflation.

Both institutional design solutions have the common feature that the objective function of the policy maker which maximises ex-ante welfare differs from ex-post welfare maximisation.[6] The general benefits of strategic delegation were outlined by Vickers (1985) who shows that in a strategic context the highest pay-off for the principal is achieved by an agent who maximises an objective function other than the principal's pay-off. Specifically, in Cournot competition profit increases when decisions are delegated to a manager who places weight on market share, not just on profits, as this commits the firm to aggressive behaviour which in turn leads rivals to cut output. Effectively, strategic delegation allows the firm to become a Stackelberg leader.

We contribute to the CB design literature in three ways. We take CB design and delegation ideas to a new setting: focusing on communication rather than actions; on financial stability rather than monetary policy. In this setting, we formalise the existence of a time inconsistency problem, which to the best of our knowledge is a novel problem. By characterising solutions to this problem, we show benefits of delegation in a strategic communication game rather than a strategic action setting. We now turn to the communication literature on which our model builds.

[5]The CB is minimising losses which are quadratic in the deviation of inflation and output from target where the target level of output exceeds the natural rate. Then, if the economy were at the target level of inflation and at the natural rate of output, the CB would strictly prefer to generate surprise inflation. Due to quadratic losses, there exists a higher level of inflation at which the CB is indifferent between further inflation and the output benefit. This is the equilibrium.

[6]In Rogoff (1985), the objective function comes from intrinsic types. In Walsh (1995), it results from a performance contract.

Strategic communication theory

Communication theory models can be grouped according to two dimensions: Is the sender restricted to telling the truth, or is he allowed to lie? Does the sender communicate strategically based on private information or does he commit to a messaging rule before he becomes privately informed? In verifiable disclosure models (Grossman, 1981; Milgrom, 1981) a privately informed sender needs to state the truth but not the full truth ("$x \geq 4$" rather than "$x = 4$" is possible). In cheap talk models (Crawford and Sobel, 1982) a privately informed sender can lie. In models of Bayesian persuasion (Kamenica and Gentzkow, 2011) the sender commits to a messaging rule before becoming privately informed. A recent survey of the literature based on commitment is Bergemann and Morris (2019).

We assume that the regulator has inside information on bank resilience when deciding on the severity of the stress test and cannot lie. Hence, our model is based on verifiable disclosure. Readers more familiar with the macroeconomic literature may equally view our model as "communication with discretion", i.e. where the regulator is not able to commit to a messaging rule. The opposite case of "communication with commitment" would be covered by a model of Bayesian persuasion. This paper contributes to the theoretical literature by solving a verifiable disclosure model with a cross-message constraint, i.e. where the privately informed sender truthfully communicates about two banks but where the pass or fail message for each bank is relative to a common level of stress.

Central bank communication

The literature on Central Bank communication can be divided according to a CB's two objectives: monetary and financial stability. For the former, which is not directly related to our paper, surveys are provided by Geraats (2002) and Blinder et al. (2008). Notable theoretical contributions include models of global games (Carlsson and van Damme, 1993; Morris and Shin, 2002; Svensson, 2006; Morris et al., 2006; Angeletos and Pavan, 2007) which show that coordination motives may lead agents to place too much weight on public relative to private messages;[7] and more recently Morris and Shin (2018)[8] who study public disclosure in "echo chambers", i.e. when the policy maker is trying to infer information from a price which is itself affected by the policy maker's communication.

We focus on CB communication about financial stability and partition this literature further into two subfields: cases where the regulator cannot commit to a communication strategy (communication with discretion, i.e. strategic communication) and cases where the regulator is assumed to be able to commit (Bayesian persuasion). We first focus on the case with discretion, then review the literature with commitment.[9]

[7] Specific applications to communication about monetary policy include Morris and Shin (2005, 2007).

[8] Similarly Goldstein and Yang (2019).

[9] A broad survey of arguments for and against disclosure of stress test results is provided by Goldstein and Sapra (2013). Relative to their review, we add a review of the more recent literature and focus in particular on

No commitment

Our key contributions to the hitherto small literature on the case without commitment are that we consider the incentive motive of stress test disclosure in addition to the reassurance motive and that we model the complexity arising from testing multiple banks simultaneously with a common stress scenario.

Bouvard et al. (2015) consider a regulator who communicates only because of the reassurance motive and faces a binary choice between full and zero disclosure. Intermediate levels of informativeness such as a hurdle rate framework are not considered. They find that regulators disclose information only in crisis when average health is so low that absent communication all banks would face a run. Disclosure then limits the run to those banks with health actually below the threshold that triggers runs.

Shapiro and Zeng (2018) argue that testing one bank repeatedly creates a trade-off: the regulator wants to acquire a reputation for being weak to encourage lending, but also for being tough to discourage excessive risk taking.[10] We focus on two different trade-offs: the time-inconsistency problem created by the intertemporal conflict between the incentive and the reassurance motives; and the conflict of testing multiple banks simultaneously against the same stress scenario. Then higher levels of stress are favourable for those banks who continue to pass but come at the cost of failing other banks.

Commitment

Several papers study stress tests when the regulator can commit to a communication rule before observing bank health (Bayesian persuasion) and when the sole purpose of stress tests is to reassure financial markets. Similar to Bouvard et al. (2015), Goldstein and Leitner (2018) find that the regulator will only release information when the average bank is perceived to be so low that, absent communication, all banks would experience a run. Goldstein and Leitner's regulator can send partially pooling messages, e.g. reporting the average health of a pool of banks. This saves at least some banks from the run.[11] However, this is at odds with current communication frameworks in which central banks have committed to regularly running and disclosing stress test results. Silence thus seems more likely to be interpreted as a bad signal which is a feature of our model. While Goldstein and Leitner's regulator can send any partially pooling message, including averages, we restrict attention to messages about the minimum health of a group. This rules out that banks with high and low health pass while medium banks fail the stress test.

Inostroza and Pavan (2018) add an explicit model of how receivers coordinate by introducing additional heterogeneous private information. Parlatore (2015) views stress tests as generating *private* signals in a Diamond and Dybvig (1983) model of bank runs.

the assumptions on communication, specifically whether it is strategic or not. This emphasize is, to the best of our knowledge, a novel contribution of our literature review.

[10]Their model is based on the implicit assumption that less well capitalised banks can lend more which is contested. For empirical evidence of the contrary see e.g. Gambacorta and Shin (2018).

[11]Goldstein and Leitner frame their paper as studying how disclosure affects risk sharing among banks (the Hirshleifer effect (1971)) which is related to our term "reassurance".

Faria-e-castro et al. (2016) study how a governments fiscal position and disclosure policy interact: When govt funds are available to bail-out weak banks, more information can be released. In their model, this reduces actual spending on bail-outs. Orlov et al. (2017) endogenise asset prices and reach the opposite conclusion of Goldstein and Leitner (2018): They argue that the regulator wants to be uninformative when bank health is low. If the regulator revealed capital shortfalls, fear of fire sales in the future would depress asset prices today, making it more costly for banks to recapitalise.

Williams (2017) also models a regulator as able to commit and as disclosing information only to avoid runs[12] but considers an extension in which banks can adjust to the communication rule. Williams (2017) finds that banks reduce their liquidity buffers when stress tests are conducted since pass grades make runs less likely and thus act as substitute for liquidity buffers. This is a response to communication to which the regulator is committed. We study the effect of anticipated *strategic* communication to which the regulator is not committed.[13]

We now turn to our model of stress tests by a regulator who cannot commit.

2.3 Model

Our model decomposes stress test design into two steps which are potentially carried out by different agents thus capturing potential benefits of delegation. First, a welfare maximising institution designer D (parliamentary committee, social planner) decides on a framework in which stress tests are conducted. Then, the CB or regulator communicates stress test results in the set framework to financial markets who update their belief on bank health. In between, banks can take a prudent action which improves their health and the CB observes health. We show that the difference in timing and information between the moves of D and CB creates a time-inconsistency problem: While a welfare maximiser wants to appear tough at the design stage to induce prudent actions by banks, a welfare maximiser wants to act soft at the communication stage to reassure financial markets as bank health is fixed. We show that welfare is maximised by delegating stress tests to a CB with a mandate that differs from welfare maximisation as this generates gains from strategic delegation. We are particularly interested in the difficulties and opportunities that arise from subjecting multiple banks to the same stress test simultaneously.

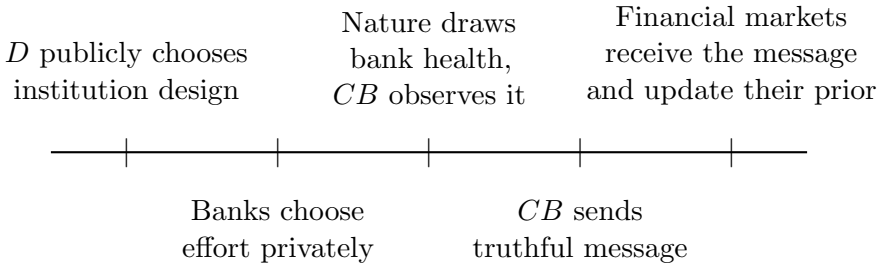
[12] He models runs explicitly based on Allen and Gale (1998).

[13] An additional difference is that Williams (2017) views stress tests as having probabilistic outcomes while we view them as deterministic. In Williams (2017), a stress test results in a public message with an error probability. The regulator commits to an error probability. In our setting stress test messages are always correct, but need not be perfectly informative.

2.3.1 Description of the Model

At the first stage, an institution designer (D) chooses the central bank's (CB) or regulator's mandate and specifies how it should communicate stress test results. Based on this stress test framework, banks decide whether to take a prudent action or not. This prudent action makes higher health realisations more likely. The health outcomes ($H_i \in [0,1]$) are privately observed by the regulator who then runs a stress test maximising his mandate and publishes results truthfully. Financial markets receive these messages and update their belief on bank stability. The sequence of moves is summarised in Figure 2.1 below:

Figure 2.1: Order of Moves



At the first stage, an institutional designer D chooses an institutional design $\mathcal{D} = \{\mathcal{C}, \mathcal{M}\}$ which consists of a communication framework $\mathcal{C}$ and a mandate for the central bank $\mathcal{M}$. The set of possible communication frameworks includes a zero disclosure framework (ZDF), in which never any information is released, a full disclosure framework (FDF), in which each bank's health H_i is always fully revealed, and a hurdle rate framework (HRF) in which all banks are judged to pass or fail relative to a common level of stress. We focus on the case with two banks which is the simplest possible case that allows us to capture three general regimes: zero hurdles (ZDF), fewer hurdles than banks (HRF), at least as many hurdles as banks (FDF).[14] A mandate $\mathcal{M}$ is a mapping from a pair of beliefs about bank health $(\mu_1, \mu_2) \in [0,1]^2$ to a value on the real line $V \in \mathbb{R}^1$, i.e. $\mathcal{M} : (\mu_1, \mu_2) \to V$. The designer is trying to maximise welfare $\mathcal{W} : (\mu_1, \mu_2) \to W$ where $W \in \mathbb{R}^1$. $\mathcal{D}$ becomes public knowledge.

Banks simultaneously make a private choice whether to take a prudent action or not. This can capture improved risk management practices or other unobserved aspects affecting bank resilience. In line with the moral hazard literature, we formalise this as a choice between high and low effort $e_i \in \{l, h\}$. High effort makes high health realisations more likely, formalised by a monotone likelihood ratio property (MLRP)[15] and is costly $C(h) > C(l)$. The bank maximises expected profit $\mathbb{E}[\pi] = \mathbb{E}[\mu_i] - C(e_i)$. A bank's strategy is thus: $\mathcal{B}_i : \mathcal{D} \to e_i$.

Bank health realisations H_1, H_2 are drawn by nature from the distributions implied by

[14] The formal proof that a framework with two hurdles results in full disclosure is given in Appendix 2.9.1.
[15] Milgrom (1981); Holmstrom (1979, 1982); Lambert (1983); Ottaviani and Sorensen (2006).

e_i and observed only by the central bank and not by financial markets. Bank health is independent across banks.

The central bank privately observes bank health $H_i \, \forall \, i$, runs a stress test and communicates the results in the framework $\mathcal{D}$. In ZDF and FDF this is a mechanical exercise. In HRF, the central bank chooses a level of stress s strategically to maximise its mandate $\mathcal{M}$ and must report truthfully whether each bank passed or failed relative to that common threshold s, which is also reported. The resulting message is a triplet $\{s, o_1, o_2\}$ where $s \in [0, 1]$ and $o_i \in \{p, f\}$ is the outcome for bank B_i. For example $\{0.2, p, f\}$ corresponds to "*Stress was* 0.2 *and* B_1 *passed while* B_2 *failed.*". This does not only reveal the number of pass and fail marks but also reveals which bank passed or failed.

In the pure communication game, where we treat $\mathcal{D}$ and e_1, e_2 as exogenously fixed, a central bank's messaging rule S is a mapping from any health state pair $(H_1, H_2) \in [0, 1]^2$ to a message $\{s, o_1, o_2\} \in [0, 1]^4$ subject to the truth-telling constraint that $o_i = p$ if and only if $H_i \geq s$, otherwise $o_i = f$, i.e. $S : (H_1, H_2) \rightarrow \{s, o_1, o_2\}$. The sender cannot lie, but he could obfuscate by sending a universally true message such as $\{0, p, p\}$. In the overall institutional design game, the central bank's strategy $\mathcal{S}$ is a profile of such messaging rules, i.e. $\mathcal{S} : \mathcal{D} \rightarrow S$.

Financial markets do not observe bank health H_i directly but observe the CB's message $\{s, o_1, o_2\}$ and update their prior belief accordingly. We denote the resulting posterior belief, i.e. expected bank health based on all information available to financial markets, as $\mu_i \in [0, 1] \, \forall \, i$.[16] Financial market participants are interested in bank health because they trade bank shares and credit default swaps or because they assess the risk of interbank loans. They are thus minimising a quadratic loss problem. In the pure communication game, we say that receivers (financial markets) follow an interpretation rule $R : \{s, o_1, o_2\} \rightarrow (\mu_1, \mu_2)$. In the overall institutional design game, the receiver's strategy $\mathcal{R}$ specifies an R for every design $\mathcal{D}$, i.e. $\mathcal{R} : \mathcal{D} \rightarrow R$. Based on the beliefs (μ_1, μ_2) pay-offs are realised.

2.3.2 Equilibrium Concept

We focus on perfect Bayesian equilibria (PBE) as solution concept. This means that we restrict our attention to those weak perfect Bayesian equilibria of the overall institutional design game which are weak perfect Bayesian equilibria in every subgame. To build an understanding of this equilibrium concept, we first define weak PBE in the pure communication game and then define PBE in the overall institutional design game.

[16] Our focus on expected bank health is in line with the stress test literature, e.g. Goldstein and Leitner (2018). However, it abstracts from potential effects of risk aversion by financial market participants. That means that in our model, as in the literature, we do not distinguish between a case where financial markets know with certainty that bank health is exactly 0.5 and a case where they know that it is anywhere on $[0, 1]$ and drawn from a uniform distribution.

Definition 2.1 *A weak perfect Bayesian equilibrium in the communication game is a set* (S, R) *such that:*

- *Given* R*;* S *maximises* V *at every health state pair* (H_1, H_2)*.*
- *Given* S*;* R *is correct in the sense of using all available information.*

In the overall game, every institutional design $\mathcal{D}$ corresponds to a subgame. The only additional subgame is the overall game. Hence, our solution concept means:

Definition 2.2 *A perfect Bayesian equilibrium (PBE) of the overall institutional design game is a set of strategy profiles* $(\mathcal{D}, \mathcal{B}_1, \mathcal{B}_2, \mathcal{S}, \mathcal{R})$ *such that*

- *Given* $\mathcal{B}_1, \mathcal{B}_2, \mathcal{S}, \mathcal{R}$*;* $\mathcal{D}$ *maximises expected welfare.*
- *For every* $\mathcal{D}$
  - *Given* $\mathcal{B}_{-i}$*,* $\mathcal{S}$*,* $\mathcal{R}$*;* $\mathcal{B}_i$ *maximises expected profits*
  - *Given* $\mathcal{B}_1, \mathcal{B}_2, \mathcal{R}$*;* $\mathcal{S}$ *maximises* V *at every health state pair* (H_1, H_2)*.*
  - *Given* $\mathcal{B}_1, \mathcal{B}_2, \mathcal{S}$*;* $\mathcal{R}$ *is correct in the sense of using all available information.*

2.3.3 Assumptions

Welfare To capture the reassurance and incentive motive of stress tests, we assume that welfare is increasing in beliefs about bank health ($\frac{\partial W}{\partial \mu_i} > 0 \forall i$) and is particularly concerned with the weak bank as the risks of belief driven phenomena such as runs, higher funding costs, or problems to roll over short term debt are particularly acute for weak banks. Thus, we assume that welfare is concave. We assume that banks enter welfare symmetrically $W(\alpha, \beta) = W(\beta, \alpha)$.

An example welfare function is $W(\mu_1, \mu_2) = \lambda \ \min(\mu_1, \mu_2) + (1 - \lambda) \ \frac{\mu_1 + \mu_2}{2}$ where $\lambda \in [0, 1[$ indexes the concern for the weak bank. While fundamentals H_i do not enter welfare directly, they enter indirectly via beliefs μ_i. The designer knows that an equilibrium property of Bayesian beliefs is $\mathbb{E}[\mu_i] = \mathbb{E}[H_i]$[17] and thus that he can only achieve high expected beliefs by increasing expected fundamentals.

Mandate The mandate is endogenously chosen by the designer. The mandate need not be identical to welfare but we restrict our attention to the same set of functions. Thus, as example we let $V(\mu_1, \mu_2) = \omega \ \min(\mu_1, \mu_2) + (1 - \omega) \ \frac{\mu_1 + \mu_2}{2}$ where $\omega \in [0, 1[$. We refer to the CB as being "tougher" than welfare if $\omega < \lambda$.

Bank's action The bank's action affects the distribution its health is drawn from. In line with the moral hazard literature, we assume that the improvement is according to

[17] This is referred to as 'Bayes plausibility'.

MLRP[18]. Denoting the pdf resulting from low effort as $g(H)$ and from high effort as $f(H)$, this assumption can be stated as:

$$\frac{f(x_0)}{g(x_0)} \leq \frac{f(x_1)}{g(x_1)} \quad \forall \quad x_0 \leq x_1 \tag{2.1}$$

This assumption is equivalent to assuming that the conditional distribution over every subinterval can be ranked according to first-order stochastic dominance (Shaked and Shanthikumar, 2007) and therefore implies that:

$$\mathbb{E}_G[H \mid a \leq H \leq b] \leq \mathbb{E}_F[H \mid a \leq H \leq b] \quad \forall \quad a \leq b \tag{2.2}$$

Thus, when recipients learn that $H_i \in [a, b]$, the resulting posterior is higher if effort is believed to be high. However, as effort is not observed, interpretations are based on conjectured effort ϵ_i not on actual effort e_i. We elaborate on this issue in section 2.5. Example CDFs which have an MLRP ordering are $G(H) = H$ and $F(H) = H^2$. For our results on the optimal communication framework , we will use that low effort leads to $G(H) = H$ and high effort to $F(H) = H^a$ where $a > 1$ ensures an MLRP ordering. Our results on the equilibrium of the communication game played in a hurdle rate framework are more general.

Central bank: no commitment, information We assume that the Central Bank chooses the level of stress strategically after observing bank health and reports results truthfully. The assumption that the CB cannot commit is in line with CBs referring to their choice of stress scenarios as judgements[19] indicating discretion, with frequent debates on how meaningful a test was[20], and with the extreme experience of the 2011 European stress test.[21] Our assumption that the regulator knows bank health when deciding on the message is motivated by regulators' access to private information in their role as bank supervisor.[22]

2.4 Reassurance

Comparing institutional designs in terms of reassurance, we show that delegating to a softer sender can be socially beneficial. First, we show that frameworks in which more information is released result in less reassurance. Second, we show that within a hurdle

[18] Milgrom (1981); Holmstrom (1979, 1982); Lambert (1983); Ottaviani and Sorensen (2006).

[19] "The FPC and PRC judge the stress scenario to be appropriate [...]." Bank of England, Key elements of the 2019 stress test, 5 March 2019.

[20] "Bank stress tests fail to tackle deflation spectre." Financial Times, 27 October 2014.

[21] "The tests were long ago branded as flawed, so the results (most banks passed) were never going to serve their market-soothing purpose. The exams refused to countenance a sovereign default, even as such an event appears imminent. Investors understood that the tests had been overtaken by events, and that a 'pass' or 'fail' was largely meaningless. Instead, they dumped bank shares [...]" Financial Times, EU bank stress tests, 18 July 2011.

[22] Moreover, stress tests as a diagnostic tool for regulators to learn about bank health existed already prior to the financial crisis. The novelty of the post-crisis regime is that results are published in a specific frequency unaffected by economic events. The private information assumption allows us to focus on the communication part of stress tests.

rate framework, reassurance is increasing in the weight the mandate places on the weak bank. Thus, delegating to a soft sender has the benefit that the softer sender provides more reassurance than a sender with social preferences credibly could. To derive this result, we characterise the equilibrium in a hurdle rate game which is also a contribution to the strategic communication literature.

To isolate the reassurance effect of different institutional designs, this section removes the incentive motive of stress test communication by taking the bank health distribution as exogenously given. This equals the assumption that bank effort is exogenous and common knowledge. It reduces the model to a communication game. We now formally define reassurance.

Definition 2.3 *Institution Design $\mathcal{D}$ is said to provide more reassurance than $\mathcal{D}'$ iff* $\mathbb{E}[\min(\mu_1, \mu_2)]$ *is higher in $\mathcal{D}$ than in $\mathcal{D}'$.*

The definition above may appear odd as it focuses purely on the weak bank and not on the average. We introduced the reassurance motive as a desire to increase the belief financial markets hold about both banks, sometimes with an added concern for the weak bank. An example functional form was $W = \lambda \ \min(\mu_1, \mu_2) + (1-\lambda) \ \frac{\mu_1+\mu_2}{2}$. The absence of the average term is explained by the ex-ante perspective this definition takes, i.e. the focus on expected reassurance across all possible health pairs. While for given bank fundamentals a regulator's choice of message can affect both the average ($\frac{\mu_1+\mu_2}{2}$) and the dispersion of beliefs ($\min(\mu_1, \mu_2)$), the ex-ante choice of an institutional design $\mathcal{D}$ only affects the expected dispersion of beliefs ($\mathbb{E}[\min(\mu_1, \mu_2)]$) and not the expected average ($\mathbb{E}[\frac{\mu_1+\mu_2}{2}]$). This follows from a property of all Bayesian beliefs, called Bayes plausibility: beliefs must be correct on average $\mathbb{E}[\mu_i] = \mathbb{E}[H_i]$. While there are different possible dispersion measures, $\mathbb{E}[\min(\mu_1, \mu_2)]$ arises from the specific functional form we will use for welfare.[23]

With bank effort fixed, it follows that the more information is released in equilibrium, the less reassurance is provided. We now turn to a result ranking designs in terms of information released in equilibrium and thus also in terms of reassurance.

Proposition 2.1 *Institutional Designs $\mathcal{D}$ can be ranked by how much reassurance they provide:*

1. *ZDF provides more reassurance than HRF which provides more than FDF.*
2. *Within HRF, reassurance is weakly increasing in the weight the mandate puts on the weak bank.*

Proof: See Appendix 2.9.4.

While the intuition behind the HRF result requires an understanding of the equilibrium in the communication game, which we turn to next, the ranking of ZDF and FDF is

[23] By Bayes plausiblity: $\mathbb{E}[\frac{\mu_1+\mu_2}{2}] = \mathbb{E}[H_i] \ \forall \ \mathcal{D}$ and thus $\mathbb{E}[W(\mu_1, \mu_2)]$ reduces to $\mathbb{E}[\min(\mu_1, \mu_2)]$.

already intuitive. ZDF means zero disclosure and no updating of beliefs. Hence, there is no dispersion of posteriors and since $\mu_i = \mathbb{E}[H_i]$ reassurance is maximised.[24] The full disclosure in FDF results in beliefs exactly reflecting fundamentals which means very low reassurance.[25]

2.4.1 Hurdle Rate: Equilibrium Communication

We show that when the regulator is asked to use a hurdle rate framework, i.e. to judge banks to pass or fail relative to a common level of stress, an equilibrium exists in the resulting communication game. In this equilibrium, the regulator passes both banks if they are of similar health and passes one but fails the other if their health is very different. Financial markets thus interpret both banks passing to mean that banks are at the level of stress or slightly above, but not so far above that the regulator wanted to let the weaker bank fail the test. We show that messages where both banks pass are used more often when the mandate places more weight on the weak bank. Even if the mandate is only concerned with the average bank, the regulator in equilibrium still uses messages where both pass some of the time. This means that cases where most banks pass but financial markets are not reassured can arise in equilibrium even if the regulator places no extra weight on the weak bank.

Solving for the equilibrium in a verifiable disclosure game where all banks are judged relative to a common level of stress (HRF) is challenging and has - to the best of our knowledge - not been done before. The difficulty comes from the regulator only having one strategic variable (s) which determines the message for both banks. This creates a trade-off in the regulator's strategic choice of s: Higher levels of stress have the benefit of increasing beliefs about those banks which continue to pass but have the cost of resulting in low beliefs about those banks who fail the stress test. Even more complexly, posteriors resulting from a messaging strategy are a function of that strategy as financial markets use all available information to infer the true state. Thus, the costs and benefits of different levels of stress depend on the strategy.

The Theorem below provides a solution to this problem, i.e. solves for the equilibrium. In this equilibrium, it is as if the regulator solved two strategic decisions sequentially: For every possible outcome pair $\{o_1, o_2\}$ and given beliefs, what is the optimal level of stress?[26] Given these candidate solutions, which is the optimal outcome pair? Financial markets correspondingly interpret both the level of stress as strategic (analogous to the first question) and the outcome pair as strategic (second question). For example, in equilibrium, $\{s, p, f\}$ is interpreted to mean bank 1 has health exactly at s and bank 2 is not just below s but far below s. If banks had been close, the regulator would have sent $\{s', p, p\}$. Theorem 2.1 describes the equilibrium mathematically and Figure 2.2 depicts

[24] In a Zero Disclosure Framework $\mathbb{E}[\min(\mu_1, \mu_2)] = \mathbb{E}[\min(\mathbb{E}[H_i], \mathbb{E}[H_i])] = \mathbb{E}[H_i]$.
[25] In a Full Disclosure Framework $\mathbb{E}[\min(\mu_1, \mu_2)] = \mathbb{E}[\min(H_i, H_i)] < \mathbb{E}[H_i]$.
[26] Outcome pairs are $\{p, p\}$, $\{p, f\}$, $\{f, p\}$, and $\{f, f\}$.

the equilibrium graphically.

Theorem 2.1 *In the Hurdle Rate Framework communication game there exists a Perfect Bayesian Equilibrium (PBE). This equilibrium is characterised by two indifference frontiers $x(H_2)$ and $y(H_2)$, which are visualised in Figure 2.2a, such that the sender plays:*

$$\begin{array}{llll} s = & H_1 & \textit{iff} \quad H_1 > x(H_2) & \rightarrow \quad \{H_1, p, f\} \\ s = & \min(H_1, H_2) & \textit{iff} \quad x(H_2) \geq H_1 \geq y(H_2) & \rightarrow \quad \{\min(H_1, H_2), p, p\} \\ s = & H_2 & \textit{iff} \quad y(H_2) > H_1 & \rightarrow \quad \{H_2, f, p\} \end{array}$$

where $x(0) = y(0) = 0$; $x(H_2) \geq y(H_2)\ \forall\ H_2$, $x(H_2) \leq 1\ \forall\ H_2$; and $x(H_2)$ and $y(H_2)$ are both continuous and monotonically increasing ($\frac{dx(H_2)}{dH_2} > 0$, $\frac{dy(H_2)}{dH_2} > 0$).

Recipients form beliefs accordingly (visualised in Figure 2.2b). Denoting the cdf of H_1 as F and of H_2 as G.

$$\begin{array}{lll} \{s, p, f\} & \textit{is interpreted as} & \mu_1 = s; \quad \mu_2 = \mathbb{E}_G[H \mid H < x^{-1}(s)] \\ \{s, f, p\} & \textit{is interpreted as} & \mu_2 = s; \quad \mu_1 = \mathbb{E}_F[H \mid H < y(s)] \\ \{s, p, p\} & \textit{is interpreted as} & \mu_1 = \alpha\ s + (1 - \alpha)\ \mathbb{E}_F[H \mid s \leq H \leq x(s)] \\ & & \mu_2 = (1 - \alpha)\ s + \alpha\ \mathbb{E}_G[H \mid s \leq H \leq y^{-1}(s)] \\ \{s, f, f\} & \textit{is interpreted as} & \mu_1 = \mu_2 = 0\ \forall\ s > 0 \end{array}$$

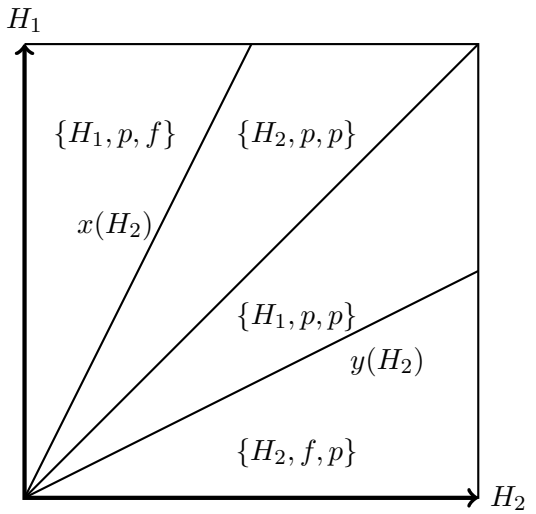
where

$$\alpha = \frac{f(s)\ [G(y^{-1}(s)) - G(s)]}{f(s)\ [G(y^{-1}(s)) - G(s)]\ +\ g(s)\ [F(x(s)) - F(s)]} \tag{2.3}$$

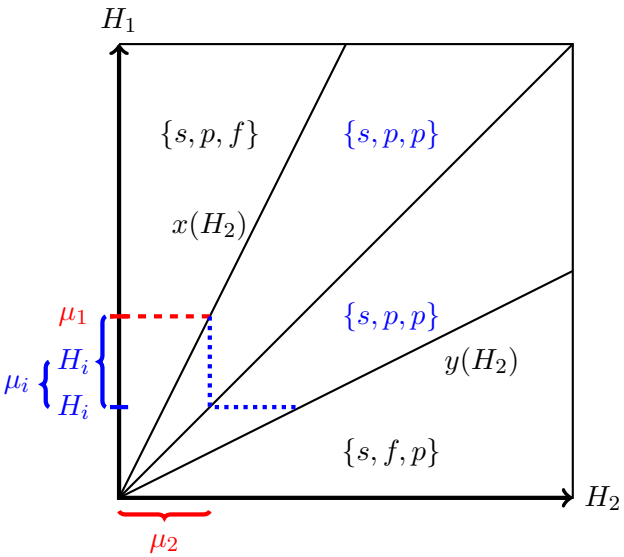
Proof: See Appendix 2.9.2.

Figure 2.2: Equilibrium in a Hurdle Rate Framework

(a) Equilibrium Communication Strategy



(b) Equilibrium Beliefs



Note: Example with $H_i \sim U[0, 1]\ \forall\ i = 1, 2$; $V(\mu_1, \mu_2) = \frac{\mu_1 + \mu_2}{2}$. Then $x(H_2) = 2H_2$; $y(H_2) = \frac{1}{2}H_2$.

The sender's actions are depicted in Figure 2.2a, the receivers interpretation is depicted in Figure 2.2b.

The two indifference frontiers $x(H_2)$ and $y(H_2)$ are defined by the following set of indifference equations where $\mu_i \mid \{s, o, o\}$ denotes the posterior formed upon receiving $\{s, o, o\}$. On $x(H_2)$:

$$V\Big(\mu_1 \mid \{H_1, p, f\},\ \mu_2 \mid \{H_1, p, f\}\Big) = V\Big(\mu_1 \mid \{H_2, p, p\},\ \mu_2 \mid \{H_2, p, p\}\Big) \tag{2.4}$$

On $y(H_2)$:

$$V\Big(\mu_1 \mid \{H_1, p, p\},\ \mu_2 \mid \{H_1, p, p\}\Big) = V\Big(\mu_1 \mid \{H_2, f, p\},\ \mu_2 \mid \{H_2, f, p\}\Big) \tag{2.5}$$

When bank health is identically distributed, the indifference frontiers are symmetric, i.e. $y(H_2) = x^{-1}(H_2)$ and $\alpha = 0.5$.

Example: Let $H_i \sim U[0,1]\ \forall\ i = 1, 2$ and let $V(\mu_1, \mu_2) = \omega\ \min(\mu_1, \mu_2) + (1 - \omega)\ \frac{\mu_1 + \mu_2}{2}$. Then the equilibrium is characterised by:

$$x(H_2) = \frac{2 - \omega}{1 - 2\omega} H_2 \quad ; \quad y(H_2) = x^{-1}(H_2) \tag{2.6}$$

This holds for $\omega < 0.5$. For $\omega \geq 0.5$ the regulator sends $\{s, p, p\}$ messages in all cases apart from those exactly on the axis, i.e. when $\min(H_1, H_2) = 0$. In these cases he sets $s = \max(H_1, H_2)$ which results in one bank passing and one bank failing the stress test.

The equilibrium described in Theorem 2.1 is not a unique weak PBE, but it is the only weak PBE which satisfies the following notion of monotonicity. The proof and a discussion of alternative equilibria which violate monotonicity is provided in Appendix 2.9.3.

Definition 2.4 *An equilibrium is said to satisfy monotonicity if beliefs satisfy the following condition:*

$$\mu_i \mid \{s, o_1, o_2\} > \mu_i \mid \{s', o_1, o_2\} \quad \forall\ s > s',\ \forall i = 1, 2,\ \forall o_1,\ \forall o_2 \tag{2.7}$$

This means that beliefs are increasing in the level of stress for every given stress test outcome, i.e. for passing and for failing the test. This seems close to commentary of stress tests where passing weak tests is seen as a relatively weak signal while passing tough tests is viewed as a strong signal. While this condition imposes a theoretical restriction, it seems plausible in the context of real world bank stress tests. For the remainder of the paper, in a HRF we focus on the equilibrium as described in Theorem 2.1.

To build an understanding of how delegation affects the equilibrium, the following proposition studies a comparative static in the sender's mandate.

Proposition 2.2 *The more weight the mandate places on the weak bank, the steeper the $x(H_2)$-frontier is in equilibrium, i.e. the larger the set of states where $\{H_2, p, p\}$ is sent relative to the states where $\{H_1, p, f\}$ is sent. Correspondingly, the $y(H_2)$-frontier is flatter.*

Proof: See Appendix 2.9.5.

Consider a sender with mandate $\mathcal{M}$ who at a specific (H_1, H_2) on his $x(H_2)$-indifference frontier is indifferent between $\{H_2, p, p\}$, which implies no dispersion of posteriors ($\mu_1 = \mu_2$) and $\{H_1, p, f\}$, which implies dispersion of posteriors ($\mu_1 > \mu_2$). Then at the same (H_1, H_2) a sender whose mandate $\mathcal{M}'$ places more weight on the weak bank, strictly prefers $\{H_2, p, p\}$. Therefore, the equilibrium under $\mathcal{M}'$ exhibits a steeper $x(H_2)$ and a flatter $y(H_2)$ than under $\mathcal{M}$.

Interestingly, even when the mandate places no extra weight on the weak bank and is only concerned with the average belief, $\{s, p, p\}$ messages are still used in equilibrium.[27] This arises because strong messages where both pass, e.g. $\{0.9, p, p\}$ are only feasible in a set of states with high H_1 and high H_2, guaranteeing high $\frac{\mu_1+\mu_2}{2}$, while strong messages where only one bank passes, e.g. $\{0.95, p, f\}$ are also feasible when one bank is extremely weak. Thus, very strong interpretations of $\{0.95, p, f\}$ cannot be an equilibrium, as then the regulator would want to send $\{0.95, p, f\}$ also at $H_1 = 0.95, H_2 = 0$ and similarly low states, undermining the high interpretation. Thus, as health state pairs with low average health cannot feasibly pool with the high average pairs that could send $\{0.9, p, p\}$, in equilibrium $\{0.9, p, p\}$ will be used even if the mandate places no extra weight on the weak bank.

The sender almost always sends $\{s, p, p\}$ messages even for mandates that place some weight on average health. An extreme equilibrium in HRF is when the sender sends $\{s, p, p\}$ in all cases apart from when $\min(H_1, H_2) = 0$. In these cases, he chooses $s = \max(H_1, H_2)$. This messaging pattern arises for a mandate which is purely concerned with the weak bank, but it also arises for mandates that place some weight on the average. In the example, $\{s, p, p\}$ is almost always send for all $\omega \geq 0.5$ which corresponds to 0.75 weight on the weak and 0.25 weight on the strong bank.

2.4.2 Hurdle Rate: Reassurance

While the previous subsection characterised what messages are sent and beliefs are formed for a given health pair, this subsection returns the focus to the ex-ante measure of expected reassurance. We show that reassurance provided in a HRF lies between reassurance in ZDF and FDF and is increasing in the softness of the mandate. This provides a rationale for delegation: Delegating stress tests to a softer sender provides more reassurance which increases welfare.

[27] This can be seen in the example used to construct Figures 2.2a and 2.2b. When $H_i \sim U[0,1] \; \forall \; i = 1, 2$; $V(\mu_1, \mu_2) = \frac{\mu_1+\mu_2}{2}$, then $x(H_2) = 2H_2$; $y(H_2) = \frac{1}{2}H_2$. Thus, a sender who places no extra weight on the weak bank sends $\{s, p, p\}$ in 50% of the states.

There are two challenges to characterising reassurance in a HRF. First, comparing reassurance in a HRF to ZDF and FDF requires computing the level of reassurance, i.e. $\mathbb{E}[\min(\mu_1, \mu_2)]$. The numerous thresholds in the HRF communication equilibrium complicate this. Second, comparing reassurance for different mandates within a HRF is not trivial as there are two opposing forces: On the one hand, a softer sender results in steeper $x(H_2)$, i.e. sends $\{s, p, p\}$ more often, which increases reassurance since $\{s, p, p\}$ creates less dispersion in posteriors than $\{s, p, f\}$. On the other hand, a steeper $x(H_2)$ frontier means that a given $\{s, p, f\}$ message results in more dispersion when it is sent.

Mathematically, reassurance in a HRF is:

$$\mathbb{E}[\min(\mu_1, \mu_2)] = \int_0^1 \int_0^1 \min(\mu_1, \mu_2)\ f(H_1)\ dH_1\ g(H_2)\ dH_2 \tag{2.8}$$

As beliefs are based on the sender's strategy which is characterised by thresholds, this becomes:

$$\begin{aligned}
\mathbb{E}[\min(\mu_1, \mu_2)] = \int_0^{x^{-1}(1)} \Bigg[& \int_0^{y(H_2)} \mathbb{E}\Big[H \mid H < y(H_2)\Big]\, f(H_1)\, dH_1 \\
& + \int_{y(H_2)}^{H_2} \frac{1}{2}H_1 + \frac{1}{2}\mathbb{E}\Big[H \mid H_1 < H < x(H_1)\Big]\, f(H_1)\, dH_1 \\
& + \int_{H_2}^{x(H_2)} \frac{1}{2}H_2 + \frac{1}{2}\mathbb{E}\Big[H \mid H_2 < H < x(H_2)\Big]\, f(H_1)\, dH_1 \\
& + \int_{x(H_2)}^{1} \mathbb{E}\Big[H \mid H < x^{-1}(H_1)\Big]\, f(H_1)\, dH_1 \Bigg]\, g(H_2)\, dH_2 \\
+ \int_{x^{-1}(1)}^{1} \Bigg[& \int_0^{y(H_2)} \mathbb{E}\Big[H \mid H < y(H_2)\Big]\, f(H_1)\, dH_1 \\
& + \int_{y(H_2)}^{H_2} \frac{1}{2}H_1 + \frac{1}{2}\mathbb{E}\Big[H \mid H_1 < H < 1\Big]\, f(H_1)\, dH_1 \\
& + \int_{H_2}^{1} \frac{1}{2}H_2 + \frac{1}{2}\mathbb{E}\Big[H \mid H_2 < H < 1\Big]\, f(H_1)\, dH_1 \Bigg]\, g(H_2)\, dH_2
\end{aligned} \tag{2.9}$$

We find that reassurance is increasing in the softness of the sender, i.e. the more frequent use of $\{s, p, p\}$ outweighs the more dispersed interpretation of $\{s, p, f\}$. We also find that reassurance in a HRF lies between ZDF and FDF for all possible mandates. These results form proposition 2.1 and have implications for institutional design.

Institutional Design: Benefit of Delegation 1 *Delegating stress tests to a softer sender can be optimal, as it has the benefit of making higher reassurance credible.*

More formally, suppose the bank health distribution is exogenous and the institutional

designer D is restricted to choosing a mandate within HRF, i.e. ZDF, FDF are not available. Then every designer D who places a strictly positive extra weight on the weak bank (e.g. $\lambda > 0$), optimally delegates stress tests to a regulator who implements the softest possible communication rule, i.e. who almost always sends $\{\min(H_1, H_2), p, p\}$.[28]

This benefit of delegation arises from a commitment problem. Ex-ante, before learning bank health, any regulator with some concern about the weak bank wants recipients to believe that he will always send $\{\min(H_1, H_2), p, p\}$ as this maximises reassurance.[29] Once the regulator learns bank health, there are cases where he is tempted to deviate and increase average posteriors at the expense of the weak bank, i.e. at the expense of creating dispersion by sending $\{H_1, p, f\}$. The larger the regulator's weight on the weak bank, the larger the set of states where sending $\{\min(H_1, H_2), p, p\}$ is actually credible and thus the larger reassurance is.

To understand why it is not credible for every sender to always send $\{\min(H_1, H_2), p, p\}$ even though this would maximise reassurance, consider the example message $\{0.2, p, p\}$. If $H_1 = 0.2, H_2 = 0.2$, this message will be sent by all regulators regardless of their preference. However, this is not true if $H_1 = 1, H_2 = 0.2$. Then, sending the feasible alternative $\{1, p, f\}$ achieves an average posterior which is higher than the average posterior resulting from sending $\{0.2, p, p\}$. Even under the worst possible interpretation of $\{1, p, f\}$, i.e. $\mu_1 = 1$, $\mu_2 = 0$, we have $\frac{\mu_1+\mu_2}{2} = 0.5$ which exceeds $\frac{\mu_1+\mu_2}{2} = 0.4$ from $\{0.2, p, p\}$.[30] Hence, at states with disperse fundamentals only senders with a strong concern for the weak bank actually send $\{\min(H_1, H_2), p, p\}$, thus accepting lower average posteriors but achieving a higher minimum belief than under the alternative of deviating to $\{H_1, p, f\}$.

Since the social planner D cannot commit to and is not credible to provide reassurance at all states, i.e. send $\{\min(H_1, H_2), p, p\}$, the social planner can benefit from delegating stress tests to a softer sender. This results in a steeper $x(H_2)$ frontier and thus in more reassurance.

While this section showed that delegation to a softer sender can be beneficial as it increases reassurance, the next section shows that delegation to a tougher sender can be beneficial as it increases incentives.

2.5 Incentives

Comparing institutional designs in terms of incentives, we show that delegating to a tougher sender can be socially beneficial as it makes the threat of letting weak banks fail

[28] This is not necessarily the same as delegating to the softest possible sender. As argued in the context of proposition 2.2, there exists a range of mandates which result in almost always sending $\{\min(H_1, H_2), p, p\}$. In our example, $\omega \in [0.5, 1]$. The designer is indifferent between these mandates as they all imply the same level of reassurance.

[29] This is the equilibrium in a HRF which achieves the highest reassurance possible. Reassurance would be even higher if the sender could commit to always sending $\{0, p, p\}$ but this is analogous to ZDF, which we have ruled out to focus on HRF.

[30] For this example, numbers are calculated assuming a uniform distribution.

the stress test more credible. This incentivizes banks to take prudent actions. To derive this, we first show that communication creates incentives for banks to take prudent actions. Then, we rank communication frameworks according to incentives created, finding that FDF implies stronger incentives than HRF which implies more than ZDF. Within HRF, tougher mandates create stronger incentives. Then, we characterise equilibria in the game with endogenous effort and find that stronger incentives ensure that the equilibrium with high effort occurs even for higher effort costs. Thus, delegating stress tests to a tougher sender can be beneficial as it makes the threat of letting weak banks fail the stress test more credible which incentivises banks to take prudent actions.

2.5.1 Costs and Benefits of Bank Effort Choice

When deciding whether to exert high or low effort, a bank weighs up the certain cost of effort $C(h) - C(l)$, against the expected benefit of higher beliefs about its health. While effort improves the distribution of bank fundamentals according to MLRP, how this improvement translates into a change in expected beliefs depends on the communication framework. We refer to a communication framework as "generating incentives" if it results in expected beliefs and not just expected fundamentals responding to effort choice. Denoting the vector of low conjectured efforts $\boldsymbol{l} = (l_1, l_2)$ and $\boldsymbol{h} = (h_1, h_2)$ correspondingly:

Definition 2.5 *A communication framework $\mathcal{C}$ generates effort incentives if and only if:*

$$\mathbb{E}\big[\mu_1(\boldsymbol{l}) \mid h_1, l_2\big] - \mathbb{E}\big[\mu_1(\boldsymbol{l}) \mid l_1, l_2\big] > 0 \tag{2.10}$$

where by Bayes plausibility of beliefs $\mathbb{E}\big[\mu_1(\boldsymbol{l}) \mid l_1, l_2\big] = \mathbb{E}[H_1 \mid l_1] = \mathbb{E}_G[H]$.

Incentives capture how expected posteriors change if a bank deviates in its effort choice from a candidate equilibrium. That means, we study the deviation when conjectured efforts and the rival's effort are held fixed. In the next subsection, we turn to the fixed point problem, i.e. characterise actual equilibria where such deviations do not occur.[31]

While in equilibrium, expected posteriors are correct ($\mathbb{E}_F[H] = \mathbb{E}[\mu(\boldsymbol{h}) \mid h_1, h_2]$, see also the discussion of Bayes plausibility in section 2.4), this need not be true in a deviation from the equilibrium. Thus, the benefit to deviating is not always $\mathbb{E}_F[H] - \mathbb{E}_G[H]$, but depends on the communication framework. We now turn to ranking frameworks by the benefit of deviating by increasing effort, i.e. by incentives.

Proposition 2.3 *Ranking of communication frameworks in terms of incentives generated:*

[31] The definition of incentives considers deviations from a low effort equilibrium. Analogously, considering deviations from a high effort equilibrium, incentives are captures by the effect of reducing effort: $\mathbb{E}_F[H] - \mathbb{E}[\mu(\boldsymbol{h}) \mid l_1, h_2] > 0$. The ordering of communication frameworks in terms of strength of incentives created is the same in both cases. The magnitudes can differ. This affects the equilibrium characterisation in the next subsection, but not the ordering which we focus on in this subsection.

1. *FDF provides more incentives than HRF which provides more than ZDF.*

2. *Within HRF, the lower the central bank's concern for the weak bank, the more incentives are generated.*

Proof: See Appendix 2.9.6.

While calculating incentives in a HRF poses several challenges, the results for the extreme frameworks are clear and help build the intuition for incentives in a HRF. ZDF provides zero incentives.[32] Since no information is released, financial markets can never update their prior. The prior is based purely on conjectured effort. Hence, in ZDF beliefs are unaffected by actual effort which creates no incentive for banks to invest in their health. FDF provides strong incentives.[33] Since true bank health is always perfectly revealed, beliefs track fundamentals one-for-one. We refer to this as full incentives.

Incentives in a HRF are difficult to understand as several partially offsetting forces interact and mathematically complex to express due to the thresholds in the communication equilibrium. The conceptual difficulty arises because of the divergence of actual and conjectured effort. While a bank's effort choice affects effort and thus the frequency of messages occurring, it does not affect conjectured effort and thus leaves the interpretation of a given message unaffected.

Fixed conjectures mute, keep unchanged, or magnify the effect of changed fundamentals. As a thought experiment, suppose that both banks are conjectured to play low effort and let H_2 be some realisation. Ex-post, how much will bank 1 benefit from having put in higher effort? The effect can be muted, i.e. less than on fundamentals. E.g. if effort leads to an improvement in H_1 from an H_1 that led to $\{H_2, f, p\}$ to a still low H_1 that also leads to $\{H_2, f, p\}$, then beliefs are unaffected. The effect can be in line with fundamentals. E.g. if initially $\{H_1, p, f\}$ is sent and effort results in a stronger $\{H_1', p, f\}$ with $H_1' > H_1$, then the belief about bank health has shifted in line with fundamentals since $\mu_i = H_i$ for this message. The effect can be magnified. If higher effort induces a switch of message from $\{H_2, f, p\}$ to $\{H_1, p, p\}$, then μ_1 has increased more strongly than H_1.

We find that incentives implied by a HRF lie between those of ZDF and FDF. Moreover, the tougher the sender, i.e. the less weight his mandate puts on the weak bank and thus the more fail grades are given out in equilibrium, the stronger the incentives. These results, which form Proposition 2.3, are proved in Appendix 2.9.6, making use of the following expression for expected beliefs in a HRF when effort is changed away from conjectured effort:

[32] Formally, in ZDF $\mathbb{E}\big[\mu_1(\boldsymbol{l}) \mid h_1, l_2\big] = \mathbb{E}_G[H]$ and thus incentives in ZDF are $\mathbb{E}_G[H] - \mathbb{E}_G[H] = 0$.
[33] Formally, in FDF $\mathbb{E}\big[\mu_1(\boldsymbol{l}) \mid h_1, l_2\big] = \mathbb{E}_F[H]$ and thus incentives in FDF are $\mathbb{E}_F[H] - \mathbb{E}_G[H] > 0$.

$$
\begin{aligned}
\mathbb{E}\big[\mu_1(\boldsymbol{l}) \mid h_1, l_2\big] = \int_0^{x^{-1}(1)} \Bigg[& \int_0^{y(H_2)} \mathbb{E}_G\Big[H \mid H < y(H_2)\Big] f(H_1) dH_1 \\
& + \int_{y(H_2)}^{H_2} \frac{1}{2} H_1 + \frac{1}{2} \mathbb{E}_G\Big[H \mid H_1 < H < x(H_1)\Big] f(H_1) dH_1 \\
& + \int_{H_2}^{x(H_2)} \frac{1}{2} H_2 + \frac{1}{2} \mathbb{E}_G\Big[H \mid H_2 < H < x(H_2)\Big] f(H_1) dH_1 \\
& + \int_{x(H_2)}^{1} H_1 f(H_1) dH_1 \Bigg] g(H_2) dH_2 \qquad (2.11) \\
+ \int_{x^{-1}(1)}^{1} \Bigg[& \int_0^{y(H_2)} \mathbb{E}_G\Big[H \mid H < y(H_2)\Big] f(H_1) dH_1 \\
& + \int_{y(H_2)}^{H_2} \frac{1}{2} H_1 + \frac{1}{2} \mathbb{E}_G\Big[H \mid H_1 < H < 1\Big] f(H_1) dH_1 \\
& + \int_{H_2}^{1} \frac{1}{2} H_2 + \frac{1}{2} \mathbb{E}_G\Big[H \mid H_2 < H < 1\Big] f(H_1) dH_1 \Bigg] g(H_2) dH_2
\end{aligned}
$$

Even the softest possible sender in a HRF creates strictly more incentives than ZDF.[34] The reason is that even if only $\{s, p, p\}$ messages are sent, then the level of s is informative and creates incentives. In ZDF, there is absolutely no information and thus no incentives.

2.5.2 Equilibrium with Bank Effort Choice

A benefit of delegating stress tests to a tougher sender is that this makes giving out fail marks more credible and therefore results in stronger incentives. We now turn to the equilibrium and show that stronger incentives ensure that equilibria with high effort occur even for higher effort costs. Welfare can be higher as a result.

Restricting our attention to pure strategy symmetric equilibria and to the case where effort is socially beneficial, i.e. costs lie below the change in expected fundamentals $\mathbb{E}_F[H] - \mathbb{E}_G[H] \geq C(h) - C(l)$ and where costs are strictly positive $C(h) - C(l) > 0$ we show the following proposition:

Proposition 2.4

1. *In ZDF, there always exists a unique pure strategy equilibrium. In this equilibrium both banks choose low effort.*
2. *In FDF, there always exists a unique pure strategy equilibrium. In this equilibrium both banks choose high effort.*

[34] For a proof of this limit result see Appendix 2.9.6, specifically Sublemma 2.9.6.3.

3. *In HRF, there exist two thresholds which are functions of the sender's mandate* $\underline{t}(\mathcal{M})$, $\bar{t}(\mathcal{M})$.

   (a) *For all* $C(h) - C(l) \leq \underline{t}(\mathcal{M})$, *there exists a unique pure strategy equilibrium. In this equilibrium, both banks choose high effort.*

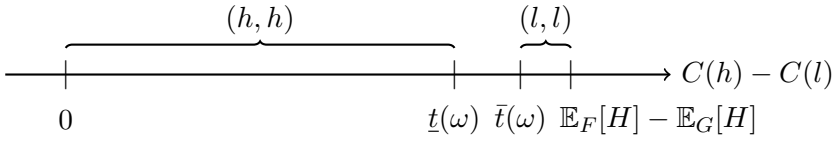
   (b) *For all* $C(h) - C(l) \geq \bar{t}(\mathcal{M})$, *there exists a unique pure strategy equilibrium. In this equilibrium, both banks choose low effort.*

   (c) *For all* $\underline{t}(\mathcal{M}) < C(h) - C(l) < \bar{t}(\mathcal{M})$, *there does not exist a pure strategy equilibrium.*

Proof: See Appendix 2.9.7.

Figure 2.3 illustrates the existence results for a HRF. The thresholds $\underline{t}(\mathcal{M})$ and $\bar{t}(\mathcal{M})$ are both increasing in the toughness of the mandate, i.e. are higher if the mandate places less concern on the weak bank. For a given mandate $\mathcal{M}$: $\bar{t}(\mathcal{M}) = \mathbb{E}[\mu_1(\boldsymbol{l}) \mid h_1, l_2] - \mathbb{E}_G[H]$ and $\underline{t}(\mathcal{M}) = \mathbb{E}_F[H] - \mathbb{E}[\mu_1(\boldsymbol{h}) \mid l_1, h_2]$.

Figure 2.3: Equilibrium in a Hurdle Rate Framework



Note: Thresholds depicted for the case where $G(H) = H, F(H) = H^2, V(\mu_i, \mu_2) = \frac{\mu_1+\mu_2}{2}$.

The extreme frameworks are again useful to build the intuition which we can later apply to the HRF. As argued in the previous subsection, ZDF provides no incentives, i.e. banks have zero benefit of effort. As costs of effort are strictly positive, banks never find it optimal to exert effort which results in the low effort equilibrium. FDF provides full incentives. Since we assume that effort is socially beneficial, full incentives will always induce effort which results in the high effort equilibrium.

The intuition for the equilibrium in HRF is best understood in two steps. First, we explain why the equilibrium has a threshold structure. Then, we investigate why there are two thresholds and not one. A given mandate implies a strength of incentives $\mathbb{E}[\mu_1(\boldsymbol{l}) \mid h_1, l_2] - \mathbb{E}_G[H]$. If costs exceed incentives, banks optimally choose low effort and a low effort equilibrium exists. For lower costs banks would prefer to deviate to high effort. Hence, a mandate implies a certain fixed benefit of effort and then the equilibrium is determined by whether costs are above or below that benefit. The tougher the mandate is, the more incentives the regulator generates, and thus the higher the threshold is below with a high effort equilibrium exists.

The intuition for why there are two thresholds instead of one, i.e. why we do not have

$\underline{t}(\mathcal{M}) = \bar{t}(\mathcal{M})$, does not follow from an understanding of the extreme frameworks. Rather, this is driven by the importance of conjectures.

Based on the equilibrium effect of incentives, we can now formulate the benefit of delegating to a tougher sender.

Institutional Design: Benefit of Delegation 2 *Delegating stress tests to a tougher sender increases incentives, as the threat of giving out fail marks is credible in more health states. Delegating to a tougher sender can thus shift the equilibrium from a low to a high effort equilibrium.*

A soft sender cannot credibly implement a tough communication rule. If financial markets believed the sender that he intends to implement a rule $\tilde{x}(H_2)$ which is tougher, i.e. flatter than his equilibrium rule $x(H_2)$, then there exist states where the sender should send $\{H_1, p, f\}$ if he follows $\tilde{x}(H_2)$, but actually prefers to send $\{H_2, p, p\}$ as he is concerned about the weak bank. Thus, a soft sender cannot credibly implement a tough communication rule as his threat to give out fail marks is not credible. He prefers to give out pass marks instead.

This section established that a benefit of delegating to a tougher sender exists as a tougher sender can induce effort and thus achieve a higher average posterior in equilibrium. This comes at the cost of less reassurance, which we showed in the previous section as a benefit of delegating to a softer sender. We now turn to the case where *both* these benefits and costs are considered, characterise when which effect dominates, and show that delegation can be optimal in either direction depending on how costly effort is.

2.6 Central Bank Design

When stress tests are used to reassure financial markets and to incentivise banks, a hurdle rate framework is the optimal design in many settings as it generates intermediate levels of both incentives and reassurance. A full disclosure framework, which generates strong incentives but no reassurance, is only optimal if two conditions hold simultaneously: effort is so costly that a hurdle rate framework cannot induce high effort, and the benefit of effort in terms of improved health fundamentals is so large that this benefit more than compensates society for foregone reassurance. A zero disclosure framework, which achieves stronger reassurance but generates no incentives, is only optimal if two conditions hold: effort has a small effect on health fundamentals, and society values reassurance considerably. Moreover, when a hurdle rate framework is chosen, the designer can exploit benefits from strategic delegation. That is, in some cases the designer sets the central bank a mandate which is tougher than welfare maximisation while in others the designer sets a softer mandate.

Our exposition of these optimal central bank design results proceeds in three steps.

First, we show that the designer's problem of choosing a design, i.e. a combination of communication framework and central bank mandate, reduces to a choice among three alternatives: ZDF, FDF, or HRF with the softest possible mandate which induces high effort.[35] Second, we characterise the designer's optimal choice for any given set of parameters and show that a hurdle rate framework is often optimal. Third, we interpret these results in terms of strategic delegation.

2.6.1 The Optimal Mandate

This subsection simplifies the designer's problem by showing that if the designer chooses a hurdle rate framework, he will always choose it in combination with the softest possible mandate which still induces high effort. This result holds for any strength of the designer's reassurance motive. Thus, the designer's choice reduces to a choice among three designs.

Proposition 2.5 *The designer's optimal design choice reduces to a choice among three alternatives: FDF, ZDF,* $\{HRF, \tilde{\mathcal{M}}\}$ *where* $\tilde{\mathcal{M}}$ *is the softest possible mandate which incentivises high effort, i.e. is defined by* $\underline{t}(\tilde{\mathcal{M}}) = C(h) - C(l)$.

Proof: See Appendix 2.9.8.

All other designs $\{HRF, \mathcal{M}\}$ where $\mathcal{M} \neq \tilde{\mathcal{M}}$ are dominated by at least one of the three alternatives. As $\underline{t}(\mathcal{M})$ is increasing in the toughness of the mandate, the mandate $\tilde{\mathcal{M}}$ is tougher when costs of effort are higher.

The intuition for this result differs depending on whether costs of effort are high or low. Cost of effort are said to be high, if there does not exist a combination of a HRF and a mandate which induces high effort. Otherwise, costs of effort are said to be low. This distinction arises because a HRF cannot provide arbitrarily high incentives, i.e. there exists an upper bound to the strength of incentives in a HRF. As depicted in Figure 2.3, even for the mandate which places no extra weight on the weak bank ($\mathcal{M}(0)$, e.g. $\omega = 0$), the strength of incentives is strictly below the benefit of effort, $\underline{t}(\mathcal{M}(0)) < \mathbb{E}_F[H] - \mathbb{E}_G[H]$. Thus, there exist high costs of effort for which no $\{\text{HRF}, \mathcal{M}\}$ can incentivise high effort, $\underline{t}(\mathcal{M}(0)) < C(h) - C(l) < \mathbb{E}_F[H] - \mathbb{E}_G[H]$.

When costs of effort are high, the designer never chooses a hurdle rate framework with any mandate as all possible $\{\text{HRF}, \mathcal{M}\}$ are welfare dominated by ZDF. By definition, high costs of effort mean that all $\{\text{HRF}, \mathcal{M}\}$ result in low effort. ZDF also results in low effort but provides more reassurance than all $\{\text{HRF}, \mathcal{M}\}$. Thus, for any welfare function which has any, potentially very small, reassurance concern, ZDF results in higher welfare than all $\{\text{HRF}, \mathcal{M}\}$. The designer's choice thus reduces to a choice between ZDF and FDF.

[35] The mandate affects the equilibrium in a HRF but has no effect in FDF or ZDF. As all combinations $\{\text{FDF}, \mathcal{M}\}$ result in financial markets learning bank health perfectly, we denote these combinations as FDF. Analogously, we denote all $\{\text{ZDF}, \mathcal{M}\}$ as ZDF.

When costs of effort are low, the designer either chooses ZDF or $\{\text{HRF}, \tilde{\mathcal{M}}\}$. The design $\{\text{HRF}, \tilde{\mathcal{M}}\}$ dominates FDF and all HRF with a tougher mandate as these alternatives provide lower reassurance than $\{\text{HRF}, \tilde{\mathcal{M}}\}$ while all these designs result in high effort. All HRF with a softer mandate than $\tilde{\mathcal{M}}$ are dominated by ZDF as ZDF provides more reassurance while all cases result in low effort.

The next section characterises which of the three candidate designs is the optimal design for a given parameter combination.

2.6.2 The Optimal Communication Framework

A hurdle rate framework is the optimal design in many settings as it provides an intermediate combination of both incentives and reassurance. More precisely, we show that it is optimal for the designer to choose a hurdle rate framework when two conditions hold simultaneously: The cost of effort is low enough that some hurdle rate framework can induce high effort and the benefit of effort is large relative to the strength of the designer's concern for the weak bank. A full disclosure framework is only optimal if two conditions hold simultaneously: effort is so costly that a hurdle rate framework cannot induce high effort, and the benefit effort creates in terms of improved health fundamentals is so large that this benefit more than compensates society for the reduction in reassurance. A zero disclosure framework is only optimal when effort has a small effect on fundamentals and simultaneously society values reassurance considerably. The results are depicted in Figure 2.4. We discuss the intuition first focusing on the case of high costs of effort, then turning to the low cost case.

For the following results on the optimal communication framework, we make functional form assumptions. We let low effort lead to $G(H) = H$ and high effort to $F(H) = H^a$ where a captures how much effort improves the bank health distribution and where $a > 1$ ensures an MLRP ordering. For welfare and the central bank's mandate, we focus on $W(\mu_1, \mu_2) = \lambda \min(\mu_1, \mu_2) + (1 - \lambda)\frac{\mu_1 + \mu_2}{2}$ and the mandate with ω respectively.

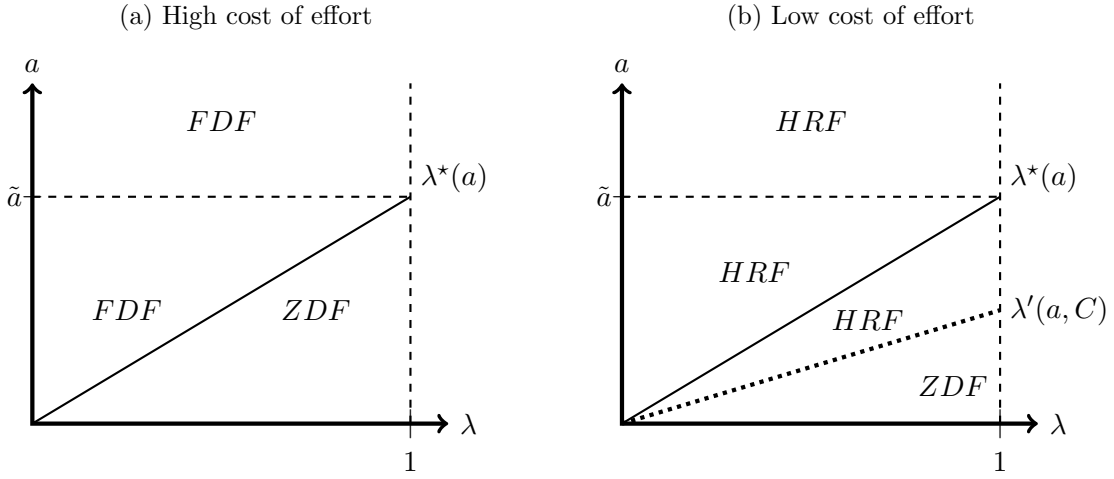
For high costs of effort, no hurdle rate framework can generate incentives which induce high effort. Thus, the optimal design is either full disclosure or zero disclosure. The designer will choose zero disclosure unless effort improves fundamentals sufficiently strongly (large a) that this compensates the designer for the reduction in reassurance caused by adopting FDF rather than ZDF. If the designer is more concerned with reassurance (large λ), he requires a larger improvement in fundamentals (large a) to be willing to adopt FDF. This gives rise to Figure 2.4a where the partitioning frontier $\lambda^*(a)$ is increasing.

For low costs of effort, the optimal design is either zero disclosure or a hurdle rate framework with the softest possible mandate which generates just enough incentives to induce high effort. Similar to the high cost of effort case, the designer will choose zero disclosure unless effort improves health fundamentals sufficiently strongly (large a) to compensate for the reduction in reassurance. When society has a stronger reassurance

motive (larger λ), the designer requires stronger benefits of effort (large a) to choose the hurdle rate framework over zero disclosure. The resulting partitioning frontier $\lambda'(a, C)$ is depicted in Figure 2.4b.

Comparing the cases of high and low costs of effort reveals that the availability of a hurdle rate framework as a design option is socially beneficial. It provides an alternative way of incentivising effort which is preferred to FDF as HRF achieves more reassurance. Moreover, a HRF is chosen in some cases where otherwise ZDF would have been chosen ($\lambda^*(a) > \lambda'(a, C)$). By reducing the cost of providing effort incentives (reduced reassurance) relative to FDF, a HRF increases the part of the parameter space in which the designer finds it optimal to induce effort.

Figure 2.4: Optimal Institution Design



Where: $G(H) = H$, $F(H) = H^a$ with $a > 1$; HRF refers to $\{\text{HRF}, \tilde{\mathcal{M}}\}$; $\tilde{a} = \mathbb{E}_F[\min(H_1, H_2)] - \mathbb{E}_G[H]$.
Note: Costs of effort are said to be high iff $C(h) - C(l) > \underline{t}(\mathcal{M}(0))$ and are otherwise said to be low.

Proposition 2.6 *Optimal Central Bank Design*

1. *When costs of effort are low, i.e. $C(h) - C(l) \leq \underline{t}(\mathcal{M}(0))$, the optimal institutional design is:*
   - *$\{HRF, \tilde{\mathcal{M}}\}$ iff $a > \lambda'(a, C) = \frac{\mathbb{E}_F[H] - \mathbb{E}_G[H]}{\mathbb{E}_F[H] - \mathbb{E}_{HRF,\tilde{\omega}}[\min(\mu_1, \mu_2)]}$. That means that the benefit of effort is sufficiently large to compensate welfare for the reduction in reassurance created when providing incentives. $\lambda'(a, C)$ is increasing in the cost of effort as $\tilde{\omega}$ is decreasing in the cost of effort.*
   - *ZDF otherwise*
2. *When costs of effort are high, i.e. $C(h) - C(l) > \underline{t}(\mathcal{M}(0))$, the optimal institutional design is:*
   - *FDF iff $a > \lambda^\star(a) = \frac{\mathbb{E}_F[H] - \mathbb{E}_G[H]}{\mathbb{E}_F[H] - \mathbb{E}_F[\min(H_1, H_2)]}$*
   - *ZDF otherwise.*

where $\lambda^\star(a) > \lambda'(a, C)$ *always holds.*

Proof: See Appendix 2.9.9.

2.6.3 Strategic Delegation

The optimal design result can be interpreted as a strategic delegation result. When a designer chooses a hurdle rate framework, he always combines it with the softest possible mandate which induces high effort. This mandate is independent of the strength of society's reassurance motive. Thus, there are cases where the designer optimally delegates to a tougher sender while in others he optimally delegates to a softer sender.

Strategic delegation in this context refers to the designer, e.g. with preferences λ, choosing a central bank mandate with different preferences, e.g. with preferences ω where $\omega \neq \lambda$. The delegation result then follows from Figure 2.4. While the optimal design and the question whether a HRF is chosen depend on three factors (cost of effort, $C(h) - C(l)$; benefit of effort, a; strength of society's reassurance motive, λ), the optimal mandate in a HRF - provided a HRF is chosen - is always the toughest possible mandate which just induces high effort $\tilde{\mathcal{M}}$. This mandate $\tilde{\mathcal{M}}$ is independent of the strength of society's reassurance motive (λ) and only depends on the cost and benefit of effort ($C(h) - C(l); a$). However, for every welfare function (λ) there exists a parameter combination of low costs of effort and high benefits (a) at which $\{\text{HRF}, \tilde{\mathcal{M}}\}$ is the optimal design. Thus, there exist cases where delegating to a tougher CB is optimal ($\tilde{\omega} < \lambda$) while in others delegating to a softer CB ($\tilde{\omega} > \lambda$) is optimal.

Institutional Design: Benefit of Delegation 3 *In a hurdle rate framework, the optimal mandate is independent of the strength of the reassurance motive in welfare. Thus, the optimal central bank mandate generically differs from welfare maximisation. There exist cases where the optimal mandate is tougher than welfare while in other cases it is softer.*

A tougher mandate than welfare provides stronger incentives and is thus optimal if these additional incentives result in an increase in effort and thus an improvement in bank health without causing a strong reduction in reassurance. In the communication game, the tougher mandate results in an equilibrium where banks fail a stress test more often. The tougher mandate is therefore a mechanism through which the designer can partially commit to being tough on weak banks. This makes the threat of revealing bank weakness credible in a larger part of the state space and thus results in strong incentives for banks.

More surprisingly, it is possible that a softer mandate than welfare is optimal. This is the case when a central bank with preferences equal to welfare would be tougher than needed to incentivise high effort. Then, a softer mandate which still achieves high effort provides more reassurance without resulting in lower health fundamentals. This is an improvement as long as welfare has at least some, potentially very small, reassurance

motive. In the communication game, a softer mandate results in an equilibrium where banks fail the stress test less often. Correspondingly, banks pass more often which in turn means that $\{s, p, p\}$ is used even if bank health is further apart. Thus, $\{s, p, p\}$ messages result in higher beliefs which means that the design achieves more reassurance overall.

In addition to delegation results, a HRF implies that stress tests become an informational contagion channel.

2.7 Informational Contagion

We show that in a HRF, stress tests become an informational contagion channel, i.e. the health of one bank affects beliefs about the health of the other bank. Changes in one bank's fundamental affect what message the regulator sends and thus what information financial markets receive. This in turn affects the beliefs financial markets hold, not just on the bank which experienced a change in its health, but also on the other bank which did not experience a change in its health. Informational contagion can take different forms: There are cases where a deterioration in the fundamental health of one bank results in lower beliefs about the other bank. More surprisingly, there are also cases where a deterioration in the health of one bank *increases* the belief about the other bank. As a HRF is the socially optimal design when costs of effort are low and the effect of effort on health is large relative to the weight society places on the weak bank, this implies that contagion can be a feature of a socially optimal design and thus need not reduce welfare.

2.7.1 Contagion Patterns

A deterioration in the health of one bank *decreases* the belief about the health of another bank when the regulator responds by lowering the level of stress. Compare (H_1, H_2) close to but above the 45-degree line[36] where $\{H_2, p, p\}$ is sent, to a (H_1, H_2') with $H_2' < H_2$. There are (H_1, H_2') at which the regulator sends $\{H_2', p, p\}$.[37] In this case, lowering the level of stress has allowed the regulator to continue to send a pass message about the weak bank instead of revealing its weakness, but has the cost of resulting in H_1 passing a weaker stress test. As a result, the beliefs about both banks are lower.

A deterioration in the health of one bank *improves* the belief about the health of another bank when the regulator responds by changing the type of message, i.e. by increasing the level of stress. For example, suppose that initially H_1 and H_2 are far apart, i.e. the health state pair is close to but below the $x(H_2)$-indifference frontier. Then already a small fall in H_2 to H_2' results in the regulator switching message from $\{H_2, p, p\}$ to $\{H_1, p, f\}$. The regulator has increased the level of stress. This amplifies the effect on H_2, i.e. beliefs μ_2

[36] $H_1 > H_2; H_1 - H_2 > 0$ but small.

[37] This is the result for small deteriorations of H_2. How small they can be depends on the sender's mandate, i.e. on how often $\{s, p, p\}$ is sent in equilibrium.

fall more strongly than fundamental health, but results in higher beliefs for the bank that continues to pass as it now passes a tougher stress test.

2.7.2 Contagion in the Optimal Design

First, we show that contagion is not socially undesirable but instead is a feature of the socially optimal design. Then, we show that the optimal design exhibits contagion for a large set of parameters.

While contagion effects are often viewed as an undesirable feature of the financial system in our model stress test becoming an information contagion channel is a feature of the socially optimal institution design. This positive interpretation of contagion as being part of a desirable design is driven by the no-commitment assumption. In our model, which assumes that the regulator cannot commit, a HRF is the unique framework which combines intermediate levels of incentives and reassurance. This makes HRF the optimal framework for many cost parameters. HRF achieves this intermediate combination by creating a trade-off for the regulator between supporting both banks or just one. Hence, the same trade-off that results in HRF combining some incentives and some reassurance also makes HRF have contagion comparative statics. To see this more clearly, suppose a regulator in the no commitment case had two hurdles available. This would remove contagion, but also no longer achieve a combination of incentives and insurance. It is in equilibrium equivalent to FDF, thus achieves full incentives but not reassurance.[38] Suppose a regulator can commit. Then he could combine partial incentives and partial reassurance without causing contagion, e.g. by choosing $s = 0.5$ for all health realisations. Hence, the communication framework which creates contagion is the socially optimal design because the mechanism that creates the contagion comparative static is the same mechanism which allows the designer ex-ante to combine intermediate levels of both reassurance and incentives.

Contagion arises in a HRF for all sender mandates. the two contagion patterns arise as long as both $\{s, p, p\}$ and $\{s, p, f\}$ messages are sent in equilibrium. We showed in section 2.4 that his is the case for all sender mandates including for a sender who places no extra weight on the weak bank. Thus, whenever HRF is the socially optimal design, contagion is a feature of this design.

2.8 Conclusion

We showed that central banks face a time inconsistency problem when publishing bank stress test results. Before a stress test, they want to *appear* tough as the threat of letting banks fail the stress test incentivizes prudent behaviour. After the stress test, they want

[38] A formal proof is provided in Appendix 2.9.1

to *act* soft by releasing only partial information in order to reassure financial markets about bank health.

We characterised an institutional design solution to this commitment problem: a social planner specifies the central bank's mandate and sets the framework within which the central bank communicates. We find that a hurdle rate framework, where all banks are judged to pass or fail relative to a common threshold, is optimal in many settings as it generates intermediate levels of both incentives and reassurance.

In a hurdle rate framework, *strategic delegation* provides additional benefits. This means the planner maximises ex-ante welfare by setting the central bank's mandate to differ from ex-post welfare maximisation. The optimal mandate can differ in either direction. If taking prudent actions has high costs for banks, the planner optimally specifies the CB's mandate to be "tougher" than welfare maximisation, so that the CB is less concerned about supporting weak banks than society. This makes it credible that the CB will reveal bank weakness by letting weak banks fail the stress test and thus generates strong incentives. If the costs of prudent actions are low, the optimal mandate is "softer" than welfare maximisation, so that the CB is very concerned about weak banks. This achieves more reassurance while still generating sufficient incentives to induce the prudent action.

Informational contagion can be a feature of an optimal design in the presence of a time-inconsistency problem and thus need not reduce welfare. This arises because, in a hurdle rate framework in some cases, the central bank's optimal response to a deterioration in the health of one bank is to lower the severity of stress so that all banks continue to pass. For all banks, the resulting pass results are interpreted less favourably than at the initial higher severity of stress. Thus, a change in one bank's health leads the CB to send a different message, which in turn affects beliefs financial markets hold about all banks. More surprisingly, there are also cases where a deterioration of one bank's health improves the beliefs about other banks. This is the case when the CB raises the level of stress. As a result, the weak bank fails the test and beliefs about its health deteriorate. However, beliefs about other banks improve as they pass a tougher test.

These results lead to a reinterpretation of past stress tests: While stress tests by the EBA or its predecessor (the CEBS) were repeatedly criticised for begin too soft,[39] which lead some to question the regulator's competence, our model suggests that these soft stress tests may have been the optimal response of a regulator facing an unhealthy banking system and subject to a time inconsistency problem. Importantly, when the banking system is unhealthy, low levels of stress are optimal for all mandates, including "tough" mandates which focus on the average health in the banking system and place no extra weight on supporting weak banks.

This paper adds a new dimension to the debate on stress test communication by

[39] "Which part of 'stress test' do the eurozone's policy makers not understand? That so many European banks passed the annual exams in July yet still had their shares trashed by investors says it all: the pass mark was too low and the questions were too narrow." Financial Times, Lex, European stress tests: a grim backdrop, 6 October 2011.

highlighting that in addition to the communication framework also the central bank's mandate affects welfare. Current central bank designs do not specify what objective the central bank should pursue when communicating stress test results. Absent an explicit mandate, central banks are likely to maximise ex-post welfare. Thus, our paper suggests that unexploited welfare gains exist: Central banks should be given an explicit mandate which differs from ex-post welfare maximisation as this difference generates gains from strategic delegation.

In practice, the optimal mandate also depends on whether the authority conducting the stress test is simultaneously the bank supervisor or whether the two tasks are carried out by different authorities. If one authority carries out both tasks, this creates implicit incentives, i.e. this authority would have additional reasons to avoid giving out fail grades as this would reflect negatively on its past performance in the role as supervisor. The authority would therefore act "softer" than implied by its mandate. Two remedies are possible. Stress testing can be assigned to an independent authority which is not the supervisor.[40] In that case, our results on the optimal mandate apply. If stress testing and supervision are to be carried out by one institution,[41] the mandate should be "tougher" than implied by our results such that the additional toughness of the mandate offsets the softness created by implicit incentives. Future research should further explore the interplay of explicit and implicit incentives in central bank communication.

[40] This is the case in the Euro Area where the EBA is responsible for stress tests while the SSM supervises banks. Our model suggests that this institutional separation has the benefit of creating stronger incentives for banks to behave prudently than stress tests run by the SSM would achieve. This is not an argument against data sharing or cooperation. Rather, our model suggests that it is beneficial if the EBA interprets its role in stress testing as an independent check on the SSM. In particular, it is important that the EBA is responsible for the severity of the adverse scenario in the stress test.

[41] This is the case in the UK where the Bank of England is responsible for stress tests and bank supervision.

2.9 Appendix

2.9.1 Two Hurdles Result in Full Disclosure

This appendix proves that a "two hurdle communication framework", where the central bank chooses *two* hurdles and truthfully communicates pass or fail for each bank relative to every hurdle, results in financial markets learning bank health perfectly and is thus analogous to the Full Disclosure Framework. The messages a CB sends in such a "two hurdle communication framework" can be expressed as $\{s^a, o_1^a, o_2^a; s^b, o_1^b, o_2^b\}$ where s^j denote a level of stress and $o_i^j \in \{p, f\}$ is the outcome (pass or fail) for bank i relative to stress level j.

Proposition 2.7 *When the central bank chooses* two *hurdles, financial markets learn the health state pair* (H_1, H_2) *perfectly in every equilibrium of this communication game.*

Proof: We prove three lemmata which together prove proposition 2.7.

Lemma 2.9.1.1 *There cannot exist a fully pooling equilibrium, i.e. an equilibrium where no information is released.*

Proof: In any fully pooling equilibrium, the CB has to always send the same message or randomise over the same messages. The only message which is feasible from all health state pairs is $\{0, p, p; 0, p, p\}$, i.e. both hurdles are set at zero and all banks pass relative to both hurdles. In a fully pooling equilibrium, these messages would result in financial market having beliefs $\mu_1 = E(H_1)$, $\mu_2 = E(H_2)$. This is the only possible candidate fully pooling equilibrium.

This candidate equilibrium can never be an equilibrium as a strictly beneficial deviation for the CB exists for any possible interpretation rule which financial markets can use. When $H_1 = 1, H_2 = 1$, there exists a message which is true and which can only be sent from $(H_1 = 1, H_2 = 1)$. This message is $\{1, p, p; 1, p, p\}$. Financial markets must interpret this as $\mu_1 = 1, \mu_2 = 1$ as that is the only health state pair for which this message is feasible. The sender with objectives $V(\mu_1, \mu_2)$ always benefits from this deviation since $\frac{\partial V}{\partial \mu_i} > 0 \; \forall \; i$ and $1 > E(H_i) \; \forall \; i$. Thus, the sender does not want to act as would be required in a fully pooling equilibrium. *q.e.d.*

Lemma 2.9.1.2 *There cannot exist a partially pooling equilibrium, i.e. an equilibrium where for some health state pairs* (H_1, H_2) *the same message is sent.*

Proof: In any partially pooling equilibrium, the central bank has to send the same message or randomise over the same messages for all health state pairs which are part of the pool. Let Ω denote the set of health state pairs (H_1, H_2) which form the pool. First, we outline

how such a pool could be implemented. Then, we show that for every pool and every implementation a strictly beneficial deviation exists.

How could a pool be implemented? Define $\bar{H}$ to be the highest health level in the pool Ω, that is:

$$\bar{H} = \max \left[\sup_{H_1 \in \Omega} H_1, \sup_{H_2 \in \Omega} H_2 \right] \tag{2.12}$$

and $\underline{H}$ to be lowest health level in the pool:

$$\underline{H} = \min \left[\inf_{H_1 \in \Omega} H_1, \inf_{H_2 \in \Omega} H_2 \right] \tag{2.13}$$

Such a pool can be implemented by choosing hurdles s^1, s^2 which are not in the interval $[\underline{H}, \bar{H}]$. If at least one hurdle was in the interval, this would reveal information and destroy the pool.

Having outlined all possible implementations of a pool, we proceed to show that for every pool and every implementation, it is always possible for the central bank to deviate and benefit strictly. Hence, no partially pooling equilibrium can exist.

We introduce the term "truthful pay-off" and define it as the pay-off the central bank would achieve if health state pairs were perfectly revealed. That is, the truthful pay-off at (H_1, H_2) is $V(H_1, H_2)$.

Sublemma 2.9.1.1 *A pool across different truthful pay-offs never exists.*

In every pool Ω, there exists a maximal truthful pay-off set M, which is the set of health state pairs that achieve the highest truthful pay-off. Denote one such health state pair (m, n). The central bank strictly benefits from deviating from the pool at (m, n) and can feasibly do so. Suppose $m > n$. The central bank can deviate by sending $\{s^1 = m, p, f; s^2 = n, p, p\}$. This must be a deviation as it cannot be the implementation of pool Ω. Financial markets know that the health state pair must be to the top-right of (m, n), i.e. that $H_1 \geq m$ and $H_2 \geq n$. For any belief financial markets hold over the states supporting this message, the central bank is strictly better off than when sending the pooling message and pooling in Ω.

Importantly, even for the worst possible interpretation about where the deviation to $\{s^1 = m, p, f; s^2 = n, p, p\}$ comes from, the deviation is still a strict improvement for the central bank. Hence, no partially pooling equilibrium across different truthful pay-off pairs can exist. *q.e.d.*

Sublemma 2.9.1.2 *A pool of different health state pairs with the same truthful pay-off never exists.*

It is not possible for health state pairs with the same truthful pay-off to form a pool because there does not exist a message which is feasible at bank health state pairs with

the same truthful pay-off and which is not feasibly at lower truthful health state pairs. Any message, with which health state pairs with the same truthful pay-off attempt to pool, would therefore attract health state pairs with lower truthful pay-offs and make it optimal for high health state pairs to separate.

Thus, to avoid pooling with health state pairs that have a lower truthful pay-off, the central bank sends $\{H_1, p, f; H_2, p, p\}$ for every (H_1, H_2) with $H_1 > H_2$. This message is sent because it is only feasible to send it from (H_1, H_2) or health states where at least one health outcome is higher. All these health state pairs have strictly higher truthful pay-offs. Thus, messages which separate from lower truthful pay-off states make pooling with equal truthful pay-off states impossible, as the messages that achieve separation from lower states are only feasible at a unique health state pair for a given truthful pay-off. *q.e.d.*

Lemma 2.9.1.3 *A fully revealing equilibrium exists.*

Proof: Given recipients believe that every health state pair results in a unique message (i.e. a message that is only send from one health state pair) and given that this is feasible (i.e. $s^1 = H_1, s^2 = H_2$ is a possible way of fully revealing the health state pair in the "two hurdle communication framework") it is optimal to send messages accordingly. *q.e.d.*

Proof of Proposition 2.7: Combining the three lemmata above proves proposition 2.7. *q.e.d.*

2.9.2 Proof of Theorem 2.1

This Appendix proves Theorem 2.1 in the general case where a bank's health can be drawn from any distribution and where different banks are potentially drawn from different distributions.

The strategy profile (S, R) specified in Theorem 2.1 is an equilibrium of the Hurdle Rate Framework communication game if and only if:

1. Given R; S maximises the sender's mandate $\mathcal{M}$ at every health state pair (H_1, H_2).
2. Given S; R is correct in the sense of using all available information.

We prove these in turn, using the logic of revealed preference.

Part 1: Given the interpretation rule R, for a given outcome pair $\{o_1, o_2\}$ higher levels of stress result in higher beliefs μ_i for both banks. Since $\mathcal{M}$ is strictly increasing in $\mu_i \, \forall \, i$, this gives rise to the following Lemma:

Lemma 2.9.2.1 *Given the interpretation rule R, the CB always prefers the highest level of stress possible for a given outcome pair to any alternative level of stress for that outcome pair.*

Formally, this means that the CB strictly prefers $\{\min(H_1, H_2), p, p\}$ to all $\{s', p, p\}$ where $s' < \min(H_1, H_2)$; the CB strictly prefers $\{H_1, p, f\}$ to all $\{s'', p, f\}$ where $s'' < H_1$; the CB strictly prefers $\{H_2, f, p\}$ to all $\{s''', f, p\}$ where $s''' < H_2$.

Proof: Since $\mathcal{M}$ is strictly increasing in μ_i and R specifies that μ_i are increasing in s for a given outcome pair, this Lemma holds. *q.e.d.*

Lemma 2.9.2.1 means that the CB chooses either $s = \max(H_1, H_2)$ or $s = \min(H_1, H_2)$ in response to interpretation rule R. We proceed by showing that the sets of health state pairs (H_1, H_2) where a given $\{s, o_1, o_2\}$ is sent can be expressed via indifference frontiers $x(H_2)$ and $y(H_2)$ as in Theorem 2.1.

Lemma 2.9.2.2 *If, at a health state pair (H_1, H_2) with $H_2 > 0$, sending $\{H_1, p, f\}$ is the message which maximises the CB's mandate $\mathcal{M}$, then for all (H_1, H_2') with $H_2' < H_2$ it is mandate maximising to send $\{H_1, p, f\}$.*

Proof: At (H_1, H_2), $\{H_1, p, f\}$ was revealed preferred to $\{H_2, p, p\}$, i.e. $V(\{H_1, p, f\}) > V(\{H_2, p, p\})$. Given the interpretation rule R specified in Theorem 2.1, $\{H_2, p, p\}$ results in higher beliefs μ_i about both banks and therefore in a higher sender pay-off than $\{H_2', p, p\}$, i.e. $V(\{H_2, p, p\}) > V(\{H_2', p, p\})$. Thus $\{H_1, p, f\}$ must be preferred to $\{H_2', p, p\}$, i.e. we must have $V(\{H_1, p, f\}) > V(\{H_2', p, p\})$. *q.e.d.*

Lemma 2.9.2.3 *If at a health state pair (H_1, H_2) sending $\{H_1, p, f\}$ is the message which maximises the CB's mandate $\mathcal{M}$, then for all (H_1', H_2) with $H_1' > H_1$ it is mandate maximising to send $\{H_1', p, f\}$.*

Proof: Analogous to the proof of Lemma 2.9.2.2: Revealed: $V(\{H_1, p, f\}) > V(\{H_2, p, p\})$. Since higher levels of stress for a given outcome pair result in higher posteriors and higher posteriors result in a higher pay-off V: $V(\{H_1', p, f\}) > V(\{H_1, p, f\})$. Hence, we must have: $V(\{H_1', p, f\}) > V(\{H_2, p, p\})$. *q.e.d.*

Lemma 2.9.2.4 *There never exists a perfectly vertical part of the $x(H_2)$ indifference frontier.*[42]

Proof by contradiction: Suppose a perfectly vertical part of the $x(H_2)$ indifference frontier existed. Then on this part there would be at least two health state pairs, (H_1^b, H_2) and (H_1^a, H_2) with $H_1^a > H_1^b$, at which the sender is both times indifferent between $\{H_2, p, p\}$ and $\{H_1, p, f\}$. At (H_1^b, H_2), the sender is indifferent between $\{H_1^b, p, f\}$ and $\{H_2, p, p\}$ if $V(\{H_1^b, p, f\}) = V(\{H_2, p, p\})$. At (H_1^a, H_2), the sender is indifferent between $\{H_1^a, p, f\}$ and $\{H_2, p, p\}$ if $V(\{H_1^a, p, f\}) = V(\{H_2, p, p\})$. This would imply that $V(\{H_1^a, p, f\}) = V(\{H_1^b, p, f\})$. But this is not possible given the interpretation rule R.

[42] A similar logic applies to $y(H_2)$.

Specifically, this would violate the monotonicity property of beliefs resulting from R which ensures that $\{s, p, f\}$ results in higher μ_i than $\{s', p, f\}$ $\forall$ $s > s'$ and $\forall \mu_i$. Hence, $\{H_1^a, p, f\}$ must be strictly preferred to $\{H_1^b, p, f\}$. Therefore, no perfectly vertical part of the $x(H_2)$ indifference frontier can exist. *q.e.d.*

Lemma 2.9.2.1 means that the CB chooses either $s = \max(H_1, H_2)$ or $s = \min(H_1, H_2)$. Lemma 2.9.2.2 - 2.9.2.3 showed that the health state pairs where $\{H_1, p, f\}$ is sent form a connected set in the top left part of the state space $(H_1, H_2) \in [0, 1]^2$. Lemma 2.9.2.4 showed that this connected set does not have a perfectly vertical boundary. The boundary also cannot be decreasing. Thus, it is without loss of generality to describe the CB's optimal strategy as sending $\{H_1, p, f\}$ for all (H_1, H_2) above an $x(H_2)$ indifference frontier which is monotonically increasing, i.e. as described in Theorem 2.1. This proves part 1. *q.e.d.*

Part 2: Given the strategy S, the interpretation rule is correct in the sense of using all available information. Naive interpretations of $\{s, p, p\}$ would result in $\mu_1 = \mathbb{E}[H \mid s \leq H \leq 1]$. Partially sophisticated interpretations would realise that $s = \min(H_1, H_2)$ and result in $\mu_1 = \alpha\ s + (1 - \alpha)\ \mathbb{E}_F[H \mid s \leq H \leq 1]$. Fully sophisticated beliefs realise that $s = \min(H_1, H_2)$ and additionally that the sender preferred $s = \min(H_1, H_2)$ to $s = \max(H_1, H_2)$. Thus, fully sophisticated beliefs result in $\mu_1 = \alpha\ s + (1 - \alpha)\ \mathbb{E}_F[H \mid s \leq H \leq x(s)]$. These beliefs correspond to the beliefs in Theorem 2.1. *q.e.d.*

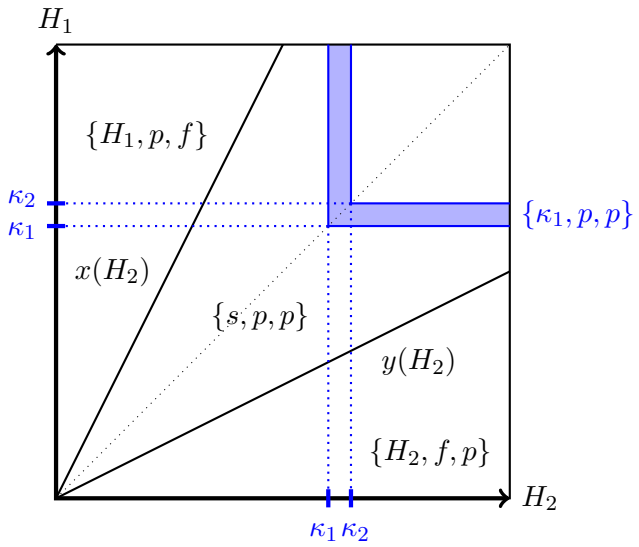
2.9.3 Equilibria in the Communication Game

The equilibrium of the hurdle rate communication game described in Theorem 2.1 is not the unique weak PBE but it is the only weak PBE which satisfies the notion of monotonicity outlined in Definition 2.4. This appendix first characterises an alternative weak PBE which exists but violates monotonicity. Then, this appendix proves that the equilibrium described in Theorem 2.1 is the unique weak PBE which satisfies monotonicity.

An Alternative Equilibrium which Violates Monotonicity

Characterisation of the alternative equilibrium:

Figure 2.5: Alternative Equilibrium: Communication Strategy



Formally, the central bank's actions are:

$s =$	H_1	iff	$H_1 > x(H_2)$		$\rightarrow$	$\{H_1, p, f\}$
$s =$	H_2	iff	$y(H_2) > H_1$		$\rightarrow$	$\{H_2, f, p\}$
$s =$	$\min(H_1, H_2)$	iff	$x(H_2) \geq H_1 \geq y(H_2)$ and			
			either $\min(H_1, H_2) < \kappa_1$ or $\min(H_1, H_2) \geq \kappa_2$		$\rightarrow$	$\{\min(H_1, H_2), p, p\}$
$s =$	κ_1	iff	$x(H_2) \geq H_1 \geq y(H_2)$ and			
			$\kappa_1 \leq \min(H_1, H_2) < \kappa_2$		$\rightarrow$	$\{\kappa_1, p, p\}$

where $x(H_2)$ and $y(H_2)$ have the same properties as in the equilibrium described in Theorem 2.1.

Recipients form beliefs accordingly. Denoting the set of health state pairs from which $\{\kappa_1, p, p\}$ is sent as Ω, and assuming that bank health distributions are identical, beliefs are given by:

$\{s, p, f\}$		is interpreted as	$\mu_1 = s; \quad \mu_2 = \mathbb{E}[H \mid H < x^{-1}(s)]$
$\{s, f, p\}$		is interpreted as	$\mu_2 = s; \quad \mu_1 = \mathbb{E}[H \mid H < y(s)]$
$\{s, p, p\}$	if $s < \kappa_1$ or $s > \kappa_2$	is interpreted as	$\mu_1 = \mu_2 = \frac{1}{2}\, s + \frac{1}{2}\, \mathbb{E}[H \mid s \leq H \leq x(s)]$
	if $\kappa_1 < s < \kappa_2$	is interpreted as	$\mu_1 = \mu_2 = s$
	if $s = \kappa_1$	is interpreted as	$\mu_1 = \mu_2 = \mathbb{E}[H \mid (H_1, H_2) \in \Omega]$
$\{s, f, f\}$		is interpreted as	$\mu_1 = \mu_2 = 0 \; \forall \; s > 0$

The messages $\{s, p, p\}$ for $\kappa_1 < s < \kappa_2$ are off the equilibrium path, i.e. they are never sent in equilibrium. A PBE therefore implies no restrictions on how these messages are interpreted other than that beliefs are formed over health state pairs from which the

message could have been sent. In the description above, these messages are interpreted in the most sceptical way possible.

Proof: This alternative equilibrium is a PBE. To economise on space and provide a clearer intuition, we show existence of such an alternative equilibrium for the case where both banks follow a uniform distribution and where $\kappa_1 = 0.6$ and $\kappa_2 = 0.65$. For these values we have $\mu_1 = \mu_2 = \mathbb{E}[H \mid (H_1, H_2) \in \Omega] = \frac{431}{600} \approx 0.718$.

Next, we verify that given these beliefs, the central bank does not want to deviate from its messaging strategy. For the indifference frontiers $x(H_2)$ and $y(H_2)$ the same logic as for the equilibrium described in Theorem 2.1 applies. This proof thus focuses on the region where $\min(H_1, H_2) \in [\kappa_1, \kappa_2[$.

For $\min(H_1, H_2) < \kappa_1$, it is technologically not possible to send $\{\kappa_1, p, p\}$.

For $\min(H_1, H_2) \in [\kappa_1, \kappa_2[$ it needs to be optimal to send $\{\kappa_1, p, p\}$. All alternative messages of the form $\{s, p, p\}$ where $s < \kappa_1$ result in a lower sender pay-off as both μ_1 and μ_2 are lower; where $s > \kappa_2$ these messages are technologically not possible to send; where $s \in [\kappa_1, \kappa_2[$ the resulting belief is $\mu_1 = \mu_2 = s$ which is lower than $\mu_1 = \mu_2 = \mathbb{E}[H \mid (H_1, H_2) \in \Omega] = \frac{431}{600} \approx 0.718$ and thus also is not a beneficial deviation. Messages of the form $\{s, p, f\}$ are available, but have been revealed to result in lower sender pay-offs than $\{s, p, p\}$ with s below but close to κ_1.

For $\min(H_1, H_2) = \kappa_2$ it needs to be optimal to send $\{s, p, p\}$ rather than $\{\kappa_1, p, p\}$. For the example values here, we have $\{0.65, p, p\}$ which results in $\mu_1 = \mu_2 = \frac{1}{2}0.65 + \frac{1}{2}\mathbb{E}[H \mid H > 0.65] = 0.7375$. Thus, no beneficial deviation exists. For $\min(H_1, H_2) > \kappa_2$ sending $\{\min(H_1, H_2), p, p\}$ is even more beneficial.

At the health state pairs where the CB's strategy is to set $s = \max(H_1, H_2)$, the message $\{\kappa_1, p, p\}$ is technologically not available. Hence, no beneficial deviation exists which confirms that this equilibrium is a PBE. *q.e.d.*

Violations of monotonicity: In this equilibrium, $\{0.6, p, p\}$ is interpreted as $\mu_1 = \mu_2 = \mathbb{E}[H \mid (H_1, H_2) \in \Omega] = \frac{431}{600} \approx 0.718$. The off-equilibrium path message $\{0.61, p, p\}$ is interpreted as $\mu_1 = \mu_2 = 0.61$. This violates monotonicity which requires that passing a tougher stress test leads to higher beliefs while in this case beliefs fall as banks pass at $s = 0.61$ relative to $s = 0.6$. Thus, despite passing a higher level of stress, beliefs are lower. That is exactly what makes this equilibrium possible and it is a violation of monotonicity.

Uniqueness Given Monotonicity

Lemma 2.9.3.1 *The equilibrium characterised in Theorem 2.1 is the unique weak PBE which satisfies monotonicity.*

Proof: The proof establishes that, for beliefs which satisfy monotonicity, the CB's unique optimal strategy is to act as outlined in Theorem 2.1. As beliefs have to be correct in

equilibrium, only the beliefs specified in Theorem 2.1 are possible equilibrium beliefs. Thus, the equilibrium (S, R), i.e. the action belief pair, is the unique equilibrium given monotonicity. Monotonicity is defined formally in Definition 2.4.

Sublemma 2.9.3.1 *Given beliefs which satisfy monotonicity, the CB always prefers the highest level of stress possible for a given outcome pair $\{o_1, o_2\}$ to any alternative level of stress for that outcome pair.*

Formally, this means that the CB strictly prefers $\{\min(H_1, H_2), p, p\}$ to all $\{s', p, p\}$ where $s' < \min(H_1, H_2)$; the CB strictly prefers $\{H_1, p, f\}$ to all $\{s'', p, f\}$ where $s'' < H_1$; the CB strictly prefers $\{H_2, f, p\}$ to all $\{s''', f, p\}$ where $s''' < H_2$.

Proof: Since $\mathcal{M}$ is strictly increasing in μ_i and by monotonicity μ_i is increasing in s for a given outcome, this Sublemma holds. *q.e.d.*

Sublemma 2.9.3.1 means that the CB chooses either $s = \max(H_1, H_2)$ or $s = \min(H_1, H_2)$. We proceed by showing that the sets of health state pairs (H_1, H_2) where a given $\{s, o_1, o_2\}$ is sent can be expressed via indifference frontiers $x(H_2)$ and $y(H_2)$ as in Theorem 2.1.

Sublemma 2.9.3.2 *If, at a health state pair (H_1, H_2) with $H_2 > 0$, sending $\{H_1, p, f\}$ is the message which maximises the CB's mandate $\mathcal{M}$, then for all (H_1, H_2') with $H_2' < H_2$ it is mandate maximising to send $\{H_1, p, f\}$.*

Proof: At (H_1, H_2), $\{H_1, p, f\}$ was revealed preferred to $\{H_2, p, p\}$, i.e. $V(\{H_1, p, f\}) > V(\{H_2, p, p\})$. Given beliefs satisfy monotonicity, $\{H_2, p, p\}$ results in higher beliefs μ_i about both banks and therefore in a higher sender pay-off than $\{H_2', p, p\}$, i.e. $V(\{H_2, p, p\}) > V(\{H_2', p, p\})$. Thus $\{H_1, p, f\}$ must be preferred to $\{H_2', p, p\}$, i.e. we must have $V(\{H_1, p, f\}) > V(\{H_2', p, p\})$.

q.e.d.

Sublemma 2.9.3.3 *If at a health state pair (H_1, H_2) sending $\{H_1, p, f\}$ is the message which maximises the CB's mandate $\mathcal{M}$, then for all (H_1', H_2) with $H_1' > H_1$ it is mandate maximising to send $\{H_1', p, f\}$.*

Proof: Analogous to the proof of Sublemma 2.9.3.2: Revealed: $V(\{H_1, p, f\}) > V(\{H_2, p, p\})$. Since by monotonicity of beliefs higher levels of stress for a given outcome pair result in higher posteriors and higher posteriors result in a higher pay-off V: $V(\{H_1', p, f\}) > V(\{H_1, p, f\})$. Hence, we must have: $V(\{H_1', p, f\}) > V(\{H_2, p, p\})$. *q.e.d.*

Sublemma 2.9.3.4 *There never exists a perfectly vertical part of the $x(H_2)$ indifference frontier.*[43]

[43] A similar logic applies to $y(H_2)$.

Proof by contradiction: Suppose a perfectly vertical part of the $x(H_2)$ indifference frontier existed. Then on this part there would be at least two health state pairs, (H_1^b, H_2) and (H_1^a, H_2) with $H_1^a > H_1^b$, at which the sender is both times indifferent between $\{H_2, p, p\}$ and $\{H_1, p, f\}$. At (H_1^b, H_2), the sender is indifferent between $\{H_1^b, p, f\}$ and $\{H_2, p, p\}$ if $V(\{H_1^b, p, f\}) = V(\{H_2, p, p\})$. At (H_1^a, H_2), the sender is indifferent between $\{H_1^a, p, f\}$ and $\{H_2, p, p\}$ if $V(\{H_1^a, p, f\}) = V(\{H_2, p, p\})$. This would imply that $V(\{H_1^a, p, f\}) = V(\{H_1^b, p, f\})$. But this is not possible given that beliefs satisfy monotonicity, i.e. this would violate the monotonicity property of beliefs. Hence, $\{H_1^a, p, f\}$ must be strictly preferred to $\{H_1^b, p, f\}$. Therefore, no perfectly vertical part of the $x(H_2)$ indifference frontier can exist. *q.e.d.*

Sublemma 2.9.3.1 means that the CB chooses either $s = \max(H_1, H_2)$ or $s = \min(H_1, H_2)$. Sublemma 2.9.3.2 - 2.9.3.3 showed that the health state pairs where $\{H_1, p, f\}$ is sent form a connected set in the top left part of the state space $(H_1, H_2) \in [0, 1]^2$. Sublemma 2.9.3.4 showed that this connected set does not have a perfectly vertical boundary. The boundary also cannot be decreasing.

Thus, only a central bank strategy as described in Theorem 2.1 is optimal when beliefs satisfy monotonicity, i.e. the CB always sends $\{H_1, p, f\}$ for all (H_1, H_2) above an $x(H_2)$ indifference frontier which is monotonically increasing. *q.e.d.*

2.9.4 Proof of Proposition 2.1: Ordering Designs by Reassurance

This Appendix proves that institution designs $\mathcal{D}$ can be ordered by how much reassurance they provide, i.e. by $\mathbb{E}[\min(\mu_1, \mu_2)]$, as described in Proposition 2.1.

Lemma 2.9.4.1 *The Zero Disclosure Framework generates* $\mathbb{E}[\min(\mu_1, \mu_2)] = \mathbb{E}[H]$.

We term this full reassurance.

Proof: Since in ZDF never any information is released, recipients can never update their prior belief $\mathbb{E}[H_i]$. Thus, we always have $\mu_i = \mathbb{E}[H_i]$ and therefore $\mathbb{E}[\min(\mu_1, \mu_2)] = \mathbb{E}[\min(\mathbb{E}(H), \mathbb{E}(H))] = \mathbb{E}[H]$. We refer to this as 'full reassurance' as it is the highest possible level of reassurance achievable. *q.e.d.*

Lemma 2.9.4.2 *The Full Disclosure Framework generates* $\mathbb{E}[\min(\mu_1, \mu_2)] = \mathbb{E}[\min(H_1, H_2)]$.

We term this zero reassurance.

Proof: Since in FDF health realisations H_i are always fully revealed, recipients always update their belief to the true fundamental: $\mu_i = H_i$. Since this is the case for all health states, the ex-ante *distribution* of fundamentals and beliefs coincide, which is stronger than Bayes plausibility which states that expected fundamental and expected belief coincide.

Mathematically: $\mathbb{E}[\min(\mu_1, \mu_2)] = \mathbb{E}[\min(H_1, H_2)]$. *q.e.d.*

ZDF always generates more reassurance than FDF as $\mathbb{E}[H] \geq \mathbb{E}[\min(\mu_1, \mu_2)]$ holds for all distributions. In the uniform example used to construct the diagrams in the paper: $\mathbb{E}[H] = \frac{1}{2} > \frac{1}{3} = \mathbb{E}[\min(H_1, H_2)]$.

Lemma 2.9.4.3 *For all mandates $\mathcal{M}$, the Hurdle Rate Framework generates a level of reassurance that lies between the levels of FDF and ZDF. The level of reassurance in a HRF is monotonically increasing in the weight the mandate places on the weak bank.*

Proof: We know from equation (2.6) that the $x(H_2)$ indifference frontier can be expressed as $x(H_2) = \frac{2-\omega}{1-2\omega} H_2$ when we assume that $V(\mu_1, \mu_2) = \omega \ \min(\mu_1, \mu_2) + (1-\omega) \ \frac{\mu_1+\mu_2}{2}$ where $\omega \in [0, 1[$ as well as that $H_i \sim U[0, 1] \ \forall \ i$. For ease of exposition, we define $m \equiv \frac{2-\omega}{1-2\omega}$ and use m to characterise the sender's preferences. Since $\omega \geq 0$ our parameter space is $m \geq 2$. For any m, we can back out the underlying ω via $\omega = \frac{m-2}{2m-1}$. Hence, there is a one-for-one mapping between m and ω. Equation (2.9) describes $\mathbb{E}[\min(\mu_1, \mu_2)]$ in a HRF. Using $x(H_2) = mH_2$ this becomes:

$$\begin{aligned}\mathbb{E}[\min(\mu_1, \mu_2)] = &\int_0^{\frac{1}{m}} \Big[\frac{H_2^2}{2m^2} + \frac{3+m}{8}(\frac{m^2-1}{m^2})H_2^2 + \frac{3+m}{4}(m-1)H_2^2 + \frac{1}{4m} - \frac{m}{4}H_2^2\Big] g(H_2) dH_2 \\ &+ \int_{\frac{1}{m}}^1 \Big[\frac{H_2^2}{2m^2} + \frac{3}{8}(\frac{m^2-1}{m^2})H_2^2 + \frac{1}{4}(\frac{m-1}{m})H_2 + \frac{1}{4} - \frac{3}{4}H_2^2 + \frac{1}{2}H_2\Big] g(H_2) dH_2\end{aligned} \tag{2.14}$$

And further to:

$$\mathbb{E}[\min(\mu_1, \mu_2)] = \frac{1}{2} + \frac{-7m^3 + m^2 + 3m - 1}{24m^4} \tag{2.15}$$

This gives rise to three sublemmata which together prove Lemma 2.9.4.3.

Sublemma 2.9.4.1 *Even when the sender's mandate is only concerned with average bank health, a hurdle rate framework implies more reassurance than the full disclosure framework.*

Proof: We know from Lemma 2.9.4.2 that in FDF $\mathbb{E}[\min(\mu_1, \mu_2)] = \mathbb{E}[\min(H_1, H_2)]$ which for these distributional assumptions is $\frac{1}{3}$. Using that the sender's mandate is only concerned with the average bank ($\omega = 0$ thus $m = 2$) in equation (2.15), we obtain $\mathbb{E}[\min(\mu_1, \mu_2)] = \frac{145}{384} \approx 0.378$. Hence, a HRF with mandate $V(\mu_1, \mu_2) = \frac{\mu_1+\mu_2}{2}$ generates more reassurance than FDF. *q.e.d.*

Sublemma 2.9.4.2 *The stronger the weight the mandate places on the weak bank, the more reassurance is provided in equilibrium.*

Proof: We aim to show that $\frac{d\mathbb{E}[\min(\mu_1,\mu_2)]}{d\omega} \geq 0 \ \forall \ \omega \in [0, 1[$. This corresponds to showing that $\frac{d\mathbb{E}[\min(\mu_1,\mu_2)]}{dm} \geq 0 \ \forall \ m \geq 2$. From (2.15) we obtain:

$$\frac{d\mathbb{E}[\min(\mu_1, \mu_2)]}{dm} = \frac{7m^3 - 2m^2 - 9m + 4}{24m^5} \tag{2.16}$$

and

$$\frac{d^2\mathbb{E}[\min(\mu_1, \mu_2)]}{dm^2} = \frac{-7m^3 + 3m^2 + 18m - 10}{12m^6} \tag{2.17}$$

Solving $\max_m \mathbb{E}[\min(\mu_1, \mu_2)]$ yields the FOC $\frac{d\mathbb{E}[\min(\mu_1,\mu_2)]}{dm} \stackrel{!}{=} 0$ which is satisfied for $m = 1$ and for two lower values of m. The SOC confirms that $m = 1$ is indeed a minimum. Thus, for all $m > 1$ we have $\frac{d\mathbb{E}[\min(\mu_1,\mu_2)]}{dm} > 0$. Hence, reassurance is monotonically increasing in m also for all $m \geq 2$. *q.e.d.*

Sublemma 2.9.4.3 *In the limit case where the sender is purely concerned with the weak bank ($m \to \infty$), i.e. when almost always $\{s, p, p\}$ is sent*[44]*, the amount of reassurance provided approaches that of the zero disclosure framework.*

Proof: Using (2.15) and applying limit theorems:

$$\lim_{m\to\infty} \mathbb{E}[\min(\mu_1, \mu_2)] = \lim_{m\to\infty} (\frac{1}{2}) + \lim_{m\to\infty} (\frac{-7m^3 + m^2 + 3m - 1}{24m^4}) = \frac{1}{2} \tag{2.18}$$

From Lemma 2.9.4.1 we know that for the zero disclosure framework $\mathbb{E}[\min(\mu_1, \mu_2)] = \mathbb{E}[\min(H_1, H_2)] = \frac{1}{2}$. Thus, $\mathbb{E}[\min(\mu_1, \mu_2) \mid HRF, m \to \infty] = \mathbb{E}[\min(\mu_1, \mu_2) \mid ZDF]$. *q.e.d.*

Intuition: Sublemma 2.9.4.3 does not mean that as $m \to \infty$ the minimum posterior is always exactly $\frac{1}{2}$. Rather, it says that the *expected* minimum posterior is $\frac{1}{2}$, or more generally equals the expected fundamental[45], but there still is a distribution of minimum posteriors. Thus, while the first moment of the distribution of $\min(\mu_1, \mu_2)$ under a HRF with $m \to \infty$ approaches the first moment under ZDF, higher moments remain different. Our definition of reassurance as $\mathbb{E}[\min(\mu_1, \mu_2)]$ focuses only on the first moment of the distribution of $\min(\mu_1, \mu_2)$. This definition gives rise to the equality in the limit result. We conjecture that a broader definition of reassurance that includes higher moments of $\min(\mu_1, \mu_2)$ would continue to find that reassurance in a HRF is below reassurance in ZDF, in this case strictly below. The ranking of frameworks in terms of reassurance would then be unaffected.

Proof of Lemma 2.9.4.3: Combining Sublemmata 2.9.4.1, 2.9.4.2, 2.9.4.3 proves Lemma 2.9.4.3. *q.e.d.*

Proof of Proposition 2.1: Combining Lemmata 2.9.4.1, 2.9.4.2, 2.9.4.3 proves Proposition 2.1. *q.e.d.*

[44] Always except when $\min(H_1, H_2) = 0$, i.e. exactly zero.

[45] $\mathbb{E}[\min(\mu_1, \mu_2)] = \mathbb{E}[H_i]$, i.e. $\mathbb{E}[\min(posterior)] = \mathbb{E}[fundamental]$

2.9.5 Proof of Proposition 2.2: Comparative Static on the Sender's Mandate

Let there be two mandates or sender pay-off functions, $V(\mu_1, \mu_2)$ and $U(\mu_1, \mu_2)$, where $U(\mu_1, \mu_2)$ is more concerned with the weak bank, i.e. is more concave.[46] Let the equilibrium in the communication game for $V(\mu_1, \mu_2)$ be characterised by indifference frontier $x(H_2)$ and the interpretation rule R while the equilibrium for $U(\mu_1, \mu_2)$ is characterised by $\tilde{x}(H_2)$ and $\tilde{R}$. On $x(H_2)$ we have, by definition of $x(H_2)$, that $V(\{H_1, p, f\}) = V(\{H_2, p, p\})$. Given the interpretation rule R, $\{H_1, p, f\}$ results in dispersed posteriors $\mu_1' \neq \mu_2'$, while $\{H_2, p, p\}$ results in the same posterior for both banks $\mu_1 = \mu_2$. Moreover, $\mu_2' < \mu_1 = \mu_2 < \mu_1'$. Therefore, a sender with mandate $U(\mu_1, \mu_2)$ which places more weight on the weak bank than $V(\mu_1, \mu_2)$ strictly prefers $\{H_2, p, p\}$ on $x(H_2)$ for interpretation R. For interpretation rule R, $U(\mu_1, \mu_2)$ strictly prefers $\{H_2, p, p\}$ to $\{H_1, p, f\}$ for all (H_1, H_2) between the 45-degree line and the $x(H_2)$-frontier. Thus, the hypothetical frontier $\bar{x}(H_2)$ at which $U(\{H_1, p, f\} \mid R) = U(\{H_2, p, p\} \mid R)$ must lie strictly above $x(H_2)$. However, $\bar{x}(H_2)$ does not characterise an equilibrium as it is based on R rather than the appropriate $\tilde{R}$. Allowing the interpretation rule to adjust to $\tilde{R}$ results in $U(\mu_1, \mu_2)$ strictly preferring $\{H_2, p, p\}$ also on $\bar{x}(H_2)$. Thus, the resulting equilibrium indifference frontier $\tilde{x}(H_2)$ must lie above $\bar{x}(H_2)$. Since $\tilde{x}(H_2) > \bar{x}(H_2) \ \forall \ H_2$ and $\bar{x}(H_2) > x(H_2) \ \forall \ H_2$, we therefore have that $\tilde{x}(H_2) > x(H_2) \ \forall \ H_2$. *q.e.d.*

2.9.6 Proof of Proposition 2.3: Ordering Designs by Incentives

This Appendix proves that institution designs $\mathcal{D}$ can be ordered by the strength of incentives they generate, i.e. by $\mathbb{E}\big[\mu_1(\boldsymbol{l}) \mid h_1, l_2\big] - \mathbb{E}_G[H] > 0$, as described in Proposition 2.3. Low effort leads to bank health being drawn from CDF $G(H)$ and high effort from $F(H)$.

Lemma 2.9.6.1 *The Zero Disclosure Framework generates no incentives.*

Proof: Since in ZDF never any information is released, recipients can never update their prior belief $\mathbb{E}_G[H]$. Thus, we always have $\mu_i = \mathbb{E}_G[H_i]$. This means that the expected posterior is fully determined by conjectured effort (low, leading to distribution G) and completely isolated from actual effort. Formally $\mathbb{E}\big[\mu_1(\boldsymbol{l}) \mid h_1, l_2\big] = \mathbb{E}\big[\mu_1(\boldsymbol{l}) \mid l_1, l_2\big] = \mathbb{E}_G[H]$. *q.e.d.*

Lemma 2.9.6.2 *The Full Disclosure Framework generates full incentives.*

Proof: Since in FDF health realisations H_i are always fully revealed, recipients always update their belief to the true fundamental: $\mu_i = H_i$. Thus, $\mathbb{E}\big[\mu_1(\boldsymbol{l}) \mid h_1, l_2\big] = \mathbb{E}[H_1 \mid h_1] = \mathbb{E}_F[H]$. Thus, incentives are $\mathbb{E}_F[H] - \mathbb{E}_G[H]$ which is strictly positive by the assumption that effort improves the bank health distribution according to an MLRP property. *q.e.d.*

[46] For our functional form example, this corresponds to $\omega_v < \omega_u$. In this proof we do not rely on these functional form assumptions.

Lemma 2.9.6.3 *For all mandates $\mathcal{M}$, the Hurdle Rate Framework generates positive but muted effort incentives, i.e. incentives with a strength that lies between the strength of incentives generated by ZDF and by FDF. The strength of incentives in a HRF is monotonically decreasing in the weight the mandate places on the weak bank.*

Proof: We know from equation (2.6) that the $x(H_2)$ indifference frontier can be expressed as $x(H_2) = \frac{2-\omega}{1-2\omega}H_2$ when we assume that $V(\mu_1, \mu_2) = \omega \ \min(\mu_1, \mu_2) + (1-\omega)\ \frac{\mu_1+\mu_2}{2}$ where $\omega \in [0, 1[$ as well as that $G(H) = H$ and $F(H) = H^2$ and the regulator conjectures that banks choose low effort. As in Appendix 2.9.4, for ease of exposition we define $m \equiv \frac{2-\omega}{1-2\omega}$ and use m to characterise the sender's preferences. Equation (2.11) describes $\mathbb{E}\big[\mu_1(\boldsymbol{l}) \mid h_1, l_2\big]$ in a HRF. Using $x(H_2) = mH_2$ this becomes:

$$\begin{aligned}\mathbb{E}\big[\mu_1(\boldsymbol{l}) \mid h_1, l_2\big] = \int_0^{\frac{1}{m}} \Bigg[&\frac{H_2}{2m}F(\frac{H_2}{m}) + \frac{3+m}{4}\int_{\frac{H_2}{m}}^{H_2} H_1 f(H_1)dH_1 \\ &+ \frac{3+m}{4}H_2[F(mH_2) - F(H_2)] + \int_{x(H_2)}^{1} H_1 f(H_1)dH_1 \Bigg] g(H_2)dH_2 \\ + \int_{\frac{1}{m}}^{1} \Bigg[&\frac{H_2}{2m}F(\frac{H_2}{m}) + \int_{\frac{H_2}{m}}^{H_2} \frac{1}{4} + \frac{3}{4}H_1 f(H_1)dH_1 \\ &+ \big(\frac{1}{4} + \frac{3}{4}H_2\big)\big(1 - F(H_2)\big) \Bigg] g(H_2)dH_2\end{aligned} \tag{2.19}$$

and

$$\mathbb{E}\big[\mu_1(\boldsymbol{l}) \mid h_1, l_2\big] = \frac{1}{48m^6} \quad \big(27m^6 + 15m^5 - 13m^4 - m^3 + 4m - 2\big) \tag{2.20}$$

This gives rise to Sublemmata which together prove Lemma 2.9.6.3.

Sublemma 2.9.6.1 *If the mandate $\mathcal{M}$ is only concerned with average beliefs, strictly positive but muted incentives are generated.*

Proof: We know from Lemma 2.9.6.2 that in FDF incentives are $\mathbb{E}_F[H] - \mathbb{E}_G[H]$ and from Lemma 2.9.6.1 that in ZDF incentives are zero. If the mandate is only concerned about average beliefs ($m = 2$) and places no extra weight on the weak bank, then (2.20) yields $\mathbb{E}\big[\mu_1(\boldsymbol{l}) \mid h_1, l_2\big] = \frac{333}{512}$. Since $\mathbb{E}_F[H] = \frac{2}{3}$ we have $\mathbb{E}_G[H] < \mathbb{E}\big[\mu_1(\boldsymbol{l}) \mid h_1, l_2\big] < \mathbb{E}_F[H]$ and thus effort incentives are indeed strictly positive and muted. *q.e.d.*

Sublemma 2.9.6.2 *Incentives are monotonically decreasing in the weight the mandate places on the weak bank.*

Proof: We aim to show that $\mathbb{E}\big[\mu_1(\boldsymbol{l}) \mid h_1, l_2\big] - \mathbb{E}_G[H]$ is decreasing in $\omega \in [0, 1[$. Since $\mathbb{E}_G[H]$ is independent of sender preferences and using notation $x(H_2) = mH_2$ to capture

the mandate ω, this is equivalent to showing that $\mathbb{E}\big[\mu_1(\boldsymbol{l}) \mid h_1, l_2\big]$ is decreasing in m for all $m \geq 2$. From (2.20) we obtain:

$$\frac{d\mathbb{E}\big[\mu_1(\boldsymbol{l}) \mid h_1, l_2\big]}{dm} = \frac{-15m^5 + 26m^4 + 3m^3 - 20m + 12}{48m^7} \tag{2.21}$$

For our entire parameter space ($m \geq 2$), this expression is negative. *q.e.d.*

Sublemma 2.9.6.3 *For all mandates $\mathcal{M}$, a Hurdle Rate Framework always creates strictly positive incentives. Even when the sender is purely concerned with the weak bank, incentives remain strictly positive and are never fully muted.*

Proof: If the mandate is purely concerned with the weak bank ($\omega = 1$), then $m \to \infty$. The limit of (2.20) thus becomes $\lim_{m\to\infty} \mathbb{E}\big(\mu_1(\boldsymbol{l}) \mid h_1, l_2\big) = \frac{9}{16}$. Thus, $\lim_{m\to\infty} \mathbb{E}\big(\mu_1(\boldsymbol{l}) \mid h_1, l_2\big) - E_G(H) > 0$ which means that incentives are strictly positive. Since by Sublemma 2.9.6.2 incentives are monotonically decreasing in ω, all $\omega \in [0, 1[$ must have even higher and thus strictly positive incentives. *q.e.d.*

Intuition for Sublemma 2.9.6.3: If the mandate is only concerned about the weak bank, then in equilibrium the CB almost always sends $\{\min(H_1, H_2), p, p\}$ except when $\min(H_1, H_2) = 0$. This is different from zero disclosure as it reveals $\min(H_1, H_2)$. Since the mandate which is only concerned about the weak bank reveals this information, it generates positive incentives. If bank 1 knew that $\min(H_1, H_2) = H_1$ will result, it would have full incentives while if it knew that $\min(H_1, H_2) = H_2$ would result it would face zero incentives. If the CB always sends $\{\min(H_1, H_2), p, p\}$ there is a chance of bank 1 being the weak bank and thus some incentives remain even in the limit.

Proof of Lemma 2.9.6.3: Combining Sublemma 2.9.6.1, 2.9.6.2, 2.9.6.3 proves Lemma 2.9.6.3. *q.e.d.*

Proof of Proposition 2.3: Combining Lemma 2.9.6.1, 2.9.6.2, 2.9.6.3 proves Proposition 2.3. *q.e.d.*

2.9.7 Proof of Proposition 2.4: Equilibrium in the Endogenous Effort Game

Proof of Part 1: Proof by backward induction: In ZDF, recipients never receive a message and thus cannot update their prior belief. The posterior belief is therefore identical to the prior belief which is based on conjectured effort. Thus, actual effort does not alter μ_i. When banks choose their effort level e_i to maximise $\mathbb{E}[\pi_i] = \mathbb{E}[\mu_i] - C(e_i)$ they anticipate that $\mathbb{E}[\mu_i]$ is not affected by e_i. Hence, there is no benefit of effort, while effort is costly. Thus, in ZDF, choosing low effort is the optimal strategy regardless of what

effort level receivers conjecture. The unique equilibrium therefore has both banks choosing low effort (l_1, l_2) and financial markets conjecturing low effort as only this case satisfies the equilibrium requirement that conjectures are correct. *q.e.d.*

Proof of Part 2: Proof by backward induction: In FDF, health realisations H_i are always fully revealed and recipients thus always update their belief to the true fundamental: $\mu_i = H_i$. The expected posterior is therefore identical to the expected fundamental $\mathbb{E}[\mu_i] = \mathbb{E}[H_i \mid e_i]$ and the expected fundamental is fully determined by actual effort and therefore is independent of conjectured effort. When banks choose their effort level e_i to maximise $\mathbb{E}[\pi_i] = \mathbb{E}[\mu_i] - C(e_i)$, they anticipate that $\mathbb{E}[\mu_i]$ is fully determined by e_i and thus face full effort incentives. They exert effort iff $\mathbb{E}_F[H] - \mathbb{E}_G[H] > C(h) - C(l)$ which is always true by our assumption that effort is socially beneficial. Hence, in FDF, high effort is the optimal action regardless of what effort level receivers conjecture. The unique equilibrium therefore has both banks choosing high effort (h_1, h_2) and financial markets conjecturing high effort as only this case satisfies the equilibrium requirement that conjectures are correct. *q.e.d.*

Proof of Part 3:

Lemma 2.9.7.1 *A low effort equilibrium exists if and only if* $C(h) - C(l) \geq \bar{t}(\mathcal{M})$

Proof: In a HRF, an equilibrium where both banks choose low effort exists only if a bank does not benefit from unilaterally deviating to high effort, i.e. $\mathbb{E}\big[\mu_1(\boldsymbol{l}) \mid h_1, l_2\big] - \mathbb{E}_G[H] \leq C(h) - C(l)$. We know from Proposition 2.3 that the strength of incentives generated in a HRF depends on the sender's mandate. Defining $\mathbb{E}\big[\mu_1(\boldsymbol{l}) \mid h_1, l_2\big] - \mathbb{E}_G[H] = \bar{t}(\mathcal{M})$, the condition above can be restated as $C(h) - C(l) \geq \bar{t}(\mathcal{M})$. All other equilibrium conditions are also satisfied, given how we constructed $\mathbb{E}\big[\mu_1(\boldsymbol{l}) \mid h_1, l_2\big]$. *q.e.d.*

Lemma 2.9.7.2 *A high effort equilibrium exists if and only if* $C(h) - C(l) \leq \underline{t}(\mathcal{M})$.

Proof: In a HRF, an equilibrium where both banks choose high effort exists only if a bank does not benefit from unilaterally deviating to low effort, i.e. $\mathbb{E}_F[H] - \mathbb{E}\big[\mu_1(\boldsymbol{h}) \mid l_1, h_2\big] \geq C(h) - C(l)$. The strength of incentives generated in a HRF depends on the sender's mandate. Defining $\underline{t}(\mathcal{M}) = \mathbb{E}_F[H] - \mathbb{E}\big[\mu_1(\boldsymbol{h}) \mid l_1, h_2\big]$, the condition can be restated as $C(h) - C(l) \leq \underline{t}(\mathcal{M})$. All othe equilibrium conditions are also satisfied given how we constructed $\mathbb{E}\big[\mu_1(\boldsymbol{h}) \mid l_1, h_2\big]$. *q.e.d.*

Lemma 2.9.7.3 $\underline{t}(\mathcal{M}) < \bar{t}(\mathcal{M})$

When both banks play high effort which results in $F(H) = H^2$, then the hurdle rate communication game for $V(\mu_1, \mu_2) = \frac{\mu_1 + \mu_2}{2}$ results in $x(H_2)$-indifference frontier $x(H_2) = \sqrt{3}\, H_2$.

In the endogenous effort game, there exists an equilibrium where both banks play high effort if and only if $C(h) - C(l) \leq \mathbb{E}_F[H] - \mathbb{E}\big[\mu_1(\boldsymbol{h}) \mid l_1, h_2\big] \approx 0.1285$. Correspondingly, both banks playing low effort results in $x(H_2) = 2H_2$ and is an equilibrium if and only if: $C(h) - C(l) \geq \mathbb{E}\big[\mu_1(\boldsymbol{l}) \mid h_1, l_2\big] - \mathbb{E}_G[H] = \frac{77}{512} \approx 0.1504$. Clearly, $0.1285 < 0.1504$ which shows that Lemma 2.9.7.3 holds. *q.e.d.*

Proof of Part 3: Since the only candidate symmetric pure strategy equilibria have (l_1, l_2) or (h_1, h_2), combining Lemma 2.9.7.1, 2.9.7.2, 2.9.7.3 proves Part 3. Lemma 2.9.7.1 and 2.9.7.2 show regions in which (l_1, l_2) or (h_1, h_2) equilibria exist and Lemma 2.9.7.3 shows that the regions of the cost space in which these equilibria exist are disjoint, leading to a region where no pure stretegy symmetric equilibrium exists between $\underline{t}(\mathcal{M})$ and $\bar{t}(\mathcal{M})$. *q.e.d.*

2.9.8 Proof of Proposition 2.5: The Optimal Mandate

Lemma 2.9.8.1 *If there exists a $\{HRF, \mathcal{M}\}$ which incentivises high effort, then $\{HRF, \tilde{\mathcal{M}}\}$, where $\tilde{\mathcal{M}}$ is the softest possible mandate which just incentivises high effort, achieves higher welfare than all tougher HRF mandates and than FDF.*

Formally, let $\tilde{\omega}$ be the softest possible mandate which in a HRF results in high effort, i.e. $\tilde{\omega}$ is defined by $C(h) - C(l) = \mathbb{E}_F[H] - \mathbb{E}\big[\mu_1(\boldsymbol{h}) \mid l_1, h_2, \tilde{\omega}\big]$. Then all tougher mandates $\omega < \tilde{\omega}$ also result in high effort. Thus, Lemma 2.9.8.1 states that

$$\mathbb{E}_{HRF,\tilde{\omega}}[W(\mu_1, \mu_2)] > \mathbb{E}_{HRF,\omega}[W(\mu_1, \mu_2)] \; \forall \; \omega < \tilde{\omega} \tag{2.22}$$

and

$$\mathbb{E}_{HRF,\tilde{\omega}}[W(\mu_1, \mu_2)] > \mathbb{E}_{FDF}[W(\mu_1, \mu_2)] \tag{2.23}$$

Proof: FDF, $\{\text{HRF}, \tilde{\omega}\}$, and all $\{\text{HRF}, \omega\}$ with $\omega < \tilde{\omega}$ result in high effort. Therefore, for $W(\mu_1, \mu_2) = \lambda \; \min(\mu_1, \mu_2) + (1 - \lambda) \; \frac{\mu_1 + \mu_2}{2}$ with any $\lambda > 0$ the welfare ranking of these designs reduces to a ranking in terms of reassurance, $\mathbb{E}[\min(\mu_1, \mu_2)]$. From Proposition 2.1 we know that FDF provides the lowest reassurance and that in a HRF, reassurance is monotonically increasing in the softness of the mandate (ω). Thus, equations (2.22) and (2.23) hold. *q.e.d.*

Lemma 2.9.8.2 *ZDF achieves higher welfare than all hurdle rate frameworks which result in low effort.*

Formally, let $\bar{\omega}$ be the toughest possible mandate which in a HRF results in low effort, i.e. $\bar{\omega}$ is defined by $C(h) - C(l) = \mathbb{E}\big[\mu_1(\boldsymbol{l}) \mid h_1, l_2, \bar{\omega}\big] - \mathbb{E}_G[H]$. Then, all softer mandates $\omega > \bar{\omega}$ also result in low effort. Thus, Lemma 2.9.8.2 states that

$$\mathbb{E}_{ZDF}[W(\mu_1, \mu_2)] > \mathbb{E}_{HRF,\omega}[W(\mu_1, \mu_2)] \; \forall \; \omega > \bar{\omega} \tag{2.24}$$

Proof: ZDF and all $\{HRF, \omega\}$ with $\omega > \bar{\omega}$ result in low effort. Therefore, for $W(\mu_1, \mu_2) = \lambda \min(\mu_1, \mu_2) + (1-\lambda) \frac{\mu_1+\mu_2}{2}$ with any $\lambda > 0$ the welfare ranking of these designs reduces to a ranking in terms of reassurance, $\mathbb{E}[\min(\mu_1, \mu_2)]$. From Proposition 2.1 we know that ZDF is the design which provides the highest reassurance among all designs. Thus, equation (2.24) holds. *q.e.d.*

Proof of Proposition 2.5: Combining Lemma 2.9.8.1 and Lemma 2.9.8.2 proves Proposition 2.5. *q.e.d.*

2.9.9 Proof of Proposition 2.6: Central Bank Design

Proof of Part 2: High Effort Costs: From Proposition 2.5 we have that in this case of high effort costs $W(\{\text{ZDF}\}) > W(\{\text{HRF}, \mathcal{M}\}) \; \forall \; \mathcal{M}$ and thus the designer's choice reduces to a choice among ZDF and FDF. We also established that ZDF results in a low effort equilibrium. Thus, if the designer chooses ZDF:

$$\mathbb{E}_{ZDF}[W(\mu_1, \mu_2)] = \mathbb{E}_G[H] \; \forall \; \lambda \tag{2.25}$$

Our focus throughout is on the case where FDF results in high effort. Thus, if the designer chooses FDF:

$$\mathbb{E}_{FDF}[W(\mu_1, \mu_2)] = \lambda \, \mathbb{E}_F[\min(H_1, H_2)] - (1-\lambda) \, \mathbb{E}_F[H] \tag{2.26}$$

The designer chooses FDF if and only if:[47]

$$\mathbb{E}_{FDF}[W(\mu_1, \mu_2)] > \mathbb{E}_{ZDF}[W(\mu_1, \mu_2)] \tag{2.27}$$

To solve for the value of λ at which the designer is indifferent, denoting it as λ^*, we solve the corresponding equality:

$$\lambda^* \, \mathbb{E}_F[\min(H_1, H_2)] - (1-\lambda^*) \, \mathbb{E}_F[H] = \mathbb{E}_G[H] \tag{2.28}$$

Which yields:

$$\lambda^\star(a) = \frac{\mathbb{E}_F[H] - \mathbb{E}_G[H]}{\mathbb{E}_F[H] - \mathbb{E}_F[\min(H_1, H_2)]} \tag{2.29}$$

Thus, for a given benefit of effort a, at $\lambda < \lambda^\star(a)$ the designer chooses FDF and otherwise ZDF.

For the full characterisation of the optimal central bank design as depicted in Figure 2.4a note that since $\mathbb{E}_F[H] - \mathbb{E}_G[H] > 0$ and $\mathbb{E}_F[H] - \mathbb{E}_F[\min(H_1, H_2)] > 0$ always hold, it must be that $\lambda^\star(a) > 0$. Moreover, if the benefit of effort is so large that $\mathbb{E}_F[\min(H_1, H_2)] > \mathbb{E}_G[H]$, then $W(\{FDF\}) > W(\{ZDF\}) \; \forall \; \lambda$ which gives rise to the top area in Figure 2.4a where FDF is chosen for all values of λ. For the functional forms

[47] Formally, FDF results in the symmetric pure strategy equilibrium which maximises welfare if and only if (2.27) holds.

$G(H) = H$, $F(H) = H^a$, $\mathbb{E}_F[\min(H_1, H_2)] > \mathbb{E}_G[H]$ holds for $a > 1.78$. Thus, we consider $\lambda^\star(a)$ only for $a \in [1, 1.78]$ as $a > 1$ is needed to ensure an MLRP ordering of distributions. Over this range, $\lambda^\star(a)$ is positive and monotonically increasing as $\lambda^\star(a) = -\frac{1}{2} - \frac{1}{2a} + a$. Thus, the optimal institutional design in the high effort case is characterised by Figure 2.4a. *q.e.d.*

Proof of Part 1: Low Effort Costs: From Proposition 2.5 we know that in this case of low effort costs $W(\{\text{HRF}, \tilde{\omega}\}) > W(\{\text{HRF}, \omega\}) \; \forall \; \omega < \tilde{\omega}$ and $W(\{\text{HRF}, \tilde{\omega}\}) > W(\{\text{FDF}\})$. We also know that $W(\{\text{ZDF}\}) > W(\{\text{HRF}, \omega\}) \; \forall \; \omega > \tilde{\omega}$. Thus, the designer's choice reduces to a choice among ZDF and $\{HRF, \tilde{\omega}\}$. As before, $\mathbb{E}_{ZDF}[W(\mu_1, \mu_2)] = \mathbb{E}_G[H] \; \forall \; \lambda$. By construction of $\tilde{\omega}$ as the softest possible mandate which induces high effort:

$$\mathbb{E}_{HRF,\tilde{\omega}}[W(\mu_1, \mu_2)] = \lambda \, \mathbb{E}_{HRF,\tilde{\omega}}[\min(\mu_1, \mu_2)] + (1 - \lambda) \, \mathbb{E}_F[H] \tag{2.30}$$

The designer chooses $\{HRF, \tilde{\omega}\}$ if and only if:[48]

$$\mathbb{E}_{HRF,\tilde{\omega}}[W(\mu_1, \mu_2)] > \mathbb{E}_{ZDF}[W(\mu_1, \mu_2)] \tag{2.31}$$

To solve for the value of λ at which the designer is indifferent, denoting it λ', we solve the corresponding equality:

$$\lambda' \, \mathbb{E}_{HRF,\tilde{\omega}}[\min(\mu_1, \mu_2)] + (1 - \lambda') \, \mathbb{E}_F[H] = \mathbb{E}_G[H] \tag{2.32}$$

Which yields:

$$\lambda'(a, C) = \frac{\mathbb{E}_F[H] - \mathbb{E}_G[H]}{\mathbb{E}_F[H] - \mathbb{E}_{HRF,\tilde{\omega}}[\min(\mu_1, \mu_2)]} \tag{2.33}$$

Thus, for a given benefit of effort a, at $\lambda < \lambda'(a, C)$ the designer chooses $\{\text{HRF}, \tilde{\omega}\}$ and otherwise ZDF.

For the full characterisation of the optimal central bank design as depicted in Figure 2.4b note that $\lambda'(a, C) > 0$ always holds. Moreover, comparing $\lambda'(a, C)$ and $\lambda^*(a)$ we always have that $\lambda'(a, C) > \lambda^*(a)$, since this equals:

$$\frac{\mathbb{E}_F[H] - \mathbb{E}_G[H]}{\mathbb{E}_F[H] - \mathbb{E}_{HRF,\tilde{\omega}}[\min(\mu_1, \mu_2)]} > \lambda^\star(a) = \frac{\mathbb{E}_F[H] - \mathbb{E}_G[H]}{\mathbb{E}_F[H] - \mathbb{E}_F[\min(H_1, H_2)]} \tag{2.34}$$

which simplifies to:

$$\mathbb{E}_{HRF,\tilde{\omega}}[\min(\mu_1, \mu_2)] > \mathbb{E}_F[\min(H_1, H_2)] \tag{2.35}$$

which holds by the reassurance ranking. Thus, the optimal institutional design in the low effort case is characterised by Figure 2.4b. *q.e.d.*

The exact value of $\lambda'(a, C)$ depends indirectly on the cost of effort. This arises because $\lambda'(a, C)$ depends directly on mandate $\tilde{\omega}$, that is the softest possible mandate which just

[48] Formally, $\{\text{HRF}, \tilde{\omega}\}$ results in the symmetric pure straetgy equilibrium which achieves the highest welfare if and only if (2.31) holds.

induces high effort, which in turn depends on the cost of effort. If costs of effort are higher, a tougher mandate is needed to incentivise effort. This corresponds to a smaller value of $\tilde{\omega}$. Then $\{\text{HRF}, \tilde{\omega}\}$ releases more information. $\lambda'(a, C)$ thus becomes steeper, i.e. rotates towards the $\lambda^*(a)$ line. When costs of effort become so high that they switch from the low cost of effort to the high cost of effort regime, then $\lambda'(a, C)$ becomes $\lambda^*(a)$ and the design chosen to induce effort switches from being $\{\text{HRF}, \tilde{\omega}\}$ to FDF. *q.e.d.*

Chapter 3

Imperfect Competition in Markets with Selection

This paper proposes a model of imperfect competition in markets with selection. While we know that in markets with advantageous selection, such as consumer credit or insurance markets, perfect competition can lead to inefficient overprovision and that monopoly results in underprovision, we do not know whether intermediate degrees of competition achieve the efficient allocation. We find that no degree of competition can restore the first best efficient allocation. Overprovision induced by advantageous selection and underprovision caused by market power do not cancel out perfectly. Instead, they co-exist for different segments of the consumer taste spectrum. However, we show that the efficient allocation can be reached through a combination of a corrective tax and tough competition policy. This suggests that if we view the build up of credit prior to the recent financial crisis as inefficient overprovision in a market with advantageous selection, then we should not conclude that competition in credit markets was "excessive", but rather conclude that we lacked the appropriate corrective taxation. Overall, our paper cautions against viewing imperfect competition as a solution to the inefficiency introduced by advantageous selection.

Keywords: *Imperfect Competition, Selection Markets, Credit Markets*

JEL Classification: D82, D43, C7, L1

3.1 Introduction

Traditionally, market failures in many markets have been attributed either to imperfect competition[1] or to imperfect information, e.g. in the market for used cars (Akerlof, 1970), various insurance markets (Einav and Finkelstein, 2011), and consumer credit markets (Karlan and Zinman, 2009). More recently, there has been growing interest in the interaction of these two sources of market failure. Many markets suffer from both simultaneously, such as health insurance (Dafny, 2010) or retail banking (European Commision, 2007; Competition and Markets Authority, 2016). A string of recent papers finds that when a market is characterised by selection, imperfect competition can remedy inefficiencies rather than exacerbate them (Lester et al., 2019; Crawford et al., 2018; Mahoney and Weyl, 2017).

Markets with advantageous selection are central in this literature and are the focus of this paper. That is markets where at high prices only consumers with low costs purchase while lower prices draw in higher cost consumers (Hemenway, 1990; De Meza and Webb, 2001). Advantageous selection has been detected in the Medigap market[2] (Fang et al., 2008) and the market for US subprime auto loans (Mahoney and Weyl, 2017). Under perfect competition, these markets can exhibit inefficient overprovision.[3] Firms undercut each other's prices even when the entering consumers are loss making as firms try to steal the rival's existing profitable consumers. This has led to competition being called "excessive", competition policy referred to as potentially "harmful" and in the context of credit markets has been blamed for contributing to the overprovision of credit prior to the recent financial crisis.[4]

While we know that in markets with advantageous selection perfect competition can result in inefficient overprovision[5] and that in markets without selection imperfect competition can lead to inefficient underprovision, we do not know whether in markets with advantageous selection *and* imperfect competition the inefficiencies cancel out. Thus, we do not know whether competition policy should tolerate intermediate degrees of competition

[1] For evidence of a recent economy wide decline in competition see Gutierrez and Philippon (2017) and De Loecker et al. (2020).

[2] Medigap is a type of private health insurance in the US which Medicare recipients can purchase. Medigap covers health related financial risks which are not covered by Medicare such as deductibles or co-insurance payments.

[3] Just as markets with adverse selection can break down completely, but do not have to, overprovision is a possibility but not a certainty in markets with advantageous selection. A detailed explanation is provided in section 3.2.

[4] "Competition policy [can be] harmful in the presence of advantageous selection." Weyl's website summarizing Mahoney and Weyl (2014), accessed May 2016. In a similar spirit, p.7 in Mahoney and Weyl (2014) states: "This suggests that policies such as the Gramm-Leach-Bliley Act, which intended to bring 'greater ... competition in the financial services industry', could have contributed to an inefficient credit boom." The quote within the quote is from "President Clinton's signing statement." The phrase "excessive competition" is common in the literature on whether a trade-off between banking competition and financial stability exists. A recent overview is Vives (2016).

[5] Seminal papers on the potential inefficiency of the competitive equilibrium in the presence of imperfect information are Stiglitz and Weiss (1981) and De Meza and Webb (1987). Stiglitz and Weiss (1981) show the possibility result that, in the presence of imperfect information, the competitive equilibrium can be characterised by inefficient *underprovision*. De Meza and Webb (1987) show the possibility result that, in the presence of imperfect information, the competitive equilibrium can be characterised by inefficient *overprovision*.

to correct the overprovision or whether alternative policies can restore efficiency.

This paper develops a model of imperfect competition in markets with selection, building on models of spatial competition (Hotelling, 1929; Thisse and Vives, 1988), to investigate whether the inefficiencies created by selection and by imperfect competition cancel out.[6] We find that the inefficiencies created by these two types of market failure never cancel out for all consumers simultaneously. Instead, overprovision for some consumers co-exists with underprovision for others. Thus, tolerating intermediate degrees of competition does not restore the first best efficient allocation.

We show that while tolerating intermediate degrees of competition cannot restore efficiency, efficiency can be restored through a combination of a corrective tax with tough competition policy. Moreover, this policy combination has the additional benefit of raising government revenue. Overall, our results caution against viewing imperfect competition as a policy remedy for selection issues as the results underlying this argument seem to depend on model specifics and apply only if other policy interventions, such as our proposed corrective tax, have been ruled out as infeasible.

While our spatial model nests the familiar results of perfect competition resulting in overprovision and monopoly in underprovision, our model provides a new perspective on how the strategic interaction of firms is affected by the presence of selection. While in markets without selection firms contemplate how many consumers a price cut would attract - either newly entering or switching from the rival - in markets with selection firms additionally consider which *type* of consumer they attract. Our model predicts that, in a market with advantageous selection, those consumers who switch to a firm in response to it cutting its price are an advantageous selection of the rival's demand. This arises because a price cut only attracts those consumers from the rival who had the largest "distance" to travel. These consumers needed to pay large transport costs, and therefore only chose to buy from the rival over the outside option if their willingness-to-pay (WTP) is high. Thus, the switching consumers have on average higher WTP than the average of the rival's demand. Given advantageous selection, this makes switching consumers lower cost than average. Hence, the switching consumers are an advantageous selection of demand. The extent to which the margin is selected depends on the degree of competition. For softer competition, that is higher transport costs, selection on the switching margin is stronger, i.e. the switching margin is less representative of demand.

After outlining our model, we characterise the efficient allocation, which is the benchmark against which we will judge equilibria,[7] and show that in markets with advantageous

[6] While this paper focuses on markets with advantageous selection and uses the case of no selection as a benchmark, our model can also be used to study imperfect competition in markets with adverse selection.

[7] Comparisons of the efficient allocation and the equilibrium under imperfect competition are confounded by transport costs, which capture the degree of competition, having multiple effects. Larger transport costs have a strategic effect, which is what we aim to capture, but also have a direct cost effect, i.e. transport costs reduce total surplus achievable in a market. To avoid biasing our results in favour of competition, we isolate the strategic effect by comparing equilibrium outcomes to the outcome a welfare maximising social planner can achieve when he faces the same transport costs.

selection, the efficient allocation can be implemented with price discrimination, but not with gate prices.[8] This arises because at locations with larger distances to travel, a given gate price attracts only higher WTP consumers. In the presence of advantageous selection, these high WTP agents have lower fundamental costs. Thus, if a gate price achieves the efficient allocation for agents with small distances to travel, it would be efficient to charge lower prices at locations with larger distances to travel. While this is possible under price discrimination, it is not possible under gate prices. Therefore, we also characterise the socially optimal gate price, that is the constrained efficient allocation where the constraint is to not be able to condition prices on location. The socially optimal gate price approximates the efficient allocation. It results in overprovision for consumers with strong brand preferences and underprovision for those who are close to indifferent among the brands.

We find that no degree of imperfect competition can restore the first best efficient allocation in the presence of advantageous selection. While softer competition results in fewer trades occurring in equilibrium, this never exactly offsets the inefficient overprovision which results under perfect competition. Instead, imperfect competition results in inefficient underprovision and overprovision co-existing for different consumer types. This result applies both when firms use gate prices and when they can price discriminate. However, the prediction of which consumers experience underprovision and which experience overprovision does depend on the firms' strategy. While if firms use gate prices there is overprovision for consumers with strong brand preferences and underprovision for consumers who are indifferent among brands, the reverse is true when firms can price discriminate.

When firms compete using gate prices, there exists an intermediate degree of competition at which the oligopoly equilibrium price coincides with the planner's optimal price. However, the planner's optimal price does not implement the efficient allocation. Due to the presence of advantageous selection, the planner's gate price is itself only constrained efficient, i.e. it approximates the first best efficient allocation under the constrained of using a gate price. Thus, at this price, consumers with strong brand preferences experience overprovision, while there is underprovision for consumers who are indifferent among the brands. That means the two types of inefficiency co-exist.

When firms use price discrimination, they compete very fiercely for consumers who are indifferent among brands and use their market power to set high prices for consumers with strong brand preferences. Since with price discrimination firms can target prices at consumers depending on brand preferences, these two motives are not conflicting, which they were under gate prices. As a result, overprovision and underprovision coexist. There is overprovision for consumers who have small or no brand preferences. While intermediate degrees of competition reduce the range of brand preferences in which overprovision arises,

[8] We use the term "gate price" for a firm setting one price to all consumers regardless of their distance travelled. The transport cost is incurred by the consumers in addition to this price. Alternative names in the literature are "uniform pricing" or "mill price".

for any degree of imperfect competition overprovision continues to exist for those consumers who are exactly indifferent between brands. There is underprovision for consumers with strong brand preferences. For softer degrees of competition, underprovision arises for a larger range of brand preferences. The extent of underprovision can be large. Already for intermediate degrees of competition, consumers with strong brand preferences experience the same degree of underprovision that arises in monopoly.

Since it is impossible to reach the first best efficient allocation by tolerating intermediate degrees of competition, we study alternative policy tools and show that the first best can be achieved through a combination of tough competition policy, which results in fierce competition, with a corrective tax. Thus, competition policy should not be lenient towards firms because they operate in a market with advantageous selection. Instead, competition policy should continue to focus on achieving perfect competition. The distortions arising from selection can be better addressed through a tax than through altered competition policy. The corrective tax has the additional benefit of generating government revenue, holds regardless of whether firms use gate prices or price discrimination, has lower informational requirements, and has a distributional effect as firms absorb some of the tax.

Additionally, we compare the corrective tax, which applies equally to all types, to type-specific policy interventions such as a type dependent subsidy or tax. Type dependent subsidies are, for example, common in the context of health insurance where they are referred to as "risk adjustment". We find that risk adjustment has higher informational requirements than the corrective tax and is vulnerable to gaming of risk scores by insurance providers.

Our results, which are in favour of competition, hold despite the fact that we make several assumptions which tend to understate the social benefit of competition. We assume that the social planner maximises total surplus. Under any alternative welfare measure which places more weight on consumer surplus relative to producer surplus, competition is even more beneficial. We focus purely on the short term strategic effect of different degrees of competition. Competition can have the additional benefit of spurring innovation and thus lowering costs. We assume that firms incur the same cost for serving a given consumer. When there are cost differences, competition can have the additional benefit of allowing the the more efficient firm to win market share from the less efficient firm.

Overall, our paper cautions against viewing imperfect competition as a solution to the inefficiencies introduced by advantageous selection. Combining tough competition policy with a corrective tax results in higher welfare than intermediate degrees of competition and has lower informational requirements. This suggests that if we view the build up of credit prior to the recent financial crisis as inefficient overprovision in a market with advantageous selection, then we should not conclude that competition in credit markets was "excessive", but rather conclude that we lacked the appropriate corrective taxation.

The remainder of this paper is structured as follows. Section 3.2 reviews the related

literature. Section 3.3 outlines our model. Section 3.4 characterises the efficient allocation and the constrained efficient allocation given the constraint of gate prices. Section 3.5 shows that our model nests the familiar results under perfect competition and monopoly. Section 3.6 characterises the equilibrium under imperfect competition. Section 3.7 outlines policy remedies other than tolerating intermediate degrees of competition. Section 3.8 discusses extension of our model and section 3.9 concludes.

3.2 Literature

This section reviews the related literature and the underlying theory. First, we provide micro-foundations for how adverse and advantageous selection can arise, describe the perfectly competitive equilibrium and show what efficiency properties it can have. Second, we sketch the existing literature on imperfect competition in selection markets. Our modelling approach departs from these papers by building on spatial models of competition and on price discrimination. Thus, we discuss the similarities and differences of the literature to this paper.

The defining feature of selection markets is that the price a firm sets not only affects how many consumers a firm attracts, but also affects which type of consumer it attracts. While the seminal literature on selection markets focused on *adverse selection* (Akerlof, 1970), where high prices attract only expensive consumers, the more recent literature has also discussed *advantageous selection* (Hemenway, 1990; De Meza and Webb, 2001), where high prices attract only low cost consumers. An excellent survey is Einav and Finkelstein (2011). The pictorial exposition in this literature section is borrowed from them, while the pictorial exposition in the remainder of the paper is our own.

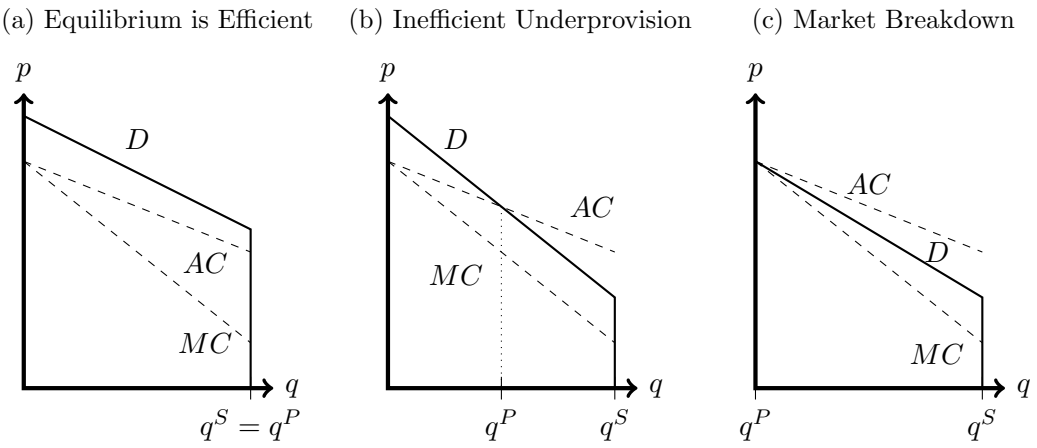
Adverse Selection

A market exhibits adverse selection if marginal costs are decreasing in output. This can be microfounded by consumers having private information on a fixed characteristic that affects the firm's cost. For example, in the market for health insurance, a consumer's health is private information and heterogeneous across agents. Consumers with low health have large expected claims and are therefore expensive for the firm to insure. In insurance markets, the WTP of rational consumers corresponds to the sum of their expected claims and a risk premium. Hence, when consumers have identical risk attitudes, low health consumers have the largest willingness-to-pay for insurance. Thus, at high prices only the low health consumers buy while lower prices draw healthier consumers into the market. Marginal costs are decreasing in output.

In markets with adverse selection, the efficiency of the perfectly competitive equilibrium differs depending on the scale of consumers' risk premia relative to the extent of cost heterogeneity. When consumers' WTP far exceed costs, e.g. because they are very risk averse, the equilibrium achieves the efficient allocation (Figure 3.1a). When consumers'

WTP are lower, the equilibrium exhibits inefficient underprovision (Figure 3.1b) and the market can even break down completely (Figure 3.1c).

Figure 3.1: Adverse Selection: Perfect Competition



D denotes demand, MC marginal cost, AC average cost
q^P is the quantity in the perfectly competitive equilibrium
q^S is the efficient quantity[9]

Advantageous Selection

A market exhibits advantageous selection if marginal costs are increasing in output.[10] This can be micro-founded by consumers having private information on a fixed characteristic which affects their decision to take an unobserved action. For example, in the market for car insurance, consumers drive on the same roads, thus have the same underlying accident probability, but differ in their risk aversion which in turn affects driving style. More risk averse agents have a higher WTP for insurance and drive more carefully, thus have lower expected insurance claims. Thus, it is possible that at high prices only the very risk averse consumers buy - they have low expected costs for the insurer - while lower prices also attract consumers with low risk aversion. Marginal costs are increasing in output.

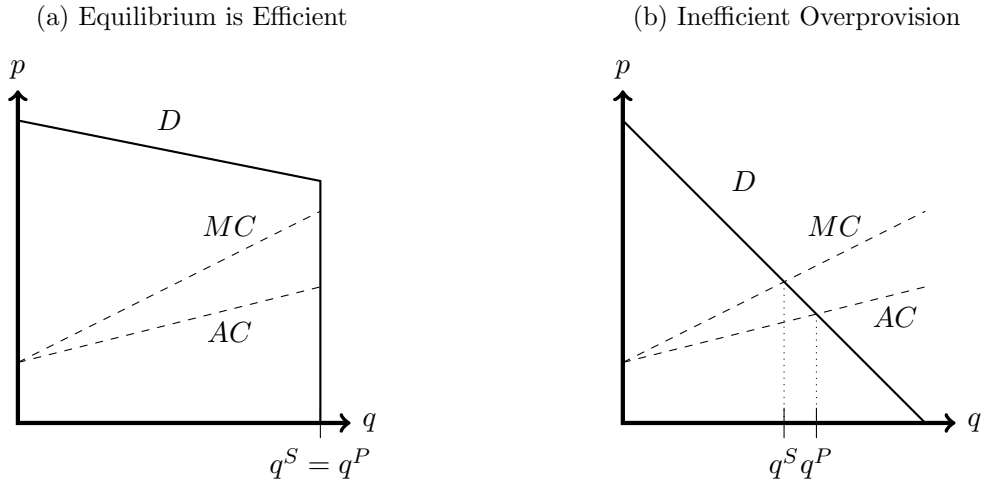
In markets with advantageous selection, the perfectly competitive equilibrium exhibits inefficient overprovision if for some consumers expected costs exceed their WTP and otherwise results in the efficient allocation. When firms face no administrative costs (then expected costs to the firm are equal to consumers expected claims) and consumers are risk averse (then their WTP exceeds their expected claims by a strictly positive risk premium), all consumers are willing to pay more for insurance than they cost the firm. In this case, the perfectly competitive equilibrium results in the efficient allocation (Figure 3.2a). When firms have administrative costs or when for behavioural reasons consumers underestimate their future claims, it is possible that some consumers' WTP is below expected cost. In

[9]The efficient quantity is denoted by 'S' to correspond to the socially optimal quantity - a concept we introduce in the context of our model later on.

[10]To the best of our knowledge, this concept was first described by Hemenway (1990), albeit under a different name, and termed "advantageous selection" by De Meza and Webb (2001).

this case, the perfectly competitive equilibrium exhibits inefficient overprovision (Figure 3.2b).

Figure 3.2: Advantageous Selection: Perfect Competition



Inefficient overprovision arises because even when the marginal consumer is loss making, it is still profitable for firms to undercut the rival's price as this steals the rival's existing consumers, who are profitable.

Two classes of empirical papers emphasize the relevance of advantageous selection: The first class of papers argue that preference heterogeneity among consumers exists and is sizable, thus opening up the possibility of advantageous selection (Cohen and Einav, 2007). The second class of papers directly shows that advantageous selection exists in a specific market by arguing that the marginal cost curve is upwards sloping. Examples of this type are: Fang et al. (2008) who find advantageous selection in the Medigap market. Using data from Einav et al. (2012) on US sub-prime auto loans, Mahoney and Weyl (2014) find evidence of strong advantageous selection. Advantageous selection thus appears to be a real world phenomenon and not just a theoretical artefact.

Imperfect Competition

A string of recent papers finds that, when a market is characterised by selection, imperfect competition can remedy inefficiencies rather than exacerbate them.

Mahoney and Weyl (2017) (henceforth, MW) find that, in markets with advantageous selection, intermediate degrees of competition are optimal as they they reduce overprovision. MW take a reduced form approach and model equilibrium prices p^* as a weighted average of the price under perfect competition p^{PC} and the price a monopolist would set p^M, i.e. $p^* = \theta p^M + (1-\theta)p^{PC}$, where $\theta \in [0,1]$ indexes the degree of competition and higher values of θ correspond to less intense competition.[11] They argue that this reduced form approach nests Cournot competition and differentiated Bertrand competition and thus

[11] This corresponds to equation (1) on p.640 in Mahoney and Weyl (2017).

yields "results that are robust to the details of the industrial organization."[12] However, this approach to selection markets[13] rests on the crucial assumption that firms always attract a random sample of consumers in industry demand (Cournot) or similarly that the consumers who switch in response to a firm cutting its price are representative of, i.e. on average equal to, consumers in the rival's demand (differentiated Bertrand). We term this the "representative switching assumption." We construct a spatial model of selection markets from first principles. In contrast to MW, our model implies that switching consumers are an advantageous selection of demand and thus not representative of demand. We discuss this in detail in section 3.3.3.

While in Mahoney and Weyl (2017) the efficiency properties of an equilibrium can be summarised by the equilibrium quantity traded, that is not the case in our model. In our spatial context, the efficiency of the equilibrium additionally depends on which consumers are allocated this quantity. This subtle difference leads to diverging results. While in MW there exists an intermediate degree of competition which achieves the first best efficient outcome, this is not the case in our model. In the version of our model where firms use gate prices, which is the version closest to MW as firms only set one price, intermediate degrees of competition can only achieve the second best efficient outcome, i.e. constrained by using gate prices. At this second best efficient outcome, overprovision for some consumers co-exists with underprovision for other consumers.

This diverging result leads us to explore different extensions. MW calibrate their model to data from health insurance markets and consumer credit markets to highlight that their theoretical possibility results apply in specific examples. Since our results indicate that tolerating intermediate degrees of competition cannot restore the first best efficient allocation when firms use gate prices, we show that this result continues to apply when firms can price discriminate. Moreover, we show that alternative policy remedies such as a corrective tax can restore the first best. Neither price discrimination nor alternative policy remedies are discussed by MW.

For markets with adverse selection, Mahoney and Weyl (2017) find that perfect competition always maximises welfare. In contrast to this, Lester et al. (2019) argue that intermediate degrees of competition or even monopoly maximises total surplus in markets with adverse selection. Their model of imperfect competition is based on the idea of "captive" consumers and combines the model by Burdett and Judd (1983) with a Rothschild and Stiglitz (1976) model of selection markets. There are two firms and a continuum of consumers. Each consumer has a probability b of receiving both offers but with probability $1 - b$ he is "captured" in the sense that he only receives one offer. Larger levels of b correspond to more intense competition. Importantly, Lester et al. (2019) allow firms to not only offer one price (as in MW) but offer a menu of contracts and thus screen consumers. Then, the degree of competition affects whether pooling contracts

[12] Quote in Mahoney and Weyl (2017) on p.638.

[13] This approach is a variant of the conduct parameter approach developed by Bresnahan (1989) and Weyl and Fabinger (2013).

can be sustained in equilibrium or not. Their result is that for severe adverse selection intermediate degrees of competition achieve the highest total surplus while for moderate adverse selection monopoly achieves the highest total surplus.

Further papers also argue that less than perfect competition is socially optimal. Crawford et al. (2018) build a structural econometric model of loan demand by small and medium companies in Italy. From their model, they construct counterfactuals which they interpret as showing that in this market with adverse selection an increase in competition can reduce welfare. Veiga and Weyl (2016) propose a model of selection markets with endogenous product quality and thus sorting incentives. They find that in insurance markets intermediate degrees of competition are socially optimal as market power acts against sorting-incentives, i.e. the problem of "cream skimming" is reduced as competition is eased. They find that this effect initially outweighs other effects of reduced competition.

Our model differs from these papers in its spatial approach. This allows us to include taste heterogeneity among consumers and thus to derive predictions on the type of switching consumers, rather than assume them. It also allows us to analyse the effect of different pricing strategies, i.e. gate prices or price discrimination. This gives rise to the important distinction between achieving the first best efficient allocation or the constraint efficient allocation. Since we find that tolerating intermediate degrees of competition can only achieve the latter, but not the former, we also study alternative policy remedies.

Price Discrimination

Our spatial approach and our focus on firm pricing strategies is related to the literature on price discrimination (for a survey, see Armstrong (2006)) and especially to the literature on price discrimination under imperfect competition.

While in monopoly and under perfect competition firms are at least weakly better off when they can price discriminate relative to when they can charge only a uniform price, the effects of price discrimination under imperfect competition are more nuanced. In imperfect competition, the constraint of being able to use only uniform prices can act as a valuable commitment device for firms. Thus, there exist cases of oligopoly where equilibrium prices at all locations are lower under price discrimination than under uniform pricing (Thisse and Vives, 1988; Corts, 1998). The empirical relevance of this possibility result is shown by Grennan (2013) in the context of the market for medical devices.

Thisse and Vives (1988) develop a spatial model of duopoly and use it to study firms' strategic choice of pricing policy. They find that firms face a Prisoner's Dilemma when choosing a pricing policy. While firms would be better off when both firms use uniform prices - which act to soften competition - the unique equilibrium is one where both firms use price discrimination.

Our model is closely related to Thisse and Vives (1988), but differs in important ways as we aim to study different questions. We focus on selection markets and thus make the cost of selling to an agent a function of his willingness-to-pay, while they do not study

selection, assume constant marginal costs and normalise these to zero. They study a market with full coverage, i.e. where all consumers always buy, while we are interested in inefficient over- and underprovision and therefore study markets where the efficient allocation has less than full provision. This is needed to create room for the possibility of inefficient over- and underprovision. They study the strategic choice of pricing policy while we focus on varying pricing strategies exogenously.

3.3 The Model

We propose a model of imperfect competition in markets with selection and derive insights on how selection affects the strategic interaction of firms. Our approach extends models of spatial competition (Hotelling, 1929; Thisse and Vives, 1988) to include selection by making the cost of selling to an agent a function of his willingness-to-pay rather than deriving costs from the firm's production function. Selection introduces new aspects to the strategic interaction of firms. While in markets without selection firms contemplate how many consumers a price cut would attract - either newly entering or switching from the rival - in markets with selection firms additionally consider which *type* of consumer they attract. A feature of our model is that, in a market with advantageous selection, those consumers who switch to a firm in response to it cutting its price are an advantageous selection of the rival's demand.

3.3.1 Specification of the Model

There are two firms which produce imperfect substitutes. This gives firms a degree of market power over consumers who differ regarding which firm's product they prefer. This theoretical construct captures several real world cases. A traditional interpretation is geographic: Consumers differ in where they live and thus face different physical distances to travel to a bank branch to apply for a loan. Alternatively, consumers might have an existing relationship with a firm and find switching costly, e.g. because only one bank can access past payment histories to judge creditworthiness,[14] or because consumers have acquired a chronic medical condition which the existing insurer must cover but based on which new insurers could refuse to offer insurance. This set-up could also capture behavioural factors, e.g. when consumers are differently attentive or find switching differently costly. Then, while some consumers switch already for small price differentials, other switch only for large price differentials. We now turn to the formal exposition.

Players

There are two firms, or plants, located at the ends of the unit interval. We name the firms L (left) and R (right). Consumers are heterogeneous along two dimensions and are

[14]This consideration was central to the Competition and Markets Authority launching the Open Banking initiative under which consumer data needs to be made accessible to rival banks. This aspect is also prominent in the literature on relationship banking (Sharpe, 1990; Boot and Thakor, 2000).

located on the unit square $(x, y) \in [0, 1]^2$. They differ in their location $x \in [0, 1]$, which can capture e.g. geographical location, and by their willingness-to-pay $y \in [0, 1]$. They have unit demand and thus purchase one good or no good. Consumers face travel costs $T(d)$ when purchasing from a firm which is distance d away, i.e. $d = x$ or $d = 1 - x$ respectively, and where $\frac{dT}{dd} > 0 \; \forall \; d$. Thus, consumers with large x prefer to purchase from R rather than L. As a normalisation, let $T(0) = 0$.

The defining feature of markets with selection is that consumers have private information on one aspect which affects both their willingness-to-pay (WTP) and their expected cost to the firm. For example in the case of health insurance, low health consumers have both a high WTP for health insurance and high expected claims.

In this model, consumers WTP (y) is private information and the expected cost to the firm $c(y)$ can depend on y. In the benchmark case of no selection, $c(y) = \alpha \; \forall y$. If $\frac{dc}{dy} < 0 \; \forall \; y$, the market exhibits advantageous selection and if $\frac{dc}{dy} > 0 \; \forall \; y$ there is adverse selection.[15]

Actions

Firms compete in prices. They simultaneously choose prices $p_i(x)$ where $i = \{L, R\}$. We distinguish between a case where firms can price-discriminate based on location x ($p_i(x)$ can vary across x) and a case of no price-discrimination ($p_i(x) = p_i \; \forall \; x$).[16] In the latter case firms are said to set a "gate price", i.e. a price everyone pays at the factory gate regardless of distance travelled. This case arises when firms do not observe x or when firms are not allowed to condition prices on it. Throughout, we assume that willingness-to-pay (y) is the consumer's private information.

Each consumer (x, y) faces a unique price pair $(p_L(x), p_R(x))$ and chooses whether to buy from L, from R, or not at all.

Pay-offs

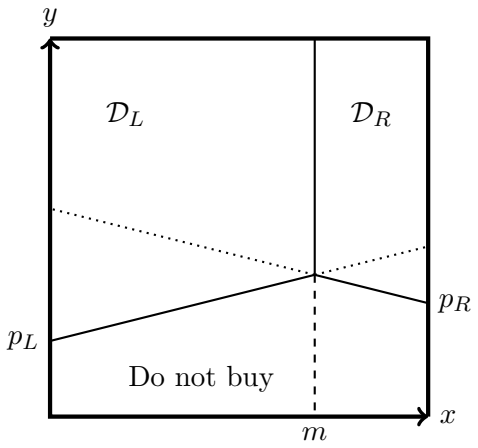
Firms maximise profit. Firm L's profit per consumer (x, y) is $\pi_L(x, y) = p_L(x) - c(y)$. Denoting the set of consumers buying from firm L as $\mathcal{D}_L$, firm L's total profit is the sum of per consumer profits across $\mathcal{D}_L$, i.e. $\Pi_L = \sum_{(x,y)\in\mathcal{D}_L} \pi_L = \sum_{(x,y)\in\mathcal{D}_L} [p_L(x) - c(y)]$.

Each consumer maximises her monetary benefit. Normalising the outside option of not purchasing to zero $u_o = 0$, purchasing from L yields $u_L = y - p_L(x) - T(x)$ while purchasing from R results in $u_R = y - p_R(x) - T(1 - x)$. The allocation resulting for an arbitrary pair of gate prices (p_L, p_R) is depicted below.

[15] Different micro-foundations for advantageous selection are possible. For the case of insurance markets, see Einav and Finkelstein (2011) or our exposition in section 3.2. In our setting, $c(y)$ would include administrative costs which is one way to rationalise why some consumers may have WTP below their expected cost.

[16] Price discrimination on distance travelled is equivalent to price discrimination based on location.

Figure 3.3: Purchasing behaviour for gate prices (p_L, p_R)



Purchasing behaviour for arbitrary prices (p_L, p_R).[17]

m is the location at which consumers are indifferent between purchasing from L or R.

We do not assume that every consumer's WTP is sufficiently high to ensure that all consumers purchase the good in equilibrium. Instead, we focus on cases where the socially optimal allocation is incomplete market coverage, i.e. some, but not all consumers, are allocated the good. This assumption is necessary to permit us to investigate whether in equilibrium inefficient over- or underprovision arises.[18]

Equilibrium concept

We focus on symmetric Nash equilibria. That means we focus on equilibria where firms set $p_L(x) = p_R(1-x)$ and consumers respond optimally to these prices. The focus on symmetric equilibria is also present in Mahoney and Weyl (2017).

3.3.2 Assumptions

Locational independence: We assume that the distribution of willingness-to-pay y, and thus also of costs $c(y)$, is independent from taste heterogeneity x. Thus, we can characterise the type distribution by two independent CDFs: $F(y)$ and $G(x)$.

Monotone selection: We assume that the entire market is characterised by the same form of selection. Thus, in adverse selection $\frac{dc}{dy} > 0 \; \forall \; y$ with strict inequality for all y, and for advantageous selection $\frac{dc}{dy} < 0 \; \forall \; y$.

Meaningful demand: As we are interested in studying competition, we restrict our attention to cases where at all x there exist benefits from trade for some y. Formally: $c(1) + T(\frac{1}{2}) \leq 1$. This ensures that the two firms are not monopolists pricing over disjoint

[17] These prices do not represent an equilibrium. In the diagram: $p_L = 0.2, p_R = 0.3, T(d) = 0.25\,d$.

[18] If every agent's WTP is so high that full coverage is efficient, i.e. allocating the good to everyone, then inefficient overprovision is impossible. As this paper studies remedies to the overprovision problem, we focus on a model where the overprovision problem can arise.

sets of demand, but that even at location $x = \frac{1}{2}$ there exists consumers with positive gains from trade. Firms will compete for these consumers.

Regularity conditions: We assume that $c(y) \geq 0 \; \forall \; y$ and that $c(y)$ is continuous and differentiable.

Cost heterogeneity at a given WTP: While our exposition treats costs as identical for all agents with the same WTP, our model is more general as we can allow cost heterogeneity at a given WTP. Then, $c(y)$ can be interpreted as the average cost at WTP y. Since all agents at a given (x, y) will make the same choice, both approaches are valid. While the approach with cost heterogeneity at a given y seems more realistic, we phrase this paper in terms of no cost heterogeneity at a given y. This makes the exposition more concise and avoids confusion between heterogeneity in costs *at* y (which is possible but not central to our model) and heterogeneity in costs *across* y (which is the defining feature of markets with selection and thus central to our model).

While the assumptions above are assumed throughout, for some results we make additional distributional and functional form assumptions. These assumptions are listed here and referred to were applicable.

- *Distributional assumption:* $F(y) = y$, i.e. the WTP distribution is uniform.
- *Distributional assumption:* $G(x) = x$, i.e. the location distribution is uniform.
- *Linear costs:* In the case of advantageous selection, we sometimes assume that $c(y)$ is linear in y. We then denote costs as $c(y) = \alpha + \beta(1 - y)$ where $\alpha \geq 0$ is the cost of the lowest cost agent and $\beta > 0$ captures the strength of advantageous selection.

3.3.3 Switching Consumers

We do not assume switching patterns, i.e. which types of consumers switch in response to a price cut, but derive them from our model. While the literature has assumed that switching consumers are representative of demand (Mahoney and Weyl, 2014), our model predicts that switching consumers are an advantageous selection of demand. This shapes the firms' strategic interaction.

The switching pattern affects whether overprovision arises. If the firm offering the lowest price attracts all consumers (perfect switching, Bertrand competition), inefficient overprovision results.[19] If, however, a price cut only attracts newly entering consumers and no switching consumers, firms never lower prices below the efficient level.[20] While Mahoney

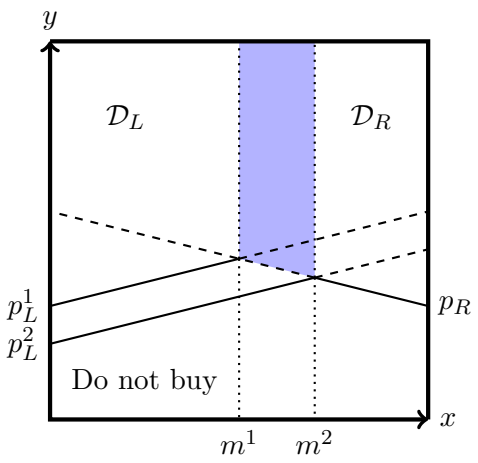
[19]This arises because in markets with advantageous selection (MC increasing, AC below MC), at the efficient allocation ($p = MC$) there still are positive industry profits ($p > AC$). Thus, when firms undercut each other until profits are zero, this results in overprovision. While the newly entering consumers are loss making, firms are compensated by stealing profitable consumers from the rival. At $p = AC$ undercutting stops.

[20]For prices below the efficient level, newly entering consumers are loss making for the firm. As there is no switching, there is no benefit of cutting prices as the firm steals no profitable consumers from the rival. Thus, it is not optimal for the firms to lower prices below the efficient level.

and Weyl (2017) assume "representative switching", i.e. that switching consumers are on average equal to the average consumer in demand, our model has the feature that switching consumers differ in a predictable way from the average consumer in demand.[21]

Our model predicts that, in a market with advantageous selection, those consumers who switch to a firm in response to it cutting its price are an advantageous selection of the rival's demand. This arises because a price cut only attracts those consumers from the rival who had the largest distance to travel, thus needed to pay large transport costs, and therefore only chose the rival over the outside option if their WTP is high. Thus, the switching consumers are higher WTP than the average of the rival's demand. Given advantageous selection, this makes switching consumers lower cost than average. Hence, the switching consumers are an advantageous selection of demand. This is depicted below:

Figure 3.4: Switching Consumers are an Advantageous Selection of Demand



Firm L's price cut from p_L^1 to p_L^2 results in consumers in the blue area switching. These consumers have on average a higher WTP than the consumers in $\mathcal{D}_R$.

The *extent* to which the margin is selected depends on the degree of competition. The softer competition (in the sense of higher transport costs), the stronger is the selection on the switching margin, i.e. the less representative it is.[22] Thus, the composition of the switching and entering margins is itself determined by the level of competition. It is not a fundamental and we cannot assume it to be fixed across different degrees of competition.

Related, but not shaping the firms' strategic interaction as strongly, is the composition of entering consumers. The consumers who in response to a price cut start purchasing initially did not buy the product and thus must have a lower WTP than the average

[21] The switching patterns in our model differ not just from the main case, but also from generalisations discussed by Mahoney and Weyl (2017). On p.640, footnote 7, they state that "Even if this assumption [representative switching] fails, so long as average switching consumers have a cost that is strictly between that of average exiting consumers and average purchasing consumers, most of our results are left unchanged." This generalises representative switching to allow switching consumers to have a *lower* average WTP than demand. Our model predicts a *higher* average WTP than demand, thus differs also from this generalisation.

[22] Only for $T(d) = 0$, i.e. perfect competition, does our model have perfect switching which is a case of representative switching. For any $T(d) = td$ with $t > 0$, switching becomes non-representative.

consumer in demand. Thus, in markets with advantageous selection, entering consumers are high cost.

3.4 Efficient Allocation

We characterise the efficient allocation, which is the benchmark against which we will judge equilibria, and show that in markets with advantageous selection, the efficient allocation can be implemented with price discrimination, but not with gate prices. Therefore, we also characterise the socially optimal gate price, that is the constrained efficient allocation where the constraint is to not be able to condition prices on location.

Comparisons of the efficient allocation and the equilibrium under imperfect competition are confounded by transport costs, which capture the degree of competition, having multiple effects. Larger transport costs have a strategic effect, which is what we aim to capture, but also have a direct cost effect, i.e. transport costs reduce total surplus achievable in a market. To avoid biasing our results in favour of competition, we isolate the strategic effect by comparing equilibrium outcomes to the outcome a welfare maximising social planner can achieve when he faces the same transport costs. This section derives this benchmark.

3.4.1 Welfare

Throughout, we focus on total surplus as our welfare measure. Thus, welfare is the sum of consumers' monetary benefit and the firms' profits.[23] Allocating the good to a consumer (x, y) generates surplus $s = y - T(d) - c(y)$. The outside option results in zero surplus.

Total surplus is the sum of s across all transactions which take place. Again denoting the set of consumers allocated the good from firm L as $\mathcal{D}_L$ and from R as $\mathcal{D}_R$, total surplus is defined as:

$$\mathcal{S} = \sum_{(x,y)\in\mathcal{D}_L} [y - T(x) - c(y)] + \sum_{(x,y)\in\mathcal{D}_R} [y - T(1-x) - c(y)] \tag{3.1}$$

3.4.2 Socially Optimal Allocation

The socially optimal allocation is to allocate the good to every agent whose WTP exceeds the sum of his costs and his transport costs, and allocate it from the firm where transport costs are lowest. Mathematically, the consumer should be allocated the good from firm L

[23] While in the main sections we abstract from the distinction between consumer surplus and producer surplus, we discuss alternatives in section 3.8.

if and only if the following two conditions both hold:

$$y \geq c(y) + T(x) \tag{3.2}$$

$$x \leq \frac{1}{2} \tag{3.3}$$

Thus, there exists a unique socially optimal allocation. This allocation has a familiar threshold structure, where $y^S(x) = c(y) + T(x)$ is a threshold. Then, all (x, y) with $y > y^S(x)$ and $x \leq \frac{1}{2}$ are allocated the good from L, all (x, y) with $y > y^S(x)$ and $x > \frac{1}{2}$ are allocated the good from R, and all others are not allocated the good.

The threshold $y^S(x)$ depends on the extent of selection. While in markets with no selection ($c(y) = \alpha \; \forall \; y$), the threshold increases in line with transport costs ($\frac{dy^S(x)}{dx} = T'(x)$), in markets with advantageous selection, the threshold is flatter than transport costs, but still increasing.

3.4.3 Socially Optimal Price

We show that while in markets with selection, the socially optimal allocation can be implemented using price discrimination, it cannot be implemented using gate prices. Thus, when the planner faces the additional constrained of using only gate prices, the socially optimal gate price can only approximate the efficient allocation. It results in overprovision for consumers with strong brand preferences and underprovision for those who are close to indifferent among the brands. Hence, while in markets without selection, there exists a gate price which implements the efficient allocation, in markets with advantageous selection, the following proposition holds.

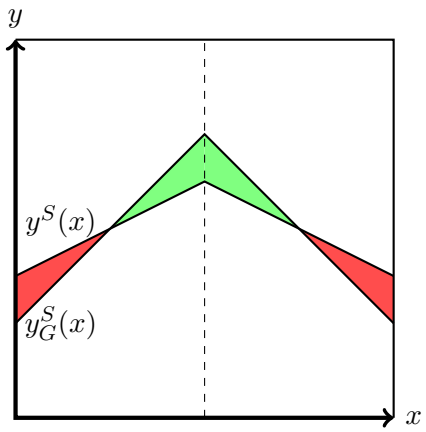
Proposition 3.1 *In markets with advantageous selection, no gate price can implement the efficient allocation.*

Proof: See Appendix 3.10.1.

Intuition: Suppose the planner sets the gate price to achieve the efficient allocation among consumers with the strongest brand preferences ($x = 0$). That means that the consumer $(0, p_G)$ is indifferent between purchasing from L and not purchasing. Thus, he generates zero surplus. At any $0 < x < \frac{1}{2}$, for this gate price the indifferent consumer will be $(x, p_G + T(x))$. That means that the WTP of the agent is by $T(x)$ higher than the WTP of the marginal agent at $x = 0$. If the market exhibits advantageous selection, the higher WTP means that fundamental costs ($c(y)$) are lower. In total, the marginal consumer $(x, p_G + T(x))$ must generate strictly positive surplus as his WTP covers his higher travel costs but the lower fundamental costs are not offset, thus creating positive surplus. Thus, agents with a slightly lower WTP than the marginal consumer at $x > 0$ would also still generate positive surplus. As they are not allocated the good, there is inefficient underprovision.

Thus, when the planner is limited to setting only gate prices, he is trying to find the constrained efficient allocation. Since gate prices result in steeper threshold than the efficient allocation, the constrained efficient allocation includes overprovision at $x = 0$ and underprovision at $x = \frac{1}{2}$. This is depicted below.[24]

Figure 3.5: Social Planner: Advantageous Selection



$y^S(x)$ is the efficient allocation, $y_G^S(x)$ is the allocation resulting from the socially optimal gate price. Green areas indicate underprovision, red areas indicate overprovision.

Example: When $c(y) = \alpha + \beta(1 - y)$, $T(d) = t\,d$, $F(y) = y$, $G(x) = x$, then:

$$p_G^S = \frac{\alpha + \beta}{1 + \beta} - \frac{\beta}{4(1 + \beta)}\, t \tag{3.4}$$

3.5 Perfect Competition and Monopoly

This section shows that our model of advantageous selection, which generates insights on imperfect competition, also nests the familiar results of perfect competition resulting in overprovision and of monopoly resulting in underprovision. Based on these results, it is tempting to believe that there exists a level of imperfect competition which corrects the overprovision caused by selection and achieves the first best efficient allocation. However, as the next section shows, this is not the case in our model - neither when firms use gate prices, nor when firms use price discrimination.

Perfect Competition

In models of spatial competition, the intensity of competition is captured by transport costs $T(d)$. Perfect competition corresponds to $T(d) = 0\ \forall\ d$. Then the firm with the lower price attracts all consumers, i.e. there is perfect switching, and captures all industry

[24] The inefficiency created by the planner not being able to price discriminate is increasing in transport costs and in the degree of selection. Moreover, selection and transport costs magnify each others effect on inefficiency.

profits. Thus, a firm wants to undercut the rival's price as long as industry profits are positive, i.e. $p > AC$.[25] In equilibrium, both firms price such that $p = AC$. That means that there is inefficient overprovision, since in markets with advantageous selection MC is increasing[26] and $AC < MC$. Overprovision arises because firms continue to undercut each other even when the newly entering consumers are loss making as firms try to steal the rival's profitable existing consumers.

Monopoly

Our model offers two ways to study monopoly, either via joint ownership or via high transport costs. In both cases the equilibrium exhibits inefficient underprovision.

In the joined ownership perspective, one firms sells both products L and R. As this firm internalises the effect that changes in the price of one product have on the demand for the other product, it will never overprovide and as it has market power it even raises prices above the efficient level and thus underprovides.

Proposition 3.2 *When $c(y)$ is linear and x and y are uniformly distributed, there is inefficient underprovision in monopoly at all locations x, regardless of the strength of selection and for all available pricing strategies.*

Proof: See Appendix 3.10.2.

When the monopolist can price discriminate and there is no selection ($c(y) = \alpha \; \forall y$), for $x \leq \frac{1}{2}$ it sets:

$$p^M(x) = \frac{1+\alpha}{2} - \frac{T(x)}{2} \tag{3.5}$$

which is the price that equates marginal revenue and marginal cost at a given location x.[27] This price is decreasing in distance travelled since the transport costs needed to travel larger distances reduce demand. The monopolist partly absorbs this effect of transport costs. As the effect is not absorbed fully, the effective price consumers face (and thus the WTP threshold above which consumers buy) is increasing in distance travelled, i.e. is $y^M(x) = \frac{1+\alpha}{2} + \frac{T(x)}{2}$. Note that $y^M(x)$ increases more slowly than $T(x)$ and is thus flatter than the threshold resulting from a gate price.

If there is advantageous selection, the monopolist sets:

$$p^M(x) = \frac{1 + c\big(p^M(x) + T(x)\big)}{2} - \frac{1}{2}T(x) \tag{3.6}$$

[25] In our model, average costs are $AC = \frac{\int_p^1 c(y)f(y)dy}{1-F(p)}$.
[26] In our model, marginal costs are $MC = c(p)$.
[27] This is the unique optimal price.

which for the linear costs example $c(y) = \alpha + \beta(1-y)$ is:[28]

$$p^M(x) \;=\; \frac{1+\alpha+\beta}{2+\beta} - \frac{1+\beta}{2+\beta}\, T(x) \tag{3.7}$$

A stronger degree of selection (larger β and equally a steeper marginal cost curve) leads to higher prices, i.e. $\frac{dp^M(x)}{d\beta} > 0$.

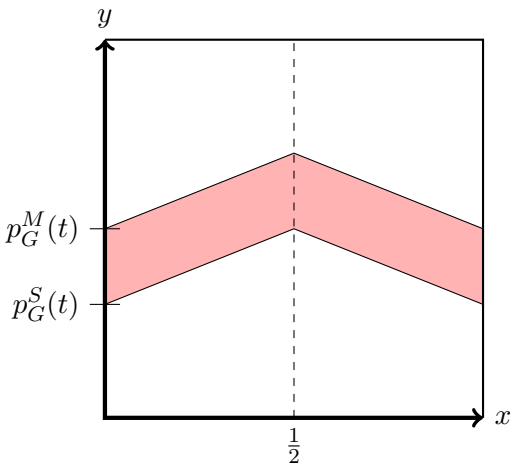
When the monopolist can only use gate prices, he cannot achieve his optimal outcome, i.e. there does not exist a gate price which implements the monopolist's optimal allocation.[29] Like in the social planner's case, this arises because the monopolist's optimal allocation has a threshold with a different slope than any gate price achieves.[30] The monopolist chooses a gate price (p_G^M) to approximate his preferred allocation. When $T(d) = t\,d$:

$$p_G^M = \frac{1+\alpha+\beta}{2+\beta} - \frac{1+\beta}{4\,(2+\beta)}\, t \tag{3.8}$$

provided that $t \leq \frac{4\,(1-\alpha)}{3+\beta}$. For $t > \frac{4\,(1-\alpha)}{3+\beta}$ there does not exist surplus generating trade at $x = \frac{1}{2}$, which violates our assumptions.

A stylised example of the allocation under monopoly is depicted below. The monopolist's gate price is always above the planner's gate price and it results in inefficient underprovision at all locations x.

Figure 3.6: Monopoly allocation for gate prices



An alternative way of studying monopoly is to focus on transport costs which are so high that no surplus generating trades at $x = \frac{1}{2}$ exist. Then, firms operate on disjoint sets

[28] This is the uniquelly optimal price. The resulting threshold describing the allocation is $y^M(x) \;=\; \frac{1+\alpha+\beta}{2+\beta} + \frac{1}{2+\beta}\, T(x)$.

[29] This means that the monopolist makes strictly larger profits with price discrimination than with gate prices.

[30] While the social planner could implement his optimal allocation with a gate price in the absence of selection, but not under selection, the monopolist can implement his preferred allocation in neither case. The monopolist's implementation problem arises because absorbing transport costs is impossible with gate prices. The planner's implementation problem arises because it is impossible to absorb the indirect effect of higher transport costs in markets with selection - i.e. lower fundamental costs.

of demand. The intuition for underprovision also applies in this case.

As this section showed that, in markets with advantageous selection, perfect competition results in inefficient overprovision and monopoly in underprovision, a natural question is whether there exists a level of imperfect competition which corrects the overprovision caused by selection and achieves the first best efficient allocation. However, while in Mahoney and Weyl (2017) such a level of competition exists, the next section shows that this is not the case in our model - neither when firms use gate prices, nor when firms use price discrimination.

3.6 Imperfect Competition

We find that no degree of imperfect competition can restore the first best efficient allocation in the presence of advantageous selection. While softer competition results in fewer trades occurring in equilibrium, this never exactly offsets the inefficient overprovision which results under perfect competition. Instead, imperfect competition results in inefficient underprovision and overprovision co-existing for different consumer types. This result is robust to different assumptions on the firm's pricing strategy, i.e. it applies both when firms use gate prices and when they use price discrimination. However, the result regarding which consumers experience underprovision and which experience overprovision does depend on the firms' strategy. While if firms use gate prices there is overprovision for consumers with strong brand preferences and underprovision for consumers who are indifferent among brands, the reverse is true when firms can price discriminate.

3.6.1 Gate Prices

When firms compete using gate prices, there exists an intermediate degree of competition at which the oligopoly equilibrium price coincides with the planner's optimal price. However, the planner's optimal price does not implement the efficient allocation. Due to the presence of advantageous selection, the planner's gate price is itself only constrained efficient, i.e. it approximates the first best efficient allocation under the constrained of using a gate price. Thus, at this price, consumers with strong brand preferences experience overprovision, while there is underprovision for consumers who are indifferent among the brands. That means the two types of inefficiency co-exist.

In the duopoly game with gate prices, firm L's profit is given by:

$$\pi_L(p_L, p_R) = \int_0^{m(p_L,p_R)} \int_{p_L+T(x)}^{1} [p_L - c(y)]\ f(y)\ dy\ g(x)\ dx \tag{3.9}$$

where $m(p_L, p_R)$ denotes the location x at which consumers are indifferent between purchasing from either firm. When transport costs are linear ($T(d) = td$ where $t \geq 0$), we get the familiar Hotelling result $m(p_L, p_R) = \frac{1}{2} + \frac{p_R - p_L}{2t}$. For this section on gate prices, we

assume linear transport costs, linear fundamental costs $c(y) = \alpha + \beta(1-y)$, and that type distributions are uniform. Then, profit maximisation gives rise to the following proposition, where p_G^C denotes the equilibrium gate price under imperfect competition, i.e. duopoly.

Proposition 3.3 *When $c(y)$ is linear and x and y are uniformly distributed, there exists a unique symmetric Nash equilibrium with equilibrium price*

$$p_G^C = \frac{1+\alpha+\beta}{2+\beta} + \frac{3+\beta}{2\,(2+\beta)}t - \frac{1}{2\,(2+\beta)}\sqrt{(2\beta^2+10\beta+13)t^2 - 4(1-\alpha)t + 4(1-\alpha)^2} \tag{3.10}$$

provided that $t \leq \frac{4\,(1-\alpha)}{3+\beta}$.

Proof: See Appendix 3.10.3.

If $t > \frac{4\,(1-\alpha)}{3+\beta}$, then firms price such that their demand sets are disjoint, i.e. at $x = \frac{1}{2}$ nobody purchases. That corresponds to behaving like a monopolist.

This proposition gives rise to a result on cost pass-through. While in itself only a minor result, our suggestion of a corrective tax as a policy tool will build on this result later on.

Corollary 3.1 *In oligopoly there is partial cost pass-through, i.e. $0 < \frac{dp_G^C}{d\alpha} < 1$, regardless of the presence and strength of selection.*

Proof: See Appendix 3.10.4.

While in the absence of selection, the standard ordering of prices under different ownership structures holds for all levels of transport costs ($p_G^M(t) > p_G^C(t) > p_G^S(t)\ \forall\ t > 0$) there is no such universal ordering under advantageous selection. When there is advantageous selection, then for very low transport costs the oligopoly equilibrium price is below the planner's optimal price. For large transport costs, the ordering is reversed.[31] This gives rise to the following proposition.

Proposition 3.4

(i) There exists a unique level of transport costs, denoted $\tilde{t}$, at which the oligopoly equilibrium price $p_G^C(t)$ and the socially optimal price $p_G^S(t)$ coincide, i.e. $p_G^S(\tilde{t}) = p_G^C(\tilde{t})$.

(ii) However, the oligopoly equilibrium allocation and the socially optimal allocation do not coincide for any level of transport costs t.

Proof: See Appendix 3.10.5.

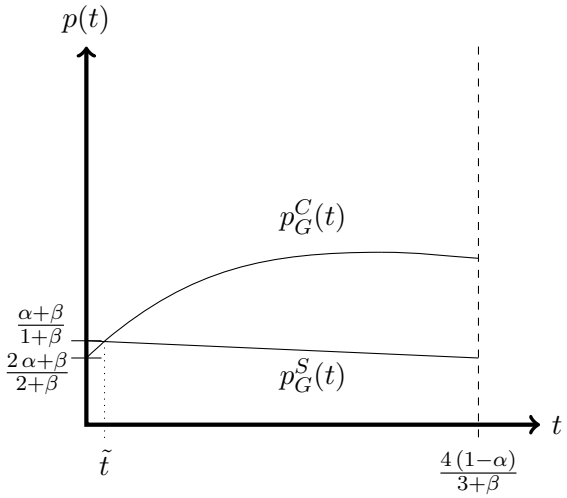
The first part of proposition 3.4 follows from a fixed point argument which builds on the limit results of perfect competition and monopoly as well as on the continuity of

[31] For all levels of transport costs $p_G^M(t)$ exceeds $p_G^C(t)$ and $p_G^S(t)$.

$p_G^C(t)$. Since we know that perfect competition ($t = 0$) results in $p_G^C(0) < p_G^S(0)$ and that $t = \frac{4\,(1-\alpha)}{3+\beta}$ results in monopoly pricing even though surplus generating trades exist, i.e. $p_G^C(\frac{4\,(1-\alpha)}{3+\beta}) > p_G^S(\frac{4\,(1-\alpha)}{3+\beta})$, there must be a level of t where $p_G^C(t) = p_G^S(t)$.

The intuition above also holds in our spatial model, but the formal exposition is complicated by larger transport costs having two effects. Larger transport costs reduce competition, which tends to raise prices, but they also alter demand which in this setting leads firms to lower their price. As a result, $p_G^C(t)$ is continuous but not monotone in t.[32] However, the intuition of $p_G^C(t)$ and $p_G^S(t)$ only crossing once is preserved in our model.

Figure 3.7: Gate Price Comparison



The second part of proposition 3.4 shows that no degree of imperfect competition can restore the efficient allocation when there is advantageous selection. This builds on proposition 3.1 that no gate price can implement the efficient allocation. The planner's optimal price is thus only constrained efficient. It is only an approximation to the efficient allocation. Hence, while for $\tilde{t}$ we have $p_G^S(\tilde{t}) = p_G^C(\tilde{t})$, the fact that $p_G^S(\tilde{t})$ is itself only an approximation to the efficient allocation means that there is overprovision for low x and underprovision for medium x (consumers who are indifferent between the brands) as in Figure 3.5.

Thus, proposition 3.4 cautions against viewing intermediate degrees of competition as desirable in markets with selection. Since $\tilde{t}$ achieves only the constrained efficient allocation, this leaves room for other policy interventions to achieve higher total surplus by reaching the first best efficient allocation. Section 3.7 explores this and finds that a corrective tax combined with perfect competition can achieve the first-best and is thus a socially more desirable policy.

Our arguments above rely on firms using gate prices. It is natural to wonder whether the impossibility result that no degree of imperfect competition achieves the first best also holds when firms can use price discrimination, i.e. are not restricted to charging all

[32]The effective price at $x = \frac{1}{2}$, $p_G^C(t) + \frac{1}{2}t$, is monotonically increasing in t.

consumers the same price. We now turn to the case of price discrimination and show that while the firms' strategies are altered and over-and underprovision occur for different types of consumers than under gate prices, the impossibility result continues to apply.

3.6.2 Price Discrimination

When firms use price discrimination, they compete very fiercely for consumers who are indifferent among brands and use their market power to set high prices for consumers with strong brand preferences. Since with price discrimination firms can target prices at consumers depending on brand preferences, these two motives are not conflicting, which they are under gate prices. As a result, overprovision and underprovision co-exist.

There is overprovision for consumers who have small brand preferences. While intermediate degrees of competition reduce the range of brand preferences in which overprovision arises, for any degree of imperfect competition overprovision continues to exist for those consumers who are exactly indifferent between brands.

There is underprovision for consumers with strong brand preferences. For softer degrees of competition, underprovision arises for a larger range of brand preferences. The extent of underprovision can be large. Already for intermediate degrees of competition, consumers with strong brand preferences experience the same degree of underprovision that arises in monopoly.

Characterisation of the Equilibrium

When firms use price discrimination, they effectively compete for separate markets at every location x. The transport cost differential $T(x) - T(1-x)$ affects how much market power a firm has at that location. Consumers with strong brand preferences for L (low x) will only buy from R if R sets a price which compensates for the extra travel, i.e. $p_R(x) \leq p_L(x) - [T(1-x) - T(x)]$, i.e. if R undercuts L by $T(1-x) - T(x)$. As firm R also needs to break even, R cannot profitably attract consumers when firm L prices no higher than break even price plus transport cost differential. For example, in the absence of selection the break even price is $p = \alpha$ and R cannot profitably undercut any $p_L(x) \leq \alpha + T(1-x) - T(x)$. This allows firm L to charge a mark-up at all $x < \frac{1}{2}$. If transport costs are sufficiently high, i.e. $T(d)$ is sufficiently steep, then at some locations x firm L can even charge the same price a monopolist would charge. Towards the middle of the unit interval, competition erodes the firm's mark-ups. The idea of transport cost differentials protecting firms' pricing power also applies when markets exhibit selection and gives rise to the following proposition:

Proposition 3.5 *When costs are linear and $F(y) = y$, then in duopoly there exists the following Nash equilibrium:*

(i) If $T(1) > \frac{1-\alpha}{2}$, then there exists a unique location x_L which partitions the location space $x \leq \frac{1}{2}$ into two regions: a region where firm L prices like a monopolist and a region where the presence of firm R causes firm L to lower prices relative to monopoly so that firm R does not sell the good. Formally, x_L is the solution to $1 - \alpha + T(x_L) = 2\,T(1 - x_L)$ and firm L prices:

$$For\ x \in [0, x_L] \quad p_L^M(x) = \frac{1+\alpha+\beta}{2+\beta} - \frac{1+\beta}{2+\beta}T(x) \tag{3.11}$$

$$For\ x \in]x_L, \frac{1}{2}] \quad p_L(x) = \frac{2\alpha+\beta}{2+\beta} + \frac{2}{2+\beta}T(1-x) - T(x) \tag{3.12}$$

For $x \in [0, x_L]$ any $p_R(x)$ which does not attract consumers is compatible with equilibrium. For $x \in]x_L, \frac{1}{2}]$, firm R prices $p_R(x) = \frac{2\alpha+\beta}{2+\beta} - \frac{\beta}{2+\beta}T(1-x)$.

(ii) If $T(1) \leq \frac{1-\alpha}{2}$, then at all locations x firm L's prices are affected by the presence of firm R. Thus, firm L prices:

$$For\ x \in [0, \frac{1}{2}] \quad p_L(x) = \frac{2\alpha+\beta}{2+\beta} + \frac{2}{2+\beta}T(1-x) - T(x) \tag{3.13}$$

and firm R sets:

$$For\ x \in [0, \frac{1}{2}] \quad p_R(x) = \frac{2\alpha+\beta}{2+\beta} - \frac{\beta}{2+\beta}T(1-x) \tag{3.14}$$

Firm R does not trade in equilibrium at any $x \leq \frac{1}{2}$.

The equilibrium prices on $x > \frac{1}{2}$ follow by the symmetric set-up of the model.[33]

Proof: See Appendix 3.10.6.

In the region $x \in]x_L, \frac{1}{2}]$ competition leads to lower prices than a monopolist would set. Firm L can cover its cost and makes an additional mark-up which equals the size of the transport cost differential. This arises because the threat of R undercutting forces L to lower prices relative to $p_L^M(x)$. L lowers prices just enough that R would make a loss from any price that attracts consumers.

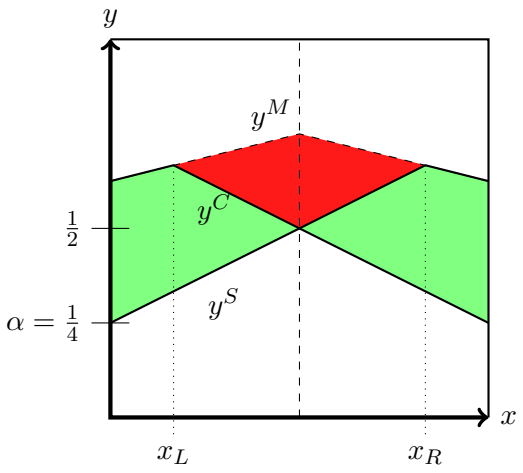
Example: Let $T(d) = t\,d$. Then iff $t > \frac{1}{2} - \frac{1}{2}\alpha$, then $x_L = \frac{2}{3} - \frac{1-\alpha}{3t}$, i.e. x_L is unaffected by the strength of selection β, and $p_L^M(x) = \frac{1+\alpha+\beta}{2+\beta} - \frac{1+\beta}{2+\beta}tx$. Iff $t \leq \frac{1}{2} - \frac{1}{2}\alpha$, then $x_L = 0$. In both cases, for $x > x_L$, $p_L(x) = \frac{2\alpha+\beta}{2+\beta} + \frac{2}{2+\beta}t(1-x) - tx$ and $p_R(x) = \frac{2\alpha+\beta}{2+\beta} - \frac{\beta}{2+\beta}t(1-x)$.

Diagrammatically:

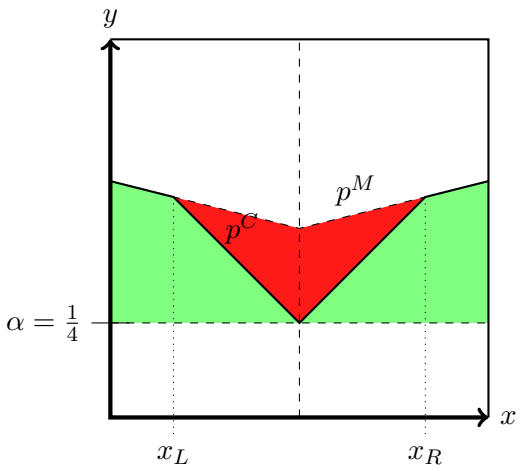
[33] This means that all consumers at $x > \frac{1}{2}$ buy from firm R, if they buy at all, and that firm R sets the monopolist's price unless transport costs are so low that firm L would undercut firm R. To prevent this, firm R sets a lower price.

Figure 3.8: Equilibrium under Duopoly: No Selection

(a) Allocation



(b) Prices



The social optimum is y^S or p^S, Monopoly results in y^M and p^M. The equilibrium outcome in imperfect competition or duopoly is y^C and p^C.
Light green shaded area is the inefficiency in duopoly relative to the socially optimal allocation.
Solid red shaded area is the efficiency gain in duopoly relative to monopoly.

The equilibrium has interesting monotonicity features:

Corollary 3.2

(i) The equilibrium price is monotonically decreasing in distance from the firm.

(ii) The number of consumers purchasing the good can be non-monotone in distance from the firm.

Proof: See Appendix 3.10.7.

The difference between price and quantity effects described in Corollary 3.2 arises because transport costs drive a wedge between the price consumers pay to the firm $p(x)$ and the total cost consumers incur when purchasing $p(x) + T(d)$. The price consumers pay to the firm, $p(x)$, is monotonically decreasing in distance travelled but the total cost incurred is not. At $x \in [0, x_L]$ the firm prices like a monopolist and thus partially absorbs transport costs. That means $p(x)$ is decreasing in x but at a slower rate than transport costs. Thus $y(x)$ is increasing. When there is effective competition ($x \in [x_L, 1 - x_L]$), prices decrease more strongly than transport costs rise. Thus, prices are decreasing in distance and quantity is decreasing, too. Hence, the non-monotonicity arises because for $x < x_L$, distance only affects a monopolist's pricing while for $x \in [x_L, \frac{1}{2}]$ it acts to intensify competition.

Comparative Static:

When competition is less fierce (transport costs are higher) firms engage in monopoly pricing at more locations, i.e. $\frac{dx_L}{dt} > 0$. This arises because the price a monopolist would set at a given x increases by less than transport costs, while the price the competitor needs to break even increases exactly by transport costs. Thus, at a given x there is less room for profitably undercutting which means that already at x closer to $\frac{1}{2}$ no room for undercutting remains and monopoly pricing occurs. Even if duopoly pricing occurs, the larger transport costs result in higher prices, $\frac{dp_L(x)}{dt} > 0$.

Efficiency Properties of the Equilibrium

While the structure of the equilibrium does not depend on the presence of selection, the efficiency properties are affected by the presence of selection.

Proposition 3.6 *In the absence of selection, when y is uniformly distributed, the duopoly equilibrium exhibits inefficient* under*provision at all x other than* $x = \frac{1}{2}$ *and efficient provision at* $x = \frac{1}{2}$.

Proof: See Appendix 3.10.8.

This proposition holds for all transport costs.[34] This result is visualised in Figure 3.8 where the inefficiency of duopoly in terms of unexploited surplus generating trades is captured by the area between $y^C(x)$ and $y^S(x)$, or correspondingly in terms of mark-ups is captured by the area between $p^C(x)$ and α (both are visualised as the light green shaded areas). The efficiency gain of duopoly relative to monopoly is captured by the gap between $y^C(x)$ and $y^M(x)$ or respectively $p^C(x)$ and $p^M(x)$ (solid red shaded area).

When a market is characterised by advantageous selection, there exists inefficient *over*provision for agents with no or only small brand preferences while simultaneously there is *under*provision for consumers with strong brand preferences. Formally:

Proposition 3.7 *In markets with advantageous selection, when costs are linear and y is uniformly distributed, the duopoly equilibrium allocation has the following efficiency properties:*

(i) If $T(1) \geq \frac{\beta}{2+2\beta}(1-\alpha)$, *there exists a unique* x_e *which solves* $\beta(1-\alpha) = (2+2\beta)\, T(1-x_e) - (2+\beta)\, T(x_e)$. *Then*

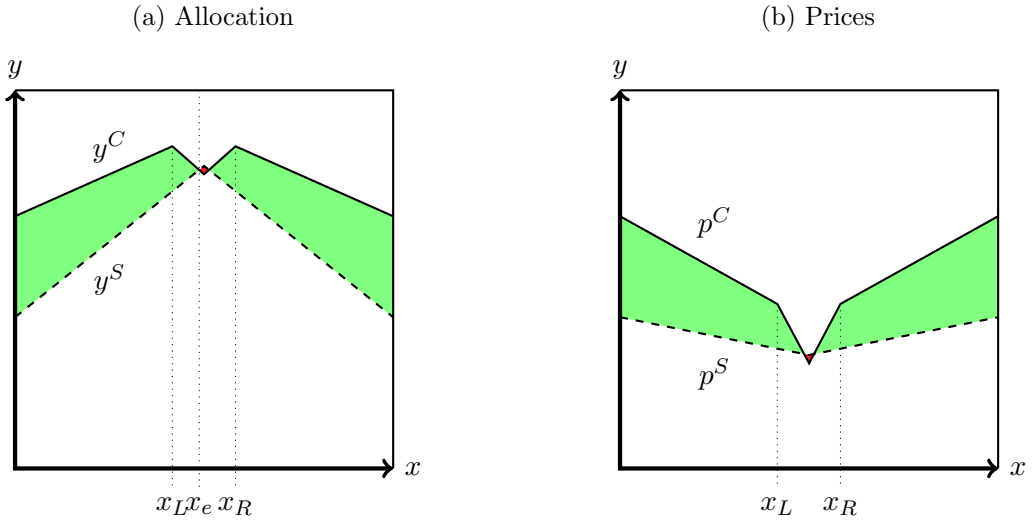
$$\begin{aligned} &\textit{For all} \quad x \in [0, x_e[\quad \textit{there is inefficient underprovision.} \\ &\textit{For all} \quad x \in]x_e, \frac{1}{2}] \quad \textit{there is inefficient overprovision.} \end{aligned}$$

(ii) If $T(1) < \frac{\beta}{2+2\beta}(1-\alpha)$, *then* $x_e = 0$ *and there is inefficient overprovision at all* x.

Proof: See Appendix 3.10.9.

[34] We only require that $T(d)$ is strictly increasing in distance travelled, i.e. that $\frac{dT}{dd} > 0\ \forall d$.

Figure 3.9: Equilibrium under Duopoly: Advantageous Selection



where y^C is the threshold characterising the duopoly equilibrium allocation, i.e. imperfect competition, y^S is the threshold which describes the socially optimal allocation.

The green areas highlight inefficient *under*provision and red areas (small, around $x = \frac{1}{2}$) highlight inefficient *over*provision.[35]

Example: When $T(d) = t\,d$ then $x_e = \frac{2}{4+3\beta} + \frac{\beta}{t}\,\frac{(2t+\alpha-1)}{(4+3\beta)}$.

Thus, inefficient overprovision and underprovision can coexist and do not cancel each other out perfectly. The inefficiency from underprovision at $x \in [0, x_e]$ arises because firms have market power over consumers with strong brand preferences (akin to "captured consumers") and exploit this to raise prices. This argument is not specific to selection markets, but our results show that it also applies in the presence of selection. The inefficiency from overprovision at $x \in [x_e, \frac{1}{2}]$ is driven by forces specific to selection markets which are familiar from our discussion of perfect competition in markets with advantageous selection: Consumers at x close to $\frac{1}{2}$ are close to indifferent between the firms' products. Thus, firms compete fiercely for these consumers until profits are completely eroded. Since we can view each x as a separate market with advantageous selection, competition eroding profits completely at an x means that at that x there is inefficient overprovision.[36]

This gives rise to two impossibility results which caution against viewing imperfect competition as an effective policy solution for the problem of inefficient overprovision in markets with advantageous selection.

[35] Under- and overprovision refer to comparisons the equilibrium under imperfect competition to the socially optimal allocation.

[36] Since in markets with advantageous selection $MC > AC$, when profits are eroded at $p = AC$ the price is below MC and thus there is inefficient overprovision.

Proposition 3.8 *Impossibility Results: When costs are linear and y is uniformly distributed:*

1. *In markets with advantageous selection, the duopoly equilibrium does* not *coincide with the socially efficient allocation for any degree of competition.*
2. *In markets with advantageous selection, no degree of imperfect competition removes overprovision at all x.*

Proof: See Appendix 3.10.10.

This proposition holds for all $T(d)$ and all strength of selection. The only condition is that selection exists, i.e. $\beta > 0$ with strict inequality.

While our model is in line with the familiar results that perfect competition leads to overprovision at all x and that monopoly leads to underprovision at all x, it yields new predictions for the intermediate case of imperfect competition. No level of competition achieves the efficient allocation at all x simultaneously. Instead, for intermediate degrees of competition there exists a cut-off location x_e and at all locations x with consumers with strong brand preferences ($x < x_e$ or $x > 1 - x_e$) there is *under*provision while at all locations x with consumers with less strong brand preferences ($x_e < x < 1 - x_e$) there is inefficient *over*provision.

Less intense competition (higher transport costs) increase the set of x in which underprovision arises (x_e increases), but there always remains some x at which overprovision exists. E.g. at $x = \frac{1}{2}$, there is overprovision for all degrees of imperfect competition. Only monopoly would remove this overprovision. Thus, imperfect competition introduces a new type of inefficiency at some x (underprovision) without removing overprovision at other x. Thus, these inefficiencies do not cancel each other out perfectly at all x, but co-exist.

This section showed that neither under gate prices, nor under price discrimination does there exist a degree of imperfect competition which achieves the first best efficient allocation. Given that it is impossible to reach the first best allocation by tolerating lower degrees of competition, a natural question is whether alternative policy tools exist which can achieve the first best. The next section focuses on such alternative policy tools.

3.7 Policy Remedies

Since no degree of imperfect competition can perfectly correct the inefficiency created by advantageous selection, this section discusses alternative policy solutions. First, we study a corrective tax, i.e. a tax or subsidy which applies equally to all trades. Then, we turn to "risk adjustment", which is a tax or subsidy conditioned on an individual's characteristics.

3.7.1 Corrective Tax

We find that the combination of a corrective tax and fierce competition achieves the first best efficient allocation and thus results in higher total surplus than intermediate degrees of competition, as these only achieve the constrained efficient allocation. Thus, competition policy should not be lenient towards firms because they operate in a market with advantageous selection. Instead, competition policy should continue to focus on achieving perfect competition. The distortions arising from selection can be better addressed through a tax than through altered competition policy. The corrective tax has the additional benefit of generating government revenue, holds regardless of whether firms use gate prices or price discrimination, has lower informational requirements, and is redistributive as firms absorb some of the tax.

Proposition 3.9 *There exists a corrective tax which achieves the first best efficient allocation.*

Proof: See Appendix 3.10.11.

A tax per unit sold (τ) is treated by firms like an upward shift in costs. If $c(y) = \alpha + \beta(1-y)$, it corresponds to an increase in α. Since corollary 3.1 established that there is imperfect cost pass-through, i.e. that $0 < \frac{dp_G^C(t)}{d\alpha} < 1$, the corrective tax leads firms to raise prices and, when correctly calculated and combined with perfect competition, the tax can restore the first-best efficient allocation.

Example: Let $c(y) = \alpha + \beta(1-y)$, $F(y) = y$, $G(x) = x$, then there is inefficient overprovision under perfect competition, i.e.

$$p_G^S(0) = \frac{\alpha+\beta}{1+\beta} > p_G^C(0) = \frac{2\,\alpha+\beta}{2+\beta} \tag{3.15}$$

Since a per unit tax (τ) results in equilibrium prices $p_G^C(0) = \frac{2\,(\alpha+\tau)+\beta}{2+\beta}$, we can solve for the tax rate which achieves the first-best efficient allocation:[37]

$$\tau^* = \frac{\beta\,(1-\alpha)}{2\,(1+\beta)} \tag{3.16}$$

Regardless of whether firms use gate prices or price discriminate, a per unit tax can always achieve the first-best efficient allocation. The optimal tax rate is the same in both cases. Thus, the policy of using a corrective tax is robust regardless of which pricing strategy firms use.

[37] More generally, for any transport cost t, there exists a tax rate or subsidy rate which achieves the constrained efficient allocation, i.e. for which $p_G^S(t) = p_G^C(t)$. This tax rate is implicitly defined by $p_G^S(\alpha,\beta,t) = p_G^C(\alpha+\tau,\beta,t)$ or equally by $\frac{\alpha+\beta}{1+\beta} - \frac{\beta}{4\,(1+\beta)}t = \frac{1+\alpha+\tau+\beta}{2+\beta} + \frac{3+\beta}{2(2+\beta)}t - \frac{1}{2(2+\beta)}\sqrt{(2\beta^2+10\beta+13)t^2 - 4(1-\alpha-\tau)t + 4(1-\alpha-\tau)^2}$. However, the first best efficient allocation is only achieved under perfect competition ($t=0$) combined with tax rate (3.16). Moreover, all cases with $t>0$ are further complicated by considerations whether firms use gate prices or price discrimination and by whether the tax rate has to apply equally to all x or can discriminate based on x.

The corrective tax not only results in larger total surplus than imperfect competition, it also has lower informational requirements. In order to calculate the level of competition which achieves the constraint efficient allocation ($\tilde{t}$), the policy maker needs to know costs $c(y)$ and both type distributions $F(y)$ and $G(x)$. In order to calculate the optimal corrective tax, which achieves the first-best efficient allocation (τ^*), the policy maker only needs to know costs $c(y)$ and the WTP distribution $F(y)$, but not the preference distribution $G(x)$. This arises because perfect competition has the additional advantage that it makes all consumers face the same total cost of purchasing $p + T(x) = p \; \forall \; x$, and thus results in consumers at all x behaving equally. This makes it unnecessary to know $G(x)$.

There are further benefits of a corrective tax not captured in our model. A corrective tax generates government revenue which can be used to reduce distortionary taxes in other markets and thus enhances efficiency. Attempts to correct overprovision through less fierce competition have no comparable benefit as instead of government revenue they increase firm profits and create a deadweight loss (transport costs).

When welfare is not defined as total surplus, but places more weight on consumer surplus than on producer surplus, a corrective tax has additional distributional benefits. Since firms absorb some of the tax, i.e. pass through is partial $0 < \frac{dp^C}{d\alpha} < 1$, the corrective tax increases total surplus but reduces producer surplus. Hence, consumer surplus and government revenue combined have increased by more than total surplus. A welfare measure emphasizing consumer surplus would increase because of the larger total surplus achieved through the corrective tax and additionally increase further because of the effect the corrective tax has on the *composition* of total surplus.

3.7.2 Risk Adjustment

An alternative policy to the corrective tax or to tolerating intermediate degrees of competition is risk adjustment. Under risk adjustment, firms receive a subsidy conditional on the type of consumers they insure such that the cost a firm incurs from insuring a consumer is independent from the consumer's type. Thus, risk adjustment effectively removes selection.

While risk adjustment is used in practice, e.g. in health insurance markets, a corrective tax has lower informational requirements and is more robust to gaming of risk scores by insurance providers. The corrective tax shifts every agent's cost equally and thus preserves heterogeneity in costs. Hence, to implement the tax, neither the policy maker nor the firm needs to know the agent's type y. Under risk adjustment, firms receive a subsidy conditional on y. Thus, both the firm and the policy maker need to know y. In practice, a consumer's riskiness is not observed and subsidy payments are conditioned on risk scores which insurers assign. This generates an incentive problem as the insurer prefers to make the consumer appear to have large expected claims in order to get a large subsidy. The

tax does not require risk scores and therefore does not generate this incentive problem.[38] Thus, we prefer the corrective tax to risk adjustment.

Gate prices: Our model predicts that non-representative switching patterns matter when calculating the effect of risk adjustment. One way of implementing risk adjustment is to adjust every agent's cost to equal the average cost of all agents purchasing. When costs are linear, this is equal to the cost of the average infra-marginal consumer. However, when using gate prices, firms' pricing decisions are driven by the average marginal consumer, which is calculated across switching and entering consumers. When the average marginal consumer is a lower type than the average infra-marginal consumer, as is the case when distributions are uniform, then in a market with advantageous selection, firms respond to this form of risk adjustment by lowering their prices.

Price discrimination: Calculations of the effect of risk adjustment, or equally comparative statics on the strength of selection, face the obstacle that a change in β corresponds to both altering the slope of marginal costs and also to altering the level of costs for a given demand. That means $\frac{dx_e}{d\beta}$ captures both a pure cost effect, i.e. an increase in total cost for a given level of demand, and an increase in the strength of selection. We need to isolate the effect of increased selection. We achieve this through an indirect approach. $\frac{dx_e}{d\alpha}$ captures a pure cost effect with no change in selection. We find that $\frac{dx_e}{d\alpha} > 0$. We also find that $\frac{dx_e}{d\beta} < 0$. Taken together, this means that the selection effect has the opposite sign of a pure cost effect and more than outweighs it. Thus, purely increasing the strength of selection while not altering average cost in the market likely leads to inefficient overprovision occurring at more locations x. Risk adjustment, i.e. a reduction in the strength of selection, has the opposing effect. Formally:

Corollary 3.3 *When costs are linear and y is uniformly distributed, then if selection is more pronounced, overprovision occurs at more locations x for given transport costs $T(d)$.*

Proof: See Appendix 3.10.12.

3.8 Discussion

Our results, which are in favour of competition, hold despite the fact that we make several assumptions which tend to understate the social benefit of competition.

We assume that the social planner maximises total surplus. Under any alternative welfare measure which places more weight on consumer surplus relative to producer surplus, competition is even more beneficial. This arises because while overprovision, triggered by fierce competition, and underprovision, triggered by soft competition, both

[38] Under the corrective tax, firms have an incentive to attract the lowest cost consumers. This is already accounted for in our model of imperfect competition. Thus, despite this effect, a tax can achieve the first best efficient outcome.

constitute a reduction in total surplus, their distributional implications differ. When there is overprovision, consumer surplus is larger than in the efficient allocation but producer surplus is reduced more strongly, hence total surplus is lower than in the efficient allocation. Hence, fierce levels of competition are socially desirable from a consumer surplus standpoint even when they lead to overprovision.[39]

We focus purely on the short term strategic effect of different degrees of competition. Competition can have the additional benefit of spurring innovation and thus lowering costs. This seems particularly relevant in markets with advantageous selection as inefficient overprovision can only occur if for some agents costs exceed their willingness-to-pay. This is rationalised, e.g. in an insurance setting, by administrative cost increasing costs above expected claims.[40] When competition has the additional benefit of eroding administrative costs, competition results in the efficient allocation (everyone purchasing insurance) even in the absence of a corrective tax.

In our model, competition has a direct cost effect and a strategic effect. Our approach was to isolate the strategic effect by judging the duopoly equilibrium relative to allocations a planner can achieve when he faces the same transport costs. An alternative approach to focus on total surplus would include the direct cost and would thus view competition as even more beneficial.

3.9 Conclusion

This paper developed a spatial model of imperfect competition in markets with selection and used it to discuss the problem of inefficient overprovision in markets with advantageous selection. The model is flexible enough to also study imperfect competition in a market with adverse selection. We aim to do this in a companion paper.

Our paper also relates to the literature on price discrimination by introducing selection and by studying a market with less than full coverage. Future research can explore these themes further by developing a game theory set up of firms' strategic choice of price strategy (as in Thisse and Vives (1988)) and study how the firms preference for one pricing strategy over the other is affected by selection.

Overall, our paper cautions against viewing imperfect competition as a solution to the inefficiencies introduced by advantageous selection. Combining tough competition policy with a corrective tax results in higher welfare than intermediate degrees of competition and has lower informational requirements. This suggests that if we view the build up of credit prior to the recent financial crisis as inefficient overprovision in a market with advantageous selection, then we should not conclude that competition in credit markets was "excessive" pre financial crisis, but rather should conclude that we lacked the appropriate corrective taxation.

[39] Industry profits are weakly positive, ensuring that it is optimal for firms to provide goods in this market.
[40] In the insurance literature, this is referred to as firms charging a loading factor.

3.10 Appendix

3.10.1 Proof of Proposition 3.1

Sketch proof: Any gate price results in an allocation which is characterised by a WTP threshold with slope $T'(x)$. Without selection, the efficient allocation is characterised by a threshold with slope $\frac{dy^S(x)}{dx} = T'(x)$ and since the slopes are equal there does exist a gate price which implements the efficient allocation ($p = \alpha$). With advantageous selection, the threshold describing the efficient allocation is flatter than transport costs. Thus, the threshold resulting from a gate price and the threshold describing the efficient allocation never coincide. *q.e.d.*

Formal proof: First, we establish three Lemmata. Then, we combine them to prove proposition 3.1.

Lemma 3.10.1.1 *For any gate price p_G, consumers purchase the good if and only if their WTP exceeds a threshold $y_G(x)$ where $\frac{dy_G(x)}{dx} = T'(x)$.*

Proof: A consumer purchases the good if and only if

$$y > p_G + T(x) \tag{3.17}$$

By definition

$$y_G(x) = p_G + T(x) \tag{3.18}$$

Thus

$$\frac{dy_G(x)}{dx} = T'(x) \tag{3.19}$$

q.e.d.

Lemma 3.10.1.2 *In markets without selection, in the efficient allocation consumers are allocated the good if and only if their WTP exceeds a threshold $y^S(x)$ where $\frac{dy^S(x)}{dx} = T'(x)$.*

Proof: It is efficient to allocate the good to a consumer if and only if

$$y > c(y) + T(x) \tag{3.20}$$

By definition

$$y^S(x) = c\big(y^S(x)\big) + T(x) \tag{3.21}$$

In the absence of selection $c(y) = \alpha \; \forall \; y$. Thus,

$$y^S(x) = \alpha + T(x) \tag{3.22}$$

Therefore,

$$\frac{dy^S(x)}{dx} = T'(x) \tag{3.23}$$

q.e.d.

Lemma 3.10.1.3 *In markets with advantageous selection, in the efficient allocation consumers are allocated the good if and only if their WTP exceeds a threshold $y^S(x)$ where $\frac{dy^S(x)}{dx} < T'(x)$.*

Proof: It is efficient to allocate the good to a consumer if and only if

$$y > c(y) + T(x) \tag{3.24}$$

By definition

$$y^S(x) = c\big(y^S(x)\big) + T(x) \tag{3.25}$$

In the presence of advantageous selection $\frac{dc}{dy} < 0$. Thus, totally differentiating and rearranging yields

$$\frac{dy}{dx} = \frac{1}{1 - \frac{dc}{dy}}\, T'(x) \tag{3.26}$$

where $\frac{1}{1-\frac{dc}{dy}} < 1$ due to the presence of advantageous selection. *q.e.d.*

By Lemma 3.10.1.1 and 3.10.1.2, in markets without selection, the equilibrium allocation and the socially optimal allocation are characterised by thresholds with the same gradient. Thus, there exists a gate price at which also the levels coincide at all x.

By Lemma 3.10.1.1 and 3.10.1.3, in markets with selection, the equilibrium allocation and the socially optimal allocation are characterised by thresholds with different gradients. Thus, if the levels of the allocation coincide at one location x, they diverge at all other locations. Hence, there does not exist a gate price which can implement the socially optimal allocation.[41] *q.e.d.*

3.10.2 Proof of Proposition 3.2

First, we prove this proposition for the case where firms use gate prices. Then, we prove it for the case where firms can price discriminate.

Lemma 3.10.2.1 *In monopoly, when firms use gate prices, there is inefficient underprovision at all locations x regardless of the strength of selection.*

[41] The socially optimal gate price is depicted in Figure 3.5 under additional distributional assumptions. It achieves the efficient allocation at $x = \frac{1}{4}$, results in inefficient overprovision at all $x \in [0, \frac{1}{4}[$ and in inefficient underprovision at all $x \in]\frac{1}{4}, \frac{1}{2}]$.

Proof: The monopolist solves $\max_{p_L} \pi$ where

$$\pi = \int_0^{\frac{1}{2}} \int_{p_L+T(x)}^{1} [p_L - c(y)]\ f(y)\, dy\ g(x)\, dx \tag{3.27}$$

using the assumptions here, i.e. that $c(y) = \alpha + \beta(1-y)$, $T(d) = t\,d$ where $t > 0$, $F(y) = y$, $G(x) = x$, profit becomes:

$$\pi = \int_0^{\frac{1}{2}} \int_{p_L+tx}^{1} [p_L - \alpha - \beta(1-y)]\ dy\ dx \tag{3.28}$$

which solves to:

$$\pi = -\Big(\frac{2+\beta}{4}\Big)p_L^2 + \frac{1}{2}(1+\alpha+\beta)p_L - \frac{1}{8}p_L t - \frac{1}{8}\beta p_L t - \frac{1}{2}\alpha - \frac{1}{4}\beta + \frac{1}{8}\alpha t + \frac{1}{8}\beta t - \frac{\beta}{48}t^2 \tag{3.29}$$

The FOC $\frac{d\pi}{dp_L} = 0$ is:

$$-\frac{2+\beta}{2}p_L + \frac{1}{2}(1+\alpha+\beta) - \frac{1}{8}t - \frac{1}{8}\beta t = 0 \tag{3.30}$$

which yields the monopolist's optimal gate price:

$$p_G^M = \frac{1+\alpha+\beta}{2+\beta} - \frac{1+\beta}{4(2+\beta)}\ t \tag{3.31}$$

This is a profit maximum as it also satisfies the SOC, i.e. $\frac{d^2\pi}{dp_L^2} = -\frac{2+\beta}{2} < 0$.

The solution was derived assuming meaningful demand. Checking that this assumption is satisfied, i.e. that at $x = \frac{1}{2}$ some consumers buy, results in the condition $t \leq \frac{4(1-\alpha)}{3+\beta}$.

The socially optimal allocation is characterised by a WTP threshold $y^S(x)$ above which consumers are allocated the good in the socially optimal allocation. $y^S(x)$ is defined by

$$y^S(x) = t\,x + c\big(y^s(x)\big) \tag{3.32}$$

which solves to

$$y^S(x) = \frac{\alpha+\beta}{1+\beta} + \frac{1}{1+\beta}\,t\,x \tag{3.33}$$

To prove Lemma 3.10.2.1, we need to show that

$$p_G^M + tx > y^S(x)\ \forall\ x \tag{3.34}$$

which is equivalent to

$$t < \frac{4(1-\alpha)}{(1+\beta)^2 - (8\beta + 4\beta^2)x} \tag{3.35}$$

For the monopolist, there is meaningful demand only if $t \leq \frac{4(1-\alpha)}{3+\beta}$. Thus if $\frac{4(1-\alpha)}{3+\beta} <$

$\frac{4(1-\alpha)}{(1+\beta)^2-(8\beta+4\beta^2)x}$, then also equation (3.35) holds. The condition

$$\frac{4(1-\alpha)}{3+\beta} < \frac{4(1-\alpha)}{(1+\beta)^2-(8\beta+4\beta^2)x} \tag{3.36}$$

simplifies to

$$\beta^2 - 8\beta x - 4\beta^2 x < 2 - \beta \tag{3.37}$$

which holds at all x for all $\beta < 1$. *q.e.d.*

Lemma 3.10.2.2 *In monopoly, when firms use price discrimination, there is inefficient underprovision at all locations x regardless of the strength of selection.*

Proof: At every location x, the monopolist prices such that $MR = MC$. This must always be a lower quantity than the efficient allocation which is characterised by $p = MC$. This argument replicates the familiar logic behind underprovision of monopoly, but while usually the argument is applied across the entire market, here it applies at every location x.

Formally, the monopolist's problem at a location x is to choose a p to maximise

$$\pi(x) = \int_{p+T(x)}^{1} [p - \alpha - \beta(1-y)]\, dy \tag{3.38}$$

which simplifies to:

$$\pi(x) = p - p^2 - T(x)p - \left[\frac{\beta}{2}p^2 + (\beta T(x) - \alpha - \beta)p + \alpha + \frac{\beta}{2} - \alpha T(x) - \beta T(x) + \frac{\beta}{2}T(x)^2\right] \tag{3.39}$$

The FOC, $\frac{d\pi}{dp} = 0$ yields

$$p^M(x) = \frac{1+\alpha+\beta}{2+\beta} - \frac{1+\beta}{2+\beta}T(x) \tag{3.40}$$

which is the unique profit maximum since the SOC holds, i.e. $\frac{d^2\pi}{dp^2} = -(2+\beta) < 0$.

There is inefficient overprovision if $p^M(x) + T(x) > y^S(x)$, which simplifies to $1 > \alpha + T(x)$ which always holds by meaningful demand. *q.e.d.*

3.10.3 Proof of Proposition 3.3

Given our assumptions that $c(y) = \alpha + \beta(1-y)$ where $\alpha \geq 0$, $\beta \geq 0$; $T(d) = t\,d$ where $t > 0$, $F(y) = y$, $G(x) = x$, we can simplify L's profit function in terms of gate prices p_L, p_R:

$$\pi_L(p_L, p_R) = \int_0^{m(p_L,p_R)} \int_{p_L+T(x)}^{1} [p_L - c(y)]\; f(y)\, dy\; g(x)\, dx \tag{3.41}$$

to:

$$\pi_L(p_L, p_R) = \int_0^{m(p_L,p_R)} \int_{p_L+tx}^{1} p_L - \alpha - \beta(1-y)\, dy\, dx \tag{3.42}$$

and to:

$$\begin{aligned}\pi_L(p_L, p_R) =& \frac{1}{2}\, t\, m(p_L, p_R)^2 \left(\alpha + \beta - (1+\beta)p_L\right) \; - \; \frac{1}{6}\, \beta\, t^2\, m(p_L, p_R)^3 \\ &+ m(p_L, p_R) \left((1+\alpha+\beta)p_L - (1+\frac{\beta}{2})p_L^2 - \alpha - \frac{1}{2}\beta\right)\end{aligned} \tag{3.43}$$

where $m(p_L, p_R)$ is the consumer who is indifferent between consuming from firm L and firm R. Thus:

$$m(p_L, p_R) = \frac{1}{2} + \frac{p_R - p_L}{2t} \tag{3.44}$$

and $\frac{dm}{dp_L} = -\frac{1}{2t}$.

Then, firm L's best response function is given by $\frac{d\pi_L}{dp_L} = 0$, i.e.

$$\begin{aligned}\frac{d\pi_L}{dp_L} =& t\; m(p_L, p_R)\; \frac{dm}{dp_L} \left(\alpha + \beta - (1+\beta)p_L\right) \; - \; \frac{1}{2}\; t\; m(p_L, p_R)^2\; (1+\beta) \\ &- \; \frac{1}{2}\; \beta\; t^2 m(p_L, p_R)^2\; \frac{dm}{dp_L} \\ &+ \; \frac{dm}{dp_L} \left((1+\alpha+\beta)p_L - (1+\frac{\beta}{2})p_L^2 - \alpha - \frac{1}{2}\beta\right) + m(p_L, p_R) \left(1+\alpha+\beta-(2+\beta)p_L\right) = 0\end{aligned} \tag{3.45}$$

This equation simplifies when using that $\frac{dm}{dp_L} = -\frac{1}{2t}$. Moreover, to solve for symmetric Nash equilibria we can use that $p_L = p_R = p$ and thus that $m = \frac{1}{2}$. This simplifies (3.45) to:

$$p_L = \frac{1+\alpha+\beta}{2+\beta} + \frac{3+\beta}{2\;(2+\beta)} t \pm \frac{1}{2\;(2+\beta)} \sqrt{(2\beta^2 + 10\beta + 13)t^2 - 4(1-\alpha)t + 4(1-\alpha)^2} \tag{3.46}$$

Thus, there are two candidate solutions as a symmetric Nash equilibrium. These candidate solutions were derived assuming that the solution satisfies the meaningful demand property. To be valid, these candidate solutions need to be consistent with meaningful demand, i.e. the price needs to be such that at all x someone buys the good. To check meaningful demand, it is sufficient to check that at $x = \frac{1}{2}$ someone buys:

$$p + T\left(\frac{1}{2}\right) < 1 \tag{3.47}$$

Using the negative square-root solution, the meaningful demand property holds for all:

$$t < \frac{4\;(1-\alpha)}{3+\beta} \tag{3.48}$$

The positive square-root solution never satisfies meaningful demand.

Thus, the negative square root solution, i.e. equation (3.10), is the unique candidate symmetric Nash equilibrium. To establish that it actually is a Nash equilibrium, we check the firm's second order condition. To derive $\frac{d^2\pi_L}{dp_L^2}$, start from a version of $\frac{d\pi_L}{dp_L}$ where we used that $\frac{dm}{dp_L} = -\frac{1}{2t}$, but did not use conditions which only hold in equilibrium i.e. that $p_L = p_R = p$ or respectively $m = \frac{1}{2}$:

$$\begin{aligned}\frac{d\pi_L}{dp_L} =& -\frac{1}{2}m(p_L,p_R)\big(\alpha+\beta-(1+\beta)p_L\big) - \frac{1+\beta}{2}tm(p_L,p_R)^2 + \frac{1}{4}\beta t m(p_L,p_R)^2 \\ &- \frac{1}{2t}\big((1+\alpha+\beta)p_L - (1+\frac{\beta}{2})p_L^2 - \alpha - \frac{1}{2}\beta\big) + m(p_L,p_R)\,\big(1+\alpha+\beta-(2+\beta)p_L\big)\end{aligned} \tag{3.49}$$

From (3.49) we can calculate the second derivative

$$\begin{aligned}\frac{d^2\pi_L}{dp_L^2} =& -\frac{1}{2}\frac{dm}{dp_L}\big(\alpha+\beta-(1+\beta)p_L\big) + \frac{1}{2}m(p_L,p_R)(1+\beta) - (1+\beta)\,tm(p_L,p_R)\,\frac{dm}{dp_L} \\ &+ \frac{1}{2}\,\beta\,t\,m(p_L,p_R)\,\frac{dm}{dp_L} \\ &- \frac{1}{2t}\big((1+\alpha+\beta)-(2+\beta)p_L\big) + \frac{dm}{dp_L}\big(1+\alpha+\beta-(2+\beta)p_L\big) - m(p_L,p_R)(2+\beta)\end{aligned} \tag{3.50}$$

which using that $\frac{dm}{dp_L} = -\frac{1}{2t}$ and that $p_L = p_R = p$ and thus $m = \frac{1}{2}$ simplifies to:

$$\frac{d^2\pi_L}{dp_L^2} = \frac{1}{4t}(-4-3\alpha-3\beta-2t-\frac{1}{2}\beta t+7p+3\beta p) \tag{3.51}$$

The second order condition is that $\frac{d^2\pi_L}{dp_L^2} < 0$ which equals

$$p < \frac{4+3\alpha+3\beta}{7+3\beta} + \frac{2+\frac{1}{2}\beta}{7+3\beta}\,t \tag{3.52}$$

which holds at $t = 0$ and for all other $t < \frac{4\,(1-\alpha)}{3+\beta}$.

Thus, there exists a unique symmetric Nash equilibrium. In this equilibrium, the price is as given in equation (3.10):

$$p_G^C = \frac{1+\alpha+\beta}{2+\beta} + \frac{3+\beta}{2\,(2+\beta)}t - \frac{1}{2\,(2+\beta)}\sqrt{(2\beta^2+10\beta+13)t^2 - 4(1-\alpha)t + 4(1-\alpha)^2}$$

q.e.d.

3.10.4 Proof of Corollary 3.1

Corollary 3.1 states that $0 < \frac{dp_G^C}{d\alpha} < 1$.

Since for all $t < \frac{4(1-\alpha)}{3+\beta}$ the equilibrium price is:

$$p_G^C = \frac{1+\alpha+\beta}{2+\beta} + \frac{3+\beta}{2\,(2+\beta)}t - \frac{1}{2\,(2+\beta)}\sqrt{(2\beta^2+10\beta+13)t^2 - 4(1-\alpha)t + 4(1-\alpha)^2}$$

it follows that

$$\frac{dp_G^C}{d\alpha} = \frac{1}{2+\beta} - \frac{t-2(1-\alpha)}{(2+\beta)\sqrt{(2\beta^2+10\beta+13)t^2 - 4(1-\alpha)t + 4(1-\alpha)^2}} \tag{3.53}$$

Lemma 3.10.4.1 $\frac{dp_G^C}{d\alpha} > 0$ *provided that* $t < \frac{4(1-\alpha)}{3+\beta}$.

Proof: $\frac{dp_G^C}{d\alpha} > 0$ is equivalent to

$$\sqrt{(2\beta^2+10\beta+13)t^2 - 4(1-\alpha)t + 4(1-\alpha)^2} > t - 2(1-\alpha) \tag{3.54}$$

Distinguish between two cases. In case 1, $t - 2(1-\alpha) < 0$, or equally $t < 2(1-\alpha)$. Then the inequality (3.54) always holds. In case 2, $t - 2(1-\alpha) > 0$. Then (3.54) simplifies to

$$(\beta+2)\,(\beta+3)\,t^2 > 0 \tag{3.55}$$

which always holds since we have assumed that $\beta \geq 0$ and $t > 0$. Thus, $\frac{dp_G^C}{d\alpha} > 0$ always holds.

q.e.d.

Lemma 3.10.4.2 $\frac{dp_G^C}{d\alpha} < 1$ *provided that* $t < \frac{4(1-\alpha)}{3+\beta}$.

Proof: $\frac{dp_G^C}{d\alpha} < 1$ is equivalent to

$$\sqrt{(2\beta^2+10\beta+13)t^2 - 4(1-\alpha)t + 4(1-\alpha)^2} > \frac{2(1-\alpha)-t}{1+\beta} \tag{3.56}$$

Distinguish between two cases. In case 1, $\frac{2(1-\alpha)-t}{1+\beta} < 0$, or equally $t > 2(1-\alpha)$. Then the inequality (3.56) always holds. In case 2, $\frac{2(1-\alpha)-t}{1+\beta} > 0$. Then (3.56) becomes

$$(2\beta^4+14\beta^3+35\beta^2+36\beta+12)\,t^2 \; -4(1-\alpha)\beta(1+\beta)\,t \; +4(1-\alpha)^2\beta(1+\beta) > 0 \tag{3.57}$$

which holds at $t = 0$. Moreover, for any β there does not exist a real t at which (3.57) holds with equality. Thus, $\frac{dp_G^C}{d\alpha} < 1$ always holds. *q.e.d.*

Proof of Corollary 3.1: Lemma 3.10.4.1 and Lemma 3.10.4.2 combined prove Corollary 3.1. *q.e.d.*

3.10.5 Proof of Proposition 3.4

Part i: *There exists a unique level of transport costs, denoted $\tilde{t}$, at which the oligopoly equilibrium price $p_G^C(t)$ and the socially optimal* price $p_G^S(t)$ *coincide, i.e. $p_G^S(\tilde{t}) = p_G^C(\tilde{t})$.*

Lemma 3.10.5.1 $p_G^S(0) > p_G^C(0)$.

Proof: $p_G^S(0) = \frac{\alpha+\beta}{1+\beta}$; $p_G^C(0) = \frac{2\alpha+\beta}{2+\beta}$. Thus, $p_G^S(0) > p_G^C(0)$ is equivalent to

$$(1-\alpha)\,\beta > 0 \tag{3.58}$$

which is true under advantageous selection ($\beta > 0$) since by the meaningful demand property we must have $\alpha < 1$. *q.e.d.*

Lemma 3.10.5.2 $p_G^S(0) < p_G^C(\frac{4(1-\alpha)}{3+\beta})$.

Proof:

$$p_G^C(\frac{4(1-\alpha)}{3+\beta}) = \frac{1+\alpha+\beta}{2+\beta} + \frac{2\,(1-\alpha)}{2+\beta} - \frac{(1-\alpha)}{(2+\beta)\,(3+\beta)}(3\beta+7) \tag{3.59}$$

Thus, $p_G^S(0) < p_G^C(\frac{4(1-\alpha)}{3+\beta})$ simplifies to

$$(\beta+2)\,(1-\beta) > 0 \tag{3.60}$$

which holds provided that $\beta < 1$. *q.e.d.*

Lemma 3.10.5.3 $p_G^S(t)$ *is decreasing in* t.

Proof: This statement follows directly from

$$p_G^S(t) = \frac{\alpha+\beta}{1+\beta} - \frac{\beta}{4\,(1+\beta)}\,t \tag{3.61}$$

q.e.d.

Lemma 3.10.5.3 implies that $p_G^S(\frac{4(1-\alpha)}{3+\beta}) < p_G^S(0)$. Thus, combined with Lemma 3.10.5.2, we have that $p_G^S(\frac{4(1-\alpha)}{3+\beta}) < p_G^C(\frac{4(1-\alpha)}{3+\beta})$.

Lemma 3.10.5.4 $p_G^C(t)$ *is strictly concave.*

Proof: Since

$$p = \frac{1+\alpha+\beta}{2+\beta} + \frac{3+\beta}{2\,(2+\beta)}t - \frac{1}{2\,(2+\beta)}\sqrt{(2\beta^2+10\beta+13)t^2 - 4(1-\alpha)t + 4(1-\alpha)^2}$$

we can calculate that

$$\frac{dp}{dt} = \frac{3+\beta}{2\,(2+\beta)} + \frac{2(1-\alpha)\;-\;(2\beta^2+10\beta+13)\,t}{2\,(2+\beta)\,\sqrt{(2\beta^2+10\beta+13)t^2 - 4(1-\alpha)t + 4(1-\alpha)^2}} \tag{3.62}$$

and defining $\phi = (2\beta^2 + 10\beta + 13)t^2 - 4(1-\alpha)t + 4(1-\alpha)^2$ we have that

$$\frac{d^2p}{dt^2} = \frac{-(2\beta^2 + 10\beta + 13)\ \phi + \Big[(2\beta^2 + 10\beta + 13)\ t - 2\ (1-\alpha)\Big]^2}{2\ (2+\beta)\ \phi^{\frac{3}{2}}} \tag{3.63}$$

Then $\frac{d^2p}{dt^2} < 0$ is equivalent to

$$\Big[(2\beta^2 + 10\beta + 13)t - 2(1-\alpha)\Big]^2 < (2\beta^2 + 10\beta + 13)\ \phi \tag{3.64}$$

or equally

$$0 < 2\beta^2 + 10\beta + 13 \tag{3.65}$$

which holds for any $\beta \geq 0$. *q.e.d.*

Proof of Proposition 3.4, Part i: Lemma 3.10.5.1 - 3.10.5.4 prove that there exists a unique crossing of $p_G^S(t)$ and $p_G^C(t)$. Figure 3.7 provides the visual intuition. *q.e.d.*

Part ii: *The oligopoly equilibrium allocation and the socially optimal allocation do not coincide for any level of transport costs t.*

Proof of Proposition 3.4, Part ii: This follows from proposition 3.1. In markets with advantageous selection, no gate price can implement the efficient allocation. Even $p_G^S(t)$ only results in the constraint efficient allocation. Thus, at $\tilde{t}$, we have $p_G^C(\tilde{t}) = p_G^S(\tilde{t})$, i.e. the duopoly equilibrium results in the constraint efficient allocation but by proposition 3.1 it does not result in the first best efficient allocation. *q.e.d.*

3.10.6 Proof of Proposition 3.5

The proof proceeds in steps. First, we characterise the price firm L would charge if firm R did not exist and term this monopolist pricing. Second, we show at which locations x firm R can profitably undercut L's monopolist prices. Then, we show how firm L responds and that the resulting equilibrium behaviour is as described in proposition 3.5.

For this proof, we use that $c(y) = \alpha + \beta(1-y)$ where $\beta > 0$, $T(d)$ where $\frac{dT}{dd} > 0$ and $T(0) = 0$, $F(y) = y$. $G(x)$ can be any CDF.

Step 1: Firm L's pricing if firm R does not exist.

At a location x, firm L's profit is

$$\pi_L(x) = \int_{p_L(x)+T(x)}^{1} [p_L(x) - c(y)]\ f(y)\ dy \tag{3.66}$$

$$\pi_L(x) = p_L - p_L^2 - p_L T(x) - \Big(\alpha + \frac{\beta}{2} - (\alpha+\beta)p_L - (\alpha+\beta)T(x) + \frac{\beta}{2}p_L^2 + \beta p_L T(x) + \frac{\beta}{2}T(x)^2\Big) \tag{3.67}$$

which by solving $\frac{d\pi_L}{dp_L} = 0$ yields

$$p_L^M(x) = \frac{1+\alpha+\beta}{2+\beta} - \frac{1+\beta}{2+\beta}\, T(x) \tag{3.68}$$

which is the unique profit maximum since $\frac{d^2\pi_L}{dp_L^2} = -(2+\beta) < 0$.

Step 2: At which locations x can firm R profitably undercut the monopolist price $p_L^M(x)$

We introduce two auxiliary concepts. $p_R^U(x)$ is the maximum undercutting price of firm R, i.e. the highest price firm R can set and still attract consumers. Formally,

$$p_R^U(x) = p_L^M(x) + T(x) - T(1-x) \tag{3.69}$$

Using (3.68) this becomes

$$p_R^U(x) = \frac{1+\alpha+\beta}{2+\beta} + \frac{1}{2+\beta}T(x) - T(1-x) \tag{3.70}$$

As second auxiliary concept, $p_R^0(x)$ is the price firm R needs to charge to make zero profit. This is the lower possible price firm R will ever offer. Solving the corresponding zero-profit condition $\pi_R(p_R^0(x), x) = 0$ results in

$$p_R^0(x) = \frac{2\alpha+\beta}{2+\beta} - \frac{\beta}{2+\beta}\, T(1-x) \tag{3.71}$$

Firm R undercuts $p_L^M(x)$ if and only if[42]

$$p_R^U(x) \geq p_R^0(x) \tag{3.72}$$

which simplifies to

$$1 - \alpha + T(x) \;\geq\; 2\, T(1-x) \tag{3.73}$$

Since, under the assumptions made, profit functions are well behaved, there exists a unique threshold x_L defined by

$$1 - \alpha + T(x_L) \;=\; 2\, T(1-x_L) \tag{3.74}$$

For all $x \leq x_L$ firm R does not undercut while for $x > x_L$ firm R does undercut.

Thus, $x_L = 0$ iff $1 - \alpha \geq 2T(1)$, i.e. transport cost are low, and otherwise x_L is the solution to $1 - \alpha + T(x_L) \;=\; 2\, T(1-x_L)$.

Step 3: Firm L's response to being undercut.

At all $x \in [x_L, \frac{1}{2}]$, firm L will respond to R undercutting its price by in turn undercutting R. Firms continue to undercut each other until a price is reached at which firm R no longer finds it profitable to undercut. Since for $x < \frac{1}{2}$, $T(x) < T(1-x)$, it follows that at

[42] If R undercuts, R uses the highest possible price which attracts consumers.

all $x \in [x_L, \frac{1}{2}]$ the equilibrium price is:

$$p_L(x) = p_R^0(x) + T(1-x) - T(x) \tag{3.75}$$

equally

$$p_L(x) = \frac{2\alpha+\beta}{2+\beta} + \frac{2}{2+\beta}T(1-x) - T(x) \tag{3.76}$$

This proves the equilibrium described in Proposition 3.5. *q.e.d.*

3.10.7 Proof of Corollary 3.2

Part i: *The equilibrium price $p_L(x)$ is monotonically decreasing in distance from the firm.*

Proof of Part i:

When $x \in [0, x_L]$: $\frac{dp_L^M}{dx} = -\frac{1+\beta}{2+\beta}\frac{dT(x)}{dd} < 0 \; \forall x$.

When $x]x_L, \frac{1}{2}]$: $\frac{dp_L}{dx} = -\frac{2}{2+\beta}\frac{dT(1-x)}{dd} - \frac{dT(x)}{dd} < 0 \; \forall \; x$. *q.e.d.*

Part ii: *The number of consumers purchasing the good can be non-monotone in distance from the firm.*

Proof of Part ii: We prove this by focusing on $p(x) + T(x)$, i.e. the total cost a consumer faces. The larger $p(x) + T(x)$, the fewer consumers purchase the good.

When $x \in [0, x_L]$, the total cost consumers face is:

$$p_L^M(x) + T(x) = \frac{1+\alpha+\beta}{2+\beta} - \frac{1+\beta}{2+\beta}T(x) + T(x) \tag{3.77}$$

$$= \frac{1+\alpha+\beta}{2+\beta} + \frac{1}{2+\beta}T(x) \tag{3.78}$$

Thus, defining total cost consumers face as $\bar{p} = p_L^M(x) + T(x)$:

$$\frac{d\bar{p}}{dx} = \frac{1}{2+\beta}\frac{dT}{dd} > 0 \; \forall \; x \tag{3.79}$$

When $x \in]x_L, \frac{1}{2}]$, the total cost consumers face is:

$$p_L(x) + T(x) = \frac{2\alpha+\beta}{2+\beta} + \frac{2}{2+\beta}\,T(1-x) \tag{3.80}$$

Thus, defining total cost consumers face as $\bar{p} = p_L(x) + T(x)$:

$$\frac{d\bar{p}}{dx} = \frac{2}{2+\beta}\Big(-\frac{dT}{dd}\Big) < 0 \; \forall \; x \tag{3.81}$$

Since for $x \in [0, x_L]$, $\frac{d\bar{p}}{dx} > 0$ and for $x \in]x_L, \frac{1}{2}]$, $\frac{d\bar{p}}{dx} < 0$, it follows that $\bar{p}$ is non-monotone in x when $x_L > 0$. We know from Proposition 3.5 that for $T(1) > \frac{1-\alpha}{2}$, $x_L > 0$ and then

the non-monotonicity arises. If $T(1) \leq \frac{1-\alpha}{2}$, then $x_L = 0$ and the non-monotonicity does not arise. This proves that the non-monotonicity is a possibility, not a certainty, and therefore proves Corollary 3.2. *q.e.d.*

3.10.8 Proof of Proposition 3.6

Part 1: *In the absence of selection, the duopoly equilibrium exhibits inefficient underprovision at all* $x < \frac{1}{2}$.

Part 2: *In the absence of selection, the duopoly equilibrium achieves the efficient allocation at* $x = \frac{1}{2}$.

In the no selection case, denote costs as $c(y) = \alpha \; \forall \; y$. We make no additional restrictions on $T(d)$ or $G(x)$ and assume that $F(y) = y$. For $x \in [0, x_L]$, the equilibrium price is $p_L^M(x) = \frac{1+\alpha-T(x)}{2}$ where x_L is the solution to $T(x_L) = 2T(1-x_L) - 1 + \alpha$. This is provided that $T(1) > \frac{1-\alpha}{2}$. If $T(1) < \frac{1-\alpha}{2}$, then $x_L = 0$. For $x \in [x_L, \frac{1}{2}]$, $p_L(x) = \alpha + T(1-x) - T(x)$.

Proof of Part 2: To show: $p_L(x) = \alpha$ at $x = \frac{1}{2} \; \forall \; T(d)$.

Substituting $x = \frac{1}{2}$ into $p_L(x) = \alpha + T(1-x) + T(x)$ yields:

$$p_L(x) = \alpha + T(\frac{1}{2}) - T(\frac{1}{2}) = \alpha \tag{3.82}$$

q.e.d.

Proof of Part 1: To show: $p_L^M(x) > \alpha \; \forall \; x \leq x_L$ and $p_L(x) > \alpha \; \forall \; x \in]x_L, \frac{1}{2}[$.

$$p_L^M(x) = \frac{1 + \alpha - T(x)}{2} > \alpha \tag{3.83}$$

solves to $1 > \alpha + T(x)$, which holds for all x by the meaningful demand property.

$$p_L(x) = \alpha + T(1-x) - T(x) > \alpha \tag{3.84}$$

solves to $T(1-x) > T(x)$ which is true for all $x < \frac{1}{2}$. *q.e.d.*

3.10.9 Proof of Proposition 3.7

In the efficient allocation, there is a WTP threshold $y^S(x)$ above which all consumers are allocated the good. The threshold is characterised by

$$y^S(x) = T(x) + c\big(y^s(x)\big) \tag{3.85}$$

which solves to

$$y^S(x) = \frac{\alpha + \beta}{1 + \beta} + \frac{1}{1 + \beta} \, T(x) \tag{3.86}$$

There is inefficient overprovision at a location x if and only if

$$p_L(x) + T(x) < y^S(x) \tag{3.87}$$

where we can focus on prices for $x > x_L$ since $x < x_L$ results in monopoly pricing which always means inefficient underprovision. Thus, (3.87) becomes

$$\frac{2\alpha+\beta}{2+\beta} + \frac{2}{2+\beta}\,T(1-x) < \frac{\alpha+\beta}{1+\beta} + \frac{1}{1+\beta}\,T(x) \tag{3.88}$$

which simplifies to

$$\beta\,(1-\alpha) > (2+2\beta)\,T(1-x) - (2+\beta)\,T(x) \tag{3.89}$$

For larger x, the left hand side of this inequality is constant while the right hand side decreases. Thus, there exists at most one crossing. Hence, there are two cases. If $T(1) < \frac{\beta}{(2+2\beta)}(1-\alpha)$, then there is inefficient overprovision at $x = 0$ and thus at all x. If $T(1) > \frac{\beta}{(2+2\beta)}(1-\alpha)$, then there exists a unique x_e defined by solving

$$\beta(1-\alpha) = (2+2\beta)\,T(1-x_e) - (2+\beta)T(x_e) \tag{3.90}$$

such that for all $x < x_e$ there is inefficient underprovision and for all $x > x_e$ there is inefficient overprovision. *q.e.d.*

3.10.10 Proof of Proposition 3.8

The two parts of Proposition 3.8 are related. We prove both parts simultaneously by showing that, in markets with advantageous selection, there is always inefficient overprovision at $x = \frac{1}{2}$, i.e. for all levels of transport costs.

Under the assumptions used to show this proposition, i.e. $c(y) = \alpha + \beta(1-y)$ where $\beta > 0$, $T(d)$ where $\frac{dT}{dd} > 0\ \forall\ d$ and $T(0) = 0$, $F(y) = y$, but $G(x)$ can be any CDF, the efficient allocation is characterised by the WTP threshold $y^S(x)$ which solves

$$y^S(x) = T(x) + c\big(y^S(x)\big) \tag{3.91}$$

which simplifies to

$$y^S(x) = \frac{\alpha+\beta}{1+\beta} + \frac{1}{1+\beta}T(x) \tag{3.92}$$

In the oligopoly equilibrium, the WTP threshold above which consumers buy is $\bar{p}(x) = p(x) + T(x)$. At $x = \frac{1}{2}$ and using the relevant price under effective competition $p_L(x)$, not $p_L^M(x)$:

$$\bar{p}\Big(\frac{1}{2}\Big) = p_L\Big(\frac{1}{2}\Big) + T\Big(\frac{1}{2}\Big) \tag{3.93}$$

$$\bar{p}\Big(\frac{1}{2}\Big) = \frac{2\alpha+\beta}{2+\beta} + \frac{2}{2+\beta}\,T\Big(\frac{1}{2}\Big) \tag{3.94}$$

There is inefficient overprovision if and only if

$$\bar{p}\Big(\frac{1}{2}\Big) < y^S\Big(\frac{1}{2}\Big) \tag{3.95}$$

$$\frac{2\alpha+\beta}{2+\beta} + \frac{2}{2+\beta}\,T\Big(\frac{1}{2}\Big) < \frac{\alpha+\beta}{1+\beta} + \frac{1}{1+\beta}\,T\Big(\frac{1}{2}\Big) \tag{3.96}$$

which simplifies to

$$\beta\left[1-\alpha-T\Big(\frac{1}{2}\Big)\right] > 0 \tag{3.97}$$

which is always true since by the meaningful demand property $1-\alpha-T\Big(\frac{1}{2}\Big) > 0$. Thus, provided there is any, possibly very small, degree of advantageous selection ($\beta > 0$), there is inefficient overprovision at $x = \frac{1}{2}$. This overprovision is not removed by any level of transport costs. *q.e.d.*

3.10.11 Proof of Proposition 3.9

We know that, in markets with advantageous selection, there is inefficient underprovision if $t = 0$, i.e. $p^S(0) > p^C(0)$. This holds when firms use gate prices (Proposition 3.4) and when they price discriminate (Proposition 3.7). In both cases, the socially optimal allocation and the equilibrium allocation are characterised by a threshold WTP above which consumers are allocated the good and the threshold is independent of x, i.e. is perfectly horizontal. The level of the threshold differs between the socially optimal allocation and the equilibrium.

Since taxes shift the equilibrium threshold without changing its slope (analogous to imperfect cost pass through described in Corollary 3.1), there exists a tax rate for which the equilibrium allocation and the socially optimal allocation coincide. *q.e.d.*

Example: Let $c(y) = \alpha + \beta(1-y)$, $F(y) = y$, $G(x) = x$, then there is inefficient overprovision under perfect competition, i.e.

$$p_G^S(0) = \frac{\alpha+\beta}{1+\beta} > p_G^C(0) = \frac{2\,\alpha+\beta}{2+\beta}$$

Since a per unit tax (τ) results in equilibrium prices $p_G^C(0) = \frac{2\,(\alpha+\tau)+\beta}{2+\beta}$, we can solve for

the tax rate which achieves the first-best efficient allocation:[43]

$$\frac{2(\alpha+\tau)+\beta}{2+\beta} = \frac{\alpha+\beta}{1+\beta} \tag{3.98}$$

which solves to

$$\tau^* = \frac{\beta\,(1-\alpha)}{2\,(1+\beta)}$$

3.10.12 Proof of Corollary 3.3

We established in Proposition 3.7 that inefficient overprovision occurs at all $x \in [x_e, \frac{1}{2}]$ where x_e is the solution to $\beta(1-\alpha) = (2+2\beta)\,T(1-x_e) - (2+\beta)\,T(x_e)$ provided that $T(1) \geq \frac{\beta}{2+2\beta}(1-\alpha)$, and otherwise $x_e = 0$.

Lemma 3.10.12.1 $\frac{dx_e}{d\alpha} > 0$.

Proof:

$$\beta(1-\alpha) = (2+2\beta)\,T(1-x_e) - (2+\beta)\,T(x_e) \tag{3.99}$$

totally differentiate

$$-\beta\,d\alpha = (2+2\beta)\,\frac{dT(1-x_e)}{dd}\,(-dx_e) - (2+\beta)\,\frac{dT(x_e)}{dd}dx_e \tag{3.100}$$

which solves to

$$\frac{dx_e}{d\alpha} = \frac{\beta}{(2+2\beta)\frac{dT(1-x_e)}{dd} + (2+\beta)\frac{dT(x_e)}{dd}} > 0\ \forall\ x \tag{3.101}$$

q.e.d.

Lemma 3.10.12.2 $\frac{dx_e}{d\beta} < 0$.

Proof:

$$\beta(1-\alpha) = (2+2\beta)\,T(1-x_e) - (2+\beta)\,T(x_e) \tag{3.102}$$

totally differentiate

$$(1-\alpha)\,d\beta = (2+2\beta)\,\frac{dT(1-x_e)}{dd}\,(-dx_e) + 2\,T(1-x_e)\,d\beta - (2+\beta)\,\frac{dT(x_e)}{dd}\,dx_e - T(x_e)\,d\beta \tag{3.103}$$

which solves to

$$\frac{dx_e}{d\beta} = -\frac{1-\alpha+T(x_e)-2T(1-x_e)}{(2+2\beta)\frac{dT(1-x_e)}{dd} + (2+\beta)\frac{dT(x_e)}{dd}} \tag{3.104}$$

[43] More generally, for any transport cost t, there exists a tax rate or subsidy rate which achieves the constrained efficient allocation, i.e. for which $p_G^S(t) = p_G^C(t)$. This tax rate is implicitly defined by $p_G^S(\alpha,\beta,t) = p_G^C(\alpha+\tau,\beta,t)$ or equally by $\frac{\alpha+\beta}{1+\beta} - \frac{\beta}{4\,(1+\beta)}t = \frac{1+\alpha+\tau+\beta}{2+\beta} + \frac{3+\beta}{2(2+\beta)}t - \frac{1}{2(2+\beta)}\sqrt{(2\beta^2+10\beta+13)t^2 - 4(1-\alpha-\tau)t + 4(1-\alpha-\tau)^2}$. However, the first best efficient allocation is only achieved under perfect competition ($t = 0$) combined with tax rate (3.16). Moreover, all cases with $t > 0$ are further complicated by considerations whether firms use gate prices or price discrimination and by whether the tax rate has to apply equally to all x or can discriminate based on x.

where $(2+2\beta)\frac{dT(1-x_e)}{dd} + (2+\beta)\frac{dT(x_e)}{dd} > 0 \; \forall \; x$ since $\frac{dT}{dd} > 0$.

We now show that $1-\alpha+T(x_e)-2T(1-x_e) > 0$, which then also proves that $\frac{dx_e}{d\beta} < 0 \; \forall \; x$. To show that $1-\alpha+T(x_e)-2T(1-x_e) > 0$, rewrite it as:

$$1-\alpha+T(x_e) > 2T(1-x_e) \tag{3.105}$$

Recall that x_L is the border between monopoly pricing and effective competition and that overprovision can only arise under effective competition, never under monopoly. Thus, we must have $x_e > x_L$. Recall that x_L was defined by

$$1-\alpha+T(x_L) = 2T(1-x_L)$$

Since $1-\alpha+T(x)$ is increasing in x and $2T(1-x)$ is decreasing in x, for all $x > x_L$ we must have that

$$1-\alpha+T(x) > 2\,T(1-x) \tag{3.106}$$

and since $x_e > x_L$, we must have

$$1-\alpha+T(x_e) > 2\,T(1-x_e) \tag{3.107}$$

Thus, we have shown that $\frac{dx_e}{d\beta} < 0 \; \forall \; x$. *q.e.d.*

Since $\frac{dx_e}{d\beta}$ captures both a pure cost effect and a selection effect and $\frac{dx_e}{d\alpha}$ captures only a pure cost effect, we interpret the results that $\frac{dx_e}{d\alpha} > 0$ but $\frac{dx_e}{d\beta} < 0$ as indicating that the selection effect has the opposite sign of the pure cost effect and more than outweighs it. *q.e.d.*

Chapter 4

Tying Contracts in Selection Markets: A Solution to the Unravelling Problem

This paper shows that, in markets with adverse selection, tying contracts can prevent unravelling and thus have a substantial positive effect on welfare. We illustrate this effect for two insurance markets. When one characteristic (such as life-expectancy) is hidden information and causes agents to have high expected cost for one product (life insurance) but low expected cost for another (pension annuity), then for tying contracts cost differences partially cancel out. Thus, cost heterogeneity is lower for the tying contract than for individual insurance contracts. This can prevent unravelling. The largest welfare benefit occurs if without tying contracts both individual markets unravel but with tying contracts all agents are insured against both risks.

Keywords: *Selection Markets, Insurance, Tying Contracts*

JEL Classification: D82, L11, G22

4.1 Introduction

We know that the presence of asymmetric information can substantially alter market outcomes. The seminal work by Akerlof (1970) and Rothschild and Stiglitz (1976) points out that markets which exhibit adverse selection can unravel, i.e. break down completely. The market for pension annuities is a good example.[1] An annuity insures agents against outliving their retirement savings. Agents differ in longevity, which is private information. It is possible that if firms price insurance at the average cost calculated across the entire population, agents with low expected cost do not want to purchase insurance. Only high cost, i.e. high longevity types, purchase the insurance policy. Firms would make a loss and raise prices to at least break even. In equilibrium, only the highest cost agents are insured. The market has "unravelled", i.e. some beneficial trades, such as insuring low longevity agents, have not occurred due to the presence of asymmetric information.[2] Thus, total surplus is lower than under perfect information.

While the effect of adverse selection in one market is understood, and is empirically well documented, less is know about how these results are altered when agents interact in multiple markets with adverse selection. Empirical evidence on unravelling includes recent developments in the annuity market in the UK[3] and US health insurance markets (Cutler and Zeckhauser, 1997).

This paper studies how the presence of a second related selection market creates possibilities to prevent unravelling in both markets. We show that tying contracts, which combine the sale of the products traded in both markets,[4] can prevent unravelling and thus have a large positive effect on welfare. The largest welfare benefit occurs if without tying contracts both individual markets unravel, but with tying contract all agents are insured against both risks. This effect arises because for tying contracts cost differences partially cancel out.

To establish the possibility result that tying contracts can prevent unravelling, we focus on the case where agents who have high expected cost in one market have low cost in the other, and vice versa. For example, consumers with a low life-expectancy have high expected cost for life-insurance but low expected cost for pension annuities, i.e. costs are negatively correlated across markets. Tying contracts combine life-insurance features, i.e.

[1] Other examples include the market for health insurance, for life-insurance, bank credit markets or the market for used cars. Pension annuities are not the most frequently used example in the literature for selection markets, but allow us to stay with one example throughout this paper.

[2] The term "unravelling" has been used with different meanings in the literature. It can refer to a complete market breakdown where nobody purchases insurance or to the case where only the agents with the highest cost purchase insurance. We use unravelling with the latter meaning. As in our model all agents have a willingness-to-pay above their expected cost, the highest degree of unravelling possible is that only high cost agents buy.

[3] Following a reform which made the conversion of pension savings to an annuity optional annuitisation rates fell from 90 percent to 12 percent (Sturrock, 2019).

[4] The concept of one contract specifying the sale of two goods has several names in the literature. While we use the term "tying", other authors prefer "bundling", or see it as a form of "cross-selling". Other papers using the term "tying" include Whinston (1990) and in a banking context Loranth and Morrison (2012) who study the desirability of banks tying lending and investment banking services.

pay out if the insured person dies, with pension features, i.e. continuing payments in retirement until death. Such contracts exist in Germany[5] and in the UK some pension plans have life insurance features.[6]

Example: Let there be two equally large groups of consumers, who differ in their life-expectancy (L = low, H = high), which is private information. Suppose that in one market (life-insurance) expected costs are decreasing in life-expectancy ($c(L) = 2, c(H) = 1$) and in the other market (pension annuity) increasing ($C(L) = 1, C(H) = 2$). Denote the average cost across the population as $E(c) = E(C) = 1.5$ respectively. Assume that all agents are risk averse and willing to pay a premium of 0.2 above their expected cost. We focus on the case of perfect competition among insurers.

When firms offer only stand alone life-insurance and pension contracts, markets unravel since $E(c) > 1.2$. If firms price at the average cost in the population $p = E(c)$, the low cost types do not want to purchase the insurance contract. Thus, firms would make a loss and have to increase the price to cover the cost of those who actually insure at $p = c(L) = 2$. In equilibrium only the highest cost agents are insured paying price 2. Not all surplus generating trades occur as types with low expected cost remain uninsured.

When firms offer tying contracts in addition to stand-alone contracts, tying contracts can be offered at a price which equals the average cost in the population $p = E(c) + E(C) = 3$. All agents are willing to buy the tying contract at this price as their willingness to pay (3.4) exceeds their expected cost (3) at which the tying contract is priced. While there is heterogeneity in costs for each individual contract, there is no heterogeneity in costs for the tying contract. Without heterogeneity in costs, the private information on life-expectancy is no longer relevant to insurance companies and thus the market does not unravel. In this equilibrium, which is unique, everyone is insured against both risks by buying a tying contract at price 3. Thus, total surplus is higher with tying contracts than without.

Generalisations: The beneficial effect of tying contracts is not limited to the special case above, where cost differences cancel out exactly. In this paper, we develop a more general model with more than two types of life-expectancy and only partial cancelling out of cost differences. We find that the socially beneficial effect of tying contracts generalises. A second channel contributes to this: tying contracts can reverse selection patterns. Let there be three types of life-expectancy. When firms are restricted to offering stand-alone contracts, there are cases where the market for each stand alone contract unravels completely. That means only the types with the highest expected cost are insured against the corresponding risk. In such an equilibrium, the medium life-expectancy types are not insured against either risk. When costs are concave in life-expectancy, these medium types are the highest cost type for the tying contract and will thus always buy the tying contract and be insured against both risks. Tying contracts thus rule out the equilibria with the lowest total surplus, i.e. where medium types are completely uninsured.

[5] They are known as "Kapitallebensversicherung".

[6] One example is the pension scheme of university employees in the UK (University Superannuation Scheme, USS).

The results extend to settings where agents have multiple dimensions of private information. We take up the recent theme in the selection market literature of allowing agents to differ not just in expected cost (life-expectancy) but also in risk aversion, with one or both being private information (Finkelstein and McGarry, 2006). We show that total surplus is higher when firms can write tying contracts, regardless of whether heterogeneity in risk preferences exists and whether risk preferences are public or private information. Moreover, we find that holding the contract space fixed, total surplus is lower when risk preferences are private rather than public information.

Throughout, we maintain the assumption that costs are negatively correlated. This is analytically convenient and allows us to derive clear possibility results. It also matches our motivating example of life-insurance and pension annuities. In these markets, tying contracts exist in some countries.[7] Our model suggests that these contracts are socially valuable without having to assume cost synergies of tying contracts such as reduced administrative costs.

Distributional effect: While introducing tying contracts increases total surplus at least weakly, there can be distributional effects and some agents can be worse off. Individuals are worse off when, following the introduction of tying contracts, they continue to be insured against only one risk but the price of this insurance has increased as other lower cost types have shifted to the tying contract.

This paper proceeds as follows. Section 4.2 reviews the related literature. Section 4.3 outlines the model of two related selection markets. Section 4.4 characterises the equilibrium, first assuming that only stand-alone contracts can be offered, and then considering the case where firms can offer both stand-alone and tying contracts. Section 4.5 derives our main results. In particular, this includes the possibility result that tying contracts can prevent unravelling and the proposition that tying contracts increase total surplus at least weakly and for some parameters strictly. Section 4.6 concludes.

4.2 Literature Review

Our main result, which states that tying contracts can solve the unravelling problem, builds on the seminal selection literature which pioneered the unravelling result. We briefly review this literature which focuses on private information on one dimension: expected costs. More recent advances in the selection literature, including papers on monopolists' use of tying contracts, have assumed private information on two dimensions: on expected cost and on risk preferences.

[7] See the earlier example on Germany and the UK.

4.2.1 One-dimensional private information

The seminal work which introduced asymmetric information to economics focused on *one* dimension of private information. This applies to both strands of the literature in information economics: Models with private information about a hidden fixed characteristic such as the quality of a used car in Akerlof's (1970) market for lemons, ability in Spence's (1973) signalling model or health risks in Rothschild and Stiglitz's (1976) model of heath insurance; Models with private information about a hidden action known as moral hazard models e.g. of incentive pay in labour markets (Holmstrom, 1979; Grossman and Hart, 1983). Our paper focuses on the first class of models, specifically on models of selection markets. Einav and Finkelstein (2011) provide a good introduction.

This literature has proven extremely influential: It won several Nobel Prizes, became a central part of the undergraduate syllabus, and shaped policy makers' thinking e.g. in the aftermath of the financial crisis and regarding health care reform. This success stems from the fact that the mechanisms described by these theoretical models are intuitively plausible and seem to be at work in the real world. Key mechanisms have been detected in the data. For example, the possibility of market unravelling, which had been pointed out by Akerlof (1970), has been described in several settings. Cutler and Zeckhauser (1997) use a policy change to document it in the market for health insurance for employees at Harvard. More recently, the market for pension annuities in the UK has seen unravelling dynamics following a reform which made it optional to convert pension savings into an annuity. Previously this had been compulsory.[8]

However, more thorough empirical tests of these theories, which aim to move beyond descriptive statistics, find mixed results. Chiappori and Salanié (2000) use data on car insurance in France and find no evidence of asymmetric information. They suggest a "positive correlation test" for asymmetric information. They point out that both pure adverse selection on riskiness and pure moral hazard predict a positive correlation between coverage and expenditure after controlling for all observable characteristics which firms are allowed to contract on. In the case of adverse selection, higher risk individuals select into more comprehensive plans while in the case of moral hazard comprehensive plans reduce incentives to mitigate risks. Chiappori and Salanié (2000) conjecture that this positive correlation is predicted more generally by models with asymmetric information. Hence, they argue that testing whether this correlation is indeed significantly positive amounts to a test for asymmetric information.[9] In their survey of empirical tests of contract theory, Chiappori and Salanié (2003) thus argue that there is a mismatch between theoretical predictions and empirical results.

[8] Following the reform, annuitization rates fell from 90 percent to 12 percent (Sturrock, 2019). See also the Financial Times article by Ralph (2017) or for an early warning Blake (2014). Technically, prior to the reform it was possible to withdraw pension savings and not convert them but this was subject to a large unfavourable tax effect. For more detail see Sturrock (2019).

[9] In their later work, Chiappori et al. (2006) formalise their earlier conjecture and show that a broader class of models implies a positive correlation. Chiappori et al. (2006) find a positive correlation indicating the presence of asymmetric information.

Both the theoretical and the empirical literature have developed substantially since then. One important avenue of research has been to model agents as having private information on *two* dimensions: their expected cost and additionally their risk preferences. These models have been successful at bringing theory and empirics closer together by providing micro-founded arguments for why a positive correlation test is not a valid test for asymmetric information. We now turn to these models of multi-dimensional private information.

4.2.2 Multi-dimensional private information

This subsection shows how shifting the focus from one to two dimensions of private information has advanced our understanding of selection markets and helped to reconcile theoretical and empirical results.

Theory: When agents have multiple-dimensions of private information, the resulting selection pattern can be the opposite of adverse selection, i.e. of what would be expected in the presence of one degree of private information. When agents have private information on their expected cost as well as their risk preferences the following pattern can emerge: Very risk averse agents select into contracts with high coverage and take precautions, thus becoming low risk, i.e. low expected expenditure. On the other hand agents with low risk aversion select into contracts with low coverage and do not take many precautions, thus becoming high risk, i.e. high expected expenditure. This form of selection, termed advantageous selection (De Meza and Webb, 2001), implies a *negative* correlation of coverage and expenditure - the exact opposite of what Chiappori and Salanié (2000) tested for. More generally, it is possible that two dimensions of asymmetric information lead to a correlation that is close to zero. Hence, we cannot rely on the positive correlation test to detect asymmetric information. Smart (2000) shows that multiple dimensions of private information also pose theoretical challenges as they can imply that the single cross property does not hold - the opposite is assumed in many one-dimensional signalling models.

While the interaction of theoretical and empirical research has rekindled the interest in multi-dimensional private information, theorists had already earlier worked on this topic, albeit from a slightly different angle. All papers mentioned above focus on a single good and multiple dimensions of private information. There exists substantial theoretical work on the case of a multi-product monopolist contracting with agents who have multiple dimensions of private information.

The idea that a multi-product monopolist could benefit from selling multiple products together in order to sort consumers was first expressed by Stigler (1963). Adams and Yellen (1976) are the first paper to formalise this problem of a multi-product monopolist facing consumers with two dimensions of private information. They characterise conditions under which it is profit maximising to sell the two goods separately, or to offer the two goods

exclusively as a bundle, or to offer "mixed bundling", i.e. offer the goods individually as well as in a bundle. They show that there exist distributions of consumers' willingness-to-pay where "mixed bundling" is the monopolist's optimal strategy. These results have been generalised by McAfee et al. (1989). For a detailed introduction with a worked example, we refer the reader to Bolton and Dewatripont (2004).

The multiple dimensions of private information generate analytical challenges (Rochet and Stole, 2003). One challenge is that with multiple dimensions it is not clear which incentive compatibility constraints bind, making it necessary to solve a problem with a larger number of constraints than in the one dimensional case. Hence, a full solution of the general monopolist's problem has not been written yet. However, substantial steps have been taken towards this goal.

Armstrong and Rochet (1999) provide a complete solution to the monopolist's problem in a simple case. One of their simplifying assumptions is to look at discrete rather than continuous dimensions of private information. They focus on the simplest possible discrete framework: a "four type framework". Moreover, they assume that preferences for the two goods are additively separable (which we will also assume). They develop a two-step approach to solving the problem, first solving a "relaxed problem" where they assume that only downwards incentive-compatibility constraints are binding, and then checking in which cases this assumption is justified, otherwise reformulating the problem. This highlights the general lesson that with multiple dimensions of private information *all* incentive compatibility constraints need to be checked and that no simplification of them remains valid for the most general preferences.

Other papers describe properties which a general solution to the monopolist's problem will have. Armstrong (1996) shows that the monopolist will want to exclude some consumers from purchasing any goods. This allows the monopolist to extract higher rents from those consumers who continue to purchase. Recently, Carroll (2017) has characterised the monopolist's optimal strategy in a "robust" setting by relaxing an assumption on the principal's knowledge. Rather than knowing the joint distribution of all dimensions of an agent's type, Carroll's principal only knows the marginal distribution. For more comprehensive overviews we refer the reader to Rochet and Stole (2003) and Bolton and Dewatripont (2004).

This paper deviates from the literature in several ways. The literature so far has focused on a monopolist, thus asking when offering specific sets of contracts, e.g. mixed bundling, is profit maximising for a monopolist. This paper investigates when bundling contracts emerge under perfect competition in a selection market. A rationale for bundling under duopoly in a non-selection market, where costs are independent of the type of consumer, is provided by Armstrong and Vickers (2010). While our papers aims to show an existence result in the spirit of Akerlof (1970) using a "simple" model similar to Armstrong and Rochet (1999), we need to deviate from their "four type framework" to a "six type framework". This has important implications: In our case full revelation of

types is impossible. While we have two dimensions of hidden information (riskiness, i.e. expected cost, and risk preferences) and two screening devices, i.e. two strategic variables (price and type of contract, i.e. individual or tying), in our case full revelation of types is *impossible* since there can be at most three different contracts in equilibrium while there are six different types of agents.

Evidence: Multiple dimensions of private information are not only of theoretical interest but have also been detected in empirical studies. Finkelstein and McGarry (2006) study the market for long-term care insurance. While they find no significant correlation of coverage and costs - which Chiappori and Salanié (2000) would have interpreted as evidence of no private information - they provide evidence that agents have two dimensions of private information: Insurance is bought by high risk individuals and high risk-aversion individuals. Overall, those buying insurance are average risk, leading to zero correlation. Hence, while it looks like there is no private information, there are two dimensions of private information. Finkelstein and McGarry's methodology is based on a novel test for asymmetric information which compares an individual's subjective probability of using a care-home in the future with the insurance company's implied probability and interprets the residual as private information. Cohen and Einav (2007) study the market for automobile insurance where they also find evidence for private information on two dimensions. However, while in Finkelstein and McGarry (2006) these two dimensions opposed each other, thus leading to a zero correlation, in Cohen and Einav (2007) the two dimensions reinforce each other. Viewed together, these papers show that multi-dimensionality exists in different real world settings, but that the exact form and pattern can differ depending on the individual market.

The positive correlation test for asymmetric information suggested by Chiappori and Salanié (2000) can not only be criticised for becoming invalid when there are multiple dimensions of private information. Even when there is only a single dimension of private information, it cannot differentiate between moral hazard and adverse selection. This would be important as both have different policy implications. Moreover, the positive correlation test gives no indication of the welfare cost of the informational asymmetry. Einav et al. (2010) propose an alternative method which improves on both these criticisms[10]: The "cost function approach". They require an additional piece of information relative to Chiappori and Salanié: exogenous variation in prices. They use this information to estimate the average cost curve (AC) and the demand curve directly. From this they can back out the marginal cost curve (MC). The slope of the marginal cost curve implies the type of selection. Moreover, with MC, AC, and demand known, the welfare costs of selection can be calculated directly.

In addition to evidence on the existence of multiple dimensions of private information, there are also papers which test for the implications of this, e.g. advantageous selection. Mahoney and Weyl (2014) find evidence of advantageous selection in the market for US

[10] A good summary is provided in Einav and Finkelstein (2011).

sub-prime auto loans using data from Einav et al. (2012). Fang et al. (2008) find evidence of advantageous selection in the Medigap market, but argue that this is more strongly driven by factors other than people having multiple dimensions of private information, e.g. by people having different cognitive abilities.

Open Questions: The empirical literature called on theorists to provide answers to the following questions: When does a second dimension of private information increase (decrease) the efficiency cost of adverse selection?[11] (Finkelstein and McGarry, 2006); How can we model a selection market with multiple dimensions of private information and allow the contract space to be determined endogenously, i.e. how can we have competition both in price and in the type of contracts offered?[12] (an "extremely important" question according to Einav and Finkelstein (2011)). This paper starts to shed light on these questions.

Related non-selection literature: In our model one piece of private information affects cost in exactly the opposite way in the different markets. Hence, if a customer was asked about his life expectancy by a life insurer, the customer would want to overstate his life expectancy. If he is asked by a pension provider, he would want to understate it. Thus, the direction in which consumers would want to distort their report in one market is exactly the opposite direction as in the other market. Lewis and Sappington (1989) study a case with opposing incentives in a very different setting. They study a principal agent problem which extends Baron and Myerson (1982). The principal (a regulator) is contracting with an agent (a firm) which has private information on its fixed and marginal costs. Lewis and Sappington (1989) impose that there is a one-for-one mapping from marginal to fixed costs, i.e. that higher marginal costs must mean fixed costs are low. Hence, the agent has "countervailing incentives": He wants to overstate marginal costs to convince the regulator to set a high per unit price, but also wants to understate marginal costs as this implies high fixed costs and thus a high lump-sum transfer. Our paper differs from Lewis and Sappington (1989) in several ways. Most importantly we focus on selection markets, not on a principal agent problem. Nevertheless, one can think of our model as the first paper to describe what we term "countervailing selection": the high cost types in one market are exactly the low cost types in another market. I.e. while in one market everyone wants to pretend to be of one extreme type (low life expectancy) to get a cheap contract, in the other market all agents want to pretend the opposite (i.e. high life expectancy).

[11] "An unanswered question, however, and an important avenue for further work, is under what conditions this added dimension of heterogeneity reduces or increases the efficiency cost of asymmetric information about risk types." p.951 in Finkelstein and McGarry (2006).

[12] "Allowing the contract space to be determined endogenously in a selection market raises challenges on both the theoretical and empirical front. On the theoretical front, we currently lack clear characterizations of the equilibrium in a market in which firms compete over contract dimensions as well as price, and in which consumers may have multiple dimensions of private information (like expected cost and risk preferences). From an empirical standpoint, the challenge is that if adverse selection greatly reduces the set of offered contracts, estimating the welfare loss from the contracts not offered may require the researcher to go quite far out of sample. While these challenges are far from trivial and may explain why there has been relatively little work of either type on this topic to date, we view this direction as an extremely important—and likely fruitful—topic for further research." p.137 in Einav and Finkelstein (2011).

Information spill-overs: The idea that negative correlation between two aspects on which agents are privately informed can mitigate the market failure caused by adverse selection is present in our model and also in Huangfu and Liu (2019), albeit in a different setting. They consider information spill-overs in a dynamic bargaining model when one seller is selling two products with the respective qualities being known only to the seller. If product qualities are negatively correlated, i.e. if one product being low quality indicates the the other is high quality, then consumers who observe that good A was sold early at a low price infer that good A was low quality and that thus good B must be high quality. Hence, consumers are willing to pay a high price for good B which reduces the market failure associated with adverse selection.

A related paper on information spill-overs is Asriyan et al. (2017) who study information spill-overs when there are multiple sellers, each selling one good, and the quality of these goods is positively correlated. The sellers in Asriyan et al. (2017) do not internalise the information spill-over. This leads to the possibility of multiple equilibria. In Huangfu and Liu (2019) and in our paper the equilibrium is unique.

While in our model, the presence of negative correlation also mitigates the welfare loss caused by adverse selection, it works through a different channel than information spill-overs. In our setting which focuses on tying contracts, cost differences can cancel out. Thus, there can be an equilibrium where all types buy the tying contract, where this prevents unravelling, but where firms have not learned the agent's type, e.g. his life-expectancy. Instead, in the simple example in our introduction, the firm does not need to infer life-expectancy. As cost differences have cancelled out, life-expectancy is irrelevant for the firm selling the tying contract.

4.3 A Model of Two Selection Markets

To formalise how tying contracts can prevent unravelling and increase welfare, this section develops a model of two selection markets where one characteristic (high life expectancy) makes agents high cost in one market (pension annuity) but low cost in the other (life insurance). Consumers have private information on their expected cost and on their risk preferences. Including private information on risk preferences allows us to highlight two mechanisms how tying contracts prevent unravelling and magnifies the welfare benefit of introducing tying contracts.

4.3.1 Outline of the Model

Agents: In the model, there are insurance firms and different types of consumers. Each consumer is characterised by his life expectancy μ_i and his risk preferences ρ_i. Life expectancy is either low (L), medium (M) or high (H): $\mu_i \in \{L, M, H\}$. All consumers are risk averse, but they differ in the strength of their risk aversion. Some consumers

are relatively at ease with risk (e) while others are more worried (w): $\rho_i \in \{e, w\}$. Thus, there are six groups of consumers: $L_e, L_w, M_e, M_w, H_e, H_w$. Both life expectancy and risk preferences are private knowledge. The insurance firms are risk neutral and profit maximising. They engage in Bertrand competition and face no administrative costs.

Preferences: Consumers demand insurance because they are risk averse and face two types of risk: Risk of an early death against which they can buy life-insurance; Risk of outliving their savings, against which they can buy a pension annuity. Throughout the paper, we work directly in terms of Willingness-To-Pay (WTP) for the respective type of insurance[13] as well as in terms of expected cost and risk premia.

We denote expected cost of life insurance for a consumer with life expectancy μ_i as $c(\mu_i)$ and for the pension as $C(\mu_i)$. Throughout, lower case letters refer to the risk faced first (early death, life insurance) while upper case letters refer to the risk at older age (outliving savings, pension). While expected costs only depend on an agent's life expectancy, an agents willingness-to-pay (WTP) for insurance also depends on his risk preferences. Hence, we denote the WTP for life-insurance of an agent with life expectancy μ_i and risk-preferences ρ_i as $v(\mu_i, \rho_i)$. As we assume that all agents are risk averse and since higher risk aversion increases a consumer's WTP for insurance, we have $c(\mu) \leq v(\mu, r) \leq v(\mu, w) \; \forall \; \mu$. The risk premium for life insurance is $r(\mu, \rho) = v(\mu, \rho) - c(\mu)$. For pensions, we analogously define WTP as $V(\mu, \rho)$ and the risk premia as $R(\mu, \rho) = V(\mu, \rho) - C(\mu)$.

Further parameter orderings arise from the specific type of insurance. Consumers with a lower life expectancy have higher expected costs for life insurance ($c(L) > c(M) > c(H)$) while they have lower expected costs for annuity providers ($C(L) < C(M) < C(H)$). For notational convenience, we define a measure of the curvature of costs δ where: $\delta = (c(M) - c(H)) - (c(L) - c(M))$. If $\delta \geq 0$ we refer to costs as concave, which is the case we focus on. $\delta < 0$ refers to convex costs. Cost structures are said to be "more concave" the larger δ is.

These parameters are summarised in the Table 4.1 below:

Table 4.1: Summary of Parameters

	WTP at ease	WTP worried	expected cost
Low life expectancy	$v(L,e), V(L,e)$	$v(L,w), V(L,w)$	$c(L), C(L)$
Medium life expectancy	$v(M,e), V(M,e)$	$v(M,w), V(M,w)$	$c(M), C(M)$
High life expectancy	$v(H,e), V(H,e)$	$v(H,w), V(H,w)$	$c(H), C(H)$

Actions: Insurance companies can offer one of three types of contracts:

[13] A rationale for purchasing a pension annuity can be derived from insurance arguments: Individuals with a concave per period utility function want to smooth their spending across time but are uncertain on their exact date of death. An annuity is insurance against living longer than retirement savings last. A rationale for life insurance can be derived from bequest motives. Thus, we see life insurance as insuring expected wealth for a family so that e.g. College education for the children could still be finance if the main bread-winner dies early.

1. Life insurance which pays out a lump sum (Z) in case of death before retirement and does not pay out if the insurance taker lives beyond retirement age.

2. A pension which pays out Y per month after retirement until the insurance taker dies.

3. A tying contract which is a promise to provide life insurance until retirement and an annuity thereafter.

Since the aim of this paper is to explore the effect of tying contracts in a simple model, we do not focus on deductibles or partial insurance. We treat the terms of the contracts (Z and Y) as exogenously fixed and as identical in the individual contracts and in the tying contract.

Timing: Firms simultaneously decide which of the three types of contracts to offer and state the price they demand.[14] Then consumers compare all contracts and choose the contract, or combination of contracts, that maximises their utility, i.e. where they achieve the highest net monetary benefit. This is the case where the difference of WTP (or the sum of both WTPs) and the price paid for the contract is the largest. Consumers also have the outside option to not buy any insurance which results in zero net monetary benefit.

Equilibrium concept: Throughout, we focus on sub-game perfect Nash equilibria (SPNE)[15]. SPNE of this screening game must have the following properties:

- In equilibrium firms make zero profit.
- From the set of contracts offered in equilibrium, agents choose the contract which maximises their net monetary benefit.
- For every type of contract, there does not exist a price to which firms could deviate and make a strictly positive profit when the profitability calculation is based on consumer purchasing behaviour which is an optimal response to this deviation.

[14]It is equally possible to think of three different pools of firms, where firms in a given pool can only offer a given type of contract. Then, firms compete on price within the pool. We prefer to think of firms as being able to choose the type of contract they offer because the costs of changing the type of contract offered in insurance markets seem low and even if costs are high, our framework captures the long-run equilibrium. Our approach and the approach sketched here yield the same result.

[15]While the first two points are also needed for Nash equilibria, the third point refines our equilibrium concept to sub-game perfect Nash equilibria. Without this refinement, we would have equilibria where marginally cutting the price of a tying contract is loss making, but cutting it substantially is profitable as this attracts lower cost consumers. Since Nash equilibria place no restriction on agents' off-equilibrium path replies, this can be sustained as a Nash equilibrium. Our focus on sub-game perfect Nash equilibria removes these implausible equilibria and delivers a unique equilibrium prediction. We discuss the unique prediction further in later sections but first turn to different types of equilibria that can exist depending on parameters.

4.3.2 Assumptions

Throughout this paper, we assume that all six types are equally likely[16] and that costs are weakly concave $\delta \geq 0$. For some results, we draw on additional assumptions. We outline these assumptions in detail here. Thereafter, when used, we just refer to their abbreviation.

Symmetry of costs For ease of exposition, we introduce notions of symmetry of costs. We do not claim that this reflects reality. However, it does simplify the exposition by reducing the number of parameters and allows us to derive cleaner conditions for existence. We distinguish between two notions of symmetry.

Weak symmetry (S1) refers to the assumption that every type's expected costs are mirror images across markets. This means that the high cost types in the respective markets are equally expensive ($c(L) = C(H)$), the medium types are equally expensive ($c(M) = C(M)$), and the low cost types are equally cheap ($c(H) = C(L)$). Risk preferences are allowed to differ, e.g. agents could be more worried about an early death than about outliving their savings.

Strong symmetry (S2) refines weak symmetry by adding the assumption that not just expected costs but also risk preferences are symmetric across markets. Hence, in addition to weak symmetry, strong symmetry adds that: $v(L,e) = V(H,e), v(M,e) = V(M,e), v(H,e) = V(L,e)$ and $v(L,w) = V(H,w), v(M,w) = V(M,w), v(H,w) = V(L,w)$.

Adverse selection among worried agents (AS1): We assume that risk preferences for worried agents do not reverse fundamental orderings ($V(H,w) > V(M,w) > V(L,w), v(L,w) > v(M,w) > v(H,w)$). This ensures adverse selection among worried agents. (AS1) does not make assumptions on relaxed agents.

Adverse selection among each type of risk-preference (AS2): (AS2) refines (AS1) by adding the assumption that also for relaxed agents risk preferences do not reverse fundamental orderings ($V(H,e) > V(M,e) > V(L,e), v(L,e) > v(M,e) > v(H,e)$). This ensures that the market has adverse selection. If risk preferences reverse fundamental orderings, the market may have advantageous selection.

4.4 Equilibrium Purchasing Behaviour

This section first characterises the equilibrium when tying contracts cannot be offered, then when they can be offered. We show how the equilibrium purchasing behaviour, i.e. amount of coverage and the price of coverage, depends on cost parameters and risk

[16] Based on this assumption, resulting equilibrium existence conditions are in terms of risk preferences and cost parameters. We conjecture that one could replicate our analysis assuming a particular cost structure, e.g. linear costs, and not imposing assumptions on the size of the six groups. This would yield existence conditions in terms of risk preferences and population shares.

preferences.[17] In all cases, there exists a unique SPNE. First, in the case where firms can only offer stand-alone insurance contracts and cannot offer tying contracts, we find that unravelling is more likely when heterogeneity in costs is large or levels of risk aversion are low. Then, we turn to the case where firms can choose between offering a tying contract or stand-alone contracts and find that tying contracts are always used in equilibrium. The number of types who are covered by a tying contract is i) lower, the more concave costs are, ii) higher, the larger levels of risk aversion are iii) higher, the larger the heterogeneity in risk aversion is.

4.4.1 Equilibrium Purchasing Behaviour: Tying Contracts cannot be Offered

When firms are restricted to offering stand alone insurance contracts, i.e. tying contracts cannot be offered, markets can be studied separately and each behaves like a selection market in Akerlof (1970) with a tendency to unravel. A small modification relative to Akerlof is that we have two rather than one dimension of private information. The strict symmetry assumption (S2) ensures that we can focus on one stand alone market and then map results across to the other.

There are five different equilibrium purchasing patterns corresponding to different degrees of unravelling. More unravelling occurs when heterogeneity in costs is larger or levels of risk aversion are lower. This follows the Akerlof intuition: Suppose everyone buys life-insurance at the average cost as calculated across the population: $\frac{c(L)+c(M)+c(H)}{3}$. For this to be an equilibrium, the lowest WTP type needs to prefer this contract to remaining uninsured. In this case, H_e needs to want to buy life-insurance at this price. Since his WTP is $v(H,e) = c(H) + r(H,e)$ he is willing to spend $\frac{c(L)+c(M)+c(H)}{3}$, which exceeds his expected cost, if and only if

$$v(H,e) - \frac{c(L)+c(M)+c(H)}{3} > 0 \tag{4.1}$$

which simplifies to

$$r(H,e) > \frac{c(L)+c(M)-2c(H)}{3} \tag{4.2}$$

Thus, (4.2) shows that H_e is willing to buy life insurance at price $\frac{c(L)+c(M)+c(H)}{3}$ if and only if his risk aversion is large or cost heterogeneity is low. A similar logic applies at all other stages of potential unravelling.

We characterise the five different equilibrium purchasing patterns in Table 4.2 below. Formal proofs are in Appendix 4.7.1. These equilibria can be ranked by total surplus from W_1 for complete coverage (no unravelling) and highest surplus to W_5 for the highest degree of unravelling where only the most expensive types are insured. Figure 4.1 illustrates the equilibrium conditions graphically.[18]

[17] Formally, the term "equilibrium purchasing behaviour" describes the mapping from agent type (μ_i, ρ_i) to contract purchased in equilibrium $(\mathcal{L}, \mathcal{P}, \mathcal{T}, \mathcal{L}\ and\ \mathcal{P}, none)$.

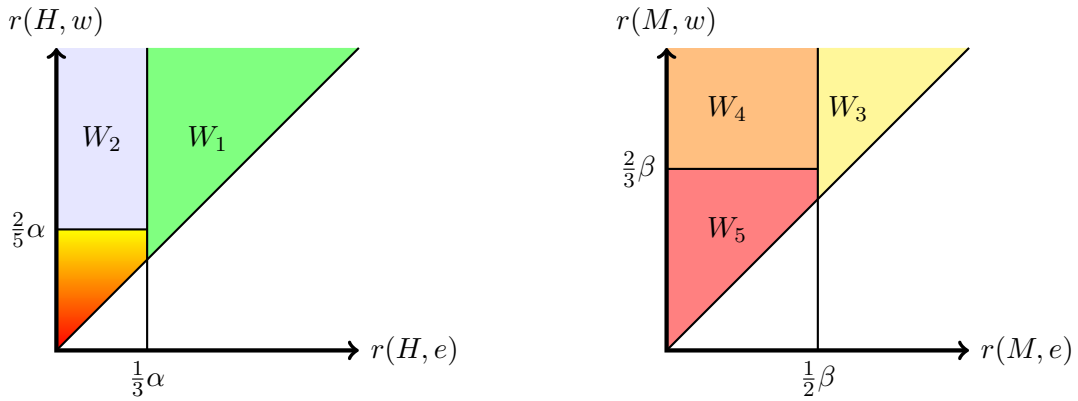
[18] Table 4.2 and Figure 4.1 depict the case where $v(M,e) > v(H,w)$. In the alternative case where $v(M,e) <$

Table 4.2: Equilibrium Purchasing Patterns: only stand-alone contracts offered

	Individual market	Both markets	SPNE if and only if
W_1	Full coverage e: L: L, M: L, H: L w: L: L, M: L, H: L	e: L: $L+P$, M: $L+P$, H: $L+P$ w: L: $L+P$, M: $L+P$, H: $L+P$	$r(H,e) > \frac{1}{3}[c(L)+c(M)-2c(H)]$
W_2	e: L: L, M: L, H: $-$ w: L: L, M: L, H: L	e: L: L, M: $L+P$, H: P w: L: $L+P$, M: $L+P$, H: $L+P$	$r(H,e) < \frac{1}{3}[c(L)+c(M)-2c(H)]$ $r(H,w) > \frac{2}{5}[c(L)+c(M)-2c(H)]$
W_3	e: L: L, M: L, H: $-$ w: L: L, M: L, H: $-$	e: L: L, M: $L+P$, H: P w: L: L, M: $L+P$, H: P	$r(H,e) < \frac{1}{3}[c(L)+c(M)-2c(H)]$ $r(H,w) < \frac{2}{5}[c(L)+c(M)-2c(H)]$ $r(M,e) > \frac{1}{2}[c(L)-c(M)]$
W_4	e: L: L, M: $-$, H: $-$ w: L: L, M: L, H: $-$	e: L: L, M: $-$, H: P w: L: L, M: $L+P$, H: P	$r(H,e) < \frac{1}{3}[c(L)+c(M)-2c(H)]$ $r(H,w) < \frac{2}{5}[c(L)+c(M)-2c(H)]$ $r(M,e) < \frac{1}{2}[c(L)-c(M)]$ $r(M,w) > \frac{2}{3}[c(L)-c(M)]$
W_5	Complete unravelling e: L: L, M: $-$, H: $-$ w: L: L, M: $-$, H: $-$	e: L: L, M: $-$, H: P w: L: L, M: $-$, H: P	$r(H,e) < \frac{1}{3}[c(L)+c(M)-2c(H)]$ $r(H,w) < \frac{2}{5}[c(L)+c(M)-2c(H)]$ $r(M,e) < \frac{1}{2}[c(L)-c(M)]$ $r(M,w) < \frac{2}{3}[c(L)-c(M)]$

Where we denoted types purchasing a pension by P, life-insurance by L, no insurance by $-$.

Figure 4.1: Equilibrium Purchasing Patterns: only stand-alone contracts offered



where $\alpha = [c(L)+c(M)-2c(H)]$ and $\beta = [c(L)-c(M)]$.

W_i is a purchasing pattern with higher total surplus than W_{i+1}.

$v(H,w)$, a similar table results where again the purchasing patterns correspond to different degrees of unravelling and where these patterns can be ordered by total surplus. A difference to Table 4.2 is that in W_3, M_e would not purchase life-insurance but H_w would. The resulting table is depicted in Appendix 4.7.2.

4.4.2 Equilibrium Purchasing Behaviour: Tying Contracts can be Offered

When firms are allowed to offer tying contracts and stand-alone contracts, three different patterns of equilibrium purchasing behaviour can arise. In each pattern, tying contracts serve different roles. First, we characterise each of these three patterns. Second, we highlight the intuition behind them which is best understood when comparing all three. Then, we derive a Theorem which covers existence, uniqueness, and properties of SPNE.

Equilibrium Pattern 1: Tying contracts act as pooling device

Definition: If the purchasing behaviour depicted below arises in equilibrium, tying contracts are said to act as a pooling device and the equilibrium is termed "equilibrium pattern 1":

Table 4.3: Equilibrium Pattern 1

Equilibrium Pattern 1	at ease	worried
Low life expectancy	tying contract	tying contract
Medium life expectancy	tying contract	tying contract
High life expectancy	tying contract	tying contract

Proposition 4.1 *(1) For some parameters there exists a sub-game perfect Nash equilibrium where tying contracts act as pooling device. (2) (S1, AS2) An equilibrium where tying contracts are a pooling device exists if and only if the following two conditions hold simultaneously:*

$$R(L, e) > \frac{1}{3}\,\delta \tag{4.3}$$

$$r(H, e) > \frac{1}{3}\,\delta \tag{4.4}$$

Proof: See Appendix 4.7.3.

Equilibrium pattern 1 is similar to the example in the introduction in the sense that all types of consumers buy the tying contract, but it adds heterogeneity in risk preferences and allows these preferences to be private knowledge.

Equilibrium Pattern 2: Tying contracts screen based on expected cost

Definition: If the purchasing behaviour depicted below arises in equilibrium, tying contracts are said to screen based on expected cost and the equilibrium is termed "equilibrium pattern 2":

Table 4.4: Equilibrium Pattern 2

Equilibrium Pattern 2	at ease	worried
Low life expectancy	only life insurance	only life insurance
Medium life expectancy	tying contract	tying contract
High life expectancy	only pension	only pension

Proposition 4.2 *(1) For some parameters, there exists a sub-game perfect Nash equilibrium where tying contracts screen based on expected cost. (2) (S1) There exists an equilibrium where tying contracts screen based on expected cost, if and only if the following conditions hold simultaneously:*

$$R(L, w) < \frac{1}{2}\,\delta \tag{4.5}$$

$$R(L, e) < \frac{1}{3}\,\delta \tag{4.6}$$

$$r(H, w) < \frac{1}{2}\,\delta \tag{4.7}$$

$$r(H, e) < \frac{1}{3}\,\delta \tag{4.8}$$

Proof: See Appendix 4.7.4.

This pattern can also be viewed as all three markets unravelling because in this equilibrium each contract is only purchased by those types with the highest cost for it: M is high cost for $\mathcal{T}$ by concavity of costs, L is high cost for $\mathcal{L}$, H for $\mathcal{P}$. Therefore, a necessary condition for this pattern to be SPNE is that M clearly have the highest cost for tying contracts (strict concavity). Like in the single contract case, the market for the tying contract unravels and thus this is an SPNE if heterogeneity in costs is large (high degree of concavity) and if risk aversion is low. Both these make L_w and H_w less willing to pay the extra cost to insure against the second risk. We now turn to an equilibrium pattern where the number of types buying the tying contract lies between pattern 1 and pattern 2.

Equilibrium Pattern 3: Tying contracts screen based on risk preferences and expected cost

Definition: If the purchasing behaviour depicted below arises in equilibrium, tying contracts are said to screen based on risk preferences and expected costs and the equilibrium is termed "equilibrium pattern 3":

Table 4.5: Equilibrium Pattern 3

Equilibrium Pattern 3	at ease	worried
Low life expectancy	only life insurance	tying contract
Medium life expectancy	tying contract	tying contract
High life expectancy	only pension	tying contract

Proposition 4.3 *(1) For some parameters, there exists a sub-game perfect Nash equilibrium where tying contracts screen based on risk preferences. (2) (S1, AS1) An equilibrium where tying contracts screen based on risk preferences exists if and only if the following conditions hold simultaneously:*

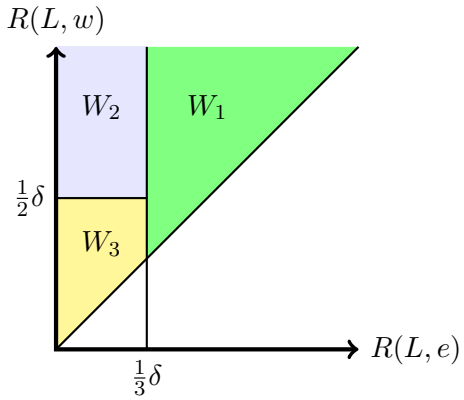
$$2\ R(L,w) > \delta > 3\ R(L,e) \tag{4.9}$$

$$2\ r(H,w) > \delta > 3\ r(H,e) \tag{4.10}$$

Proof: See Appendix 4.7.5.

To conclude the outline of purchasing patterns, the purchasing patterns when tying contracts are offered are summarised in Figure 4.2 below, which is analogous to the corresponding Figure in section 4.4.1.

Figure 4.2: Equilibrium Purchasing Patterns: Tying and stand-alone contracts offered



W_i is a purchasing pattern with higher total surplus than W_{i+1}.

Intuition

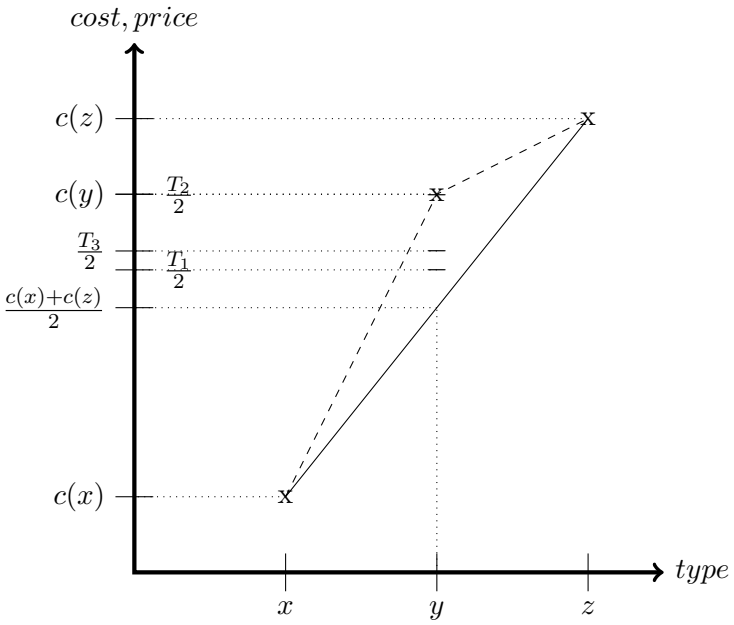
This section provides a diagrammatic intuition for how concavity of costs, levels of risk aversion and heterogeneity of risk aversion affect which equilibrium purchasing pattern is the unique SPNE. We also formalise these forces in a Theorem.

The intuition builds on considering potential deviations from a conjectured equilibrium. Suppose the economy starts in pattern 2. Thus, L-types buy stand alone life insurance for a price $c(L)$ and make a net monetary benefit of $r(L, \rho)$. They will deviate to a tying contract if and only if it offers them a higher net monetary benefit. Mathematically, they deviate if and only if:

$$\frac{1}{2}R(L, \rho) > \frac{T_k}{2} - \frac{c(L) + c(H)}{2} \tag{4.11}$$

where T_k is the price of a tying contract in equilibrium pattern $k \in \{1, 2, 3\}$ and $T_1 < T_3 < T_2$. To visualise this condition, we focus on one selection market in which x is low cost (for pensions: L, for life-insurance: H), y is medium cost (M) and z is high cost. The assumption that costs are concave ($\delta > 0$) ensures that $c(y) < \frac{c(x)+c(z)}{2}$ and yields the diagram:

Figure 4.3: Costs in one selection market



Condition (4.11) states that L_ρ wants to deviate to T_k if and only if the risk premium $R(L, \rho)$ is sufficiently large to cover the gap between T_k and $\frac{c(x)+c(z)}{2}$. This gap depends on the equilibrium pattern k and on concavity of costs: In equilibria where more low cost types buy the tying contract, the price for the tying contract is lower by the firms' zero profit condition, making the potential deviation more attractive ($T_1 < T_3 < T_2$). A higher degree of concavity increases $T_k - \frac{c(x)+c(z)}{2}$ for every k.

There are three possible cases:

1. Both worried and relaxed agents have such high risk premia that we have $\frac{1}{2}R(L, \rho) > \frac{T_1}{2} - \frac{c(L)+c(H)}{2}$ $\forall$ ρ. Agents are so risk averse that they do want to pay the extra cost of insuring the second risk. In this case, pattern 1 is the unique SPNE.

2. The relaxed types have such low risk premia that $\frac{1}{2}R(L,r) < \frac{T_1}{2} - \frac{c(L)+c(H)}{2}$ and the worried have such low risk premia that $\frac{1}{2}R(L,w) < \frac{T_3}{2} - \frac{c(L)+c(H)}{2}$. Agents risk aversion is so low that neither of them wants to pay the extra cost of insuring the second risk. L_w does not want to insure at price T_3 and L_e does not even want to insure at T_1, the lowest possible T. In this case, pattern 2 is the unique SPNE.

3. There is large heterogeneity in risk premia. The relaxed have low risk premia such that $\frac{1}{2}R(L,e) < \frac{T_1}{2} - \frac{c(L)+c(H)}{2}$, but the worried have high risk premia $\frac{1}{2}R(L,w) > \frac{T_3}{2} - \frac{c(L)+c(H)}{2}$. Heterogeneity in risk preferences is so large hat while L_e is not willing to insure at T_1, L_w is willing to insure even at T_3. In this case, pattern 3 is the unique SPNE.

Combining the results for these three cases, which are the only patterns that ever constitute an SPNE, yields the following Theorem:

Existence and uniqueness results

Theorem 4.1 *Existence and Uniqueness Theorem (S2, AS2): There always exists a unique sub-game perfect Nash equilibrium. In this equilibrium, tying contracts are always used or all agents are insured against both risks. The number of types who are covered by a tying contract in equilibrium is i) lower, the more concave costs are (higher δ), ii) higher, the larger levels of risk aversion are (higher $r(\mu,\rho)$), iii) higher, the larger the heterogeneity in risk aversion is (higher $r(\mu,w) - r(\mu,e)$).*

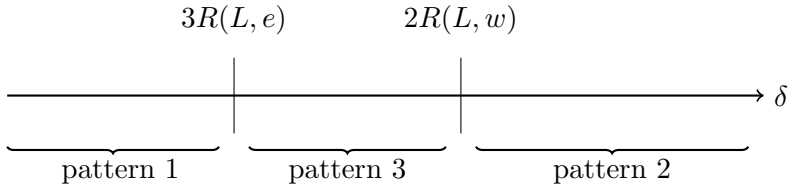
Proof: See Appendix 4.7.6.

Sketch proof: First, we establish that always exactly one of the three patterns is the unique SPNE. Then, we show that across these three patterns, the comparative static results $i) - iii)$ hold.

To show that always exactly one of the three patterns is SPNE we distinguish between two cases and visualise existence conditions under (S2):

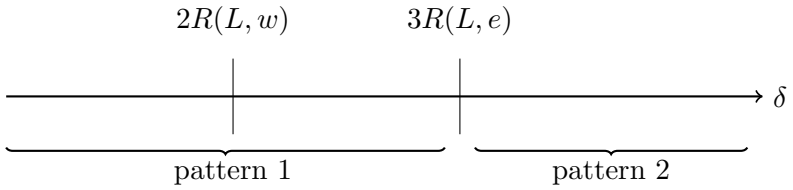
Case 1: $3R(L,e) < 2R(L,w)$

Figure 4.4: Existence of SPNE if $3R(L,e) < 2R(L,w)$



Case 2: $3R(L,e) > 2R(L,w)$

Figure 4.5: Existence of SPNE if $3R(L,e) > 2R(L,w)$



To see that the comparative static results $i) - iii)$ hold, note that the number of groups buying a tying contract is smallest for pattern 2, larger for pattern 3 and highest for pattern 1. *The number of types who are covered by a tying contract in equilibrium is*

- *i) lower, the more concave costs are (higher δ):* follows directly from the diagrams above.
- *ii) higher, the larger levels of risk aversion are (higher $r(\mu,\rho)$):* also visible in the diagram above.
- *iii) higher, the larger the heterogeneity in risk aversion is (higher $r(\mu,w) - r(\mu,e)$):* For given costs and given $R(\mu,e)$, higher $R(\mu,w)$ can lead to a change of SPNE from pattern 2 to 3 thus lead to more people being covered by a tying contract.

As there always exists a unique SPNE which is one of these three patterns, every SPNE has properties which are characterised by this Corollary:

Corollary 4.1 *Every SPNE has the following properties:*

1. *Every agent is insured against at least one risk.*
2. *Medium life-expectancy types always buy the tying contract and thus insured against both risks.*
3. *The extreme types are always insured at least against the risk for which they are high cost.*
4. *Worried agents buy at least as much insurance as agents who are at ease with risk.*

Proof: See Appendix 4.7.7.

4.5 Welfare

To investigate the social benefit of introducing tying contracts, we adopt total surplus as welfare measure and show that introducing tying contracts increases welfare regardless

of whether risk preferences are public or private information, that risk preferences being private rather than public information reduces welfare for the case with and without tying contracts, and that the welfare gain from the former outweighs the loss from the latter.

We adopt total surplus as welfare measure because it captures the social cost of markets unravelling and is convenient to use. Surplus is the benefit generated from a trade: $v(\mu_i, \rho_i) - c(\mu_i)$ or $V(\mu_i, \rho_i) - C(\mu_i)$. Total surplus sums the benefits across all trades which occur in equilibrium: $\sum_{i\in\Omega} v(\mu_i, \rho_i) - c(\mu_i) + \sum_{i\in\Lambda} V(\mu_i, \rho_i) - C(\mu_i)$ where Ω is the set of agents buying life-insurance in equilibrium and Λ is the set of agents buying the pension. Total surplus can equally be expressed as the sum of risk premia across all trades that occur: $\sum_{i\in\Omega} r(\mu_i, \rho_i) + \sum_{i\in\Lambda} R(\mu_i, \rho_i)$.

4.5.1 Welfare Effects of Introducing Tying Contracts

This subsection shows formally that tying contracts are socially beneficial. A specific case of this more general result is the example in the introduction, where individual markets unravel completely and introducing tying contracts leads to full coverage, i.e. every agent being insured against both risks.

Proposition 4.4 *When firms can offer tying contracts and stand alone contracts, rather than being restricted to stand alone contracts:*

1. *total surplus is at least weakly higher and for some parameters strictly higher.*
2. *surplus of some agents is strictly lower for some parameters.*

This holds regardless of whether risk preferences are public or private information.

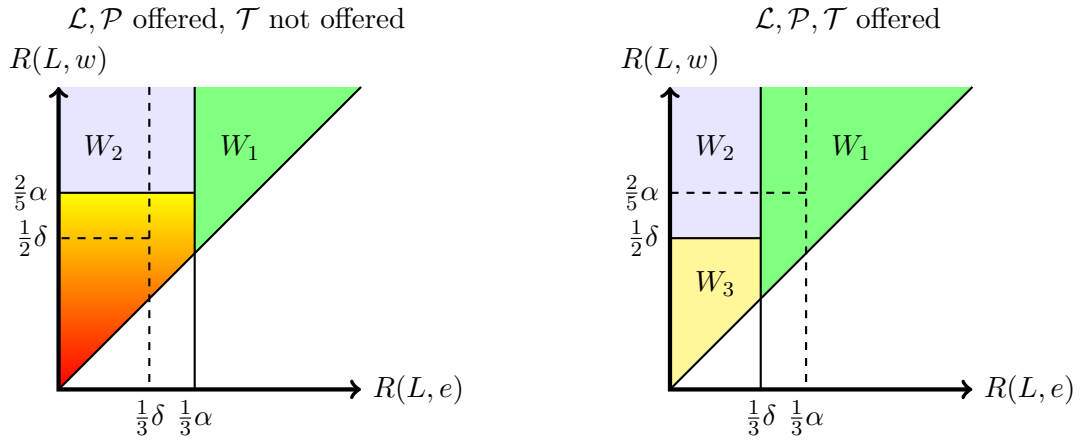
Proof: See Appendix 4.7.8.

Corollary 4.2 *There exist cases where individual markets unravel but introducing tying contracts prevents unravelling and hence leads to a substantial welfare improvement.*

Proof: See Appendix 4.7.9.

Sketch proof: Ordering equilibrium purchasing patterns from high to low welfare and comparing the conditions under which they are SPNE reveals that for all parameters, when tying contracts and stand-alone contracts are offered, equilibrium welfare is always at least as high as when only stand-alone contracts and no tying contracts are offered; for some parameters welfare is strictly higher, e.g. $\frac{1}{3}[c(L) + c(M) - 2c(H)] > R(L, e) > \frac{1}{3}\delta$.

Figure 4.6: Total surplus is weakly higher when tying contracts are offered



where $\alpha = [c(L) + c(M) - 2c(H)]$ and $\delta = (c(M) - c(H)) - (c(L) - c(M))$.

W_i is a purchasing pattern with higher total surplus than W_{i+1}.

Intuition: Tying contracts increase welfare because they cancel out cost differences at least partially and thus prevent or reduce unravelling. The example in the introduction showed that when cost differences cancel out *exactly*, then the private information on life-expectancy is irrelevant for the provider of the tying contract. In equilibrium, all agents purchase the tying contract. In this outcome, everyone is insured against both risks, thus total surplus is maximised. In the model introduced in section 4.3, the analogous case of costs cancelling out exactly is the case of linear costs ($\delta = 0$).

Importantly, tying contracts still increase welfare even if cost differences do *not* perfectly cancel out. In this case, there are two channels at work creating the beneficial welfare effect of tying contracts. The countervailing incentive channel and a channel that tying contracts reverse selection patterns.

The countervailing incentive mechanism continues to operate, though in a modified form. When cost differences do not cancel out exactly, private information remains relevant. Nevertheless, tying contracts still increase welfare as they reduce the marginal cost of purchasing a second type of insurance. To see this, compare the marginal consumer's incentives to deviate when all types are insured against both risks, i.e. all types buy the tying contract, to his incentives when tying contracts are not offered and all agents buy both stand alone contracts.

First, focus on the case when no tying contracts are offered and agents buy both stand alone contracts. L_e is the marginal consumer for the pension annuity. When weighing up the benefit of the pension against its cost, L_e will buy the pension if and only if:

$$V(L,e) > \mathcal{P} \quad where \quad \mathcal{P} = \frac{c(L) + c(M) + c(H)}{3} \tag{4.12}$$

Regardless of his pension decision, he continues to purchase life-insurance at $\mathcal{L} = \frac{c(L)+c(M)+c(H)}{3}$.

Second, consider the case where all agents purchase tying contracts. Then, L_e faces a choice between being insured against both risks through the tying contract or being insured against only one risk through a stand-alone life-insurance contract. L_e compares $v(L,e) + V(L,e) - \mathcal{T}$ to $v(L,e) - \mathcal{L}$. Thus, he will buy the tying contract if and only if:

$$V(L,e) > \mathcal{T} - \mathcal{L} \quad where \quad \mathcal{T} = \frac{2c(L) + 2c(M) + 2c(H)}{3} \; and \; \mathcal{L} = c(L) \tag{4.13}$$

The crucial point is that while the total cost of insuring both risks is the same in both cases ($\frac{2c(L)+2c(M)+2c(H)}{3}$), the marginal cost the agent faces in (4.13) is *always* below that in (4.12). Intuitively, one can think of T as offering *both* types of insurance, each at $E(c)$. Then, when L_e deviates to only stand alone insurance he not only trades of the cost $E(c)$ against the benefit of the pension (this applies equally in both (4.12) and (4.13)) but additionally also in (4.13) includes that life-insurance becomes more expensive (from $E(c)$ to $max(c)$). This second chanel exists because deviating from T reveals the agent to have a low life-expectancy and thus to be high cost for the life-insurance company. This second channel is only present in (4.13) and makes deviating less attractive. As deviating is less attractive in (4.13) than (4.12), the welfare maximising allocation, where everyone is insured against both risks, is a SPNE for a strictly larger set of parameters when tying contracts are offered than when they are not offered.

Similarly, we can view the tying contract as lowering the marginal payment an agent needs to make in order to buy a pension annuity. One can view T as providing life-insurance at the price of stand-alone life-insurance ($c(L)$). Then, T provides the pension for L_e at $2E(c) - c(L)$. This is less than the marginal payment when no tying contracts are offered ($E(c)$).

This logic applies not only to the purchasing pattern where all types insure against both risks, but also to the marginal types in other purchasing patterns, e.g. L_w in pattern 3 (tying contracts screen based on risk-preferences and expected cost).

An additional mechanism for how tying contracts increase welfare stems from tying contracts reversing selection patterns. Tying contracts are a new product, with a different underlying cost structure. When costs are concave, medium life-expectancy types have the highest expected cost for the tying contract. Thus, even if the new market for tying contracts unravels completely, i.e. only the highest cost agents are insured, then tying contracts still are beneficial as the highest cost agents are exactly the medium life-expectancy agents, who would have been completely uninsured under stand-alone contracts. For example, when risk premia are low relative to concavity, the SPNE without tying contracts has only the highest cost types insuring: L buy life-insurance, H buy a pension, M remain uninsured. With tying contracts, in the SPNE each contract also insures only those agents with the highest cost for it, but there is now an additional contract: L buy life-insurance, H buy the pension, M buy the tying contract. Thus, even though in each market only the highest cost agents purchase, tying contracts have increased

welfare because the highest cost agents for tying contracts were previously uninsured.

Distributional effect: While introducing tying contracts never reduces total surplus, some agents can be worse off. This applies to agents who purchase stand-alone insurance in both cases but who face a higher price for stand-alone insurance once tying contracts are introduced. Tying contracts have this effect when they are bought by medium life-expectancy types, who previously held down the cost for stand alone contracts. For example, there exist parameters when in the absence of tying contracts L_e, L_w, M_e, M_w purchase life-insurance and once tying contracts are introduced only L_e and L_w purchase life-insurance while M_e and M_w buy the tying contract. Then, the price of life-insurance has increased from $\mathcal{L} = \frac{c(L)+c(M)}{2}$ to $\mathcal{L} = c(L)$. Thus, L_e and L_w pay a higher price for the same coverage and thus are worse off.

4.5.2 Welfare Effects of Changes in the Information Structure

The literature on selection markets has increasingly focused on agents having private information on both expected cost and risk preferences, rather than just on the former. This has raised the question what effect such a change of informational assumptions has on welfare (Finkelstein and McGarry, 2006). In our setting, welfare is weakly lower when risk preferences are private information compared to when they are public information. With this comparison, we aim to contribute to the broader literature on multiple dimensions of private information in selection markets. The comparison yields an interesting benchmark to which the next subsection compares the benefits of tying contracts.

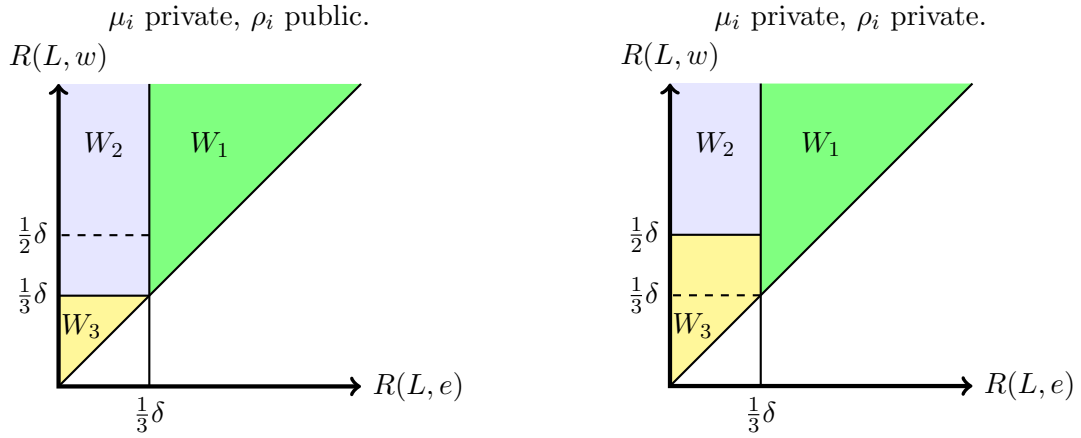
Proposition 4.5 *Welfare is weakly lower for all parameters, and strictly lower for some parameters, when expected costs and risk preferences are both private information compared to when only the former are private information and the latter are public information.*

Proof: See Appendix 4.7.10.

For the comparison in proposition 4.5, we hold the contract space fixed. The proposition is true regardless of which contract space assumption we hold fixed, i.e. whether only stand-alone contracts, or tying and stand-alone contracts, or only tying contracts are offered.

Sketch proof: Ordering equilibrium purchasing patterns from high to low welfare and comparing the conditions under which they are SPNE for the case where ρ_i is private information to the case where it is public information reveals that the level of insurance coverage coincides for all parameters except when $R(L,e) < \frac{1}{3}\delta$ and $\frac{1}{3}\delta < R(L,w) < \frac{1}{2}\delta$. For these parameters, the market unravels fully if ρ_i is private information and only partially if ρ_i is public information.

Figure 4.7: Total surplus is weakly lower when risk preferences are private information



where $\delta = (c(M) - c(H)) - (c(L) - c(M))$.

W_i is a purchasing pattern with higher total surplus than W_{i+1}.

Intuition: When risk preferences are public information, prices can depend on risk preferences and unravelling can occur only within a risk preference group. When risk preferences are private information, all agents face the same price. Then, when consumers who are low cost and at ease stop to buy insurance, this affects the price low cost worried consumers face and unravelling can therefore spread from consumers who are at ease with risk to worried consumers.

For example, suppose that when risk preferences are observed, purchasing pattern 3 is SPNE (tying contracts screen based on risk preferences and expected cost), i.e. all worried agents insure against both risks by buying a tying contract at $T(w) = \frac{2c(L)+2c(M)+2c(H)}{3}$ while among the relaxed agents only M_e buys a tying contract at $T(e) = 2c(M)$. The other relaxed agents insure against only one risk. L_e buys life-insurance at $\mathcal{L} = c(L)$ and H_e buys a pension at $\mathcal{P} = C(H) = c(L)$. For this to be a SPNE, L_w needs to prefer $T(w)$ to $\mathcal{L}(w) = c(L)$, i.e.

$$V(L,w) > T(w) - \mathcal{L} \tag{4.14}$$

Thus, the marginal cost of insuring the second risk is $T(w) - \mathcal{L} = 2E(c) - max(c)$.

Compare this to the case where risk preferences are private knowledge. Again, take purchasing pattern 3 as starting point (tying contracts screen based on risk preferences and expected cost), i.e. all worried agents and M_e buy the tying contract at $T = \frac{c(L)+2c(M)+c(H)}{2}$, L_e buys $\mathcal{L} = c(L)$, and H_e buys $\mathcal{P} = C(H) = c(L)$. The marginal agent in the tying contract is again L_w. L_w wants to purchase T if and only if:

$$V(L,w) > T - \mathcal{L} \tag{4.15}$$

Thus, the marginal cost of insuring the second risk is $T - \mathcal{L} = \frac{c(L)+2c(M)+c(H)}{2} - c(L)$.

Crucially, the marginal cost of insuring the second risk is *always* lower in (4.14) than in (4.15). This is because when the agent is known to be worried (4.14), purchasing T reveals

nothing on his expected costs, while when the agent could be at ease or worried (4.15), then purchasing T makes the agent higher expected cost for the tying contract as he now behaves identical to M_e but not to L_e and H_e. Therefore, the set of risk preferences which satisfy (4.14) is strictly larger than (4.15). This means that for some parameters (4.14) holds and (4.15) does not but the reverse is not true. When (4.15) is violated, the SPNE is a purchasing pattern with lower welfare. Hence, there exist parameters for which (4.14) results in strictly higher welfare than (4.15) while the reverse is never true.

4.5.3 Comparing Welfare Effects

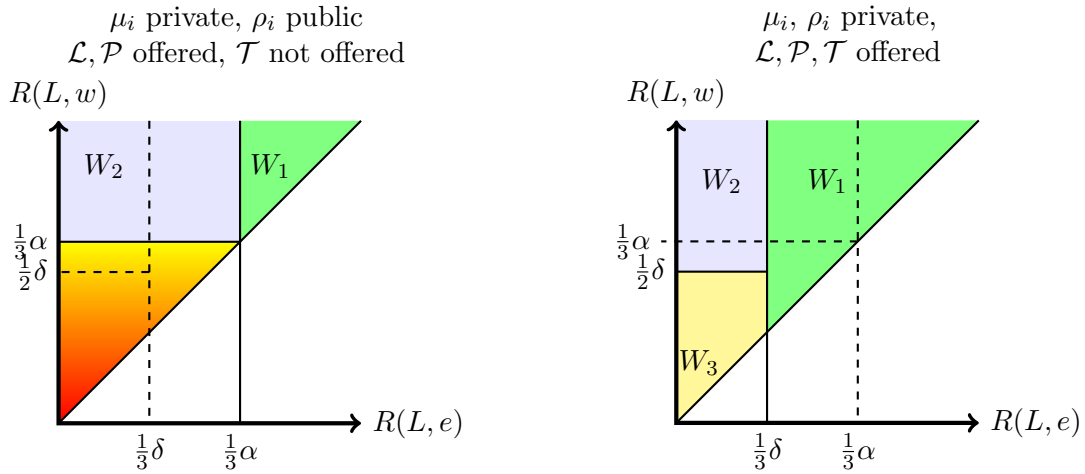
Comparing the welfare effects from the previous two sub-sections indicates the relative size of the welfare effect. The welfare benefit from introducing tying contracts more than offsets the welfare cost of risk preferences being private rather than public information. Thus, while private information on risk preferences makes unravelling of stand alone markets more likely, it also substantially magnifies the benefit of tying contracts.

Proposition 4.6 *The welfare gain from allowing tying contracts outweighs the welfare loss from risk preferences being private rather than public information. When expected costs and risk preferences are public information and tying contracts are offered, welfare is at least as high as when only expected costs are private information and not tying contracts are offered. For some parameters welfare is strictly higher.*

Proof: See Appendix 4.7.11.

Sketch proof: Ordering equilibrium purchasing patterns from high to low welfare and comparing the conditions under which they are SPNE proves the proposition: For all parameters, equilibrium welfare is at least weakly higher when there are two dimensions of private information and tying contracts are allowed compared to the case where there is one dimension of private information and tying contracts are banned. For some parameters welfare is strictly higher.

Figure 4.8: Comparing welfare effects



The welfare benefit of introducing tying contracts outweighs
the loss from risk preferences being private information.
where $\alpha = [c(L) + c(M) - 2c(H)]$ and $\delta = (c(M) - c(H)) - (c(L) - c(M))$.
W_i is a purchasing pattern with higher total surplus than W_{i+1}.

Intuition: The benefit of tying contracts outweighs the welfare cost of added private information due to two mechanisms: First, countervailing incentives make agents low cost which is always more beneficial than hiding high costs through a lack of information. Second, tying contracts present an additional product which increases welfare even if it is only purchased by the highest cost agents and thus rules out the SPNE which achieve the lowest total surplus.

First, to see the strength of countervailing incentives, focus on the case where all agents apart from L_e and H_e insure against both risks. L_e and H_e insure only against the risk for which they have high costs. When risk preferences are public information and no tying contracts are offered, the marginal person purchasing life-insurance (H_w) faces a cost of adding life-insurance to his pension of $\frac{c(L)+c(M)+c(H)}{3}$. When risk preferences are private information and no tying contracts are offered, H_w faces a cost of $\frac{2c(L)+2c(M)+c(H)}{5}$. As in section 4.5.2, this increase in costs of adding life-insurance arises because L now pools with L_e, M_e, L_w, M_w, H_w which is a pool with higher expected cost than pooling with L_w, M_w, H_w. When risk preferences are private information and tying contracts are offered, then H_w is the marginal person purchasing the tying contract and faces a cost of adding life-insurance to the pension, i.e. of buying T instead of P, as given by:

$$v(H, w) > T - \mathcal{P} \quad where \quad T = \frac{c(L) + 2c(M) + c(H)}{2} \; and \; \mathcal{P} = c(L) \tag{4.16}$$

and thus the cost of adding life-insurance to the pension is $T - \mathcal{P} = \frac{1}{2}[2c(M) + c(H) - c(L)]$. Crucially, the cost $T - \mathcal{P}$ is *always* below the cost of adding insurance in the other two cases outlined above. This is because insuring against both risks through the tying contract implies lower cost insurance also for the risk which the agent would have insured anyway. For H_w this is the pension. Hence, the logic described in section 4.5.1 is at work and outweighs the effect of changes in information outlined in section 4.5.2. As the marginal agent faces lower costs, the parameter space for which this purchasing pattern, or higher welfare patterns, are SPNE is larger. Hence, introducing tying contracts increases the set in the parameter space for which a given level of welfare or higher welfare is reached.

Second, tying contracts are a new product which at least insures those agents who have high cost for it. When costs are concave, medium life-expectancy types have the highest sum of expected costs and thus are the most expensive for the tying contract. Thus, M types will *always* buy T when it is offered. Importantly, when T is not offered and markets unravel, M types do not insure against either risk. Thus, introducing tying contracts increases welfare for those parameters which in the absence of tying contracts would result in the lowest welfare. Thus, a channel through which tying contracts increase welfare is that they make the SPNE which results in the lowest welfare impossible.

4.6 Conclusion

This paper showed that, in markets with adverse selection, tying contracts can prevent unravelling and thus have a substantial positive effect on welfare. We illustrated this effect for two insurance markets. When one characteristic (such as life-expectancy) is hidden information and causes agents to have high expected cost for one product (life insurance) but low expected cost for another (pension annuity), then for tying contracts cost differences partially cancel out. Thus, cost heterogeneity is lower for the tying contract than for individual insurance contracts. This can prevent unravelling. The largest welfare benefit occurs if without tying contracts both individual markets unravel but with tying contracts all agents are insured against both risks.

This paper raises broader research questions, which are interesting avenues for future research. For example: What types of risks should be pooled together in one insurance contract and which types should be insured separately?

This question can be viewed as investigating the optimal boundaries of an insurance contract or of an insurance firm. Our paper showed that these boundaries depend on whether costs are private or public information, not just on the level of costs. Alternative explanations for why tying contracts are observed in some markets are based on cost synergies. When providing one combined contract has lower administrative costs than providing two stand alone contracts, tying contracts arise as a cost saving device. This paper showed that tying contracts play an additional role. Even when no administrative costs and no cost synergies exist, tying contracts, i.e. wide boundaries of an insurance contract, are socially valuable as they mitigate the welfare loss from information asymmetries.

Insurance markets are often heavily regulated. For some insurance products, such as health insurance, the breadth of coverage is regulated to include certain services but exclude others. Our model indicates that defining the breadth of risks insured as too narrow foregoes the social benefit of wide definitions, which can alleviate selection issues. It should be possible for firms to offer contracts which combine life-insurance and pension annuity features. Our results indicate that tying contracts across these two markets could solve the unravelling problem at no cost, i.e. without requiring public funds.

4.7 Appendix

4.7.1 Proof of Statements in Section 4.4.1

Statement 1: Less unravelling occurs when levels of risk aversion are higher.

Proof: In Figure 4.1 we can see that an increase in risk aversion and the resulting increase in risk premia always weakly increases welfare by resulting in an equilibrium with less unravelling. Similarly, the corresponding table shows that there are four inequalities which determine the degree of unravelling: $r(H,e) < \frac{1}{3}[c(L) + c(M) - 2c(H)]$, $r(H,w) < \frac{2}{5}[c(L) + c(M) - 2c(H)]$, $r(M,e) < \frac{1}{2}[c(L) - c(M)]$, $r(M,w) < \frac{2}{3}[c(L) - c(M)]$. For all these inequalities, if risk premia are larger than their respective threshold, the equilibrium exhibits less unravelling. *q.e.d.*

Statement 2: More unravelling occurs when heterogeneity in costs is larger.

Proof: If there is no heterogeneity, i.e. if $c(L) = c(M) = c(H)$, then in equilibrium there is full coverage. If we keep $\frac{c(L)+c(H)}{2}$ centred on $c(M)$, but spread out $c(L)$ and $c(H)$, which is a mean preserving spread, then the thresholds which risk premia need to clear increase and for given risk premia there is more unravelling. *q.e.d.*

4.7.2 Equilibrium Purchasing Patterns

The equilibrium purchasing patterns for the case where firms can offer only stand-alone contracts and where $v(M,e) > v(H,w)$ were described in Figure 4.1. This Appendix shows the purchasing patterns when firms can offer only stand-alone contracts and $v(M,e) < v(H,w)$.

Table 4.6: Equilibrium Purchasing Patterns: only stand-alone contracts offered

	Individual market	Both markets	SPNE if and only if
W_1	Full coverage e: L: L, M: L, H: L w: L: L, M: L, H: L	e: L: $L+P$, M: $L+P$, H: $L+P$ w: L: $L+P$, M: $L+P$, H: $L+P$	$r(H,e) > \frac{1}{3}[c(L)+c(M)-2c(H)]$
W_2	e: L: L, M: L, H: $-$ w: L: L, M: L, H: L	e: L: L, M: $L+P$, H: P w: L: $L+P$, M: $L+P$, H: $L+P$	$r(H,e) < \frac{1}{3}[c(L)+c(M)-2c(H)]$ $r(M,e) > \frac{1}{5}[2c(L)+c(H)-3c(M)]$
W_3	e: L: L, M: $-$, H: $-$ w: L: L, M: L, H: L	e: L: L, M: $-$, H: P w: L: $L+P$, M: $L+P$, H: $L+P$	$r(H,e) < \frac{1}{3}[c(L)+c(M)-2c(H)]$ $r(M,e) < \frac{1}{5}[2c(L)+c(H)-3c(M)]$ $r(H,w) > \frac{1}{4}[2c(L)+c(M)-3c(H)]$
W_4	e: L: L, M: $-$, H: $-$ w: L: L, M: L, H: $-$	e: L: L, M: $-$, H: P w: L: L, M: $L+P$, H: P	$r(H,e) < \frac{1}{3}[c(L)+c(M)-2c(H)]$ $r(M,e) < \frac{1}{5}[2c(L)+c(H)-3c(M)]$ $r(H,w) < \frac{1}{4}[2c(L)+c(M)-3c(H)]$ $r(M,w) > \frac{2}{3}[c(L)-c(M)]$
W_5	Complete unravelling e: L: L, M: $-$, H: $-$ w: L: L, M: $-$, H: $-$	e: L: L, M: $-$, H: P w: L: L, M: $-$, H: P	$r(H,e) < \frac{1}{3}[c(L)+c(M)-2c(H)]$ $r(M,e) < \frac{1}{5}[2c(L)+c(H)-3c(M)]$ $r(H,w) < \frac{1}{4}[2c(L)+c(M)-3c(H)]$ $r(M,w) < \frac{2}{3}[c(L)-c(M)]$

Where we denoted types purchasing only a pension by P, only life-insurance by L, no insurance by $-$.

4.7.3 Proof of Proposition 4.1

This Appendix proves the existence of an equilibrium where tying contracts act as a pooling device (pattern 1). The purchasing behaviour is described by the table below:

Equilibrium Pattern 1	at ease	worried
Low life expectancy	tying contract	tying contract
Medium life expectancy	tying contract	tying contract
High life expectancy	tying contract	tying contract

The proof proceeds in two steps: First, we derive conditions which are necessary and sufficient for the existence of this type of equilibrium. Then, we show that it is possible that all these conditions are satisfied simultaneously.

Zero profit condition For this to be an equilibrium, the following zero-profit conditions

need to hold:

$$\mathcal{T} = \frac{2c(L) + 2c(M) + 2c(H)}{3} \tag{4.17}$$

Potentially binding incentive compatibility conditions on the equilibrium path:[19] Since tying contracts are chosen in equilibrium, we must have:

$$\mathcal{T} \leq \mathcal{L} + \mathcal{P} \tag{4.18}$$

Since all agents buy the tying contract, all potential deviations are to buying less insurance. We thus need to check the following conditions.

I) Deviating to buying only life insurance needs to be suboptimal. This means giving up the pension. Among the worried agents, L_w values the pension the least. This follows from $V(H, w) > V(M, w) > V(L, w)$. Among the at ease agents, L_e values the pension the least since $V(H, e) > V(M, e) > V(L, e)$. Comparing L_e and L_w, L_e values the pension less. Hence, L_e would be the first agent to deviate from the tying contract to only buying life insurance and thus give up the pension. Thus a potentially binding IC is:

$$V(L, e) + \mathcal{L} > \mathcal{T} \tag{4.19}$$

II) Deviating to only the pension needs to be suboptimal. This would mean giving up life-insurance. By a similar logic as above: among the worried agents H_w would be the first to give up life-insurance, among the at ease agents, H_e would be the first, and comparing H_w and H_e, H_e would be the first to give up life insurance. Thus, a potentially binding IC is:

$$v(H, e) + P > T \tag{4.20}$$

III) Deviating to no insurance needs to be suboptimal. Among agents with low life expectancy, L_e is the first to deviate. Among agents with medium life expectancy M_e is the first to deviate. Among agents with high life expectancy, H_e is the first to deviate. Hence, we have the following three potentially binding IC:

$$v(L, e) + V(L, e) > \mathcal{T} \tag{4.21}$$

$$v(M, e) + V(M, e) > \mathcal{T} \tag{4.22}$$

$$v(H, e) + V(H, e) > \mathcal{T} \tag{4.23}$$

[19] A full explanation of our approach via potentially binding incentive compatibility constraints is provided in Appendix 4.7.4.

Using weak symmetry, (4.21) becomes:

$$c(L) + r(L,e) + C(L) + R(L,e) > \frac{2C(L) + 2C(M) + 2C(H)}{3} \tag{4.24}$$

$$C(H) + r(L,e) + C(L) + R(L,e) > \frac{2C(L) + 2C(M) + 2C(H)}{3} \tag{4.25}$$

$$r(L,e) + R(L,e) > \frac{1}{3}\delta \tag{4.26}$$

Condition (4.3), which is the result of later steps in this derivation, is stricter. Hence, this condition here is not binding and does not enter the proposition.

By a similar logic, equation (4.23) is guaranteed to hold by (4.4).

Using weak symmetry, equation (4.22) can be written as:

$$2C(M) + r(M,e) + R(M,e) > \frac{2C(L) + 2C(M) + 2C(H)}{3} \tag{4.27}$$

$$r(M,e) + R(M,e) > \frac{1}{3}[2C(L) + 2C(H) - 4C(M)] \tag{4.28}$$

$$r(M,e) + R(M,e) > -\frac{2}{3}[(C(M) - C(L)) + (C(M) - C(H))] \tag{4.29}$$

which always holds due to our focus on concave costs $\delta \geq 0$.

No profitable deviation based on optimal response of consumers: For equilibrium type 1 to be SPNE we need to rule out that deviating to offering stand alone life-insurance could be profitable at any price $\mathcal{L}$.

- $\mathcal{L} < c(H)$ can never be profitable as every potential customer has higher costs.
- $\mathcal{L} = c(H)$: Then $\mathcal{L} + \mathcal{P} < \mathcal{T}$ and thus everyone buys the stand alone life-insurance with some consumers additionally purchasing a stand-alone pension annuity. Costs for the firm providing life-insurance are $\frac{c(L)+c(M)+c(H)}{3}$ and the firm makes a loss.
- $c(H) < \mathcal{L} < \mathcal{T} - \mathcal{P}$: On this range, all agents continue to purchase the stand alone life insurance, thus costs remain at $\frac{c(L)+c(M)+c(H)}{3}$. Thus, on this range there would exist a profitable stand alone life-insurance contract if and only if: $\mathcal{T} - \mathcal{P} > \frac{c(L)+c(M)+c(H)}{3}$. This simplifies to $\frac{c(L)+c(M)+c(H)}{3} > c(L)$ which never holds. Hence, no additional conditions are needed to rule out that a profitable $\mathcal{L}$ exists on this range.
- $\mathcal{L} = \mathcal{T} - \mathcal{P}$. Note that in equilibrium all agents buy the tying contract. An agent of type (μ, ρ) deviates to buying the stand alone life insurance if and only if $\mathcal{T} - \mathcal{L} > V(\mu, \rho)$. If $\mathcal{L} = \mathcal{T} - \mathcal{P}$ this becomes $C(H) > V(\mu, \rho)$. Thus, neither H_e nor H_w would deviate. Hence, costs to the firm are $\frac{c(L)+c(M)}{2}$ which is always less than $\mathcal{L} = \mathcal{T} - \mathcal{P} = \frac{2}{3}c(M) + \frac{2}{3}c(H) - \frac{1}{3}c(L)$.

- $\mathcal{L} = \frac{c(L)+c(M)}{2}$. A type (μ, ρ) would deviate to buying a tying contract if and only if $\mathcal{T} - \mathcal{L} > V(\mu, \rho)$ which here becomes $\frac{1}{6}c(L) + \frac{1}{6}c(M) + \frac{2}{3}c(H) > V(\mu, \rho)$. Since the left term in this inequality does not exceed $c(M)$, neither M_e nor M_w will deviate. Thus costs are $c(L)$ and firms make a loss.
- $\mathcal{L} = c(L)$. Then $\mathcal{T} - \mathcal{L} > V(\mu, \rho)$ becomes $\frac{2c(M)+2c(H)-c(L)}{3} > V(\mu, \rho)$. As the first type to deviate would be L_e this condition can be simplified to $R(L, e) > \frac{1}{3}\ \delta$. By symmetry we also have $r(H, e) > \frac{1}{3}\ \delta$. These are exactly the conditions for equilibrium type 1 to be SPNE given in the main text: (4.3) and (4.4).
- $\mathcal{L} > c(L)$. The conditions above rule out that $\mathcal{L} = c(L)$ attracts any consumers. Thus, higher prices will also not attract any consumers and thus not result in strictly positive profits.

This concludes the proof of part (2) of Proposition 4.1.

We prove part (1) of Proposition 4.1 by example. Consider the following parameters:

Table 4.7: Parameters

Life expectancy	WTP at ease	WTP worried	expected cost
Low	$v(L, r) = 7$ $V(L, r) = 2.5$	$v(L, w) = 7.5$ $V(L, w) = 4.5$	$c(L) = 6$ $C(L) = 1$
Medium	$v(M, r) = 5.2$ $V(M, r) = 5.2$	$v(M, w) = 5.6$ $V(M, w) = 5.6$	$c(M) = 5$ $C(M) = 5$
High	$v(H, r) = 2.5$ $V(H, r) = 7$	$v(H, w) = 4.5$ $V(H, w) = 7.5$	$c(H) = 1$ $C(H) = 6$

These parameters are consistent with all restrictions implied by our setting and satisfy the conditions for the existence of equilibrium type 1. Hence, this example proves that there exist parameters for which an equilibrium type 1 exists. *q.e.d.*

4.7.4 Proof of Proposition 4.2

This Appendix proves the existence of an equilibrium where tying contracts screen based on expected cost (pattern 2). The purchasing behaviour is described by the table below:

Equilibrium Pattern 2	at ease	worried
Low life expectancy	only life insurance	only life insurance
Medium life expectancy	tying contract	tying contract
High life expectancy	only pension	only pension

The proof proceeds in two steps: First, we derive conditions which are necessary and sufficient for the existence of this type of equilibrium. Then, we show that it is possible

that all these conditions are satisfied simultaneously.

Zero profit conditions For this to be an equilibrium, the following zero-profit conditions need to hold:

$$\mathcal{L} = c(L) \tag{4.30}$$
$$\mathcal{P} = C(H) \tag{4.31}$$
$$\mathcal{T} = c(M) + C(M) \tag{4.32}$$

Potentially binding incentive compatibility conditions on the equilibrium path: With multiple dimensions of private information, we need to think about all incentive compatibility constraints (ICC). In this case, there are 24 ICC. For this particular type of equilibrium, we can reduce the number of relevant ICC by the following arguments:

For a tying contract to be bought in equilibrium we must have:

$$\mathcal{T} \leq \mathcal{L} + \mathcal{P} \tag{4.33}$$

Moreover, we can focus on the set of agents buying a particular contract and ask: "Who would be the first to deviate in a given direction?", then focusing on this ICC.

L_e and L_w only buy life-insurance. The deviation to buying more insurance is relevant for L_w. This then implies that L_e will not want to deviate either. L_w could buy more insurance by either buying a stand-alone pension, or by switching from stand-alone life insurance to a tying contract. Since we have $\mathcal{T} < \mathcal{L} + \mathcal{P}$, L_w would not add a stand alone pension but might switch to a tying contract. Hence, a potentially binding incentive compatibility constraint is that L_w prefers stand alone life-insurance to the tying contract:

$$v(L, w) - \mathcal{L} > v(L, w) + V(L, w) - \mathcal{T} \tag{4.34}$$

The deviation to less insurance is relevant for L_e and then implies that L_w will not want to deviate in this direction either. As life insurance is available at a price equal to expected cost, L_e will always buy stand alone life-insurance. Essentially, the zero profit condition guarantees that this ICC holds.

Similarly, H_e and H_w only buy the pension. The relevant deviations are for H_w to deviate to the tying contract and for H_e to deviate to not buying any insurance. The latter is again guaranteed to hold by the zero profit condition. The former is a potentially binding ICC: H_w needs to prefer the stand alone pension to the tying contract:

$$V(H, w) - \mathcal{P} > V(H, w) + v(H, w) - \mathcal{T} \tag{4.35}$$

Both M types insure against both risks by buying a tying contract. Since $\mathcal{T} \leq \mathcal{L} + \mathcal{P}$

individual contracts insure against only one risk at more than half the cost. Hence, M types never want to deviate.

Thus, there are three potentially binding incentive compatibility constraints:

$$v(L,w) - \mathcal{L} > v(L,w) + V(L,w) - \mathcal{T} \tag{4.36}$$

$$V(H,w) - \mathcal{P} > V(H,w) + v(H,w) - \mathcal{T} \tag{4.37}$$

$$\mathcal{T} \leq \mathcal{L} + \mathcal{P} \tag{4.38}$$

The zero-profit conditions ensure that (4.38) always holds and simplify the other two conditions to:

$$c(M) + C(M) - c(L) > V(L,w) \tag{4.39}$$

$$c(M) + C(M) - C(H) > v(H,w) \tag{4.40}$$

Assuming weak symmetry and no further assumptions, this becomes:

$$\big(C(M) - C(L)\big) + \big(C(M) - C(H)\big) > R(L,w) \tag{4.41}$$

and:

$$\big(C(M) - C(L)\big) + \big(C(M) - C(H)\big) > r(H,w) \tag{4.42}$$

No profitable deviation based on optimal response of consumers: It is never profitable for firms to offer a contract with prices above equilibrium prices as no consumer will buy this new contract. To establish that this pattern is indeed an SPNE, we need to rule out that offering lower prices is profitable, where profitability is calculated based on consumers responding optimally to the new contract.

Lower $\mathcal{T}$ could be profitable when a) it attracts L_w, H_w and when b) it is even lower and in addition to L_w, H_w also attracts L_e, H_e . We focus first on a), then turn to b).

a) L_w buys the tying contract if and only if:

$$V(L,w) > \mathcal{T} - \mathcal{L} \tag{4.43}$$

A tying contract attracting L_w and H_w is profitable if and only if:

$$T > \frac{c(L) + 2c(M) + c(H) + C(L) + 2C(M) + C(H)}{4} \tag{4.44}$$

A condition for SPNE is that (4.43) and (4.44) never hold simultaneously, i.e. that:

$$V(L,w) + \mathcal{L} < \frac{c(L) + 2c(M) + c(H) + C(L) + 2C(M) + C(H)}{4} \tag{4.45}$$

which solves to (4.5). Applying the same approach to H_w solves to (4.7).

b) L_e buys the tying contract if and only if:

$$V(L,e) > \mathcal{T} - \mathcal{L} \tag{4.46}$$

A tying contract which attracts L_e, L_w, H_e, H_w is profitable if and only if:

$$T > \frac{2c(L + 2c(M) + 2c(H) + 2C(L) + 2C(M) + 2C(H)}{6} \tag{4.47}$$

A condition for SPNE is that (4.46) and (4.47) never hold simultaneously. This is identical to a condition in Appendix 4.7.5 where it is shown to simplify to the following two conditions: $R(L,r) < \frac{1}{3}\delta$ and $r(H,e) < \frac{1}{3}\delta$, i.e. to equations (4.6) and (4.8).

SPNE also requires that offering stand-alone life-insurance at a lower price must be suboptimal. While the formal argument that this condition holds is long, it adds no further conditions for the SPNE.

- $\mathcal{L} < c(H)$ is never profitable since all types have higher costs.
- $\mathcal{L} = c(H)$: Then $\mathcal{L} + \mathcal{P} < \mathcal{T}$. Thus, everyone buys the stand alone life insurance contract. This leads to costs above price and makes this contract loss making.
- $c(H) < \mathcal{L} < v(H,e)$: There does not exist a profitable $\mathcal{L}$ in this range if and only if costs exceed $v(H,e)$. This can be written as $\frac{c(L)+c(M)-2c(H)}{3} > r(H,e)$, which we can also write as $C(H) + C(M) - 2C(L) > 3r(H,e)$. After algebraic manipulation, this becomes: $\frac{1}{3}\delta + \frac{1}{3}[2C(H) - C(M) - C(L)] > r(H,e)$. Since $C(H) > C(M) > C(L)$ we have $[2C(H) - C(M) - C(L)] > 0$. Thus, this condition is less tight than (4.8).
- $v(H,e) < \mathcal{L} < v(H,w)$: Costs are $\frac{2c(L)+2c(M)+c(H)}{5}$. Such a contract would be profitable only if there exists a price above costs but below $v(H,w)$, i.e. $\frac{2c(L)+2c(M)+c(H)}{5} < r(H,w)+c(H)$. After algebraic manipulation this becomes: $\frac{2}{5}\delta + \frac{1}{5}\kappa < r(H,w)$ where $\kappa = 4c(L) - 2c(M) - 2c(H)$. Note that $\kappa > 2\delta$, which follows from $c(L) > c(M)$, $\frac{2}{5}\delta + \frac{1}{5}\kappa > \frac{4}{5}\delta$. Hence, a price in this range would be profitable only if $r(H,w)$ exceeded $\frac{2}{5}\delta + \frac{1}{5}\kappa$ and thus also exceeded $\frac{4}{5}\delta$. Condition (4.7) is even tigher and thus already rules this out. Hence, this adds no further restriction on SPNE.
- $v(H,w) < \mathcal{L} < \mathcal{T} - \mathcal{P}$: Costs are $\frac{c(L)+c(M)}{2}$ or even higher. These costs are above $\mathcal{T} - \mathcal{P}$ and hence setting a price in this range is never profitable.
- $\mathcal{L} = \mathcal{T} - \mathcal{P}$. Suppose that L_e, L_w, M_e, M_w deviate to buying the life insurance contract. Then costs are $\frac{c(L)+c(M)}{2}$. In all other possible cases, costs are even higher. $\mathcal{L} = \mathcal{T} - \mathcal{P} = 2c(M) - c(L) < \frac{c(L)+c(M)}{2}$. Hence, this contract is loss making.
- $\mathcal{L} = \frac{c(L)+c(M)}{2}$: A type (μ,ρ) deviates from the tying contract to stand alone life insurance if and only if: $\mathcal{T} - \mathcal{L} > V(\mu,\rho)$ which can also be written as $\frac{3}{2}C(M) -$

$\frac{1}{2}C(H) > V(\mu, \rho)$. Since $V(M, w) > V(M, e) > C(M)$ and since $C(H) > C(M)$ the left hand side is less than $C(M)$ and the right hand side is larger than $C(M)$. Hence, the inequality never holds for M_e or M_w. Thus, only the L types buy. Which means that costs are $c(L)$ and thus above the price. Hence, this contract is not profitable.

- $\mathcal{L} = c(L)$: This is the equilibrium price. Here, firms make zero profit.

This concludes the proof of Part (2) of Proposition 4.2.

We prove part (1) of Proposition 4.2 by example. Consider the following set of parameters:

Table 4.8: Parameters

Life expectancy	WTP at ease	WTP worried	expected cost
Low	$v(L, e) = 7$ $V(L, e) = 1.5$	$v(L, w) = 7.5$ $V(L, w) = 2.3$	$c(L) = 6$ $C(L) = 1$
Medium	$v(M, e) = 5.2$ $V(M, e) = 5.2$	$v(M, w) = 5.6$ $V(M, w) = 5.6$	$c(M) = 5$ $C(M) = 5$
High	$v(H, e) = 1.5$ $V(H, e) = 7$	$v(H, w) = 2.3$ $V(H, w) = 7.5$	$c(H) = 1$ $C(H) = 6$

These parameters are consistent with all restrictions implied by our setting[20] and satisfy the conditions for the existence of equilibrium type 2. Hence, this example proves that there exist parameters for which an equilibrium type 2 exists. *q.e.d.*

4.7.5 Proof of Proposition 4.3

This Appendix proves the existence of a sub-game perfect Nash equilibrium where tying contracts screen based on risk preferences and expected costs (pattern 3). The purchasing behaviour is described by the table below:

Equilibrium Pattern 3	at ease	worried
Low life expectancy	only life insurance	tying contract
Medium life expectancy	tying contract	tying contract
High life expectancy	only pension	tying contract

The proof proceeds in two steps: First, we derive conditions which are necessary and sufficient for the existence of this type of equilibrium. Then, we show that it is possible that all these conditions are satisfied simultaneously.

[20] For example, for risk averse agents WTP needs to exceed their expected cost to the insurer. That is the case for these parameters.

Zero-Profit conditions: For this to be an equilibrium, the following zero-profit conditions need to hold:

$$\mathcal{L} = c(L) \tag{4.48}$$
$$\mathcal{P} = C(H) \tag{4.49}$$
$$\mathcal{T} = \frac{c(L) + 2c(M) + c(H) + C(L) + 2C(M) + C(H)}{4} \tag{4.50}$$

Potentially binding incentive compatibility conditions on the equilibrium path: There is a large number of incentive compatibility constraints. Within the set of agents buying a contract, we can focus on those agents who would be the first to deviate in a specific direction. This yields eight potentially binding incentive compatibility constraints:

$$v(L,e) - \mathcal{L} > v(L,e) + V(L,e) - \mathcal{T} \tag{4.51}$$
$$V(H,e) - \mathcal{P} > v(H,e) + V(H,e) - \mathcal{T} \tag{4.52}$$
$$\mathcal{T} < \mathcal{L} + \mathcal{P} \tag{4.53}$$
$$v(L,w) + V(L,w) - \mathcal{T} > v(L,w) - \mathcal{L} \tag{4.54}$$
$$v(M,e) + V(M,e) - \mathcal{T} > v(M,e) - \mathcal{L} \tag{4.55}$$
$$v(H,w) + V(H,w) - \mathcal{T} > V(H,w) - \mathcal{P} \tag{4.56}$$
$$v(M,e) + V(M,e) - \mathcal{T} > V(M,e) - \mathcal{P} \tag{4.57}$$
$$v(M,e) + V(M,e) - \mathcal{T} > 0 \tag{4.58}$$

Assuming weak symmetry and that all groups are equally large, risk preferences for worried agents do not reverse fundamentals, no extra assumptions on risk preferences of relaxed agents, we can make the following simplifications:

Combining equations (4.51) and (4.54) and using the zero-profit conditions yields:

$$R(L,w) > \frac{C(M) - C(L)}{2} + \frac{C(M) - C(H)}{2} > R(L,e) \tag{4.59}$$

Hence, a necessary but not sufficient condition for the existence of this equilibrium is that costs are concave in the sense that: $C(M) - C(L) > C(H) - C(M)$, i.e. $\delta > 0$.

Equation (4.55) becomes:

$$R(M,e) > \frac{C(L) - C(H)}{2} \tag{4.60}$$

which always holds. Hence, given our specific assumptions this incentive compatibility constraint is not binding.

Similar to the cases above, equations (4.52) and (4.56) yield:

$$r(H,w) > \frac{C(M)-C(L)}{2} + \frac{C(M)-C(H)}{2} > r(H,e) \tag{4.61}$$

while (4.57) is:

$$r(M,e) > \frac{C(L)-C(H)}{2} \tag{4.62}$$

which always holds. Equation (4.58) can be rewritten as:

$$r(M,e) + R(M,e) > -\Big[\frac{C(M)-C(L)}{2} + \frac{C(M)-C(H)}{2}\Big] \tag{4.63}$$

Note that if costs are concave, then this equation always holds as the right hand side is always negative, while the left hand side must always be positive due to risk aversion. Since by the conditions (4.59) and (4.61) concavity is already a necessary but not sufficient condition for the existence of this equilibrium, this condition adds no restriction on the existence of the equilibrium.

Equation (4.53) becomes:

$$C(L) + 2C(M) + C(H) < 4C(H) \tag{4.64}$$

which always holds.

Hence, given our specific assumptions, this Nash equilibrium exists if and only if the following two conditions both hold:

$$R(L,w) > \frac{C(M)-C(L)}{2} + \frac{C(M)-C(H)}{2} > R(L,e) \tag{4.65}$$

$$r(H,w) > \frac{C(M)-C(L)}{2} + \frac{C(M)-C(H)}{2} > r(H,e) \tag{4.66}$$

No profitable deviation based on optimal response of consumers: For this Nash equilibrium to also be a sub-game perfect Nash equilibrium (SPNE), additional conditions apply. For SPNE, we require that any price other than the equilibrium prices leads to non-positive profits where profits are calculated based on consumers responding optimally to this new price. For a firm, it is never optimal to deviate to higher $\mathcal{T}, \mathcal{L}, \mathcal{P}$. Since rivals insurance companies continue to offer the contract at the equilibrium price, this will never attract consumers and thus makes zero-profit.

In our selection market setting it is possible that cutting $\mathcal{T}$ marginally leads to losses, but that a larger cut to $\mathcal{T}' < \mathcal{T}$ is profitable as $\mathcal{T}'$ attracts lower cost customers to buy the tying contract. For this type of equilibrium, the only lower cost consumers a tying

contract could attract are L_e. L_e start buying the tying contract if and only if:

$$V(L,e) > \mathcal{T}' - \mathcal{L} \tag{4.67}$$

This contract is profitable if and only if costs are below the price, i.e.

$$\mathcal{T}' > \frac{2c(L) + 2c(M) + 2c(H)}{3} \tag{4.68}$$

A condition for SPNE is thus that (4.67) and (4.68) never hold simultaneously. This is the case if and only if:

$$V(L,e) + \mathcal{L} < \frac{2c(L) + 2c(M) + 2c(H)}{3} \tag{4.69}$$

which simplifies to the conditions:

$$R(L,e) < \frac{1}{3}\delta \tag{4.70}$$

and

$$r(H,e) < \frac{1}{3}\delta \tag{4.71}$$

Having established conditions under which it is not profitable to offer a cheaper tying contract, we now establish conditions under which it is not optimal to deviate to a lower price for the life-insurance contract. By symmetry, it is then also not profitable to deviate to a lower price for the pension annuity. For equilibrium type 3, ruling out deviations to lower $\mathcal{L}$ does not result in additional existence conditions, i.e. given the assumptions and conditions above, there does not exist a $\mathcal{L}' < \mathcal{L}$ which is profitable. We show this, by ruling out different ranges of $\mathcal{L}'$.

- $\mathcal{L} < c(H)$ is definitely loss making since $c(L) > c(M) > c(H)$.
- $\mathcal{L} = c(H)$: Then $\mathcal{L} + \mathcal{P} < \mathcal{T}$ and thus all types who buys the tying contract in conjectured SPNE switch to buying the life insurance contract. H_e also buys the life insurance contract since by risk aversion $v(H,e) > c(H)$. Hence, at this price everyone buys the life insurance contract. Thus, costs are $\frac{c(L)+c(M)+c(H)}{3} > c(H)$ and therefore firms make a loss.
- $c(H) < \mathcal{L} < \mathcal{T} - \mathcal{P}$: The life insurance contract is bought by L_e and L_w, M_e, M_w, H_w. Whether it is bought by H_e depends on his preferences. Suppose that H_r does buy the life-insurance contract. The costs are $\frac{c(L)+c(M)+c(H)}{3}$ which exceeds $\mathcal{T} - \mathcal{P}$ and is thus loss making. If H_e does not buy the contract, costs are even higher.
- $\mathcal{L} \geq \mathcal{T} - \mathcal{P}$. The cheapest way to insure both risks is the tying contract. An agent of type (μ, ρ) will deviate from $\mathcal{T}$ to buying only $\mathcal{L}$ if and only if $\mathcal{T} - \mathcal{L} > V(\mu, \rho)$. At $\mathcal{L} = \mathcal{T} - \mathcal{P}$ this becomes $C(H) > V(\mu, \rho)$ which cannot hold for H_w. Suppose

that L_w, L_e, M_w, M_e deviate to buying the life insurance contract at this price. Then, costs are $\frac{c(L)+c(M)}{2}$ which exceeds $\mathcal{L} = \mathcal{T} - \mathcal{P}$. If fewer types deviate, costs are even higher. Hence, $\mathcal{L} = \mathcal{T} - \mathcal{P}$ is never profitable.

- $\mathcal{L} = \frac{c(L)+c(M)}{2}$. Types choose this contract if and only if $\mathcal{T} - \mathcal{L} > V(\mu, \rho)$ which becomes $\frac{1}{2}C(M) + \frac{1}{2}C(L) > V(\mu, \rho)$. Since $V(M, w) > V(M, e) > C(M) > C(L)$, neither M_e nor M_w want to buy the contract at this price. Thus, costs to the firm at this price are $c(L)$ and the firm is making a loss.
- $\mathcal{L} = c(L)$. This is the equilibrium price. Here, firms make zero-profits.

Hence, moving from Nash equilibrium to a sub-game perfect Nash equilibrium imposes the additional conditions (4.70) and (4.71) and no further conditions. Combining these conditions with the NE conditions yields the conditions for SPNE existence given in equation (4.9) and (4.10).This concludes the proof of Part (2) of Proposition 4.3.

We prove part (1) of Proposition 4.3 by example. Consider the following set of parameters:

Table 4.9: Parameters

Life expectancy	WTP at ease	WTP worried	expected cost
Low	$v(L, e) = 7$ $V(L, e) = 1.5$	$v(L, w) = 7.5$ $V(L, w) = 4.5$	$c(L) = 6$ $C(L) = 1$
Medium	$v(M, e) = 5.2$ $V(M, e) = 5.2$	$v(M, w) = 5.6$ $V(M, w) = 5.6$	$c(M) = 5$ $C(M) = 5$
High	$v(H, e) = 1.5$ $V(H, e) = 7$	$v(H, w) = 4.5$ $V(H, w) = 7.5$	$c(H) = 1$ $C(H) = 6$

These parameters are consistent with all restrictions implied by our setting (e.g. agents being risk averse means that their WTP must exceed costs to the insurer) and satisfy the conditions for the existence of equilibrium type 3. Hence, this example proves that there exist parameters for which an equilibrium type 3 exists. *q.e.d.*

4.7.6 Proof of Theorem 4.1: Existence and Uniqueness

This proof proceeds in two steps: First, we show that no equilibrium other than type 1, 2, or 3 can ever exist. Then, we show that exactly one of these three types of equilibria always exists.

Lemma 4.7.6.1 *No equilibrium other than type 1, 2, 3 can ever exist.*

First, we establish two features which any equilibrium must satisfy. Then, we show that only equilibrium types 1, 2, 3 satisfy these features.

Sublemma 4.7.6.1 *In any equilibrium, the extreme life-expectancy types (L_e, L_w, H_e, H_w) are always insured at least against the risk for which they have high expected cost.*

Proof of Sublemma 4.7.6.1 Proof by contraction. Take any candidate equilibrium where low life-expectancy types (L) do *not* buy life insurance. There are two subcases:

Case 1: Other types buy life-insurance:

If some other type buys life insurance, then this reveals that this other agent (call him type J) has $v(J, \rho) > \mathcal{L}$ where $\mathcal{L}$ denotes the price of life-insurance and ρ could be worried or relaxed. However, since low life-expectancy types (L-types) have the highest WTP for life-insurance ($v(L, w) > v(J, \rho)\ \forall\ J$), L-types must then either: be buying a tying contract that implies life insurance, or they would want to deviate and also start buying life-insurance. Hence, either low life-expectancy types are already insured via the tying contract, or the candidate equilibrium breaks down since L types would want to deviate, i.e. the incentive compatibility constraint for L types is violated.

Case 2: No other type buys life insurance:

In this case it must be profitable for a firm to offer stand alone life insurance at a price $\mathcal{L}$ where $\mathcal{L} \geq c(L)$ because the L-types' willingness-to-pay (WTP) exceeds expected cost: $v(L, w) > v(L, r) > c(L)$. Competition then drives the price to $\mathcal{L} = c(L)$. If L types are not already insured against an early death through a tying contract, they would start buying this stand alone life insurance contract. Hence, there are profitable deviations for firms, ruling this out as equilibrium. If other types also started buying stand-alone life insurance at this price, this would only increase profits since these other types have lower expected costs than L-types.

The same logic applies to high life-expectancy (H) agents who will always be insured at least against longevity, i.e. at least buy the pension. *q.e.d.*

Sublemma 4.7.6.2 *The medium life-expectancy types always buy the tying contract in equilibrium.*

Proof of Sublemma 4.7.6.2 This proof uses the assumption that we focus solely on cases with concave costs and on strong symmetry.

a)Suppose neither L nor H types buy the tying contract.
Then competition will drive the price of the tying contract $\mathcal{T}$ down to equal the expected cost of M types, i.e. $\mathcal{T} = c(M) + C(M)$. In this case, both medium life expectancy types (M_e and M_w) buy the tying contract.

b)Suppose L_w, H_w also buy the tying contract, but L_e, H_e do not.

Then by concavity of parameters, expected costs of those types buying the tying contract is now lower than in case a). Hence, competition drives the price $\mathcal{T}$ to lower values than in a), i.e. $\mathcal{T} = \frac{c(L)+2c(M)+c(H)+C(L)+2C(M)+C(H)}{4} < c(M) + C(M)$ where the inequality follows from the concavity assumption. Thus, M types get full insurance for a lower price than in a), i.e. for less than expected cost. Therefore, it is optimal for M types to buy the tying contract.

c)Suppose all types buy the tying contract, i.e. also L_e, L_w, H_e, H_w By concavity expected costs in $\mathcal{T}$ has fallen even further than in case b). Thus, by the same logic as in case b): M types optimally always buy the tying contract. *q.e.d.*

Proof of Lemma 4.7.6.1:
From Sublemma 4.7.6.2 we know that M_e and M_w always buy the tying contract. Since the tying contract is purchased in equilibrium, we must have $\mathcal{L} + \mathcal{P} > \mathcal{T}$. Hence, nobody insures against both risks by buying two individual contracts. All agents wanting to insure against both risks buy the tying contract.

From Sublemma 4.7.6.1 we know that L_e, L_w, H_e, H_w never purchase zero insurance but insure at least against the risk for which they are high expected cost. We also know that L_w has a higher WTP for insurance than L_e. Hence, if only one agent deviates from stand alone life-insurance to the tying contract, it must be L_w. Then by symmetry we must also have H_w buying the tying contract. The alternative is that both L_e and L_w deviate to buying the tying contract. Then by symmetry, H_e and H_w must also deviate to buying the tying contract. Hence, there are only three possible types of equilibria left to consider:

I)L_e, L_w insure only against an early death, i.e. only buy life insurance. By symmetry H_e and H_w buy only the pension. This corresponds to equilibrium type 2: tying contract screen based on expected cost.

II)L_w buys the tying contract; L_e continues to buy only life insurance. By symmetry H_w also buys the tying contract and H_e buys only the pension. This is equilibrium type 3: tying contracts screen based on risk preferences and expected cost.

III)L_e and L_w both buy the tying contract. By symmetry H_e and H_w also buy the tying contract. This corresponds to equilibrium type 1: tying contracts act as pooling device.

Hence, we have shown that no other candidate equilibria than types 1, 2, 3 exist. This concludes the proof of Lemma 4.7.6.1. *q.e.d.*

Lemma 4.7.6.2 *Of the types of equilibria 1, 2, 3, exactly one always exists.*

Sublemma 4.7.6.3 *If equilibrium type 3 does not exist because concavity is too low (δ too low), then equilibrium type 1 must exist.*

Proof: Recall that equilibrium type 3 exists if and only if $2R(L,w) > \delta > 3R(L,e)$ where $\delta = [C(M) - C(L)] + [C(M) - C(H)]$; also recall that equilibrium type 1 exists if and only if $\delta < 3R(L,e)$. Equilibrium type 3 does not exist because concavity is too low if $\delta < 3R(L,e)$ which is exactly the condition for existence of equilibrium type 1. Hence, whenever equilibrium type 3 does not exist because concavity is too low, equilibrium type 1 exists. *q.e.d.*

Sublemma 4.7.6.4 *If equilibrium type 3 does not exist because concavity is too large, then equilibrium type 2 must exist.*

Proof: Recall that equilibrium type 3 exists if and only if $2R(L,w) > \delta > 2R(L,e)$; also recall that equilibrium type 2 exists if and only if $\delta > R(L,w)$. Equilibrium type 3 does not exist because concavity is too high if $\delta > 2R(L,w)$. Then we must also have $\delta > R(L,w)$ and hence equilibrium type 2 exists. *q.e.d.*

Combining the Sublemmata: We either have equilibrium type 3 or one of the other two types. Hence, there always exists exactly one of these equilibria. This concludes the proof of Lemma 4.7.6.2. *q.e.d.*

All these three types of equilibria have the properties stated in Theorem 4.1. Thus, Lemma 4.7.6.1 and Lemma 4.7.6.2 combined prove Theorem 4.1. *q.e.d.*

4.7.7 Proof of Corollary 4.1

Corollary 4.1 consists of four statements. We prove each one in turn.

Statement 1: *Every agent is insured against at least one risk.*

Proof: By Sublemma 4.7.6.1 L_e, L_w, H_e, H_w are insured against at least one risk. By Sublemma 4.7.6.2, M_e, M_w are insured against both risks. Thus, all agents are insured against at least one risk. *q.e.d.*

Statement 2: *Medium life-expectancy types always buy the tying contract and thus insured against both risks.*

Proof: Statement 2 is analogous to Sublemma 4.7.6.2. Thus, the proof of Sublemma 4.7.6.2 also applies here. *q.e.d.*

Statement 3: *The extreme types are always insured at least against the risk for which they are high cost.*

Proof: Statement 3 is analogous to Sublemma 4.7.6.1. Thus, the proof of Sublemma 4.7.6.1 also applies here. *q.e.d.*

Statement 4: *Worried agents buy at least as much insurance as agents who are at ease with risk.*

Proof: By Lemma 4.7.6.1 any equilibrium is either type 1,2, or 3. All three types have the property that worried agents buy at least as much insurance as agents who are at ease with risk where the comparison is made holding life-expectancy constant. *q.e.d.*

4.7.8 Proof of Proposition 4.4: Effect of Introducing Tying Contracts

The proof shows that the purchasing pattern which results in the highest total surplus exists in a larger set of parameters when tying contracts can be offered compared to the case where they cannot be offered. Then, we show that the set of parameters in which the purchasing patterns with the highest or second highest total surplus are achieved is a larger set of parameters when tying contracts can be offered compared to the case where they cannot be offered. The process continues for all five purchasing patterns.

For ease of exposition, we summarise the existence conditions underlying these comparisons in tables. In the tables, T refers to an agent being insured against both risks. When tying contracts are available, this is achieved by purchasing a tying contract. When tying contracts are not available, this is achieved by purchasing both stand-alone contracts.

Table 4.10: Case 1: Expected Costs are Private Information, Risk Preferences are Public Information

Purchasing Pattern	*L allowed* *P allowed* *T banned* SPNE iff:	*L allowed* *P allowed* *T allowed* SPNE iff:
T T T T T T	$R(L,e) > \frac{1}{3}[c(L) + c(M) - 2c(H)]$	$R(L,e) > \frac{1}{3}\delta$
L T T T P T	$R(L,e) < \frac{1}{3}[c(L) + c(M) - 2c(H)]$ $R(L,w) > \frac{1}{3}[c(L) + c(M) - 2c(H)]$	$R(L,e) < \frac{1}{3}\delta$ $R(L,w) > \frac{1}{3}\delta$
L L T T P P	$R(L,e) < \frac{1}{3}[c(L) + c(M) - 2c(H)]$ $R(L,w) < \frac{1}{3}[c(L) + c(M) - 2c(H)]$ $R(M,e) > \frac{1}{2}[c(L) - c(M)]$	$R(L,e) < \frac{1}{3}\delta$ $R(L,w) < \frac{1}{3}\delta$
L L – T P P	$R(L,e) < \frac{1}{3}[c(L) + c(M) - 2c(H)]$ $R(L,w) < \frac{1}{3}[c(L) + c(M) - 2c(H)]$ $R(M,e) < \frac{1}{2}[c(L) - c(M)]$ $R(M,w) > \frac{1}{2}[c(L) - c(M)]$	never
L L – – P P	$R(L,e) < \frac{1}{3}[c(L) + c(M) - 2c(H)]$ $R(L,w) < \frac{1}{3}[c(L) + c(M) - 2c(H)]$ $R(M,e) < \frac{1}{2}[c(L) - c(M)]$ $R(M,w) < \frac{1}{2}[c(L) - c(M)]$	never

note: $c(L) + c(M) - 2c(H) = \delta + 2c(L) - c(M) - c(H) > \delta$

Table 4.11: Case 2: Expected Costs and Risk Preferences are Private Information

Purchasing pattern	*L, P allowed, T banned* SPNE iff:	*L, P, T allowed* SPNE iff:
T T T T T T	$R(L,e) > \frac{1}{3}[c(L)+c(M)-2c(H)]$	$R(L,e) > \frac{1}{3}\delta$
L T T T P T	$R(L,e) < \frac{1}{3}[c(L)+c(M)-2c(H)]$ $R(L,w) > \frac{2}{5}[c(L)+c(M)-2c(H)]$	$R(L,e) < \frac{1}{3}\delta$ $R(L,w) > \frac{1}{2}\delta$
L L T T P P	$R(L,e) < \frac{1}{3}[c(L)+c(M)-2c(H)]$ $R(L,w) < \frac{2}{5}[c(L)+c(M)-2c(H)]$ $R(M,e) > \frac{1}{2}[c(L)-c(M)]$	$R(L,e) < \frac{1}{3}\delta$ $R(L,w) < \frac{1}{2}\delta$
L L – T P P	$R(L,e) < \frac{1}{3}[c(L)+c(M)-2c(H)]$ $R(L,w) < \frac{2}{5}[c(L)+c(M)-2c(H)]$ $R(M,e) < \frac{1}{2}[c(L)-c(M)]$ $R(M,w) > \frac{2}{3}[c(L)-c(M)]$	never
L L – – P P	$R(L,e) < \frac{1}{3}[c(L)+c(M)-2c(H)]$ $R(L,w) < \frac{2}{5}[c(L)+c(M)-2c(H)]$ $R(M,e) < \frac{1}{2}[c(L)-c(M)]$ $R(M,w) < \frac{2}{3}[c(L)-c(M)]$	never

note: $c(L)+c(M)-2c(H) > \delta$, $\frac{2}{5}[c(L)+c(M)-2c(H)] > \frac{1}{2}\delta$

Analogous to Table 4.2, this table depicts the case where $v(M,e) > v(H,w)$. As in the context of Table 4.2, the case where $v(M,e) < v(H,w)$ results in a similar table.

4.7.9 Proof of Corollary 4.2

Proof by example: Consider the parameters listed in Appendix 4.7.5. Appendix 4.7.5 showed that with tying contracts, the equilibrium does not imply full unravelling. This Appendix shows that for the exact same parameters, individual markets unravel, i.e. when firms cannot write tying contracts welfare is substantially reduced.

Start: Let the price of life-insurance equal the cost of the consumer who has the lowest expected cost: $\mathcal{L} = c(H) = 1$. At this price all six groups of agents $(L_e, L_w, M_e, M_w, H_e, H_w)$ buy life insurance. Thus, the average expected cost of this contract, calculated across all consumers who buy at this price, is $\frac{2*1+2*5+2*6}{6} = 4$. Thus, firms make a loss and raise their price to $\mathcal{L} = 4$ in order to at least break even.

Then: At $\mathcal{L} = 4$ high-life expectancy relaxed (H_e) stop to buy since $v(H,e) = 1.5 < 4$. WTP below price makes it optimal to not purchase the contract. Thus, the costs to the firm are now $\frac{1+2*5+2*6}{5} = 4.6$. Hence, firms make a loss and raise their price to $\mathcal{L} = 4.6$.

Then: At $\mathcal{L} = 4.6$ the high life-expectancy worried (H_w) stop to buy since $v(H,w) = 4.5 < 4.6$. Thus, costs are now $\frac{2*5+2*6}{4} = 5.5$. Hence, firms make a loss and need to raise price to $\mathcal{L} = 5.5$.

Then: At $\mathcal{L} = 5.5$ the medium life expectancy at ease (M_e) stop to buy since $v(M, e) = 5.2 < 5.5$. Thus, costs are now $\frac{1*5+2*6}{3} = 5\frac{2}{3}$. Firms make a loss and raise price to $\mathcal{L} = 5\frac{2}{3}$.

Then: At $\mathcal{L} = 5\frac{2}{3}$ the medium life expectancy worried (M_w) stop to buy since $v(M, w) = 5.6 < 5\frac{2}{3}$. Thus, costs are now $\frac{2*6}{2} = 6$. Firms thus make a loss and raise price to $\mathcal{L} = 6$.

Then: At $\mathcal{L} = 6$ only low life expectancy agents buy (L_w, L_e). Both at ease and worried low life expectancy agents continue to buy at $\mathcal{L} = 6$ and the firm makes zero profit since at $\mathcal{L} = 6$ the costs are also 6.

Thus, the equilibrium in the individual market is $\mathcal{L} = 6$ and only low life expectancy agents buy, i.e. those agents with the highest cost. Hence, there is *full* unravelling, i.e. only the agents with the highest expected costs are insured.

Above, we started from $\mathcal{L} = c(H) = 1$. Any $\mathcal{L} < 1$ is loss making for any customer and hence cannot be an equilibrium. Firms may start at some $\mathcal{L} > 1$. In that case, one can replicate the unravelling logic from above, starting at a later step. The end point of unravelling will be reached regardless of the $\mathcal{L}$ at which firms start. *q.e.d.*

4.7.10 Proof of Proposition 4.5: Effect of Changes in the Information Structure

The table below shows that for all parameters the equilibrium purchasing pattern achieves weakly higher and sometimes strictly higher welfare when risk preferences are public information rather than private information.

Table 4.12: Effects of Changes in the Information Structure

	Pattern 1	Pattern 3	Pattern 2	
	T T T T T T	L T T T P T	L L T T P P	other patterns
SPNE iff: μ_i *private info* ρ_i *private info*	$R(L, e) > \frac{1}{3}\delta$	$R(L, e) < \frac{1}{3}\delta$ $R(L, w) > \frac{1}{2}\delta$	$R(L, e) < \frac{1}{3}\delta$ $R(L, w) < \frac{1}{2}\delta$	never
SPNE iff: μ_i *private info* ρ_i *public info*	$R(L, e) > \frac{1}{3}\delta$	$R(L, e) < \frac{1}{3}\delta$ $R(L, w) > \frac{1}{3}\delta$	$R(L, e) < \frac{1}{3}\delta$ $R(L, w) < \frac{1}{3}\delta$	never

4.7.11 Proof of Proposition 4.6 : Comparing Welfare Effects

The table below contains contains the conditions underlying Figure 4.8, which proves the proposition.

Table 4.13: Comparing Welfare Effects

Purchasing pattern	μ_i *private info*, ρ_i *public info* L, P *allowed*, T *banned* SPNE iff:	μ_i, ρ_i *private info* L, P *allowed*, T *banned* SPNE iff:	μ_i, ρ_i *private info* L, P, T *allowed* SPNE iff:
Pattern 1 T T T T T T	$R(L,e) > \frac{1}{3}[c(L)+c(M)-2c(H)]$	$R(L,e) > \frac{1}{3}[c(L)+c(M)-2c(H)]$	$R(L,e) > \frac{1}{3}\delta$
Pattern 3 L T T T P T	$R(L,e) < \frac{1}{3}[c(L)+c(M)-2c(H)]$ $R(L,w) > \frac{1}{3}[c(L)+c(M)-2c(H)]$	$R(L,e) < \frac{1}{3}[c(L)+c(M)-2c(H)]$ $R(L,w) > \frac{2}{5}[c(L)+c(M)-2c(H)]$	$R(L,e) < \frac{1}{3}\delta$ $R(L,w) > \frac{1}{2}\delta$
Pattern 2 L L T T P P	$R(L,e) < \frac{1}{3}[c(L)+c(M)-2c(H)]$ $R(L,w) < \frac{1}{3}[c(L)+c(M)-2c(H)]$ $R(M,e) > \frac{1}{2}[c(L)-c(M)]$	$R(L,e) < \frac{1}{3}[c(L)+c(M)-2c(H)]$ $R(L,w) < \frac{2}{5}[c(L)+c(M)-2c(H)]$ $R(M,e) > \frac{1}{2}[c(L)-c(M)]$	$R(L,e) < \frac{1}{3}\delta$ $R(L,w) < \frac{1}{2}\delta$
Pattern 4 L L − T P P	$R(L,e) < \frac{1}{3}[c(L)+c(M)-2c(H)]$ $R(L,w) < \frac{1}{3}[c(L)+c(M)-2c(H)]$ $R(M,e) < \frac{1}{2}[c(L)-c(M)]$ $R(M,w) > \frac{1}{2}[c(L)-c(M)]$	$R(L,e) < \frac{1}{3}[c(L)+c(M)-2c(H)]$ $R(L,w) < \frac{2}{5}[c(L)+c(M)-2c(H)]$ $R(M,e) < \frac{1}{2}[c(L)-c(M)]$ $R(M,w) > \frac{2}{3}[c(L)-c(M)]$	never
Pattern 5 L L − − P P	$R(L,e) < \frac{1}{3}[c(L)+c(M)-2c(H)]$ $R(L,w) < \frac{1}{3}[c(L)+c(M)-2c(H)]$ $R(M,e) < \frac{1}{2}[c(L)-c(M)]$ $R(M,w) < \frac{1}{2}[c(L)-c(M)]$	$R(L,e) < \frac{1}{3}[c(L)+c(M)-2c(H)]$ $R(L,w) < \frac{2}{5}[c(L)+c(M)-2c(H)]$ $R(M,e) < \frac{1}{2}[c(L)-c(M)]$ $R(M,w) < \frac{2}{3}[c(L)-c(M)]$	never

where the proposition follows since

$$c(L) + c(M) - 2c(H) = \delta + 2c(L) - c(M) - c(H) > \delta \tag{4.72}$$

and

$$\frac{2}{5}[c(L) + c(M) - 2c(H)] > \frac{1}{2}\delta \tag{4.73}$$

as well as

$$\frac{1}{3}[c(L) + c(M) - 2c(H)] > \frac{1}{2}\delta \tag{4.74}$$

q.e.d.

Chapter 5

Signalling Through Costly Effort Provision: Theory and Evidence from the Field

with Lukas Hensel, Johannes Hermle, Christopher Roth, Anselm Hager

We propose a model of signalling which endogenizes agents' decisions on whether to communicate their costly effort to signal their unobserved type. The model predicts that types that have high reputational concerns and high motivation *increase* their effort and the likelihood of sharing their effort choices, when the action becomes *more* costly. We test the predictions of the model using novel field data from an online application which records canvassing effort and the communication of canvassing effort among a sample of political activists. As predicted by our model, activists who revealed to have reputational concerns by having previously communicated their effort level on social media, increase their canvassing effort if effort becomes more costly, while the opposite pattern holds for individuals without prior communication of effort choices. Furthermore, canvassers with prior communication of effort are more likely to communicate their effort if it was more costly. Our paper provides theoretical and empirical insights on agents' strategic decision on whether to communicate their costly effort choices in order to signal their unobserved types.

Keywords: Signalling, Costly Effort Provision, Political Activism, Natural Field Data, Beliefs, Canvassing, Quasi-experiment

JEL Classification: D74, D8, P0

5.1 Introduction

Signalling models strongly shape our understanding of economic behaviour, including people's consumption decisions (Bursztyn et al., 2017a; Bernheim, 1994), educational choices (Spence, 1973), political behaviour (DellaVigna et al., 2017), health behaviours (Karing, 2018; Karing and Naguib, 2017), and pro-social behaviour (Bénabou and Tirole, 2006; Dubé et al., 2017; Bénabou et al., 2018; Falk, 2017). An important open theoretical and empirical question for the literature on signalling is how signalling behaviour differs when the signal is neither perfectly observed, nor observed with an exogenous probability, but when instead agents strategically decide whether to communicate their signalling actions or not.

We build a simple model of signalling behaviour which is reliant on the following intuition: An agent's action is shaped not only by the costs and benefits of her actions, but also by their signalling value. Furthermore, the likelihood that an agent signals depends on the costliness of the signal. Our model predicts that there are types who have both high motivation and high reputational concerns, who will *increase* their effort choices when the action becomes *more* costly and only communicate their effort choices when the action is more costly. Agents with lower reputational concerns or with lower motivation, on the other hand, decrease their effort choices in response to increases in costs. Thus, our theoretical model predicts that changes in costs can have a "surprising" effect on effort choices. While in a full information setting increases in costs lead agents to reduce effort, we show that in a signalling model the sign can be reversed for a part of the population. This arises because the increase in costs shifts the economy from a pooling to a separating equilibrium.

Testing the predictions of signalling models is difficult due to data constraints. Field evidence testing the predictions of signalling models is rare. Even more so, naturally-occurring behavioural data on signalling from field settings is very scarce. We overcome these data-constraints by leveraging a large data-set on the campaigning behaviour of political activists gathered in the context of a high-stakes national election among a sample of political activists. Our dataset contains unique data on canvassing and sharing of the canvassing effort on social media. We combine this data with quasi-experimental variation in the costliness of their effort choices. We match fine-grained geo-coded real time data on canvassing effort with high resolution temporal and spatial data on temperatures which allows us to measure whether the weather is good or bad. This data-set allows us to directly test the implications our model.

Pooling our sample we find that higher temperatures (good weather) increases people's canvassing effort. As predicted by our model, for individuals without prior communication of their effort choices, higher temperatures increase canvassing effort. However, for individuals with prior sharing of their effort choices, *lower* temperatures (bad weather) *increase* canvassing effort. Our unique dataset also allows us to test a key prediction of

our model regarding the sharing of costly effort choices: We show that those consumers who increase effort in response to higher costs (worse weather) are also more likely to communicate their effort when costs are high. We interpret this as further evidence for our signalling interpretation. We confirm these patterns using a variety of different specifications and definitions of the temperature data.

Our paper relates to a large literature in economics on signalling models (Spence, 1973). Most closely related are papers on social and self-signalling (Bénabou and Tirole, 2006; Dubé et al., 2017). The main contribution of our paper is to leverage novel data on signalling to test some of the predictions of simple signalling models. Our evidence is consistent with the predictions of influential signalling models (Bénabou and Tirole, 2006) which predict that effort can decrease when effort becomes less costly as the signalling value of effort has fallen. However, these signalling models treat the probability of revealing one's costly effort choice as exogenously determined and therefore make no prediction about the likelihood of signalling costly effort. Our model extends the work by Bénabou and Tirole (2006) by endogenizing agents' decision on whether to signal their effort level. We derive the perfect Bayesian equilibria of this game and subsequently test the predictions of the model regarding effort choices and regarding the communication of effort choices on social media.

We also contribute to a growing literature studying social image concerns and peer pressure (DellaVigna et al., 2012, 2017; Gerber et al., 2008; Bursztyn and Jensen, 2015, 2017; Bursztyn et al., 2017a,b,c). Previous papers have varied the visibility of economic behaviour to study the effects of social image concerns. Our approach differs from the previous literature as we employ unique data on the sharing of effort choices of individuals both as an outcome variable and in order to shed light on heterogeneity in behavioural responses to changes in the costliness and effectiveness of actions. By measuring people's sharing of effort choices we can provide direct evidence on the relevance of reputational concerns in shaping people's decision to share their effort choices.

A large literature in behavioural economics has generated a body of evidence where economic incentives counter-intuitively reduce the supply of prosocial behaviour (Bowles and Polania-Reyes, 2012; Frey and Jegen, 2001; Falk and Kosfeld, 2006; Gneezy and Rustichini, 2000a,b). Similarly, psychologists have gathered evidence where extrinsic incentives lead to motivational crowd-out (Deci, 1971; Greene and Lepper, 1978). However, the empirical evidence in the field for the crowding out effect of economic incentives on prosocial behaviour has been mixed (Gneezy et al., 2011; Kamenica, 2012).

Our results differ from this previous literature in several important ways: First, and most importantly, unlike the previous literature we can study whether people share their effort choices which allows us to more directly test for the relevance of signalling motives. Second, we have evidence that signalling behaviour is more prevalent by individuals who have stronger reputational concerns and for individuals with lower social identification with the party. In our setting, agents motivated to make a career in their party communicate

their costly effort in order to signal their commitment to the party.

Finally, our experiment contributes to a growing body of literature studying the motivation and behaviour of party supporters (Enos and Hersh, 2015; Perez-Truglia and Cruces, 2017; Hensel et al., 2018a,b) and protesters (Cantoni et al., 2019; González, 2020; Enikolopov et al., 2016, 2017; Cantoni et al., 2016; Passarelli and Tabellini, 2017). Our evidence highlights the importance of social image concerns for shaping people's strategic decision to choose and signal their effort choices.

5.2 Signalling Model

We study the behaviour of political activists in a model of pro-social action à la Bénabou and Tirole (2006), but make the visibility of the pro-social action a strategic choice.

5.2.1 Model Outline

Political activists are heterogeneous along two dimensions: commitment to the party (v_i) and concern for their reputation (μ_i). The vector (v_i, μ_i) fully describes an agent's type and is private information. For simplicity we focus on the binary case: $v_i \in \{v_L, v_H\}$ where $v_L < v_H$, $\mu_i \in \{\mu_L, \mu_H\}$ where $\mu_L < \mu_H$. Each day, agents choose a level of canvassing effort $e \in \{e_L, e_M, e_H\}$ where $e_L < e_M < e_H$. Effort yields a direct benefit $B(e, v)$, which depends on the agent's commitment to the party, and is costly $C(e)$. This cost $C(e)$ captures the opportunity cost of time and the disutility from walking from door to door. Effort e is the agent's private knowledge but is truthfully recorded on a mobile phone app. At the end of a day, the app allows agents to truthfully disclose their effort e by posting it on facebook. This technological constraint ensures truthful communication. As in games with verifiable disclosure, agents can report the truth or remain silent but they cannot lie. Sending a signal ($s = 1$) is costly ($S > 0$ but small). This captures the cognitive cost of remembering to signal, the cost of pushing the signalling button in the app, and the negative effect of alienating friends who support a different political party. If the agent does not signal ($s = 0$) he incurs no cost. Other party members follow the activist's facebook activities and update their perception of the agent's commitment to the party ($\hat{v}$) based on the signal (or lack thereof). Activists value the reputation their peers hold of them with $R(\hat{v}, \mu)$ as this will positively affect their standing in the party and improve chances when running for office in the future. An agent's utility is thus given by:

$$u = B(e, v) - C(e) + R(\hat{v}, \mu) - S(s) \tag{5.1}$$

We focus on a one-shot game[1] and restrict our attention to Perfect Bayesian Equilibria

[1]While this focus on one-shot games abstracts from learning from repeated interaction, it can be rationalized by several behavioural assumptions: signal recipients may not recall previous periods or may not use previous periods in forming their expectations, activists may not realize the implications of repeated signalling, there may be an exogenous probability of canvassing not being possible on a given day, e.g. due to illness or professional

(PBE) which survive tests for equilibrium dominance (Cho and Kreps, 1987).

Our empirical approach uses that weather conditions affect an agent's cost of canvassing. Letting weather conditions be binary (good or bad), the corresponding assumption is that $C_{good}(e_H) < C_{bad}(e_H)$ and $C_{good}(e_M) < C_{bad}(e_M)$. This captures that the costs of going outdoors and canvassing (e_M, e_H) is higher if the weather conditions are worse. This contrasts to the outside option of not canvassing and potentially staying indoors which is unaffected by the weather. Formally $C_{good}(e_L) = C_{bad}(e_L)$. More strongly, we normalise the outside option of not canvassing to be zero at all weather conditions, i.e. $B(e_L, v_i) - C(e_L) = 0 \forall i \in \{L, H\}$ for both weather conditions.

5.2.2 Theory Results

If commitment to the party (v_i) is common knowledge, then equilibrium effort is weakly decreasing in costs. However, when commitment to the party is private knowledge, we show that the reverse can be true: There exist types for whom equilibrium effort can be increasing in costs - consistent with our empirical findings.

Observable Benchmark: *When commitment to the party is common knowledge, equilibrium effort is weakly decreasing in costs.*

If v_i is common knowledge, agents will never send a costly signal as they cannot affect their reputation. Thus, agents solve $max_e B(e, v) - C(e)$. In a continuous case with regularity conditions to ensure second order conditions hold, agents choose e such that: $B'(e, v) = C'(e)$. If canvassing becomes more costly, e.g. due to bad weather, then equilibrium effort falls. In the discrete case our assumptions on how weather shifts costs ensure that we get the same result as in the continuous case.

We are primarily interested in the case where commitment to the party is private knowledge. When v_i is private information it may be optimal for agents to signal their effort in order to improve their reputation. This makes the following two types of equilibria possible:

Proposition 5.1 *Equilibrium Type 1: There can exist an equilibrium where (v_L, μ_H) and (v_H, μ_H) types pool, i.e. both signal the same effort level.*

Equilibrium Type 1:

Equilibrium Type 1	μ_L	μ_H
v_L	$e_L, s = 0$	$e_M, s = 1$
v_H	$e_M, s = 0$	$e_M, s = 1$

commitments, or an interaction of these factors.

When the corresponding beliefs are that not signalling is interpreted as $\hat{v} = E(v \mid \mu_L)$, signalling e_M is interpreted as $\hat{v} = E(v \mid \mu_H)$ and the off-equilibrium path beliefs are: signalling e_L leads to $\hat{v} = v_L$ and signalling e_H to $\hat{v} = E(v \mid \mu_H)$.

Equilibrium type 1 exists if the following conditions are satisfied:

1. $S > 0$
2. The full information choice of intrinsically not motivated agents v_L is e_L:

$$B(e_L, v_L) - C(e_L) > B(e_M, v_L) - C(e_M) > B(e_H, v_L) - C(e_H)$$

3. (v_L, μ_L) does not want to mimic (v_H, μ_H):

$$B(e_L, v_L) - C(e_L) + R(E(v \mid \mu_L), \mu_L) > B(e_M, v_L) - C(e_M) + R(E(v \mid \mu_H), \mu_L) - S$$

To show the existence result of proposition 5.1, it is sufficient to assume that $R(\hat{v}, \mu_L) = 0 \forall \hat{v}$, i.e. the types which have little concern for their reputation do not value their reputation at all. Then, (1) and (2) ensure that (3) holds.

4. The full information choice of intrinsically motivated agents v_H is e_M: $B(e_M, v_H) - C(e_M) > B(e_L, v_H) - C(e_L)$ and $B(e_M, v_H) - C(e_M) > B(e_H, v_H) - C(e_H)$.
5. (v_H, μ_L) does not want to mimic (v_H, μ_H). The assumption made in (3) ensures this holds.
6. $\frac{dR(\hat{v}, \mu_H)}{d\hat{v}} > 0$
7. (v_L, μ_H) does not want to give up signalling:

$$R\big(E(v \mid \mu_H), \mu_L\big) - R\big(E(v \mid \mu_L), \mu_L\big) - S > B(e_L, v_L) - C(e_L) - \big(B(e_M, v_L) - C(e_M)\big)$$

Proof: Appendix 5.5 contains an example ("good weather") which satisfies all these conditions. That constitutes the proof that an equilibrium of type 1 can exist.

Proposition 5.2 *Equilibrium Type 2: There can exist an equilibrium where (v_L, μ_H) and (v_H, μ_H) types separate.*

Equilibrium Type 2:

Equilibrium Type 2	μ_L	μ_H
v_L	$e_L, s = 0$	$e_L, s = 0$
v_H	$e_M, s = 0$	$e_H, s = 1$

Where the corresponding beliefs are that not signalling is interpreted as $\hat{v} = E(v \mid not(v_H, \mu_H))$, signalling e_H is interpreted as $\hat{v} = v_H$ and the off-equilibrium path beliefs

are: signalling e_L leads to $\hat{v} = v_L$ and signalling e_M leads to $\hat{v} = E(v \mid \mu_H)$.

Equilibrium type 2 exists if the following conditions hold:

1. Condition 1-6 are the same as for equilibrium type 1.[2]

7. (v_L, μ_H) does not want to mimic (v_H, μ_H), i.e. the reputational gain is not sufficient to cover the effort and signalling costs:

$$B(e_L, v_L) - C(e_L) + R\big(E(v \mid not(v_H, \mu_H)), \mu_H\big) > B(e_H, v_L) - C(e_H) + R(v_H, \mu_H) - S$$

8. (v_H, μ_H) does not want to give up separating, i.e. the reputational gain is worth the costs from signalling and high effort:

$$B(e_H, v_H) - C(e_H) + R(v_H, \mu_H) - S > B(e_M, v_H) - C(e_M) + R\big(E(v \mid not(v_H, \mu_H)), \mu_H\big)$$

Proof: Appendix 5.5 contains an example ("bad weather") which satisfies all these conditions. That constitutes the proof that an equilibrium of type 2 can exist.

Proposition 5.3 *There exist types for whom equilibrium effort can be increasing in costs.*

This effect can arise because an increase in the costs can shift the economy from a type 1 equilibrium to a type 2 equilibrium.

Proof: We decompose the proof of proposition 5.3 into three steps:

Lemma 5.2.2.1 *There exist parameters which we term "good weather parameters" for which equilibrium type 1 exists but equilibrium type 2 does not exist.*

Proof: By example. The "good weather" parameters in appendix 5.5 satisfy Lemma 5.2.2.1.

Lemma 5.2.2.2 *There exist parameters which we term "bad weather parameters" for which equilibrium type 2 exists but equilibrium type 1 does not exist.*

Proof: By example: The "bad weather" parameters in appendix 5.5 satisfy Lemma 5.2.2.2.

Lemma 5.2.2.3 *There exist parameters which satisfy Lemma 5.2.2.1 and after adjusting only the common costs of effort satisfy Lemma 5.2.2.2.*

[2] In equilibrium type 2, a less tight version of condition (2) is sufficient to ensure existence of the equilibrium: For equilibrium type 2 we do not require a ranking of $B(e_M, v_L) - C(e_M)$ and $B(e_H, v_L) - C(e_H)$. We just require both to be less than $B(e_L, v_L) - C(e_L)$. This distinction is not important for our purposes.

Proof: By example. The "bad weather" parameters can be obtained as a transformation of the "good weather" parameters by adjusting only costs (subtracting 3 from $C(e_M)$ and $C(e_H)$).

Thus, for activists of type (v_H, μ_H), equilibrium effort can be increasing in costs. *q.e.d.*

Intuition: As the weather deteriorates, canvassing becomes more costly for everyone. This makes (v_L, μ_H) activists less willing to signal and opens the possibility for (v_H, μ_H) agents to separate by putting in even higher effort. Hence, the equilibrium increase in effort is explained by a shift from a pooling to a separating equilibrium.

5.3 Field Evidence

5.3.1 Setting and Data

Setting: Our outcome data on political activism was collected in the run-up to a recent general election in a Western European country.[3] During the electoral campaign almost all political parties set a novel focus on door-to-door canvassing as a means of campaigning. The party with which we cooperated strongly promoted canvassing as a campaigning tool both through internal communication channels and to the wider public through mass media. All canvassing volunteers were instructed to record all canvassed doors in a smartphone application to help the party organize this and future campaigns. The actual canvassing activities were organized as follows. Party members in charge of the local campaign assigned dates and locations for canvassing events and then asked local supporters to join these events.[4] Individual members were then faced with the decision in how many of these events they would participate. The party established this structure to guide canvassing activity to the most promising areas and to lower the barriers for supporters to engage in canvassing.

We draw on unique behavioural outcome data from the smartphone application distributed by the party. We use this data to construct a daily panel of individual level canvassing and signalling activity. Our sample comprises all party supporters who registered at least one canvassed door using the party's smartphone application. In total we observe 3474 activists over the the study period of 233 days totalling 809,442 individual-day observations.

Data on canvassing: To construct measures of canvassing activity we use the fact that the application recorded the location and time for each visited door. This data spans the 233 days before the federal election and includes several local and state elections. We construct two main measures of canvassing: whether an activist conducted any canvassing and the number of doors canvassed on a given day (winsorized at the 90th percentile of

[3]The time period also spans several state and local elections.

[4]Local campaign organizers were a mix of paid staff and local party officials who volunteered for the position.

positive observations to deal with positive outliers).

Data on communication of effort: A unique feature of our data is that we directly observe whether activists communicate their effort choices on facebook. Activists received points and could achieve certain 'titles'[5] in the application as reward for their canvassing activity. Activists could then share their points or 'title' on facebook using the application. We use this data to construct an indicator of whether an activist communicated her contribution on a given day at least once in any form, i.e. we combine the sharing of points and of 'titles' in one variable.

5.3.2 Descriptives

Descriptive statistics: Figure 5.1 presents the average fraction of activists in our sample who canvass and signal on each day. The fraction of individuals canvassing on a given day ranges from close to zero at the beginning of the campaign to about 13% right before the election.[6] We observe that both canvassing and signalling increase as the federal election approaches.

Table 5.1 displays basic summary statistics averaged across time. The individual-day statistics in Panel A show that, on average, activists canvass only on 1.5% of days and knock on 0.83 doors per day. Most of the time activists communicate their total number of points rather than their title (0.5% vs 0.1%).

Panel B displays summary statistics at the individual level. On average, activists canvassed on 3.59 days knocking on 180 doors. 29% of activists communicated their canvassing behaviour at least once. On average, activists communicated their contribution 2.25 times during the study period. Columns 2 and 3 show that overall canvassing activity is higher for activists who communicated their effort at any point during the campaign.

Who communicates effort choices on social media? Next we provide descriptive facts on who communicates their effort. To do so, we leverage unique survey data which we collected a few months after the campaign. The survey contains a large number of variables, capturing economic preferences, psychological variables, information on political ambitions as well as standard demographics.

Table 5.5 summarizes the descriptive results. We find that those activists who communicate their effort, i.e. who signal, provide more effort, identify more strongly with the party and have substantially larger career concerns. Moreover, we find that younger individuals, and party members are more likely to signal their effort choices. Our data highlights that those people who signal have both (i) higher commitment to the party as captured by people's identification with the party, (ii) higher reputational concerns as they wish to pursue a political career within the party. This is exactly in line with a

[5]There are in total seven different ranks that canvassers can reach. The first level is for people with less than 20 points in the app. The second level is reached when reaching 2500 points. The final level is reached after 200,000 points. Individuals receive 200 points per door.

[6]The intermediate spikes are due to local and state elections.

key prediction of our model: that those people who signal both have high reputational concerns and high motivation (high commitment to the party).

5.3.3 The Effect of the Costliness of Effort

Temperature data: We use the average temperature in the twelve hours between 9am and 9pm as exogenous variation in the cost of canvassing.[7] The temperature data is constructed using hourly observations from weather stations from the county's official weather agency. Each postal code is matched to the five weather stations closest to its center point (there are a total of 382 stations with no missing values during the study period). We use the inverse square-distance weighted mean average temperature of these five stations as our main temperature measure.[8]

Average effects of temperature on canvassing: We first assess the reduced form effect of temperature on canvassing effort. For this purpose, we estimate the following main empirical specification:

$$e_{izt} = \alpha_i + \beta_t + \gamma_1 \text{temperature}_{zt} + \varepsilon_{izt}$$

where e_{ist} is canvassing effort of individual i in zipcode z at day t; α_i are individual fixed effects; β_t are day fixed effects; temperature_{zt} is average temperature between 9am and 9pm in postal code z on day t. Thus, γ_1 identifies the effect of differences in temperature demeaned at the individual and day level. To account for the temporal correlation of temperature, we cluster standard errors at the postal code level.

Higher temperatures result in higher canvassing activity on average (see Column 1 of Table 5.2). A one degree increase in the temperature leads to an average increase in the fraction of canvassing individuals by 0.08 percentage points (about 5% of the mean). Figure 5.2 displays non-parametric regressions which highlight that the relationship between canvassing and temperature is approximately linear.[9]

Heterogeneous effects by prior communication of effort: To test the implications of our model we investigate how the effect of temperature shocks varies by types. Because we do not directly observe an activist's type, we use past sharing of effort choices as proxy for having both high commitment to the party and having high reputational concerns. Since past sharing of canvassing activity is strongly correlated with canvasing activity, we need to control for past canvassing activity and its interaction with temperature. This allows us to look at heterogeneous effects by past sharing of effort choices holding the past effort level constant.

[7] 84% of registered canvassing activity takes place during this time period.

[8] When postal codes consist of multiple disjoint areas we further average the final measure across the geographic centers of the different areas.

[9] The election took place in autumn in a European country. Average temperature between 9am and 9pm in the high-activity phase of the campaign was at around 15 degrees Celsius. Therefore, higher temperatures indicate better weather and reduced the costliness of canvassing

We estimate the following main empirical specification:

$$e_{izt} = \alpha_i + \beta_t + \gamma_1 \text{temperature}_{zt} + \gamma_2 \text{temperature}_{izt} \times \text{anysignal}_{iz,t-1}$$
$$+\gamma_3 \text{anysignal}_{iz,t-1} + \gamma_4 \text{temperature}_{izt} \times \text{doorscumul}_{iz,t-1} + \gamma_5 \text{doorscumul}_{iz,t-1} + \varepsilon_{izt}$$

where $\text{anysignal}_{iz,t-1}$ denotes whether the individual has ever signalled in the past and $\text{doorscumul}_{iz,t-1}$ is the cumulative number of doors canvassed in the past (measured in 100 doors).[10]

Heterogeneous effects of effort choices on communication: A prediction of our signalling model is that some activists increase their effort as effort becomes more costly. In line with this rationale for the increase in effort as driven by signalling motives, we find that for the signalling activists, i.e. those who previously communicated canvassing effort, higher costs of effort also increase the likelihood that these activists share communicate their canvassing statistics. The effect of costliness on signalling is significantly stronger than the effect on any canvassing (Table 5.4). For agents driven by signalling motives, results on signalling are robust to controlling for their contemporaneous effort choices. This suggests that the effect of temperature on signalling is not mechanical to the fact that people just put in more effort choices. Our evidence suggests that people signal more as the effort was more costly.

On the other hand, for activists without prior sharing of effort choices, the likelihood of engaging in any signalling increases in response to higher temperatures, but less so than the likelihood of engaging in any canvassing. Indeed, once we control for contemporaneous canvassing the effect of temperature on signalling are very close to zero, consistent with our model prediction that these types are not primarily driven by reputational concerns but rather by intrinsic motivation, i.e. alignment with the party.

Robustness of results: Our results are robust to using a large variety of different model specifications. For example, our results are robust to using different definitions of the temperature variable. For example, using the maximum temperature instead of the average temperature leaves our results unchanged (Tables A1, A2, A3). The results are also robust to using fixed effects for the first two digits of postal codes instead of individual fixed effects (Tables A4, A5, A6) and to allowing for spatial correlation at the two-digit postal code level (Tables A7, A8, A9).[11]

5.4 Conclusion

We propose a novel model of signalling which endogenizes agents' decision on whether to share their costly effort in order to signal their unobserved type. The model predicts that

[10]The daily measures of doors knocked used to construct this measure are winsorized at the 99^th percentile of positive observations to account for outliers.

[11]There are 93 different first two-digit combinations among the postal codes.

types that have high reputational concerns and high commitment to the party decrease their effort and the likelihood of sharing their action to signal their commitment, when the action becomes less costly. Agents with either low reputational concerns or low commitment to the party, on the other had, increase their effort when costs decrease.

We test these predictions combining unique geo-coded real time data on canvassing effort and signalling behaviour with quasi-experimental variation in temperatures. In line with the predictions of the model, we find that for individuals without prior signalling of their effort choices higher temperatures increase canvassing effort. However, for individuals with prior signalling of their effort choices, higher temperatures *decrease* canvassing effort. Moreover, we show that people's decision to signal their effort choices increases significantly with the costliness of the signal for respondents with prior signalling behaviour.

5.5 Theoretical Appendix

5.5.1 Numerical example

The examples parameters below show that it is possible that parameters satisfy all conditions for the existence of equilibrium type 1 ("good weather"), that it is possible that parameters satisfy all conditions for the existence of equilibrium type 2 ("bad weather"). Moreover, the parameters show that a cost changed (weather change) can shift the equilibrium from type 1 to type 2. For these parameters, bad weather shifts pay-offs conditional on canvassing by -3.

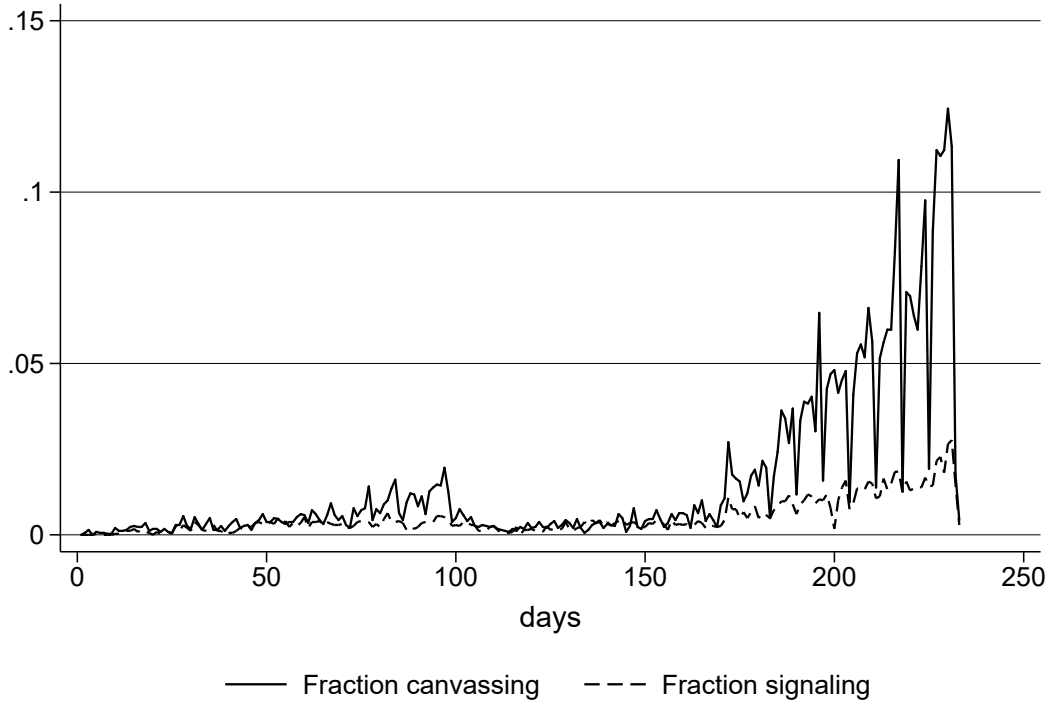
For ease of computation we suppose that the population consists of agents of the following proportions: $(v_L, \mu_L) = 20$, $(v_H, \mu_L) = 10$, $(v_L, \mu_H) = 30$, $(v_H, \mu_H) = 40$. The other parameters for this example are summarized in the table below.

	Good Weather	Bad Weather
$B(e_L, v_L) - C(e_L)$	0	0
$B(e_M, v_L) - C(e_M)$	-1	-4
$B(e_H, v_L) - C(e_H)$	-6	-9
$B(e_L, v_H) - C(e_L)$	0	0
$B(e_M, v_H) - C(e_M)$	10	7
$B(e_H, v_H) - C(e_H)$	6	3
S	1	1

To conclude the example, suppose that $R(\hat{v}, \mu_H)$ is linear in $\hat{v}$ and that $R(v_H, \mu_H) = 10$ and $R(v_L, \mu_H) = 0$. Then all equilibrium conditions are met.

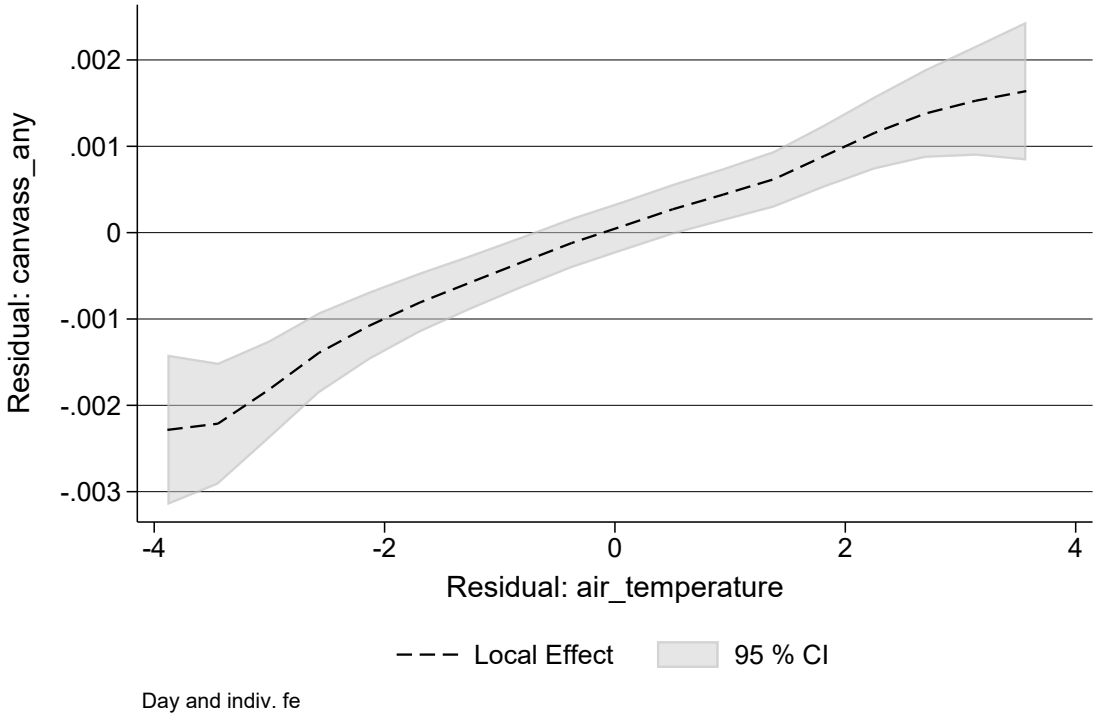
5.6 Main Figures

Figure 5.1: Evolution of Canvassing and Signalling over time



Notes: Figure 5.1 shows the fraction of activists engaging in canvassing and signalling over time. The election was held on day 233.

Figure 5.2: Non-parametric effect of average temperature on any canvassing



Notes: Figure 5.2 shows the non-linear effect of temperature on any canvassing conditional on day and individual fixed effects. Estimates are obtained using local linear regressions on residualized temperature and a dummy indicating any canvassing activity. Top and bottom 1% of the residual temperature distribution are omitted to avoid biased closer to the limits.

5.7 Main Tables

Table 5.1: Summary statistics

	Pooled	Signaled	Did not signal
Panel A: day-individual level variables			
Canvassed	0.015	0.027	0.011
Doors visited (winsorized)	0.830	1.724	0.461
Signaled	0.005	0.017	0.000
Times signaled (winsorized)	0.548	1.875	0.000
Signaled points	0.005	0.016	0.000
Signaled title	0.001	0.004	0.000
Temperature (degrees Celsius)	15.219	15.214	15.221
Observations	809442	236728	572714
Panel B: individual level variables			
Days canvassed	3.592	6.273	2.485
Doors canvassed (winsorized)	179.598	352.040	108.320
Signaled	0.292	1.000	0.000
Times signaled (winsorized)	2.252	7.700	0.000
Signaled points (winsorized)	1.964	6.716	0.000
Signaled title (winsorized)	0.277	0.947	0.000
Observations	3474	1016	2458

Notes: Table 5.1 displays summary statistics of the individual-day panel. Panel A presents summary statistics at the day-individual level. Panel B presents summary statistics at the individual level.Column 1 displays variable means for all activists. Column 2 displays variable means for the subsample of individuals who signaled at any point of the electoral campaign. Column 3 displays variable means for the subsample of individuals who did not signal during electoral campaign. Winsorized variables at the day-individual are winsorized at the 99^th percentile of non-negative observations. Winsorized variables at the individual are winsorized at the 99^th percentile of all observations.

Table 5.2: Impact of temperature on any canvassing

	Dependent variable is any canvassing		
	(1)	(2)	(3)
Temperature	0.0008***	0.0008***	0.0010***
	(0.0001)	(0.0001)	(0.0001)
Temperature × Has signaled before			-0.0016***
			(0.0003)
Has signaled before		0.0355***	0.0646***
		(0.0039)	(0.0074)
Temperature × Days canvassed			-0.0003
			(0.0002)
Days canvassed		-0.0004	0.0060
		(0.0014)	(0.0046)
Outcome mean	0.0154	0.0154	0.0154
Number obs	809442	809442	809442

Notes: Table 5.2 displays the effect of temperature on a dummy for any canvassing activity. The level of analysis is individual-day. All specifications include individual and day fixed effects. Standard errors are clustered at the postal code level. Column 1 displays the average impact of temperature on any canvassing. Column 2 further includes the cumulative number of days canvassed until a given day and whether an activist has signaled their effort on social media at any point before a given day. Column 3 includes interactions of these variables interacted with temperature.

Table 5.3: Impact of temperature on doors (winsorized)

	Dependent variable is doors canvassed		
	(1)	(2)	(3)
Temperature	0.0547***	0.0570***	0.0687***
	(0.0178)	(0.0178)	(0.0192)
Temperature × Has signaled before			-0.1787***
			(0.0351)
Has signaled before		3.4041***	6.5915***
		(0.5575)	(1.0761)
Temperature × Days canvassed			0.0191
			(0.0157)
Days canvassed		0.4283***	0.1107
		(0.1334)	(0.3608)
Outcome mean	0.8304	0.8304	0.8304
Number obs	809442	809442	809442

Notes: Table 5.3 displays the effect of temperature on doors canvassed. Doors canvassed is winsorized at the 99^th percentile of non-negative observations. The level of analysis is individual-day. All specifications include individual and day fixed effects. Standard errors are clustered at the postal code level. Column 1 displays the average impact of temperature on doors canvassed. Column 2 further includes the cumulative number of days canvassed until a given day and whether an activist has signaled their effort on social media at any point before a given day. Column 3 includes interactions of these variables interacted with temperature.

Table 5.4: Impact of temperature on any signalling

	Dependent variable is any signaling			
	(1)	(2)	(3)	(4)
Temperature	0.0002**	0.0002**	0.0004***	0.0003***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Temperature × Has signaled before			-0.0016***	-0.0014***
			(0.0003)	(0.0003)
Has signaled before		0.0307***	0.0599***	0.0508***
		(0.0033)	(0.0066)	(0.0062)
Temperature × Days canvassed			-0.0002*	-0.0001
			(0.0001)	(0.0001)
Days canvassed		0.0013	0.0050**	0.0042*
		(0.0008)	(0.0026)	(0.0024)
Contemporaneous canvassing				0.1409***
				(0.0066)
Outcome mean	0.0049	0.0049	0.0049	0.0049
Number obs	809442	809442	809442	809442

Notes: Table 5.4 displays the effect of temperature on any signalling. Any signalling is defined as sharing one's points or title on facebook. The level of analysis is individual-day. All specifications include individual and day fixed effects. Standard errors are clustered at the postal code level. Column 1 displays the average impact of temperature on any signalling. Column 2 further includes the cumulative number of days canvassed until a given day and whether an activist has signaled their effort on social media at any point before a given day. Column 3 includes interactions of these variables interacted with temperature. Column 4 further controls for any canvassing activity on a given day.

Table 5.5: Summary statistics by signalling

	Signaled	Did not signal	Δ	se(Δ)	p(Δ=0)	Observations
Canvassing outcomes variables						
Any canvassing	0.887	0.179	0.708	(0.043)	0.000	767
Days canvassed	0.887	0.204	0.683	(0.043)	0.000	767
Doors visited (winsorized)	522.452	47.661	474.791	(142.273)	0.001	767
Psychological variables						
Trust index	0.095	-0.008	0.103	(0.141)	0.465	715
Self-efficacy index	0.261	-0.023	0.284	(0.122)	0.020	712
Expectations index	0.152	-0.013	0.166	(0.143)	0.247	721
Identification with party (z-scored)	0.218	-0.019	0.237	(0.102)	0.020	729
Identification with party platform (z-scored)	0.092	-0.008	0.100	(0.110)	0.364	726
Activities in party						
Considers political post in party	0.508	0.314	0.195	(0.068)	0.004	744
Considers running for office	0.729	0.429	0.300	(0.061)	0.000	744
Has political post in party	0.034	0.109	-0.076	(0.026)	0.004	744
Has run for office	0.051	0.193	-0.142	(0.032)	0.000	744
Perceived cohesion in local party	0.325	-0.028	0.353	(0.131)	0.007	731
Number of party members known personally	37.569	39.747	-2.178	(4.071)	0.593	731
Demographics						
Female	0.210	0.227	-0.017	(0.054)	0.750	767
Age	31.274	45.460	-14.185	(1.897)	0.000	767
Participated in door-to-door workshop	0.435	0.312	0.123	(0.065)	0.060	767
Is party member	0.984	0.929	0.055	(0.019)	0.004	767
Years of party membership	5.210	15.146	-9.936	(1.010)	0.000	767
Correct knowledge questions (out of 5)	2.140	2.142	-0.002	(0.192)	0.993	629
Risk preference (z-scored)	0.108	-0.009	0.117	(0.146)	0.421	717
Time preference (z-scored)	-0.007	0.001	-0.008	(0.139)	0.956	717

Notes: Table 5.5 compares the characteristics of individuals who signaled and who did not signal during the electoral campaign. The data was obtained two months after the election using a post-election survey sent out to all party supporters that expressed interest in supporting the party.

5.8 Appendix Tables

Table A1: Impact of maximum temperature on any canvassing

	Dependent variable is any canvassing		
	(1)	(2)	(3)
Max temperature	0.0006***	0.0006***	0.0008***
	(0.0001)	(0.0001)	(0.0001)
Max temperature × Has signaled before			-0.0015***
			(0.0003)
Has signaled before		0.0355***	0.0654***
		(0.0039)	(0.0075)
Max temperature × Days canvassed			-0.0003
			(0.0002)
Days canvassed		-0.0004	0.0056
		(0.0014)	(0.0047)
Outcome mean	0.0154	0.0154	0.0154
Number obs	809442	809442	809442

Notes: Table A1 displays the effect of daily maximum temperature on a dummy for any canvassing activity. The level of analysis is individual-day. All specifications include postal code and day fixed effects. Standard errors are clustered at the 2-digit postal code level (93 clusters). Column 1 displays the average impact of temperature on any canvassing. Column 2 further includes the cumulative number of days canvassed until a given day and whether an activist has signaled their effort on social media at any point before a given day. Column 3 includes interactions of these variables interacted with temperature.

Table A2: Impact of maximum temperature on doors (winsorized)

	Dependent variable is doors canvassed		
	(1)	(2)	(3)
Max temperature	0.0429***	0.0446***	0.0549***
	(0.0132)	(0.0131)	(0.0148)
Max temperature × Has signaled before			-0.1594***
			(0.0323)
Has signaled before		3.4043***	6.6195***
		(0.5576)	(1.1016)
Max temperature × Days canvassed			0.0180
			(0.0142)
Days canvassed		0.4271***	0.0839
		(0.1334)	(0.3635)
Outcome mean	0.8304	0.8304	0.8304
Number obs	809442	809442	809442

Notes: Table A2 displays the effect of daily maximum temperature on doors canvassed. Doors canvassed is winsorized at the 99^th percentile of non-negative observations. The level of analysis is individual-day. All specifications include postal code and day fixed effects. Standard errors are clustered at the 2-digit postal code level (93 clusters). Column 1 displays the average impact of temperature on doors canvassed. Column 2 further includes the cumulative number of days canvassed until a given day and whether an activist has signaled their effort on social media at any point before a given day. Column 3 includes interactions of these variables interacted with temperature.

Table A3: Impact of maximum temperature on any signalling

	Dependent variable is any signaling			
	(1)	(2)	(3)	(4)
Max temperature	0.0002**	0.0002**	0.0003***	0.0002***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Max temperature × Has signaled before			-0.0015***	-0.0013***
			(0.0003)	(0.0003)
Has signaled before		0.0307***	0.0608***	0.0516***
		(0.0033)	(0.0068)	(0.0063)
Max temperature × Days canvassed			-0.0002	-0.0001
			(0.0001)	(0.0001)
Days canvassed		0.0013	0.0048*	0.0040*
		(0.0008)	(0.0025)	(0.0024)
Contemporaneous canvassing				0.1409***
				(0.0066)
Outcome mean	0.0049	0.0049	0.0049	0.0049
Number obs	809442	809442	809442	809442

Notes: Table A3 displays the effect of daily maximum temperature on any signalling. Any signalling is defined as sharing one's points or title on facebook. The level of analysis is individual-day. All specifications include individual and day fixed effects. Standard errors are clustered at the 2-digit postal code level (93 clusters). Column 1 displays the average impact of temperature on any signalling. Column 2 further includes the cumulative number of days canvassed until a given day and whether an activist has signaled their effort on social media at any point before a given day. Column 3 includes interactions of these variables interacted with temperature. Column 4 further controls for any canvassing activity on a given day.

Table A4: Impact of temperature on any canvassing: two-digit postal code fixed effects

	Dependent variable is any canvassing		
	(1)	(2)	(3)
Temperature	0.0008***	0.0008***	0.0011***
	(0.0001)	(0.0001)	(0.0001)
Temperature × Has signaled before			-0.0015***
			(0.0003)
Has signaled before		0.0281***	0.0551***
		(0.0026)	(0.0064)
Temperature × Days canvassed			-0.0005**
			(0.0002)
Days canvassed		0.0037***	0.0126***
		(0.0010)	(0.0045)
Outcome mean	0.0154	0.0154	0.0154
Number obs	809442	809442	809442

Notes: Table A4 displays the effect of temperature on a dummy for any canvassing activity. The level of analysis is individual-day. All specifications include first two digit of postal code and day fixed effects. Standard errors are clustered at the postal code level. Column 1 displays the average impact of temperature on any canvassing. Column 2 further includes the cumulative number of days canvassed until a given day and whether an activist has signaled their effort on social media at any point before a given day. Column 3 includes interactions of these variables interacted with temperature.

Table A5: Impact of temperature on doors (winsorized): two-digit postal code fixed effects

	Dependent variable is doors canvassed		
	(1)	(2)	(3)
Temperature	0.0607***	0.0630***	0.0735***
	(0.0170)	(0.0170)	(0.0179)
Temperature × Has signaled before			-0.1641***
			(0.0328)
Has signaled before		2.3780***	5.3000***
		(0.3415)	(0.8433)
Temperature × Days canvassed			0.0186
			(0.0160)
Days canvassed		0.2884***	-0.0315
		(0.0769)	(0.3288)
Outcome mean	0.8304	0.8304	0.8304
Number obs	809442	809442	809442

Notes: Table A5 displays the effect of temperature on doors canvassed. Doors canvassed is winsorized at the 99^th percentile of non-negative observations. The level of analysis is individual-day. All specifications include first two digit of postal code and day fixed effects. Standard errors are clustered at the postal code level. Column 1 displays the average impact of temperature on doors canvassed. Column 2 further includes the cumulative number of days canvassed until a given day and whether an activist has signaled their effort on social media at any point before a given day. Column 3 includes interactions of these variables interacted with temperature.

Table A6: Impact of average temperature on any signalling: two-digit postal code fixed effects

	Dependent variable is any signaling			
	(1)	(2)	(3)	(4)
Temperature	0.0002*	0.0002**	0.0004***	0.0002**
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Temperature × Has signaled before			-0.0016***	-0.0014***
			(0.0003)	(0.0003)
Has signaled before		0.0325***	0.0608***	0.0529***
		(0.0024)	(0.0066)	(0.0062)
Temperature × Days canvassed			-0.0001	-0.0000
			(0.0001)	(0.0001)
Days canvassed		-0.0005	0.0018	0.0000
		(0.0006)	(0.0023)	(0.0022)
Contemporaneous canvassing				0.1445***
				(0.0071)
Outcome mean	0.0049	0.0049	0.0049	0.0049
Number obs	809442	809442	809442	809442

Notes: Table A6 displays the effect of temperature on any signalling. Any signalling is defined as sharing one's points or title on facebook. The level of analysis is individual-day. All specifications include first two digit of postal code and day fixed effects. Standard errors are clustered at the postal code level. Column 1 displays the average impact of temperature on any signalling. Column 2 further includes the cumulative number of days canvassed until a given day and whether an activist has signaled their effort on social media at any point before a given day. Column 3 includes interactions of these variables interacted with temperature. Column 4 further controls for any canvassing activity on a given day.

Table A7: Impact of temperature on any canvassing: cluster at two-digit postal code level

	Dependent variable is any canvassing		
	(1)	(2)	(3)
Temperature	0.0008***	0.0008***	0.0010***
	(0.0002)	(0.0002)	(0.0002)
Temperature × Has signaled before			-0.0016***
			(0.0003)
Has signaled before		0.0355***	0.0646***
		(0.0041)	(0.0074)
Temperature × Days canvassed			-0.0003
			(0.0004)
Days canvassed		-0.0004	0.0060
		(0.0021)	(0.0076)
Outcome mean	0.0154	0.0154	0.0154
Number obs	809442	809442	809442

Notes: Table A7 displays the effect of temperature on a dummy for any canvassing activity. The level of analysis is individual-day. All specifications include postal code and day fixed effects. Standard errors are clustered at the first two digit of postal code level (93 clusters). Column 1 displays the average impact of temperature on any canvassing. Column 2 further includes the cumulative number of days canvassed until a given day and whether an activist has signaled their effort on social media at any point before a given day. Column 3 includes interactions of these variables interacted with temperature.

Table A8: Impact of temperature on doors (winsorized): cluster at two-digit postal code

	Dependent variable is doors canvassed		
	(1)	(2)	(3)
Temperature	0.0547***	0.0570***	0.0687***
	(0.0205)	(0.0201)	(0.0221)
Temperature × Has signaled before			-0.1787***
			(0.0340)
Has signaled before		3.4041***	6.5915***
		(0.5895)	(1.0738)
Temperature × Days canvassed			0.0191
			(0.0253)
Days canvassed		0.4283**	0.1107
		(0.1781)	(0.5386)
Outcome mean	0.8304	0.8304	0.8304
Number obs	809442	809442	809442

Notes: Table A8 displays the effect of temperature on doors canvassed. Doors canvassed is winsorized at the 99^th percentile of non-negative observations. The level of analysis is individual-day. All specifications include postal code and day fixed effects. Standard errors are clustered at the first two digit of postal code level (93 clusters). Column 1 displays the average impact of temperature on doors canvassed. Column 2 further includes the cumulative number of days canvassed until a given day and whether an activist has signaled their effort on social media at any point before a given day. Column 3 includes interactions of these variables interacted with temperature.

Table A9: Impact of average temperature on any signalling: cluster at two-digit postal code level

	Dependent variable is any signaling			
	(1)	(2)	(3)	(4)
Temperature	0.0002*	0.0002**	0.0004***	0.0003**
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Temperature × Has signaled before			-0.0016***	-0.0014***
			(0.0004)	(0.0003)
Has signaled before		0.0307***	0.0599***	0.0508***
		(0.0035)	(0.0074)	(0.0067)
Temperature × Days canvassed			-0.0002	-0.0001
			(0.0001)	(0.0001)
Days canvassed		0.0013	0.0050*	0.0042*
		(0.0008)	(0.0027)	(0.0024)
Contemporaneous canvassing				0.1409***
				(0.0094)
Outcome mean	0.0049	0.0049	0.0049	0.0049
Number obs	809442	809442	809442	809442

Notes: Table A9 displays the effect of temperature on any signalling. Any signalling is defined as sharing one's points or title on facebook. The level of analysis is individual-day. All specifications include individual and day fixed effects. Standard errors are clustered at the first two digit of postal code level (93 clusters). Column 1 displays the average impact of temperature on any signalling. Column 2 further includes the cumulative number of days canvassed until a given day and whether an activist has signaled their effort on social media at any point before a given day. Column 3 includes interactions of these variables interacted with temperature. Column 4 further controls for any canvassing activity on a given day.

Chapter 6

Conclusion

This thesis used information theory to study policy questions in banking, finance, and industrial organization. We showed that central banks face a time inconsistency problem when publishing bank stress test results, characterised institutional design solutions, and found that a hurdle rate framework, where all banks are judged to pass or fail relative to a common threshold, is optimal in many settings. We showed that, in a hurdle rate framework, stress tests become an informational contagion channel. This suggests that it can be socially desirable that central banks do not release all their information.

We studied imperfect competition in markets with selection and analysed whether lax competition policy is socially desirable in markets with advantageous selection. While we acknowledge that markets with advantageous selection can experience inefficient overprovision under perfect competition, we showed that competition policy is not the most targeted and most suitable tool to correct this inefficiency. We found that tolerating intermediate levels of competition never achieves the first best efficient allocation. However, the first best efficient allocation can be reached through a combination of a corrective tax and tough competition policy.

We showed that, in markets with adverse selection, tying contracts can prevent unravelling and thus have a substantial positive effect on welfare. This suggests that while tying contracts may sometimes be socially harmful, e.g. when a monopolist uses tying contracts to exploit its market power, they can be socially beneficial when they cover two markets with adverse selection and prevent unravelling.

In a short fourth paper, we showed that information theory can rationalize patterns of behaviour which are observed in a field experiment. Specifically, we showed that some agents (political activists) increase their effort (door-to-door canvassing) when it becomes more costly (bad weather) and interpret this in a signalling model.

Chapter 7

Bibliography

Adams, William James and Janet L Yellen, "Commodity Bundling and the Burden of Monopoly," *The Quarterly Journal of Economics*, 1976, *90* (3), 475–498.

Akerlof, George A, "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism," *The Quarterly Journal of Economics*, 1970, *84* (3), 488–500.

Allen, Franklin and Douglas Gale, "Optimal Financial Crises," *The Journal of Finance*, August 1998, *53* (4), 1245–1284.

Angeletos, George-Marios and Alessandro Pavan, "Efficient Use of Information and Social Value of Information," *Econometrica*, 2007, *75* (4), 1103–1142.

Armstrong, Mark, "Multiproduct Nonlinear Pricing," *Econometrica*, 1996, *64* (1), 51–75.

_ , "Recent Advances in the Economics of Price Discrimination," in Richard Blundell, Whitney K. Newey, and Torsten Persson, eds., *Advances in Economics and Econometrics: Theory and Applications*, Vol. 2 of *Ninth World Congress* Cambridge University Press 2006, pp. 97–141.

_ **and Jean-Charles Rochet**, "Multi-dimensional screening: A user's guide," *European Economic Review*, 1999, *43*, 959–979.

_ **and John Vickers**, "Competitive Non-linear Pricing and Bundling," *The Review of Economic Studies*, January 2010, *77* (1), 30–60.

Asriyan, Vladimir, William Fuchs, and Brett Green, "Information Spillovers in Asset Markets with Correlated Values," *American Economic Review*, 2017, *107* (7), 2007–2040.

Baron, David P. and Roger B. Myerson, "Regulating a Monopolist with Unknown Costs," *Econometrica*, 1982, *50* (4), 911–930.

Barro, Robert J. and David B. Gordon, "A Positive Theory of Monetary Policy in a Natural Rate Model," *Journal of Political Economy*, 1983, *91* (4), 589–610.

_ **and** _ , "Rules, discretion and reputation in a model of monetary policy," *Journal of Monetary Economics*, 1983, *12* (1), 101–121.

Bénabou, Roland and Jean Tirole, "Incentives and Prosocial Behavior," *The American Economic Review*, 2006, *96* (5), 1652–1678.

Bénabou, Roland, Armin Falk, and Jean Tirole, "Narratives, Imperatives and Moral Reasoning," Technical Report, Working Paper 2018.

Bergemann, Dirk and Stephen Morris, "Information Design: A Unified Perspective," *Journal of Economic Literature*, 2019, *57* (1), 1ß57.

Bernheim, B Douglas, "A Theory of Conformity," *Journal of Political Economy*, 1994, *102* (5), 841–877.

Blake, David, "The consequences of not having to buy an annuity," *Discussion Paper*, 2014.

Blinder, Alan S, Michael Ehrmann, Marcel Fratzscher, Jakob De Haan, and David-Jan Jansen, "Central Bank Communication and Monetary Policy: A Survey of Theory and Evidence," *Journal of Economic Literature*, 2008, *46* (4), 910–945.

Boleslavsky, Raphael and Christopher Cotton, "Grading Standards and Education Quality," *American Economic Journal: Microeconomics*, May 2015, *7* (2), 248–279.

Bolton, Patrick and Mathias Dewatripont, *Contract Theory*, Cambridge, Massachusetts: The MIT Press, 2004.

Boot, Arnoud W.A. and Anjan V. Thakor, "Can Relationship Banking Survive Competition?," *The Journal of Finance*, April 2000, *55* (2), 679–713.

Bouvard, Matthieu, Pierre Chaigneau, and Adolfo De Motta, "Transparency in the Financial System: Rollover Risk and Crises," *The Journal of Finance*, 2015, *70* (4), 1805–1837.

Bowles, Samuel and Sandra Polania-Reyes, "Economic incentives and social preferences: substitutes or complements?," *Journal of Economic Literature*, 2012, *50* (2), 368–425.

Bresnahan, Timothy F, "Empirical Studies of Industries with Market Power," *Handbook of Industrial Organization*, 1989, *II*, 1012–1057.

Burdett, Kenneth and Kenneth L. Judd, "Equilibrium Price Dispersion," *Econometrica*, July 1983, *51* (4), 955–969.

Bursztyn, Leonardo and Robert Jensen, "How Does Peer Pressure Affect Educational Investments?," *The Quarterly Journal of Economics*, 2015, *130* (3), 1329–1367.

_ **and** _ , "Social Image and Economic Behavior in the Field: Identifying, Understanding, and Shaping Social Pressure," *Annual Review of Economics*, 2017, *9*, 131–153.

_ **, Bruno Ferman, Stefano Fiorin, Martin Kanz, and Gautam Rao**, "Status Goods: Experimental Evidence from Platinum Credit Cards," *The Quarterly Journal of Economics*, 2017.

_ **, Georgy Egorov, and Robert Jensen**, "Cool to be Smart or Smart to be Cool? Understanding Peer Pressure in Education," *The Review of Economic Studies*, 2017.

_ **, Thomas Fujiwara, and Amanda Pallais**, "'Acting Wife': Marriage Market Incentives and Labor Market Investments," *American Economic Review*, 2017, *107* (11), 3288–3319.

Cantoni, Davide, David Y Yang, Noam Yuchtman, and Y Jane Zhang, "The Fundamental Determinants of Anti-Authoritarianism," *Working Paper*, 2016.

_ **, David Yang, Noam Yuchtman, and Jane Zhang**, "Protests as Strategic Games: Experimental Evidence from Hong Kong's Antiauthoritarian Movement," *The Quarterly Journal of Economics*, 2019, *134* (2), 1021–1077.

Carlsson, Hans and Eric van Damme, "Global Games and Equilibrium Selection," *Econometrica*, 1993, *61* (5), 989–1018.

Carroll, Gabriel, "Robustness and Separation in Multidimensional Screening," *Econometrica*, 2017, *85* (2), 453–488.

Chiappori, Pierre-André and Bernard Salanié, "Testing for Asymmetric Information in Insurance Markets," *Journal of Political Economy*, 2000, *108* (1), 56–78.

_ **and** _ , "Testing Contract Theory: A Survey of Some Recent Work," *Advances in Economics and Econometrics: Theory and Applications, Eighth World Congress.*, 2003, *1*, 115–149.

_ **, Bruno Jullien, Bernard Salanié, and François Salanié**, "Asymmetric information in insurance: General testable implications," *RAND Journal of Economics*, 2006, *37* (4), 783–798.

Cho, In-Koo and David M Kreps, "Signaling Games and Stable Equilibria," *The Quarterly Journal of Economics*, 1987, *102* (2), 179–222.

Cohen, Alma and Liran Einav, "Estimating Risk Preferences from Deductible Choice," *The American Economic Review*, 2007, *97* (3), 745–788.

Competition and Markets Authority, "Making banks work harder for you: Final Report of the Market Investigation into Retail Banking," August 2016.

Corts, Kenneth S., "Third-Degree Price Discrimination in Oligopoly: All-Out Competition and Strategic Commitment," *The RAND Journal of Economics*, 1998, *29* (2), 306–323.

Crawford, Gregory S., Nicola Pavanini, and Fabiano Schivardi, "Asymmetric Information and Imperfect Competition in Lending Markets," *American Economic Review*, July 2018, *108* (7), 1659–1701.

Crawford, Vincent P and Joel Sobel, "Strategic Information Transmission," *Econometrica*, 1982, *50* (6), 1431–1451.

Cutler, David M. and Richard J. Zeckhauser, "Adverse Selection in Health Insurance," *NBER Working Paper*, 1997.

Dafny, Leemore S., "Are Health Insurance Markets Competitive?," *American Economic Review*, September 2010, *100*, 1399–1431.

De Loecker, Jan, Jan Eeckhout, and Gabriel Unger, "The Rise of Market Power and the Macroeconomic Implications," *Quarterly Journal of Economics*, 2020, *135* (2), 561–644.

De Meza, David and David C. Webb, "Too Much Investment: A Problem of Asymmetric Information," *The Quarterly Journal of Economics*, 1987, *102* (2), 281–292.

_ **and** _ , "Advantageous Selection in Insurance Markets," *The RAND Journal of Economics*, 2001, *32* (2), 249–262.

Deci, Edward L, "Effects of Externally Mediated Rewards on Intrinsic Motivation," *Journal of Personality and Social Psychology*, 1971, *18* (1), 105.

DellaVigna, Stefano, John A List, and Ulrike Malmendier, "Testing for Altruism and Social Pressure in Charitable Giving," *The Quarterly Journal of Economics*, 2012, *127* (1), 1–56.

_ **, John A. List, Ulrike Malmendier, and Gautam Rao**, "Voting to Tell Others," *The Review of Economic Studies*, 2017, *84* (1), 143–181.

Diamond, Douglas W. and Philip H. Dybvig, "Bank Runs, Deposit Insurance, and Liquidity," *Journal of Political Economy*, 1983, *91* (3), 401.

Dubé, Jean-Pierre, Xueming Luo, and Zheng Fang, "Self-signaling and prosocial behavior: A cause marketing experiment," *Marketing Science*, 2017, *36* (2), 161–186.

Dubey, Pradeep and John Geanakoplos, "Grading exams: 100, 99, 98, ... or A, B, C?," *Games and Economic Behavior*, 2010, *69*, 72–94.

Einav, Liran, Amy Finkelstein, and Mark R. Cullen, "Estimating Welfare in Insurance Markets Using Variation in Prices," *The Quarterly Journal of Economics*, 2010, *125* (3), 877–921.

_ **and** _ , "Selection in Insurance Markets: Theory and Empirics in Pictures," *Journal of Economic Perspectives*, 2011, *25* (1), 115–138.

_ **, Mark Jenkins, and Jonathan Levin**, "Contract Pricing in Consumer Credit Markets," *Econometrica*, 2012, *80* (4), 1387–1432.

Enikolopov, Ruben, Alexey Makarin, and Maria Petrova, "Social Media and Protest Participation: Evidence from Russia," *Working Paper*, 2016.

_ **,** _ **,** _ **, and Leonid Polishchuk**, "Social Image, Networks, and Protest Participation," *Working Paper*, 2017.

Enos, Ryan D and Eitan D Hersh, "Party Activists as Campaign Advertisers: The Ground Campaign as a Principal-Agent Problem," *American Political Science Review*, 2015, *109* (2), 252–278.

European Commision, "Sector Inquiry on Retail Banking: Final Report," January 2007.

Falk, Armin, "Facing Yourself-A Note on Self-image," *Working Paper*, 2017.

_ **and Michael Kosfeld**, "The Hidden Costs of Control," *American Economic Review*, 2006, *96* (5), 1611–1630.

Fang, Hanming, Michael P Keane, and Dan Silverman, "Sources of Advantageous Selection: Evidence from the Medigap Insurance Market," *The Journal of Political Economy*, 2008, *116* (2), 303–350.

Faria-e-castro, Miguel, Joseba Martinez, and Thomas Philippon, "Runs versus Lemons: Information Disclosure and Fiscal Capacity," *Review of Economic Studies*, 2016, pp. 1–25.

Finkelstein, Amy and Kathleen McGarry, "Multiple Dimensions of Private Information: Evidence from the Long-Term Care Insurance Market," *The American Economic Review*, 2006, *96* (4), 938–958.

Frey, Bruno S and Reto Jegen, "Motivation Crowding Theory," *Journal of Economic Surveys*, 2001, *15* (5), 589–611.

Gambacorta, Leonardo and Hyun Song Shin, "Why bank capital matters for monetary policy," *Journal of Financial Intermediation*, 2018, *35*, 17–29.

Geraats, Petra M, "Central Bank Transparency," *The Economic Journal*, 2002, *112*, F532–F565.

Gerber, Alan S., Donald P. Green, and Chrisopher W. Larimer, "Social Pressure and Voter Turnout: Evidence from a Large-Scale Field Experiment," *American Political Science Review*, 2008, *102* (1), 33–48.

Gneezy, Uri and Aldo Rustichini, "A Fine is a Price," *The Journal of Legal Studies*, 2000, *29* (1), 1–17.

_ **and** _ , "Pay Enough or Don't Pay At All," *The Quarterly Journal of Economics*, 2000, *115* (3), 791–810.

_ **, Stephan Meier, and Pedro Rey-Biel**, "When and Why Incentives (don't) Work to Modify Behavior," *Journal of Economic Perspectives*, 2011, *25* (4), 191–210.

Goldstein, Itay and Haresh Sapra, "Should Banks' Stress Test Results be Disclosed? An Analysis of the Costs and Benefits," *Foundations and Trends in Finance*, 2013, *8* (1), 1–54.

_ **and Liyan Yang**, "Good disclosure, bad disclosure," *Journal of Financial Economics*, 2019, *131*, 118–138.

_ **and Yaron Leitner**, "Stress tests and information disclosure," *Journal of Economic Theory*, 2018, *177*, 34–69.

González, Felipe, "Collective Action in Networks: Evidence from the Chilean Student Movement," *Journal of Public Economics*, 2020, *188.*

Greene, David and Mark R Lepper, *The hidden costs of reward: New perspectives of the psychology of human motivation*, Hillsdale, NJ: L. Erlbaum Associates; New York, 1978.

Grennan, Matthew, "Price Discrimination and Bargaining: Empirical Evidence from Medical Devices," *American Economic Review*, February 2013, *103* (1), 145–177.

Grossman, Sanford J, "The Informational Role of Warranties and Private Disclosure about Product Quality," *The Journal of Law and Economics*, 1981, *24* (3), 461–483.

_ **and Oliver D Hart**, "An Analysis of the Principal-Agent Problem," *Econometrica*, 1983, *51* (1), 7–45.

Gutierrez, German and Thomas Philippon, "Declining Competition and Investment in the U.S.," *Working Paper*, 2017.

Hemenway, David, "Propitious Selection," *The Quarterly Journal of Economics*, 1990, *105* (4), 1063–1069.

Hensel, Lukas, Johannes Hermle, Anselm Rink, and Christopher Roth, "Competition and Political Engagement: Evidence From a Natural Field Experiment," *Working Paper*, 2018.

_ **,** _ **,** _ **, and** _ , "Political Activists as Free-Riders: Evidence From a Natural Field Experiment," *Working Paper*, 2018.

Hirshleifer, Jack, "The Private and Social Value of Information and the Reward to Inventive Activity," *American Economic Review*, September 1971, *61* (4), 561–574.

Holmstrom, Bengt, "Moral Hazard and Observability," *The Bell Journal of Economics*, 1979, *10* (1), 74–91.

_ , "Moral Hazard in Teams," *The Bell Journal of Economics*, 1982, *13* (2), 324–340.

Hotelling, Harold, "Stability in Competition," *The Economic Journal*, 1929, *39* (135), 41–57.

Huangfu, Bingchao and Heng Liu, "Information Spillover in Multi-good Adverse Selection," *Working Paper*, 2019.

Inostroza, Nicolas and Alessandro Pavan, "Persuasion in Global Games with Application to Stress Testing," *Working Paper*, January 2018.

Kamenica, Emir, "Behavioral Economics and Psychology of Incentives," *Annual Review of Economics*, 2012, *4* (1), 427–452.

_ **and Matthew Gentzkow**, "Bayesian Persuasion," *The American Economic Review*, 2011, *101* (6), 2590–2615.

Karing, Anne, "Social Signaling and Childhood Immunization: A Field Experiment in Sierra Leone," *Working Paper*, 2018.

_ **and Karim Naguib**, "Social Signaling and Prosocial Behavior: Experimental Evidence in Community Deworming in Kenya," *Working Paper*, 2017.

Karlan, Dean and Jonathan Zinman, "Observing Unobservables: Identifying Information Asymmetries with a Consumer Credit Field Experiment," *Econometrica*, November 2009, *77* (6), 1993–2008.

Kydland, Finn E. and Edward C. Prescott, "Rules Rather than Discretion: The Inconsistency of Optimal Plans," *Journal of Political Economy*, June 1977, *85* (3), 473–492.

Lambert, Richard A., "Long-Term Contracts and Moral Hazard," *The Bell Journal of Economics*, 1983, *14* (2), 441–452.

Lester, Benjamin, Ali Shourideh, Venky Venkateswaran, and Ariel Zetlin-Jones, "Screening and Adverse Selection in Frictional Markets," *Journal of Political Economy*, 2019, *127* (1).

Lewis, Tracy R. and David E.M. Sappington, "Countervailing Incentives in Agency Problems," *Journal of Economic Theory*, 1989, *49* (2), 294–313.

Loranth, Gyöngyi and Alan D. Morrison, "Tying in Universal Banks," *Review of Finance*, 2012, *16*, 481–516.

Lucas, Robert E., "Expectations and the Neutrality of Money," *Journal of Economic Theory*, 1972, *4*, 103–124.

_ , "An Equilibrium Model of the Business Cycle," *Journal of Political Economy*, December 1975, *83* (6), 1113–1144.

Mahoney, Neale and E Glen Weyl, "Imperfect Competition in Selection Markets," *NBER Working Paper*, 2014.

_ **and** _ , "Imperfect Competition in Selection Markets," *Review of Economics and Statistics*, 2017, *99* (4), 637–651.

McAfee, R. Preston, John McMillan, and Michael D. Whinston, "Multiproduct Monopoly, Commodity Bundling, and Correlation of Values," *The Quarterly Journal of Economics*, 1989, *104* (2), 371–383.

Milgrom, Paul R., "Good News and Bad News: Representation Theorems and Applications," *The Bell Journal of Economics*, 1981, *12* (2), 380.

Morris, Stephen and Hyun Song Shin, "Social Value of Public Information," *The American Economic Review*, 2002, *92* (5), 1521–1534.

_ **and** _ , "Central Bank Transparency and the Signal Value of Prices," *Brookings Papers on Economic Activity*, 2005, *1* (September), 1–66.

_ **and** _ , "Optimal Communication," *Journal of the European Economic Association*, 2007, *5* (2/3), 594–602.

_ **and** _ , "Central Bank Forward Guidance and the Signal Value of Market Prices," *American Economic Association: Papers and Proceedings*, May 2018, *108*, 572–577.

_ **,** _ **, and Hui Tong**, "Social Value of Public Information: Morris and Shin (2002) Is Actually Pro-Transparency, Not Con: Reply," *The American Economic Review*, 2006, *96* (1), 453–455.

Orlov, Dmitry, Pavel Zryumov, and Andrzej Skrzypacz, "Design of Macroprudential Stress Tests," *Working Paper*, May 2017.

Ottaviani, Marco and Peter Norman Sorensen, "Professional advice," *Journal of Economic Theory*, 2006, *126*, 120–142.

Parlatore, Cecilia, "Transparency and Bank Runs," *Working Paper*, 2015.

Passarelli, Francesco and Guido Tabellini, "Emotions and Political Unrest," *Journal of Political Economy*, 2017, *125* (3), 903–946.

Perez-Truglia, Ricardo and Guillermo Cruces, "Partisan Interactions," *Journal of Political Economy*, 2017, *125* (4), 1208–1243.

Ralph, Oliver, "Pension freedoms leave people with 'little choice'," *Financial Times*, nov 2017.

Reis, Ricardo, "Central Bank Design," *The Journal of Economic Perspectives*, 2013, *27* (4), 17–43.

Rochet, Jean-Charles and Lars A. Stole, "The Economics of Multidimensional Screening," in Mathias Dewatripont, Lars Peter Hansen, and Stephen J. Turnovsky, eds., *Advances in Economics and Econometrics: Theory and Applications*, Cambridge: Cambridge University Press, 2003, chapter 5, pp. 150–198.

Rogoff, Kenneth, "The Optimal Degree of Commitment to an Intermediate Monetary Target," *The Quarterly Journal of Economics*, November 1985, *100* (4), 1169–1189.

Rothschild, Michael and Joseph Stiglitz, "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information," *The Quarterly Journal of Economics*, 1976, *90* (4), 629–649.

Shaked, Moshe and George J. Shanthikumar, *Stochastic Orders* Springer Series in Statistics, Springer, 2007.

Shapiro, Joel and Jing Zeng, "Stress Testing and Bank Lending," *Working Paper*, July 2018.

Sharpe, Steven A., "Asymmetric Information, Bank Lending and Implicit Contracts: A Stylized Model of Customer Relationships," *The Journal of Finance*, September 1990, *45* (4), 1069–1087.

Smart, Michael, "Competitive Insurance Markets with Two Unobservables," *International Economic Review*, 2000, *41* (1), 153–169.

Spence, Michael, "Job Market Signaling," *The Quarterly Journal of Economics*, 1973, *87* (3), 355–374.

Stigler, George J., "United States v. Loew's Inc.: A Note on Block-Booking," *The Supreme Court Review*, 1963, pp. 152–157.

Stiglitz, Joseph E. and Andrew Weiss, "Credit Rationing in Markets with Imperfect Information," *The American Economic Review*, 1981, *71* (3), 393–410.

Sturrock, David, "The role of pessimism about survival in explaining the 'annuity puzzle'," *Institute for Fiscal Studies: Observation*, January 2019.

Svensson, Lars E O, "Social Value of Public Information: Comment: Morris and Shin (2002) Is Actually Pro-Transparency, Not Con," *The American Economic Review*, 2006, *96* (1), 448–452.

Thisse, Jacques-Francois and Xavier Vives, "On The Strategic Choice of Spatial Price Policy," *American Economic Review*, March 1988, *78* (1), 122–137.

Veiga, André and E. Glen Weyl, "Product Design in Selection Markets," *The Quarterly Journal of Economics*, 2016, pp. 1007–1056.

Vickers, John, "Delegation and the Theory of the Firm," *The Economic Journal*, 1985, *95*, 138–147.

Vives, Xavier, *Competition and Stability in Banking*, Princeton University Press, 2016.

Walsh, Carl E., "Optimal Contracts for Central Bankers," *The American Economic Review*, March 1995, *85* (1), 150–167.

Weyl, E Glen and Michal Fabinger, "Pass-Through as an Economic Tool: Principles of Incidence under Imperfect Competition," *Journal of Political Economy*, 2013, *121* (3), 528–583.

Whinston, Michael D., "Tying, Foreclosure, and Exclusion," *American Economic Review*, September 1990, *80* (4), 837–859.

Williams, Basil, "Stress Tests and Bank Portfolio Choice," *Working Paper*, 2017.