ESSAYS IN INFORMATIONAL ECONOMICS

INAUGURAL-DISSERTATION

Robert Ulbricht

Essays in Informational Economics

INAUGURAL-DISSERTATION

zur Erlangung des Grades Doctor oeconomiae publicae (Dr. oec. publ.)
an der Ludwig-Maximilians-Universität München

2012

vorgelegt von

ROBERT ULBRICHT

Referent: Prof. Dr. Klaus M. Schmidt
Korreferent: Prof. Matthias Doepke, Ph.D.
Promotionsabschlussberatung: 15. Mai 2013

Für meine Eltern
Andrea, Andreas, Michael & Steffi

CONTENTS OF THE DISSERTATION

PREFACE

Informational economics

Information is an important aspect of many economic situations: buyers typically inform themselves about product characteristics and wish to search for the most favorable price before buying; employers wish to ensure that workers do a satisfactory job, but are often unable to directly monitor workers' efforts; and firms typically do not perfectly know the profitability of investing, prompting them to infer about business conditions by carefully evaluating a variety of market indicators before making their production choices.

All of these examples have in common that information is not freely available, but is subject to *frictions* that make it a valuable resource. The pioneering study asserting the importance of informational frictions is Stigler (1961). In this study, Stigler analyzes the decision problem of a buyer who does not know the prices charged by nearby stores. Just like in the first of the aforementioned examples, in this case the optimal strategy of the buyer amounts to sampling a certain number of stores and buying the good once he finds an offer that is *sufficiently* cheap. This informational friction of buyers not knowing the full distribution of prices is a first example of how the lack of knowledge may explain a systematically different behavior than under a full-information benchmark.

Half a century later, informational frictions have become a core part of many economist's toolboxes. As of today, the field of informational economics can be broadly divided into three areas.

Search Theory First, problems in the spirit of Stigler (1961) have become known as *search theory*. Early groundbreaking works on search theory, such as McCall (1970) and Mortensen (1970), have formalized the ideas of Stigler in a mathematically tractable framework that laid the foundation for later applications of search theory. In particular, the study of the labor market has been strongly influenced by the idea that the matching of firms and workers is subject to search frictions (e.g., Diamond, 1982; Mortensen and Pissarides, 1994). The success of these models in shaping our understanding of unemployment as well as the impact of search frictions on the economic discipline as a whole has been reflected in the 2010 Nobel price being awarded to Peter A. Diamond, Dale T. Mortenson, and Christopher A. Pissarides "for their analysis of markets with search frictions".

Agency Conflicts Second, problems of asymmetric information as in the second of the aforementioned examples have been studied in the literature on *agency conflicts*. Early contributions to this literature have introduced frameworks to think about asymmetric information in games (e.g., Harsanyi, 1967, 1968a,b) and in market settings (e.g., Akerlof, 1970).

Later developments have studied the question of how either an outside observer interested in efficiency or an agent interested in maximizing his payoffs should design the "rules of the game" in order to optimally deal with informational asymmetries. Important directions of this field that has become known as *contract theory* are the study of moral hazard problems in static settings (e.g., Mirrlees, 1976; Hölmstrom, 1979; Rogerson, 1985; Grossman and Hart, 1983) as well as in dynamic settings (e.g., Hölmstrom and Milgrom, 1987; Spear and Srivastava, 1987; Phelan and Townsend, 1991; Sannikov, 2008); the study of incomplete contracts (e.g., Grossman and Hart, 1986; Hart and Moore, 1990; Nöldeke and Schmidt, 1995); the study of relational contracts (e.g., MacLeod and Malcomson, 1989; Baker et al., 1993, 2002; Levin, 2003); and the study of adverse selection (e.g., Mirrlees, 1971; Mussa and Rosen, 1978; Baron and Myerson, 1982) as well as mechanism design problems in more general (e.g. Myerson and Satterthwaite, 1983; Moore and Repullo, 1988).

The methods developed in these papers have been employed in a great many of applications. Some of the most impacted areas are the theory of the organization of firms; the analysis of credit lending; the design of optimal auction mechanisms; and most recently, the introduction of financial frictions into macroeconomics. Reflecting the impact of agency conflicts on the discipline of economics, the 2001 Nobel price has been awarded to George A. Akerlof, Michael Spence, Joseph E. Stiglitz "for their analyzes of markets with asymmetric information"; and the 2007 Nobel price has been awarded to Leonid Hurwicz, Eric S. Maskin, and Roger B. Myerson "for having laid the foundations of mechanism design theory".

Dispersed information Lastly, the third strand of the informational economics literature studies how in complex environments, where *information is dispersed* across a large number of agents, these agents aggregate information from available sources and how this affects equilibrium dynamics.

Early works within this strand have focused on showing that Fama's (1970) efficient market hypothesis does not hold for dispersed information settings (Hellwig, 1980; Grossman and Stiglitz, 1980). Later works on *global games*, in particular Carlsson and Van Damme (1993) and Morris and Shin (1998), have shown how dispersion information in combination with strategic complementarity may pin down a unique equilibrium in games that are characterized by multiple equilibria under full information. Recently, the interest in this strand of the literature has been newly sparked by studies, such as Woodford (2003), Lorenzoni (2009), and Angeletos and La'O (2012), that have shown how learning from dispersed information can explain many macroeconomic dynamics and comovements observed in the data and can also give a theoretical underpinning to the popular view that business cycles are driven by expectations and sentiments.

Overview

This dissertation consists of four chapters. The chapters are based on self-contained research papers that contribute to the three areas of informational economics outlined above.

Chapter 1 The first chapter of my dissertation explores the idea that increases in uncertainty among agents may be a reason why the recent financial crisis has been so persistent. To this end, the chapter develops a dynamic macroeconomic model with the unique feature that it combines two types of informational frictions: (1) Informational asymmetries that give rise to a financial friction, constraining firms' production choices during financial crises. (2) Agents in the model do not observe business conditions, but have to endogenously infer about them by extracting information from market prices in a dispersed information setting.

The model explains why large financial crises have a disproportionately severe and long-lasting impact on the real economy. At the same time, the model also explains why high-frequency fluctuations on the financial market during "normal times" have only a little impact on the economy. That is, the model gives rise to an asymmetry between small and large shocks; or put differently, the model explains why the transmission of financial shocks is inherently nonlinear.

Underlying these findings is a novel mechanism that makes uncertainty among agents endogenous. The mechanism is that during financial crises the production choices of financially constrained firms are not governed anymore by business conditions, but are dictated by financial constraints. Accordingly, production choices of firms stop reflecting economic fundamentals, so that less can be learned by observing the economy during financial crises. This means that uncertainty among agents is endogenously increasing during financial crises.

As a corollary to these results, the model also make a number of predictions that are consistent with empirical data and that lend support to the theoretical mechanism discussed above: First, risk premia are high during financial crises; a natural consequence of increasing uncertainty. Second, asset price movements become endogenously amplified during financial crises, explaining the high volatility of asset prices in crisis times. Third, the cross-sectional diversity of opinions is increasing during financial crises.

Chapter 2 The second chapter of my dissertation aims at filling a gap in the political economics literature. While the literature has studied various transition *mechanism*, the properties of these mechanisms, in particular the resulting distribution of political systems, has been largely unexplored. To this end, Chapter 2 develops a model of political transitions in which the types of political systems and the likelihood of transition events are determined endogenously.

Importantly, the study rests on the co-existence of reforms (changes to a political system that are initiated from within a regime) and revolts (changes to the political system that are enforced by outsiders of the regime) along the equilibrium path. To ensure this co-existence, the model is based on the assumption that potential insurgents against a political regime are imperfectly informed about the regime's ability to defend itself against upheavals. This gives rise to a signaling game which limits a regime's ability to reduce revolutionary pressure by making political concessions, ensuring the prevalence of revolts along the equilibrium path. This is because, in equilibrium, concessions would be interpreted as a sign of weakness, helping potential insurgents to coordinate their protest, and thereby undermining the efficiency of concessions in reducing revolutionary pressure.

Within this framework, Chapter 2 of my dissertation answers two questions: First, what types of regimes will endogenously emerge in equilibrium? Second, which types of regimes will be political stable?

Regarding the types of political systems that arise in equilibrium, it is shown that revolts generally result in autocracies, whereas political reforms generally enfranchise the majority of the population. This suggests that only peaceful reforms can lead to sustainable democratization, lending theoretical support to a long-lasting view in political science according to which members of former autocracies are key actors in the establishment of democracies (for references, see the introduction of Chapter 2).

Regarding the stability of regimes, it is shown that democracies are intrinsically stable, leading to long episodes without political change. In contrast, autocracies are subject to frequent regime changes. Yet, autocratic systems are persistent over time as they are frequently overthrown by small groups of insurgents, resulting in political systems similar to their predecessors.

Taken together, these results imply both a polarization of regimes into extreme types *during* transitions and a *persistence* of extreme political systems once they have emerged. Accordingly, the long-run distribution of political systems is double hump-shaped with mass concentrated on the extremes.

These findings are consistent with cross-country data on political transitions and regime types for nearly 100 years of recent history. To show this, Chapter 2 combines three recent datasets, merging information on political systems with data on political transitions.

CHAPTER 3 The third chapter of my dissertation contributes to both the literature on search theory and agency conflicts. The chapter is motivated by the observation that searching is not only an important friction by itself, but is often delegated to an agent. For instance, recruiting agencies are hired to search for job candidates. Real estate agents are contracted to search for prospective buyers or, alternatively, to search for attractive houses. And insurance brokers are often hired to find new clients.

Chapter 3 analyzes such a situation when search is delegated to an agent. The interaction between the agent who realizes the benefits from searching (i.e., the "principal") and the agent who conducts the search is governed by two kinds of informational asymmetries. First, samples are drawn from a distribution that is only known by the agent, giving rise to an adverse selection problem. Second, search itself cannot be observed by the principal. Thus, the principal's problem is to bring the agent to reveal the optimal search policy and, simultaneously, to induce him to actually search according to this policy.

Chapter 3 shows that search in this case is optimally delegated through the use of a screening menu, which is exclusively comprised of simple bonus contracts (as are widespread in many industries). Moreover, search policies are almost surely inefficient; either search is terminated prematurely, or it is completely undirected. In contrast, if either of the two informational asymmetries is resolved, the first-best outcome can be supported in equilibrium.

Chapter 4 The final chapter of my dissertation is motivated by a well-known regularity of airline pricing. Namely that prices for airline tickets rise as the scheduled departure date approaches. Previous studies that aim at explaining this fact heavily rely on the assumption that airlines can commit to state-independent pricing schemes. This assumption appears, however, to be undermined by recent developments in the airline industry. First, the usage of modern databases and computer systems made it easy for airlines to collect and process information about current demand conditions and to project them into the future. Second, the emergence of online booking platforms have led today's airlines to make use of these information by employing sophisticated dynamic pricing schemes that condition on all available information.

Chapter 4 aims at filling the theoretical gap arising from these developments. To this end, it develops a novel theory of dynamic pricing in industries that are characterized by short-term capacity constraints. The key mechanism is that when supply is fixed in the short term, firms price more aggressively in earlier periods in order to relax competition in the future. Accordingly, prices are increasing over time precisely because firms do employ state-contingent pricing schemes. Applied to the airline industry, this explains why ticket prices rise close to the scheduled arrival date, even if airlines are unable to commit to future prices. Importantly, while the costs of pricing aggressively is born by individual airlines, the benefits of altering the market structure are enjoyed by all competitors. Accordingly, dynamic airline pricing constitutes a public goods problem from the perspective of individual airlines, implying that prices are intertemporally less dispersed on more competitive routes.

The second contribution of Chapter 4 is to introduce a novel, hand-collected dataset of 1.4 million airline ticket prices on 92 intra-European routes. Using this dataset, the theoretical predictions are tested successfully.

Organization

To retain the self-contained character of the individual chapters of this dissertation, each chapter has its own appendix and list of references. All formal proofs in this dissertation are deferred to these appendices. I use the usual Halmos symbol $\Box$ to mark the completion of a proof; the conclusion of important sub-steps or lemmas that are a part of larger proofs are marked by a variation thereof $\diamond$. An in-depth overview over the contents of each chapter is provided at their respective beginnings.

Acknowledgments

I am extremely grateful to my advisers Matthias Doepke and Klaus M. Schmidt, and to Mirko Wiederholt for their invaluable guidance and support. I also would like to thank my coauthors Lukas Buchheim, Caspar Siegert, and Ludwig Straub for the fruitful collaboration without which this dissertation would not have been possible. Over the course of my dissertation I have profited from many discussions with too many people to name them all. Some persons who had a particular strong impact on my work are listed at the beginnings of the individual chapters. Finally, I owe the greatest debt to my parents Andrea, Andreas, Steffi, and Michael, and to my sisters Anne, Laura, and Rhea for all their support.

ROBERT ULBRICHT
Evanston, Illinois
December 2012

References

Akerlof, G. (1970), "The market for "lemons": Quality uncertainty and the market mechanism," The quarterly journal of economics, 488–500.
Angeletos, G.-M. and J. La'O (2012), "Sentiments," mimeo.
Baker, G., R. Gibbons, and K. Murphy (1993), "Subjective performance measures in optimal incentive contracts," Tech. rep., National Bureau of Economic Research.
——— (2002), "Relational Contracts and the Theory of the Firm," The Quarterly Journal of Economics, 117:39–84.
Baron, D. and R. Myerson (1982), "Regulating a monopolist with unknown costs," Econometrica: Journal of the Econometric Society, 911–930.
Carlsson, H. and E. Van Damme (1993), "Global games and equilibrium selection," Econometrica: Journal of the Econometric Society, 989–1018.
Diamond, P. (1982), "Aggregate demand management in search equilibrium," The Journal of Political Economy, 881–894.
Fama, E. (1970), "Efficient capital markets: A review of theory and empirical work," The journal of Finance, 25:383–417.
Grossman, S. and O. Hart (1983), "An analysis of the principal-agent problem," Econometrica: Journal of the Econometric Society, 7–45.
——— (1986), "The costs and benefits of ownership: A theory of vertical and lateral integration," The Journal of Political Economy, 691–719.

Grossman, S. and J. Stiglitz (1980), "On the impossibility of informationally efficient markets," The American Economic Review, 393–408.
Harsanyi, J. (1967), "Games with Incomplete Information Played by "Bayesian" Players, Part I. The Basic Model," Management science, 14:159âĂŞ–182.
——— (1968a), "Games with Incomplete Information Played by "Bayesian" Players, Part II. Bayesian Equilibrium Points," Management Science, 14:320–334.
——— (1968b), "Games with Incomplete Information Played by "Bayesian" Players, Part III. The Basic Probability Distribution of the Game," Management Science, 14:486–502.
Hart, O. and J. Moore (1990), "Property Rights and the Nature of the Firm," Journal of political economy, 1119–1158.
Hellwig, M. F. (1980), "On the aggregation of information in competitive markets," Journal of Economic Theory, 22:477–498.
Hölmstrom, B. (1979), "Moral hazard and observability," The Bell Journal of Economics, 74–91.
Hölmstrom, B. and P. Milgrom (1987), "Aggregation and linearity in the provision of intertemporal incentives," Econometrica: Journal of the Econometric Society, 303–328.
Levin, J. (2003), "Relational incentive contracts," The American Economic Review, 93:835–857.
Lorenzoni, G. (2009), "A Theory of Demand Shocks," The American Economic Review, 99:2050–2084.
MacLeod, W. and J. Malcomson (1989), "Implicit contracts, incentive compatibility, and involuntary unemployment," Econometrica: Journal of the Econometric Society, 447–480.
McCall, J. (1970), "Economics of information and job search," The Quarterly Journal of Economics, 113–126.
Mirrlees, J. (1971), "An exploration in the theory of optimum income taxation," The review of economic studies, 175–208.
——— (1976), "The optimal structure of incentives and authority within an organization," The Bell Journal of Economics, 105–131.
Moore, J. and R. Repullo (1988), "Subgame perfect implementation," Econometrica: Journal of the Econometric Society, 1191–1220.
Morris, S. and H. Shin (1998), "Unique equilibrium in a model of self-fulfilling currency attacks," American Economic Review, 587–597.
Mortensen, D. (1970), "Job search, the duration of unemployment, and the Phillips curve," The American Economic Review, 60:847–862.
Mortensen, D. and C. Pissarides (1994), "Job creation and job destruction in the

theory of unemployment," The review of economic studies, 61:397–415.

Mussa, M. and S. Rosen (1978), "Monopoly and product quality," Journal of Economic theory, 18:301–317.

Myerson, R. and M. Satterthwaite (1983), "Efficient mechanisms for bilateral trading," Journal of economic theory, 29:265–281.

Nöldeke, G. and K. Schmidt (1995), "Option contracts and renegotiation: a solution to the hold-up problem," The RAND Journal of Economics, 163–179.

Phelan, C. and R. Townsend (1991), "Computing multi-period, information-constrained optima," The Review of Economic Studies, 58:853–881.

Rogerson, W. (1985), "The first-order approach to principal-agent problems," Econometrica: Journal of the Econometric Society, 1357–1367.

Sannikov, Y. (2008), "A continuous-time version of the principal-agent problem," The Review of Economic Studies, 75:957–984.

Spear, S. and S. Srivastava (1987), "On repeated moral hazard with discounting," The Review of Economic Studies, 54:599–617.

Stigler, G. (1961), "The economics of information," The journal of political economy, 213–225.

Woodford, M. (2003), "Imperfect common knowledge and the effects of monetary policy," in Knowledge, Information and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps, ed. by P. Aghion, R. Frydman, J. Stiglitz, and M. Woodford, Princeton: Princeton University Press, 25–58.

Chapter 1

CREDIT CRUNCHES, INFORMATION FAILURES, AND THE PERSISTENCE OF PESSIMISM*

Contents

* This chapter is based on joint work with Ludwig Straub. We would like to thank Matthias Doepke, Klaus M. Schmidt, and Mirko Wiederholt for extensive feedback and guidance throughout this project. We are also grateful to Klaus Adam, George-Marios Angeletos, Marco Bassetto, Markus K. Brunnermeier, Larry Christiano, Martin Eichenbaum, Christian Hellwig, Guido Lorenzoni, Alessandro Pavan, Giorgio Primiceri, Aleh Tsyvinski, and seminar participants at Munich and Northwestern Universities for many helpful comments and discussions. This research started while I was visiting Northwestern University; I thank the economics department for their hospitality.

Abstract

This chapter of my dissertation examines how financial crises affect the ability of agents to learn about economic fundamentals, and how this in turn affects the transmission of financial shocks through the economy. To this end, we introduce a model where noise in the financial market drives business cycles. Agents endogenously learn about fundamentals from market prices, but financial constraints systematically destroy the informational capacity of prices in financial

crises. This is because financially constrained agents stop responding to available information, reducing the efficiency of prices in aggregating information that is dispersed across the economy. As a result, times of financial crisis are marked by both endogenously increasing uncertainty and increasingly persistent pessimism, providing a powerful amplification mechanism for financial shocks. Importantly, this mechanism is inherently nonlinear. Whereas small or positive financial shocks have only a little influence on the economy, unusually adverse shocks virtually shut down market learning and result in disproportionately severe and persistent crashes—characterized by substantial losses in employment, output, and asset prices; and high levels of uncertainty, volatility, and risk premia.

Keywords

Credit crises, endogenous uncertainty, financial frictions, heterogeneous information, asymmetric and nonlinear business cycles.

JEL CODES: D83, E32, E44, G01.

1. Introduction

Observers of the recent financial crisis often emphasize the role of uncertainty for the transmission and amplification of financial shocks. In particular, a widespread idea is that increasingly uncertain business conditions are a key factor for the persistence of the crisis. For instance, IMF chief economist Olivier Blanchard argues that: "(Financial) crises feed uncertainty. And uncertainty affects behavior, which feeds the crisis. Were a magic wand to remove uncertainty … the crisis would largely go away."[1]

Understanding these ideas requires thinking about uncertainty as being endogenous to the state of the economy. This chapter of my dissertation examines how financial distress affects the ability of agents to learn about economic fundamentals, and how this in turn affects the transmission of financial shocks through the economy. Specifically, we study a dynamic macroeconomic model where agents learn about economic fundamentals from market prices. The presence of financial frictions endogenously determines the efficiency of the pricing mechanism in aggregating available information and thereby governs uncertainty among agents.

The model highlights a novel mechanism that explains the characteristic persistence of financial crises. In contrast to other types of crises, *financially* constrained firms cannot step up their investments when they become more confident in their business outlooks. As a result, in states of financial distress, real business activity does *not* reflect actual business conditions, leaving market observers uncertain about the state of the economy. With high uncertainty feeding back into the financial market, this perpetuates financial distress and creates a persistent cycle of uncertainty and financial constraints. As this chapter shows, this feedback loop has important implications for the behavior of financial markets and the behavior of the production sector in response to financial shocks.

Preview of the model The analysis is based on a stylized two-sector economy. In a production sector, entrepreneurs produce a single consumption good, using labor provided by workers as the only input factor; and in a financial sector, both workers and entrepreneurs trade an asset whose returns correlate with the economy's average productivity (the exogenous fundamental of the economy).

The model is built on two assumptions. First, fluctuations in the financial market affect the real sector via financial frictions. Specifically, we assume that

[1] This quote is taken from a guest article written by Olivier Blanchard for The Economist, January 31, 2009.

entrepreneurs must borrow in order to pay their workers and that, as in Kiyotaki and Moore (1997), all debt must be secured with collateral. Having entrepreneurs use the financial asset as collateral, this creates a cap on hiring which tightens as asset prices tumble. Second, we assume that agents cannot observe the economy's average productivity directly. Instead they endogenously learn about productivity by observing the market-clearing prices in both sectors, which aggregate further exogenously available information that is dispersed across the economy.

Results The results in this chapter of my dissertation are driven by the interaction of informational and financial frictions. This combination leads adverse financial shocks to systematically destroy the real economy's "informational capacities" (i.e., less can be learned from observing the production sector). The reason is that when entrepreneurs become financially constrained, they cease to respond to available information. Wages—or, any other production-based source of information[2]—therefore become less efficient in aggregating information that is dispersed across the production sector. This reduces overall learning and increases uncertainty among agents in times of financial distress.

The endogenous nature of the real economy's informational capacities gives rise to two *information*-based mechanisms that amplify adverse financial shocks.[3] To see this, note that when less can be learned from today's production sector, agents place more weight on other sources of information. In particular, agents' opinions are more affected by information contained in asset prices as they tumble. This amplifies pessimism among agents, tightens financial constraints even more, and creates a harmful feedback loop. At the same time, with little to learn from today's economy, agents are also affected more by prior information in forming their opinions. This causes pessimism in the financial market to become inherently persistent, thereby increases future financial distress, and in turn inhibits future learning. Metaphorically speaking, the financially constrained economy gets stuck in a "pessimism trap".

Importantly, both of these mechanisms are inextricably tied to the endogenous

[2] In our model, the only statistic that is (directly) affected by entrepreneurs' production choices are wages. In a richer model, however, the efficiency in aggregating information would not only be reduced for wages, but for all observable prices (or quantities) that vary with the economy's production.

[3] To sharpen our results, we abstract from any risk-related mechanisms or any other mechanisms that directly translate increases in uncertainty into real effects. However, including any such mechanism only amplifies our findings, which we illustrate in Section 7.3 where we extend the model to study the effects of risk-aversion.

nature of uncertainty—were one to remove uncertainty during a financial crisis (using the "magic wand" imagined by Olivier Blanchard), the crisis would indeed largely go away. Moreover, while the bulk of the literature on financial frictions discusses amplification mechanisms that are symmetric,[4] the *information*-based amplification mechanisms just described are inherently asymmetric. That is, whereas adverse financial shocks inhibit learning and have an amplified and persistent impact on the economy, positive financial shocks improve learning and have a *de*-amplified and *non*-persistent impact.

More generally, we find that the more negative a financial shock, the greater are both the amplification and the persistence of the shock. In particular, while the amplification and persistence of small shocks is negligible, rare adverse ("tail") shocks virtually destroy the real economy's informational capacities and entail highly amplified and persistent crashes. This non-linearity (or convexity) provides a novel theoretical explanation for why during normal times the day-to-day fluctuations in the financial market appear to have only a small impact on the real economy, whereas unusually adverse financial shocks propagate persistently throughout the whole economy.

Taken together, the aforementioned results imply that financial crises are characterized by (i) amplified and (ii) persistent losses in output, employment, and asset prices; (iii) as well as high uncertainty. In addition, the collapse in the informational capacities explains three further key characteristics of financial crises, namely: (iv) highly diverse views on the state and fate of the economy; (v) volatile asset prices; and (vi) large risk premia. Highly diverse views result from agents placing more weight on private sources of information—which are inherently diverse—when less can be learned from today's production sector. At the same time, as discussed above, agents also pay more attention to asset prices. As in any rational expectations equilibrium, this implies that asset prices become more exposed to noisy demand shocks within the financial market, increasing volatility. Lastly, when agents are risk-averse, risk premia on asset prices naturally rise as uncertainty increases.

[4] This is, for instance, true for the seminal contributions by Kiyotaki and Moore (1997) and Bernanke et al. (1999), and most of the literature thereafter. An important exception are Brunnermeier and Sannikov (2012), who consider an economy with a financial sector, in which shocks have similar nonlinear effects. See the literature review for details.

Methodological contribution The information loss in our model is inextricably tied to endogenous learning governed by *nonlinear* laws of motion.[5] That is, agents learn about the fundamental not only through linear signals but also through nonlinear ones. More specifically, we show that, in virtue of financial frictions, real sector prices (in the presence of noisy demand) are informationally equivalent to observing a perturbed, *concave* function of the fundamental. The slope of the function is decreasing in the "constrainedness" of the economy. In a general theorem, we then prove that "well-behaved" concave signals generally result in higher uncertainty when the signal realizes in flatter regions. This theorem applies to a large class of information structures and holds independent of the specifics of our model.

One technical challenge in analyzing the dynamic properties of our model is that nonlinear Gaussian signal structures generally do not pair with conjugate prior distributions. To address this problem, we develop a *quasi*-Gaussian framework, departing from the assumption that the small additive noise terms included in the nonlinear signals are normally distributed. In particular, we construct the noise terms in such a way that the nonlinear signals behave *as if* they were linear normal signals with a *state-dependent* signal precision. Within this framework, our general theorem then maps every state of the economy to a unique signal precision, which is decreasing in the economy's constrainedness. Because, in the limit of signals becoming linear, the state-dependent signal precision becomes state-*in*dependent, quasi-Gaussian signals can be understood as a natural extension of the standard linear Gaussian framework to the case of nonlinear signals.

Related literature At a methodological level, the two building blocks of our model relate this study to two strands of modern macroeconomics. First, there is a large literature on financial frictions that demonstrates how small shocks can get amplified through the financial system. In particular, our formalization of credit constraints is based on Kiyotaki and Moore (1997).[6] We contribute to

[5] Mertens (2011) and Hassan and Mertens (2011) also analyze a model with heterogeneous information and nonlinear laws of motion. However, in their paper, the information structure is such that nonlinearities are transformed away and agents actually update according to linear prices, avoiding endogenous uncertainty.

[6] More recent studies based on credit constraints include, e.g., Krishnamurthy (2003); Iacoviello (2005); Kiyotaki and Moore (2008); Gertler and Karadi (2011); and Gertler and Kiyotaki (2011). Important contributions that are based on other financial frictions include, e.g., Bernanke and Gertler (1989); Carlstrom and Fuerst (1997); and Bernanke et al. (1996,

this literature by identifying a new *informational* role of financial frictions that complements the *constraining* role known from previous works. Throughout this chapter, we highlight the consequences of this new informational mechanism by comparing our model to the counterfactual case where the constraining role of credit constraints remains intact, but all informational effects of frictions are shut down. As with the majority of the financial frictions literature, we find that in this counterfactual case, shocks *symmetrically* and *linearly* affect the economy. That is, both adverse and positive shocks are amplified through financial frictions in exactly the same way. In contrast, our findings that amplification is asymmetric and nonlinear is a novel feature of the *information*-based mechanism introduced in this chapter.[7]

Second, this study closely connects to an emerging literature on heterogeneous information in macroeconomics and finance (see, e.g., Morris and Shin, 2002 and Woodford, 2003).[8] From a methodological perspective, we contribute to this literature by showing how learning from nonlinear signals gives rise to endogenous uncertainty and by embedding nonlinear signals in a conjugate prior framework. From an applied perspective, our study is similar to Angeletos et al. (2010) who also study how learning from the real sector affects the financial market, but do not consider how in the presence of financial constraints the financial market feeds back to the information aggregation and how that causes learning to collapse during financial crises. Perhaps closest to our model is a framework by La'O (2010), which also combines informational with financial frictions. However, because La'O resolves all dispersion of information at the time agents learn from financially constrained markets, posterior uncertainty in her model is completely determined by the exogenous amount of information available to the economy, ruling out the informational mechanism that drives

---

1999), which are based on Townsend's (1979) costly state verification approach; and Kurlat (2010); Bigio (2011, 2012); and Boissay et al. (2012), which consider frictions originating in adverse selection. See Appendix B.2 for a discussion how the ideas developed in this chapter can be applied to the costly-state-verification and adverse selection approaches.

7 An important exception are Brunnermeier and Sannikov (2012), who consider an economy with a financial sector, in which shocks have similar nonlinear effects. While the findings of Brunnermeier and Sannikov are similar in spirit, their mechanism is, however, not. In particular, information in their model is perfect and uncertainty is constant over time.

8 More recent contributions to the dispersed information literature with a macroeconmic focus include Adam (2007); Angeletos and Pavan (2004, 2007, 2009); Amato and Shin (2006); Morris and Shin (2006); Amador and Weill (2008); Lorenzoni (2009, 2010); Hellwig and Veldkamp (2009); Hassan and Mertens (2011); Goldstein et al. (2011); and Angeletos and La'O (2012a,b).

our results.[9]

At a more applied level, this chapter also relates to a recent literature following Bloom (2009) that puts forth the idea of uncertainty-driven business cycles resulting from *exogenous* uncertainty shocks.[10] Our approach relates to these works in two ways. First, we provide a microfoundation for why uncertainty increases specifically during times of financial distress, which is also when empirical measures of uncertainty are highest. An important insight from our microfoundation is that the endogenous nature of uncertainty unleashes a powerful feedback loop, which is absent in business cycles that are driven by exogenous uncertainty shocks. Therefore, in contrast to Bloom, who finds that uncertainty shocks give rise to rapid drops and rebounds in economic activity, we find that high uncertainty goes along with amplified and persistent crises. Second, the literature on exogenous uncertainty shocks complements our findings in that it discusses a number of additional channels, absent in our model, by which increases in uncertainty may propagate through the economy. In particular, Christiano et al. (2009) and Gilchrist et al. (2010) illustrate how fluctuations in uncertainty are amplified through a combination of risk-aversion and financial frictions and have strong effects on the real sector.[11]

Finally, our finding that financial frictions destroy information relates to a small and closely related literature that studies endogenous fluctuations in uncertainty. Van Nieuwerburgh and Veldkamp (2006) explore the idea that learning about total factor productivity is slow in recessions when total business activity is low. The reason is that if output is perturbed by an additive noise term, then this noise term contributes relatively more to output when output is low, leading to higher uncertainty during recessions (see also, Veldkamp, 2005 and Ordoñez, 2010). The mechanism studied in this chapter differs from the mechanisms in these papers in that the efficiency of learning is governed by the degree to which the economy is constrained rather than the level of output. As outlined above, this difference leads to a number of important implications for the transmission of financial shocks. Apart from this, our study also differs in that it considers

[9] We also differ from La'O (2010) in that we focus on business cycle dynamics, whereas La'O studies the (static) composition of output and price volatility in fundamental and noise shocks in a single period model.

[10] See also Sim (2008); Bachmann and Bayer (2009); and Bloom et al. (2012).

[11] While we abstract from risk-related transmission mechanism of uncertainty in our model to sharpen our results, these mechanisms are clearly important and strongly amplify the role of the key mechanism identified in this study. For a illustration, see Section 7.3 where we consider an extension of our model to the case where agents are risk averse.

learning from price signals and provides an explicit foundation for why prices vary in their informational content in times of financial distress. This approach is shared with two related contributions by Yuan (2005) and Albagli (2011). However, both papers focus on one-shot financial market settings and do not include a real sector, preventing them from analyzing the transmission of financial shocks through the economy, which is at the core of our contribution.[12]

OUTLINE The plan for the rest of this chapter is as follows. The next section introduces the model economy. Section 3 examines how financial frictions affect the ability of agents to learn from market prices. Section 4 characterizes the full equilibrium. Section 5 then explores how shocks are transmitted through the economy. Section 6 illustrates our theoretical results with a numerical example. Section 7 points out some further empirical predictions, and Section 8 concludes. All proofs are deferred to the appendix.

2. The model

Our model is based on two ingredients: financial frictions and endogenous learning. In the interest of analytical tractability, we make a number of simplifying assumptions. In particular, we focus on labor as a single input good, so that learning from the real sector takes the form of extracting information from the market clearing wage. Nonetheless, our analysis can be applied to any other price that varies with entrepreneurs' optimal production choices, and is meant to more generally capture the idea that agents learn about business conditions by observing the real sector. Also, we focus on a single, stylized financial friction to model spillovers from the financial market. However, while the model heavily rests on the constraining role of asset prices, it is irrelevant by which financial friction this is explained (see Appendix B.2).

[12] Another paper sharing the broad theme is Bachmann and Moscarini (2011), which looks at a mechanism that increases the cross-sectional dispersion of beliefs during crises, but in which posterior uncertainty remains constant. Also, there is a growing literature on rational inattention, which is based on the idea that learning is costly, effectively leading agents to endogenously pick their desired level of uncertainty (see, e.g., Sims, 2003, 2006; Maćkowiak and Wiederholt, 2009, 2010; and Woodford, 2009).

Preferences and technologies Consider a discrete time, infinite horizon economy with a continuum 1 of risk-neutral, one-period lived agents.[13] A proportion m of each generation's agents are farmers, while the remaining $1 - m$ are gatherers. With a slight abuse of notation, we use $\mathcal{F}$ and $\mathcal{G}$ to denote the set of farmers and gatherers at a given date, respectively. Both farmers and gatherers consume a single consumption good, a perishable fruit, which is produced using two distinct technologies. First, there is an entirely exogenous production unit. In reference to Lucas (1978) it is helpful to think about this unit as a "tree" (or asset), which bears a random number $\tilde{A}_t$ of fruits and comes in a total supply of 1, equally distributed across each generation. Second, farmers have access to a "field" which transforms labor input n_{it} into additional fruits. The production function for field work is given by

$$F(\tilde{A}_{it}, n_{it}) = \tilde{A}_{it} \log(n_{it}),$$

where $\tilde{A}_{it}$ is an idiosyncratic random productivity parameter of farmer $i \in \mathcal{F}$ at date t.

For simplicity, farmers are excluded from doing fieldwork themselves, but may employ gatherers for the purpose of cultivating their fields. Gatherer i's disutility of working is given by $v : \mathbb{R}_+ \to \mathbb{R}_+$, a twice differentiable, increasing, and strictly convex function, with $v'(0) = 0$ and $\lim_{n\to\infty} v'(n) = \infty$. Gatherer i thus wishes to maximize the quantity

$$\mathrm{E}\left\{\tilde{c}_{it} - v(n_{it}) | \mathcal{I}_{it}\right\}, \tag{1}$$

and farmer i wishes to maximize

$$\mathrm{E}\left\{\tilde{c}_{it} | \mathcal{I}_{it}\right\}, \tag{2}$$

where $\tilde{c}_{it}$ represents consumption of fruits, and $\mathrm{E}\left\{\cdot | \mathcal{I}_{it}\right\}$ is an expectations operator given information set $\mathcal{I}_{it}$.[14]

Field productivities $\{\tilde{A}_{it} : i \in \mathcal{F}\}$ are taken to be lognormally distributed, so that $\log(\tilde{A}_{it}) \equiv \tilde{\theta}_{it}$ has a normal distribution with mean $\tilde{\theta}_t$ and variance $1/\tau_\xi$,

[13] Agents in our model are one-period lived to induce a common prior among agents at all times. This ensures that heterogeneously informed agents in our model do not run into Townsend's (1983) infinite regress problem, allowing us to derive all our results in an analytic fashion.

[14] Throughout, we differentiate between stochastic variables and their realizations by accentuating the stochastic version with a tilde ("~").

and where the average log productivity $\tilde{\theta}_t$ follows a first-order autoregressive process:

$$\tilde{\theta}_t = \rho\tilde{\theta}_{t-1} + \tilde{\epsilon}_t,$$

where $\tilde{\epsilon}_t$ is Gaussian noise with variance $1/\tau_\epsilon$. The dividend from fruit trees is assumed to be positively correlated with the average productivity and is given by

$$\log(\tilde{A}_t) = \gamma_0 + \gamma_1\tilde{\theta}_t + \tilde{u}_t,$$

where $(\gamma_0, \gamma_1) \in \mathbb{R} \times \mathbb{R}_+$ and $\tilde{u}_t$ is an independent (of $\tilde{\theta}_t$) random variable that possibly introduces additional noise to dividend payments.

Markets and credit constraints There are two types of markets operating at date t. First, a competitive labor market matches demand and supply for field work and determines the market clearing wage w_t. Second, a competitive stock market determines ownership of fruit trees and pins down an asset price q_t. Shares on trees are assumed to be perfectly divisible and entitle its owners to claim all $\tilde{A}_t$ fruits falling from the corresponding tree. In both markets, current period consumption serves as the unit of account. Furthermore, we simplify the analysis by ruling out margin trading and short selling of trees, effectively restricting asset holdings of agent i to $0 \leq x_{it} \leq 1$.[15]

We now describe the financial friction in our economy. Following Kiyotaki and Moore (1997), we assume that farmers lack the means to commit to paying their wage debt after production is sunk. As a consequence, gatherers refuse to do field work unless they are provided a security by farmers in exchange for their labor.[16] We assume that fruits harvested from fields are nontradeable, so

[15] This specification is adopted from Albagli et al. (2011) and in combination with our assumptions on noisy asset demand (see below), it keeps the law of motion of asset prices tractable within a conjugate prior framework. Further note, that these no-borrowing constraints are consistent with the lack of commitment power that we impose as a key friction on the labor market.

[16] Our assumption that farmers cannot commit to paying their wage bill is based on theoretical arguments developed by Hart and Moore (1994, 1998). In their 1994 paper, such commitment problem arises from the possibility to renegotiate wages at any point during the production process. Accordingly, if farmers are *indispensable* for reaping the benefits of field work (e.g., because fruits harvested on fields are nontransferable, as we assume in our setting), then the outside option of gatherers is reduced to the value of collateral, and farmers could renegotiate a smaller wage whenever the wage bill exceeds the value of collateral. Alternatively, if, as in their 1998 paper, farmers can "run away" after production is harvested, gatherers are likewise left with only the value of collateralized assets. In both cases, in anticipation of the

that fields itself cannot be used as an security.[17] Instead, farmers may use trees as collateral to pay gatherers. Also, we simplify the problem of how to account the value of collateral by assuming that the asset market operates at least twice: A first time parallel to the labor market, ensuring that all information that is possibly aggregated by trading trees is already available when the value of collateral is determined; and a second time after production is realized and wages are paid, so that the value of collateral is based on the current market price q_t for trees. These assumptions jointly imply that the wage debt of a farmer i at date t is bounded from above by the market value of his asset holdings $x_{it}q_t$, constraining labor demand to satisfy

$$n_{it} \leq (q_t/w_t)\, x_{it}. \tag{3}$$

Information The average productivity takes the role of the "fundamental" in our economy. More generally, $\tilde{\theta}_t$ is meant to reflect the "profitability" of investments, comprising, e.g., technology shocks and aggregated business conditions. Agents base their expectations about $\tilde{\theta}_t$ on the information

$$\mathcal{I}_{it} = \{s_{it}\} \cup \{w_s, q_s\}_{s=1}^{t},$$

which, in addition to the publicly observable history of prices $\{w_s, q_s\}_{s=1}^{t}$, contains a private signal $\tilde{s}_{it}$, which reveals the true average productivity $\tilde{\theta}_t$ perturbed by some independent Gaussian noise $\tilde{\xi}_{it}$ with variance $1/\tau_\xi$:

$$\tilde{s}_{it} = \tilde{\theta}_t + \tilde{\xi}_{it}.$$

For simplicity, $\{\tilde{s}_{it}\}_{i\in\mathcal{F}}$ is assumed to be perfectly correlated with farmers' productivities $\{\tilde{\theta}_{it}\}_{i\in\mathcal{F}}$, so that by learning the realization of $\tilde{s}_{it}$ a farmer also learns the productivity of his field.

Furthermore, to prevent $\tilde{\theta}_t$ from being perfectly revealed by the market, prices w_t and q_t are perturbed by noise traders with stochastic asset demand $\Phi(\sqrt{\tau_\xi}\,(\tilde{\eta}_t - \mu))$ and labor demand $\Psi_{\theta_t,q_t}(\tilde{\omega}_t)$. Here $\tilde{\eta}_t$ and $\tilde{\omega}_t$ are independent Gaussian noise with variances $1/\tau_\eta$ and $1/\tau_\omega$, Φ is the cumulative standard normal distribution, μ is a constant which we conveniently set to offset the risk related

---

moral hazard, gatherers will not accept an outstanding wage debt that exceeds the value of collateral.

[17] This assumption is again based on Hart and Moore's (1994) rationale for why farmers cannot commit to paying their wages in the first place: If farmers are indispensable for operating field production, then fields are naturally worthless to gatherers (see also, Footnote 16).

components contained in $\tilde{q}_t$,[18] and $\Psi_{\theta_t,q_t} : \mathbb{R} \to \mathbb{R}$ is a function that transforms $\tilde{\omega}_t$ into a random variable $\tilde{\Psi}_t$ which may depend on the realizations of $\tilde{\theta}_t$ and $\tilde{q}_t$. Noisy asset demand $\Phi\left(\sqrt{\tau_\xi}\,(\tilde{\eta}_t - \mu)\right)$ is divided between the two occurrences of the asset market in a fixed ratio of $1 - m$ to m.

Note that to illustrate the precise conditions for which our main theorem holds, we deliberately keep Ψ_{θ_t,q_t} as general as possible for now, restricted only by the assumptions below. In addition, this generality, later grants us the freedom that is necessary to generalize the standard conjugate Gaussian framework and to extend it to the case of nonlinear learning.

Distributional assumptions As it will be seen, one key feature of this model is that the amount of information that is aggregated from the real sector through the labor market varies with the state of the economy. To define the conditions for which our characterization of learning holds, it is convenient to first introduce normalized versions of labor demand and supply, χ_s^d and χ_t^s, which describe them relative to the upper bound on labor as given by equation (3),[19]

$$\chi_t^d = \log\left(\frac{m\int_{\mathcal{F}} n_{it}\,\mathrm{d}i}{q_t/w_t}\right) \quad \text{and} \quad \chi_t^s = \log\left(\frac{(1-m)\int_{\mathcal{G}} n_{it}\,\mathrm{d}i}{q_t/w_t}\right);$$

and define

$$\chi_t^m = \log\left(\frac{\exp(\theta_t)/w_t}{q_t/w_t}\right) = \theta_t - \log(q_t),$$

which corresponds to the unconstrained relative demand of the median-productivity farmer. Intuitively, χ_t^d measures the fraction of farmers operating at their collateral constraint and provides a useful proxy for the constrainedness

[18] Although agents are risk-neutral in the model economy, the lognormal distribution of dividends implies that the asset price $\tilde{q}_t$ behaves as if agents were risk-seeking with respect to the fundamental $\tilde{\theta}_t$. By setting μ to $\gamma_1/2$ times the cross-sectional information dispersion (i.e., $\mu = \gamma_1/(2\tau_\xi)$), the risk-discount is exactly offset by the bias in noise traders' demand, yielding an asset price that behaves as if agents were risk-neutral with respect to $\tilde{\theta}_t$ and noise traders were unbiased. Also note that by transforming $\tilde{\eta}_t$ into $\Phi\left(\sqrt{\tau_\xi}\,(\tilde{\eta}_t - \mu)\right) \in [0, 1]$, noise traders' demand matches the support of endogenous asset demand, ensuring the existence of a market clearing price. The unbiased version of this specification (with $\mu = 0$) is adopted from Albagli et al. (2011) and keeps the law of motion of asset prices tractable within a conjugate prior framework (see also, Footnote 15).

[19] Here we anticipate that in equilibrium all farmers will hold $x_{it} = 1$ assets at the time the labor market operates, so that the collateral constraint is given by $n_{it} \leq q_t/w_t$.

of the economy. Note that, even after observing w_t and q_t, $\tilde{\chi}_t^d|(w_t, q_t)$ is a non-degenerate random number that, via optimal labor demand, depends on $\tilde{\theta}_t$. In the next section, it will be seen that forming the posterior $\tilde{\chi}_t^d|\chi_t^s$ conveniently summarizes the information that can be extracted from the labor market.

To develop an intuition for the conditions under which our main results hold, it is helpful to define them in terms of the posterior $\tilde{\chi}_t^d|\chi_t^s$. Would there be no market noise, then labor demand would equal labor supply and we would have that $\tilde{\chi}_t^d|\chi_t^s = \chi_t^s$. Accordingly, the properties of the random variable $\tilde{\chi}_t^d|\chi_t^s$ reflect how the market noise $\tilde{\Psi}_t$ correlates with the state of the economy. In the following, we impose two restrictions on these properties. Importantly, even though $\tilde{\chi}_t^d|\chi_t^s$ arises endogenously, it is possible to map these restrictions on $\tilde{\chi}_t^d|\chi_t^s$ back into assumptions about the exogenous noise term $\tilde{\Psi}_t$ (for details, see Appendix B.3).[20]

With this in mind, we impose the following key restriction:[21]

Property 1: *$Var\{\tilde{\chi}_t^d|\chi_t^s\}$ is constant for all $\chi_t^s \in \mathbb{R}$.*

Property 1 ensures that the amount of information about the normalized labor demand $\tilde{\chi}_t^d$ that is contained in the market clearing wage $\tilde{w}_t$ is constant throughout all states of the economy. In Proposition 1, we show that this specification is equivalent to requiring that, in the absence of credit constraints, uncertainty about $\tilde{\theta}_t$ behaves exactly like in a standard economy where it is constant over time. This ensures that there is *no* time dependency of uncertainty inherent to the stochastic process $\tilde{\Psi}_t$, so that *any* variation of uncertainty will be the result of credit constraints.

Additionally, we shall also require the following regularity condition:

Property 2: *It holds that*

(i) $\tilde{\chi}_t^d|\chi_t^s$ *satisfies the monotone likelihood ratio property (MLRP) with respect to* χ_t^s, or

(ii) $\tilde{\chi}_t^m|\chi_t^s$ *belongs to a location-scale family of distributions; i.e.,* $\tilde{\chi}_t^m|\chi_t^s = \alpha_1(\chi_t^s) + \alpha_2(\chi_t^s)\tilde{X}$ *where* $\tilde{X}$ *is a non-degenerate, square-integrable random variable with mean zero and* $\alpha_1 : supp(\tilde{\chi}_t^s) \to \mathbb{R}$ *increasing.*

[20] The bottom line is that there exists a monotone transformation of $\tilde{\Psi}_t$ that directly enters the updating problem of agents which gives rise to $\tilde{\chi}_t^d|\chi_t^s$. By "backward-engineering" Bayes' law, any assumption on the posterior distribution, can therefore also be traced back to an assumption in terms of the signal structure defined by $\tilde{\Psi}_t$.

[21] Note that in Properties 1 and 2, the conditional distributions $\tilde{\chi}_t^d|\chi_t^s$ and $\tilde{\chi}_t^m|\chi_t^s$ are meant to denote posterior distributions that result from a flat prior.

This property states that observing a higher labor supply allows for the statistical inference that also the corresponding fundamental labor demand (net of market noise $\tilde{\Psi}_t$) is higher in the sense of stochastic ordering. This is the natural analogue to the case without market noise, where fundamental demand *exactly* equals supply. More specifically, Property 2 specifies two alternative ordering criterion. In case (i), we adopt the commonly used monotone likelihood ratio property.[22] In case (ii), we state an alternative distributional assumption which gives rise to a specific class of "location-scale" posteriors. Here, $\alpha_1(\chi_t^s)$ is the mean of the posterior and $(\alpha_2(\chi_t^s))^2$ is proportional to the posterior variance, where the ordering takes the form of assuming that α_1 is increasing. While it will be seen that tighter financial constraints imply an higher uncertainty in both cases, introducing the more specific location-scale setting allows us later to focus on a conjugate Gaussian framework for analyzing the dynamics of this economy.

Timing The timing of events within one period can be summarized as follows:

1. The random variables $\{\tilde{\epsilon}_t, \{\tilde{\xi}_{it} : i \in [0,1]\}\}$ are realized and agents learn the realizations of $\tilde{s}_{it}$.

2. Noise traders' demand and supply $\{\tilde{\eta}_t, \tilde{\omega}_t\}$ realize, the labor and asset market operate.

3. Field production takes place, farmers choose whether or not to pay their wage bill, and gatherers seize collaterals if farmers default on their wage debt.

4. The asset market operates again.

5. Fruits from trees are gathered and consumption takes place.

6. A new generation replaces the old one and period $t+1$ begins.

Equilibrium definition Because of the assumption that agents cannot trade on margin, farmers are prevented from sidestepping collateral constraints by buying additional trees. Accordingly, the only benefit of holding trees that is reflected in the market clearing price q_t is the expected dividend payoff. Moreover, a simple arbitrage argument then implies that trees are traded at the same price in both openings of the asset market at any date t. In appendix B.1, we show

22 Formally, MLRP states that $\chi_t^s < \hat{\chi}_t^s$ implies that $\text{Prob}(\tilde{\chi}_t^d | \chi_t^s)/\text{Prob}(\tilde{\chi}_t^d | \hat{\chi}_t^s)$ is decreasing in $\tilde{\chi}_t^d$.

that without loss of generality we can treat the two asset markets as if all assets were traded in a single pooled market that operates parallel to the labor market, and where labor demand of all farmers is constrained to satisfy $n_{it} \leq q_t/w_t$. Accordingly, the information $\{\mathcal{I}_{it} : i \in [0,1]\}$ defined in the preceding paragraphs is the basis for all labor supply $\{n_{it} : i \in \mathcal{G}\}$, labor demand $\{n_{it} : i \in \mathcal{F}\}$, and asset demand choices $\{x_{it} : i \in [0,1]\}$ at date t. Given these considerations, a competitive rational expectations equilibrium is then defined in the usual manner.

Definition: Given a stochastic process of shocks $\{\tilde{\epsilon}_t, \tilde{\eta}_t, \tilde{\omega}_t, \{\tilde{\xi}_{it} : i \in [0,1]\}\}$, an equilibrium in this economy is a stochastic process of choices $\{\tilde{x}_{it}, \tilde{n}_{it} : i \in [0,1]\}$ and prices $\{\tilde{w}_t, \tilde{q}_t\}$, such that:

1. $\{\tilde{x}_{it}, \tilde{n}_{it} : i \in [0,1]\}$ maximize expected utility (1) and (2) given $\{\tilde{w}_t, \tilde{q}_t\}$ and $\{\tilde{s}_{it} : i \in [0,1]\}$;

2. markets clear, i.e.,

$$(1-m)\int_{\mathcal{G}} \tilde{n}_{it}\,\mathrm{d}i = m\int_{\mathcal{F}} \tilde{n}_{it}\,\mathrm{d}i + \Psi_{\tilde{\theta}_t,\tilde{q}_t}(\tilde{\omega}_t) \tag{4}$$

   and

$$\int_0^1 \tilde{x}_{it}\,\mathrm{d}i + \Phi\left(\sqrt{\tau_\xi}\,(\tilde{\eta}_t - \mu)\right) = 1; \tag{5}$$

3. expectations in (1) and (2) are formed optimally given $\{\tilde{w}_t, \tilde{q}_t\}$ and $\{\tilde{s}_{it}\} = \{\tilde{\theta}_t + \tilde{\xi}_{it}\}$.

3. Learning with financial frictions

In this section, we explore the key mechanism of this study. It will be seen how learning from the real sector breaks down when financial constraints are tight.

Agents learn from the real sector via the *endogenous* history of market prices $\{\tilde{w}_t, \tilde{q}_t\}$. Hereby, asset prices play a dual role. On the one hand, changes in $\tilde{q}_t$ tighten financial constraints and thereby affect the problem of extracting information from $\tilde{w}_t$. On the other hand, $\tilde{q}_t$ is also a source of information on its own. In the next section, it will be seen that this latter problem of extracting information from $\tilde{q}_t$ is standard. For now, we therefore focus on the novel problem of extracting information from $\tilde{w}_t$, by studying how an exogenously given asset price $\bar{q}$—without any informational content on its own—constrains farmers' choices and how this affects the information aggregation.

From the market clearing condition (4), we have that $\tilde{w}_t$ solves

$$(1-m)v'^{-1}(\tilde{w}_t) = n(\tilde{w}_t, \bar{q}, \tilde{\theta}_t) + \tilde{\Psi}_t,$$

where $v'^{-1}(\tilde{w}_t)$ is the optimal labor supply of a single gatherer, $\min\{A_{it}, \bar{q}\}/w_t$ is the optimal labor demand of a single farmer with productivity A_{it}, and

$$n(\tilde{w}_t, \bar{q}, \tilde{\theta}_t) = \frac{m}{\tilde{w}_t} \int_{-\infty}^{\infty} \min\{\exp(z), \bar{q}\} \, \mathrm{d}\Phi(\sqrt{\tau_\xi}\,(z - \tilde{\theta}_t)), \tag{6}$$

is the aggregated labor demand. Transforming noisy labor demand from an additive perturbation $\tilde{\Psi}_t$ to a multiplicative perturbation $\tilde{\psi}_t$ of farmers' labor demand,[23] the market clearing condition can be rewritten as

$$\tilde{\chi}_t^s = \tilde{\chi}_t^d + \tilde{\psi}_t, \tag{7}$$

where $\tilde{\chi}_t^s$ and $\tilde{\chi}_t^d$ are normalized labor supply and demand as defined in the previous section.

Conditional on any realization of $(\tilde{w}_t, \tilde{q}_t)$, $\tilde{\chi}_t^s$ is a publicly known number. Learning from the real sector is therefore equivalent to observing a signal $\tilde{\chi}_t^s$ which communicates the true value of $\tilde{\chi}_t^d$ perturbed by $\tilde{\psi}_t$. This amounts to a nonlinear signal structure. To see this, note that

$$\tilde{\chi}_t^d = H(\tilde{\theta}_t - \log(\bar{q})),$$

where $H : \mathbb{R} \to \mathbb{R}_-$ is defined by $H(x) = \log(n(1,1,x))$. The key observation is that H is increasing and concave. Intuitively, aggregated labor demand is obviously increasing in the average productivity. However, as an increasing number of farmers is operating at their collateral constraint (i.e., as $(\theta_t - \log(\bar{q}))$ increases), fewer farmers respond to changes in their productivities. Aggregating thus implies that aggregated labor demand $\tilde{\chi}_t^d = H(\tilde{\theta}_t - \log(\bar{q}))$ is also less responsive to changes in the fundamental $\tilde{\theta}_t$. Hence the concavity of H.

[23] Formally, $\tilde{\psi}_t = \log(\tilde{\Psi}_t/n_t + 1)$; e.g., $\psi_t = 0.01$ refers to approximately an one percent amplification of fundamental labor demand $n_t = n(w_t, q_t, \theta_t)$.

3.1. *Signal extraction* without *credit constraints*

Before proceeding to our main theorem, it is insightful to first consider the limit case where all farmers are unconstrained. Formally, let $\bar{q} \to \infty$. Then $H' \to 1$ for all $\theta_t \in \mathbb{R}$, so that $\mathrm{Var}\{\tilde{\theta}_t | \chi_t^s\} = \mathrm{Var}\{\tilde{\chi}_t^d | \chi_t^s\}$. Property 1 therefore exactly ensures that learning in the absence of credit constraints yields a posterior uncertainty that is constant over time.

Proposition 1: *Absent credit constraints, $Var\{\tilde{\theta}_t | \chi_t^s\} = Var\{\tilde{\chi}_t^d | \chi_t^s\}$; i.e., $Var\{\tilde{\theta}_t | \chi_t^s\}$ is constant if and only if Property 1 holds.*

By Proposition 1, uncertainty is constant in any unconstrained version of our economy. Any variations in uncertainty are therefore the exclusive result of learning in the presence of credit constraints.

3.2. *Signal extraction* with *credit constraints*

We now address how credit constraints affect learning from the real sector. We have already discussed that financial constraints give rise to a concave signal structure. The following theorem states that independently from the specific properties of our model, such a signal structure always leads to signals with a precision that decreases as the signal realizes in flatter regions.[24]

Theorem 1: *Let $\tilde{s}$, $\tilde{\theta}$ and $\tilde{\varepsilon}$ be three non-degenerate random variables, and let f be an increasing function defined on the convex hull of the support of $\tilde{\theta}$, such that $\tilde{s} = f(\tilde{\theta}) + \tilde{\varepsilon}$ and $\tilde{\theta}$, $f(\tilde{\theta})$ square-integrable. Furthermore suppose that either $f(\tilde{\theta})|s$ satisfies the monotone likelihood ratio property with respect to s;* or *$\tilde{\theta}|s = \alpha_1(s) + \alpha_2(s)\tilde{X}$ for some non-degenerate, square-integrable random variable $\tilde{X}$ and some functions α_1, α_2 defined on the support of $\tilde{s}$ and α_1 increasing. Then*

1. *$Var\{\tilde{\theta}|s\}$ is increasing in s if f is concave and $Var\{f(\tilde{\theta})|s\}$ is nondecreasing in s,*
2. *$Var\{\tilde{\theta}|s\}$ is decreasing in s if f is convex and $Var\{f(\tilde{\theta})|s\}$ is nonincreasing in s.*

[24] For readability, we state the theorem for increasing functions f only, but it is straightforward to generalize it to all monotonic f, leading to the opposite predictions for $\mathrm{Var}\{\tilde{\theta}|s\}$ if f is decreasing.

In both cases, the monotonicity of $Var\{\tilde{\theta}|s\}$ *is strict whenever f is strictly concave or convex.*

The intuition behind this theorem is quiet straightforward. When a signal $\tilde{s}$ realizes in flatter regions, then a "well-behaved" signal structure allows for the posterior belief that also the fundamental $\tilde{\theta}$ takes values for which f is flat.[25] But then, a Bayesian must also believe that the realization of $\tilde{s}$ is largely driven by noise $\tilde{\epsilon}$ rather than by the fundamental $\tilde{\theta}$. Hence the increase in posterior uncertainty.

Framed in terms of our model, this reasoning translates into wage signals that are largely driven by noisy demand fluctuations rather than fundamentals whenever the economy is in a constrained state. Formally, we have:

Proposition 2: *$Var\{\tilde{\theta}_t|\chi_t^s\}$ is strictly increasing in χ_t^s.*

The role of Proposition 2 for this study cannot be overstated. It precisely tells us how and when uncertainty about $\tilde{\theta}_t$ is fluctuating in the economy. In particular, it establishes that uncertainty rises after negative shocks to $\tilde{q}_t$, which is at the core of all our results.

3.3. *Quasi-Gaussian signal structure*

Proposition 2 is inextricably tied to learning through the nonlinear function H. Unfortunately, learning from nonlinear Gaussian signals generally leads to posterior distributions that do not conjugate with the corresponding priors, making the dynamic analysis highly intractable. To address this problem, we henceforth restrict ourselves to information structures that satisfy case (ii) of Property 2 with $\tilde{X}$ being a standard normal random variable. Under this assumption, beliefs of agents evolve *as if* agents observed a Gaussian signal with an exogenously given *state-dependent* signal precision τ_v. Using this approach of informationally equivalent Gaussian signals, we are able to embed the idea of time-varying uncertainty in a convenient framework with conjugate Gaussian priors.

[25] By "well-behaved", we refer to the assumption that either $f(\tilde{\theta}|s)$ are ordered according to the MLRP property—which then in Milgrom's (1981) language implies that "good news" for $\tilde{s}$ is also "good news" for $f(\tilde{\theta})$ and, hence, for $\tilde{\theta}$—or, alternatively, that the first moment of the posterior distribution of $\tilde{\theta}$ is increasing in s—similarly implying that good news for $\tilde{s}$ is good news for $\tilde{\theta}$.

LEMMA 1: *Suppose Property 2(ii) holds with $\tilde{X}$ being standard normally distributed. Then, given a normal prior over $\tilde{\theta}_t$, observing $\tilde{\chi}_t^s = \chi_t^s$ is equivalent to observing a signal $\tilde{\theta}_t + \tilde{v}_t$ with realization $\alpha_1(\chi_t^s) + \log(\bar{q})$, where $\tilde{v}_t$ is Gaussian noise with variance $1/\tau_v(\theta_t + v_t - \log(\bar{q}))$ and $\tau_v : \mathbb{R} \to \mathbb{R}_+$ is strictly decreasing,* $\lim_{z\to-\infty} \tau_v(z) = 1/Var\{\tilde{\chi}_t^d|\chi_t^s\}$, *and* $\lim_{z\to\infty} \tau_v(z) = 0$.

The "quasi-Gaussian" signal structure that follows from Lemma 1 effectively decomposes the inference problem of agents into a straightforward interpretation of a Gaussian signal and a computationally intensive calculation of the relevant signal precision. The benefit of this decomposition is that agents' beliefs can be computed straightforwardly within a conjugate prior framework. The computationally intensive part, on the other hand, only has to be solved by the model analyst. However, applying the results from Proposition 2, we already know that τ_v is decreasing as the economy gets more constrained, enabling us to derive all of our results analytically without the need to solve the exact inference problem. (For providing a numerical example, we simulate the exact mapping in Section 6.)

Figure 1 illustrates the properties of τ_v that follow from Lemma 1: (i) τ_v is decreasing as the economy gets more constrained (i.e., as $\alpha_1(\chi_t^s) = \theta_t + v_t - \log(\bar{q})$ increases, reflecting either tighter credit conditions, or an increased labor demand relative to existing credit conditions); (ii) τ_v converges to $1/\mathrm{Var}\{\tilde{\chi}_t^d|\chi_t^s\} = const$ as the economy gets completely unconstrained (i.e, no information is lost when all farmers are unconstrained); and (iii) τ_v converges towards zero, so that the signal gets completely uninformative, as the economy becomes fully constrained.

4. Equilibrium characterization

Equipped with the quasi-Gaussian representation, we are now ready to characterize the equilibrium. We proceed in two steps. First, we fix the information structure (by fixing τ_v) and analyze the resulting signal extraction problem taking into account all signals. Because parts of agents' information is extracted from endogenous asset prices, this step involves finding the "usual" fixed point between a perceived law of motion of $\tilde{q}_t$ and its actual behavior. Second, allowing the information structure to vary with the state of the economy, we then establish the full informational equilibrium where information is simultaneously aggregated from labor and asset markets and where prices in both markets are consistent

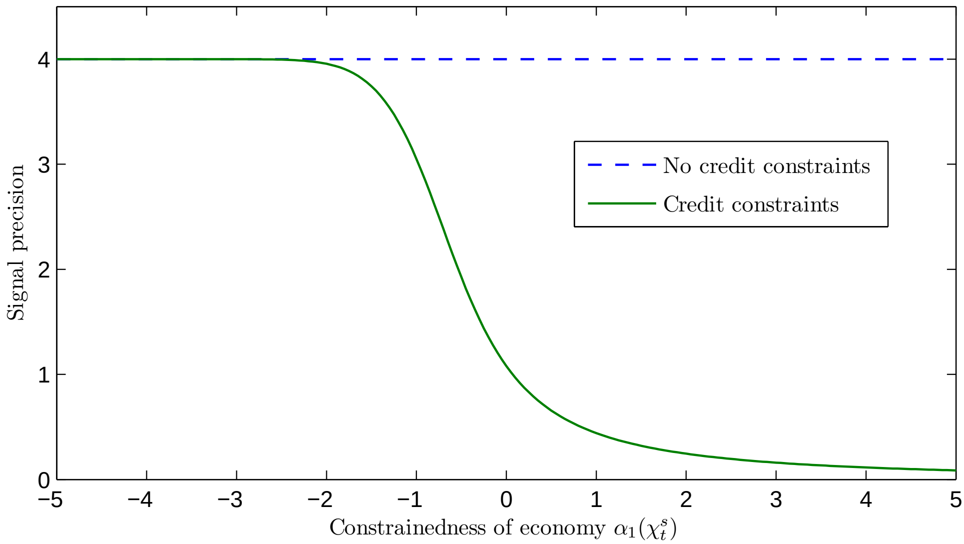

Figure 1. Endogenous signal precision.

with the resulting beliefs.[26]

4.1. Signal extraction from the financial market

From Lemma 1 it follows that we can focus on Normal posteriors. Accordingly, let $b_t \equiv \mathrm{E}\{\tilde{\theta}_t | \mathcal{I}_{it} \smallsetminus s_{it}\}$ and $1/\pi_t \equiv \mathrm{Var}\{\tilde{\theta}_t | \mathcal{I}_{it} \smallsetminus s_{it}\}$ denote the posterior mean and variance that result from observing the *public* history of market prices up to date t. Further let $b_{it} \equiv \mathrm{E}\{\tilde{\theta}_t | \mathcal{I}_{it}\}$ and $1/\bar{\pi}_t \equiv \mathrm{Var}\{\tilde{\theta}_t | \mathcal{I}_{it}\}$ denote the first two posterior moments given information set $\mathcal{I}_{it}$ (i.e., including the *private* signal $\tilde{s}_{it}$), where in anticipation of the results below, we drop the subscript i from the posterior precision $\bar{\pi}_t$. Then optimal asset demand is given by

$$x_{it} = \begin{cases} 1 & \text{if } \mathrm{E}\{\tilde{A}_t | b_{it}, \bar{\pi}_t\} > q_t, \\ [0,1] & \text{if } \mathrm{E}\{\tilde{A}_t | b_{it}, \bar{\pi}_t\} = q_t, \\ 0 & \text{if } \mathrm{E}\{\tilde{A}_t | b_{it}, \bar{\pi}_t\} < q_t, \end{cases}$$

[26] Formally, agents post labor and asset demand schedules that are fully contingent on both prices. Agents therefore learn from observing prices in the labor and asset market simultaneously and the resulting beliefs have to be consistent with the market clearing prices on both markets.

where $\mathrm{E}\{\tilde{A}_t|b_{it}, \bar{\pi}_t\} = \exp(\gamma_0 + \gamma_1 b_{it} + \gamma_1^2/(2\bar{\pi}_t))$.[27] Aggregating, adding noise traders' demand, and rearranging yields the market clearing condition

$$\Phi(\sigma_t^{-1}(b_{mt}(q_t) - \bar{b}_t)) = \Phi(\sqrt{\tau_\xi}\,(\tilde{\eta}_t - \mu)),$$

where $b_{mt}(q_t)$ is the belief of the marginal trader with $\mathrm{E}\{\tilde{A}_t|b_{mt}, \bar{\pi}_t\} = q_t$, $\bar{b}_t$ is the average belief $\int b_{it}\,\mathrm{d}i$, and $\sigma_t^2 = \mathrm{Var}\{b_{it}\}$ is the cross-sectional dispersion of beliefs. This pins down the marginal traders' belief $b_{mt} = \sigma_t\sqrt{\tau_\xi}\,(\tilde{\eta}_t - \mu) + \bar{b}_t$ and the equilibrium price. Given any conjectured law of motion for the market clearing price, this price also serves as an endogenous signal. In equilibrium the beliefs resulting from interpreting this signal have to give rise to optimal asset demands that yield an actual law of motion equal to the conjectured one. This fixed-point problem has a log-linear solution, which is established in the following lemma (for a detailed derivation, see, e.g., Hellwig (1980)).

Lemma 2: *Fix any τ_v. Then there exists a unique log-linear asset price*

$$q_t = \exp\{\gamma_1(\bar{b}_t + \bar{\pi}_t^{-1}\tau_\xi\eta_t) + \gamma_0\} \tag{8}$$

which clears the asset market, where

$$\bar{b}_t = \bar{\pi}_t^{-1} \times \begin{bmatrix} \tau_\xi & \tau_\eta & \tau_v & \hat{\pi}_{t-1} \end{bmatrix} \times \begin{bmatrix} \theta_t \\ \theta_t + \eta_t \\ \theta_t + v_t \\ \rho b_{t-1} \end{bmatrix},$$

$$\bar{\pi}_t = \tau_\xi + \tau_\eta + \tau_v + \hat{\pi}_{t-1},$$

$$\hat{\pi}_{t-1} = \frac{\pi_{t-1}\tau_\epsilon}{\pi_{t-1} + \rho^2\tau_\epsilon},$$

and where (b_t, π_t) are given by $(\bar{b}_t, \bar{\pi}_t)$ after setting $\tau_\xi = 0$.

Given a public history summarized by b_{t-1} and π_{t-1}, and given a signal precision τ_v, this lemma characterizes the unique log-linear relationship between the asset price at date t and the stochastic variables $\tilde{\theta}_t$, $\tilde{\eta}_t$, and $\tilde{v}_t$. The equilibrium

[27] Here we assume without loss of generality that $\tilde{u}_t$ is constant zero. Suppose it is not. Then by the independence of $\tilde{u}_t$ we have that $\mathrm{E}\{\tilde{A}_t|b_{it}, \bar{\pi}_t\} = \mathrm{E}\{\exp(\tilde{u}_t)\}\exp(\gamma_0+\gamma_1 b_{it}+\gamma_1^2/(2\bar{\pi}_t)) = \exp(\gamma_0'+\gamma_1 b_{it}+\gamma_1^2/(2\bar{\pi}_t))$, where $\gamma_0' = \gamma_0+\log(\mathrm{E}\{\exp(\tilde{u}_t)\})$ and, hence, any nonzero random noise $\tilde{u}_t$ can be absorbed by the constant γ_0.

price increases in all three variables, not just in the fundamental. The reason is that positive noisy demand realizations on the asset and labor market falsely suggest that the fundamental increased. The exact weight that is put on these sources of information, however, depends on the signal precision τ_v. Intuitively, as the real sector aggregates less information, less weight is placed on real sector prices, and vice versa. We also see that posterior uncertainty $1/\bar{\pi}_t$ is increasing as the signal precision τ_v of the labor market decreases.

4.2. *General informational equilibrium*

To characterize the full informational equilibrium that takes into account the mutual dependence of signal precision and asset price, define $g_q : \mathbb{R}_+^2 \times \mathbb{R}^4 \to \mathbb{R}$, such that $\log(q_t) = g_q(\tau_v, \Omega_t)$ as given by (8), and define $g_\tau : \mathbb{R}_+ \times \mathbb{R}^5 \to \mathbb{R}_+$, such that $\tau_v(\theta_t + v_t - \log(q_t)) = g_\tau(\log(q_t), \Omega_t)$ is the signal precision defined in Lemma 1. Here, $\Omega_t \equiv (\pi_{t-1}, b_{t-1}, \theta_t, \eta_t, v_t)$ is the vector of state variables in period t. Then the informational equilibrium can be computed pointwise for any state of the economy Ω_t, by solving the fixed-point problem[28]

$$g_\tau(\cdot, \Omega_t) - g_q^{-1}(\cdot, \Omega_t) = 0. \tag{9}$$

The following proposition establishes the existence of a solution to this problem for all possible states Ω_t and, hence, the existence of an informational equilibrium.

Proposition 3: *For all $\Omega_t \in \mathbb{R}_+ \times \mathbb{R}^4$, there exists a solution to the fixed-point problem* (9). *The solution is unique for all Ω_t inside a set $\Xi \subset \mathbb{R}_+ \times \mathbb{R}^4$. In particular, it is unique for all Ω_t that satisfy*

$$(\gamma_1 - 1)(\theta_t + v_t) < M_1$$

or

$$\theta_t + v_t < M_2 + \gamma_1^{-1} \log(q_t|_{\tau_v \to 0})$$

where $M_1, M_2 \in \mathbb{R}$ are parameters defined by the primitives of the model. In contrast, if $\Omega_t \notin \Xi$, the economy is in a "sunspot" state where (9) *has multiple solutions.*

Proposition 3 implies that for any initial state Ω_0 the equilibrium dynamics of the model economy can be computed recursively by computing a solution to (9)

[28] For expositional reasons, we abstract from the non-generic case where g_q is flat throughout the main body of this chapter. The case where g_q^{-1} does not exist is carefully treated in all formal proofs.

given Ω_t and then using Lemma 2 to determine Ω_{t+1}. As long as $\{\Omega_s\}_{s=1}^t \in \Xi^t$, this pins down a unique equilibrium path. This is the case as long as the information aggregated in the labor and asset market is not too conflicting. In states where the labor market signals a realizations for $\tilde{\theta}_t$ that is sufficiently more optimistic than what is suggested by the asset market, the equilibrium gives rise to self-fulfilling sunspots: If agents coordinate on interpreting the conflicting information in an optimistic way, asset prices will be little constraining, so that observing the "good" labor market news will be sufficiently informative to justify an optimistic interpretation. On the other hand, if agents interpret the evidence in a pessimistic way, the resulting financial constraints will obscure the "good" news from the labor market and the pessimistic interpretation is indeed justified.

Because the findings in this study generally carry over to sunspot regimes (but require a more subtle distinction between cases), we focus throughout most of this study on paths where $\{\Omega_s\}_{s=1}^t \in \Xi^t$. A brief discussion of sunspot regimes can be found in Appendix B.4.

5. Transmission of shocks

We now explore how random shocks to the economy are propagated through the informational equilibrium. In the next two subsections, it will be seen how shocks are statically amplified or de-amplified depending on their size and composition. In Section 5.3, it will be seen that a similar distinction divides shocks in those that are dynamically persistent, and others that are non-persistent.

5.1. Static asymmetries: Amplifying and de-amplifying shocks

Consider the solution to the fixed point problem (9). We say that a shock is *amplified* through the endogenous information structure if and only if its absolute impact on q_t is larger than in the hypothetical benchmark in which τ_v is fixed at the level it would attain in the absence of shocks. By that definition, which shocks are amplified and which shocks are de-amplified upon impact?

For answering this question, it is useful to define $\mathring{\tau}_v \equiv \tau_v(-\gamma_0)$ to be the precision of v_t in the absence of shocks.[29] In the benchmark case, the market-clearing price then reads $\mathring{q}_t \equiv \exp(g_q(\mathring{\tau}_v, \Omega_t))$. Contrast this with the case, where τ_v adjusts endogenously to the asset price. Then the endogeneity of the

[29] I.e., for $\theta_t = \eta_t = v_t = 0$ and $b_{t-1} = 0$, implying $\log(q_t) = \gamma_0$ and, hence, $\mathring{\tau}_v = \tau_v(-\gamma_0)$.

information structure unfolds a feedback loop. Intuitively, the benchmark price $\mathring{q}_t$ can be seen as the "initial" impact of the shock. In response to this change in $\tilde{q}_t$ and also in response to the shock itself (compare Lemma 1), τ_v is now changing to $\tau_v^1 \equiv g_\tau(\log(\mathring{q}_t), \Omega_t)$. However, unless $\tau_v^1 = \mathring{\tau}_v$, agents who optimally form expectations respond to such changes in τ_v by re-weighting the available sources of information, leading to an asset price q_t^1, and so on.

To answer which shocks are amplified, we need to compare the equilibrium of this feedback loop—i.e., the solution q_t^* to (9)—with $\mathring{q}_t$. Observe that the two directions of the feedback loop can be summarized by (i) the sign of $\partial g_q / \partial \tau_v$ (how do changes in τ_v feed back to q_t), and (ii) the sign of $\{\tau_v^1 - \mathring{\tau}_v\}$ (after a single "cycle" through the feedback loop, how does τ_v change in response to q_t^0).[30] Combining, we have four cases to consider, summarized by Figure 2. If g_q is increasing in τ_v and $\tau_v^1 < \mathring{\tau}_v$, then g_q^{-1} intersects with g_τ to the left of $\log(\mathring{q}_t)$ and we have that $q_t^* < \mathring{q}_t$ (see Panel a). Likewise, if g_q is increasing in τ_v and $\tau_v^1 > \mathring{\tau}_v$, then $q_t^* > \mathring{q}_t$ (see Panel b). A similar reasoning establishes the opposite relationship if g_q is decreasing in τ_v (see Panels c and d).

So what determines the signs of $\partial g_q / \partial \tau_v$ and $\{\tau_v^1 - \mathring{\tau}_v\}$? From Lemma 2, we have that

$$\operatorname{sign}\left\{\frac{\partial g_q}{\partial \tau_v}\right\} = \operatorname{sign}\left\{\frac{\partial}{\partial \tau_v}\left(\bar{\pi}_t^{-1}\begin{bmatrix}\tau_\xi + \tau_\eta & \tau_v & \hat{\pi}_{t-1}\end{bmatrix} \times \begin{bmatrix}\theta_t + \eta_t \\ \theta_t + v_t \\ \rho b_{t-1}\end{bmatrix}\right)\right\}$$
$$= \operatorname{sign}\left\{-(\theta_t + \eta_t) + \delta_1(\theta_t + v_t)\right\},$$

where $\delta_1 : \mathbb{R} \to \mathbb{R}$ is defined by

$$\delta_1(x) = (\tau_\xi + \tau_\eta)^{-1}(-\check{b}_{t-1} + \check{\pi}_{t-1}x)$$

with $\check{b}_{t-1} = \rho b_{t-1}\hat{\pi}_{t-1}$ and $\check{\pi}_{t-1} = \tau_\xi + \tau_\eta + \hat{\pi}_{t-1}$. Here, $\delta_1^{-1}(\theta_t + \eta_t)$ is the residual asset price, which would obtain if there were no labor market signal. It includes all information that is inferred from *other* sources than w_t—i.e., the prior b_{t-1}, the idiosyncratic signals $\{s_{it}\}$, and news extracted from q_t itself. Weighting these sources against $\theta_t + v_t$—the information extracted from w_t—then pins down the equilibrium price. Intuitively, if $\theta_t + v_t$ is larger than $\delta_1^{-1}(\theta_t + \eta_t)$, then if agents increase the weight on w_t, they become more optimistic. As a result, g_q is increasing in τ_v exactly if $\theta_t + \eta_t < \delta_1(\theta_t + v_t)$.

[30] Here we exploit that, as formally shown in the proof to Proposition 4, $\operatorname{sign}\{\tau_v^1 - \mathring{\tau}_v\} = \operatorname{sign}\{\tau_v^* - \mathring{\tau}_v\}$, allowing us to use $\operatorname{sign}\{\tau_v^1 - \mathring{\tau}_v\}$ to determine the qualitative effects of the feedback loop on q_t^*.

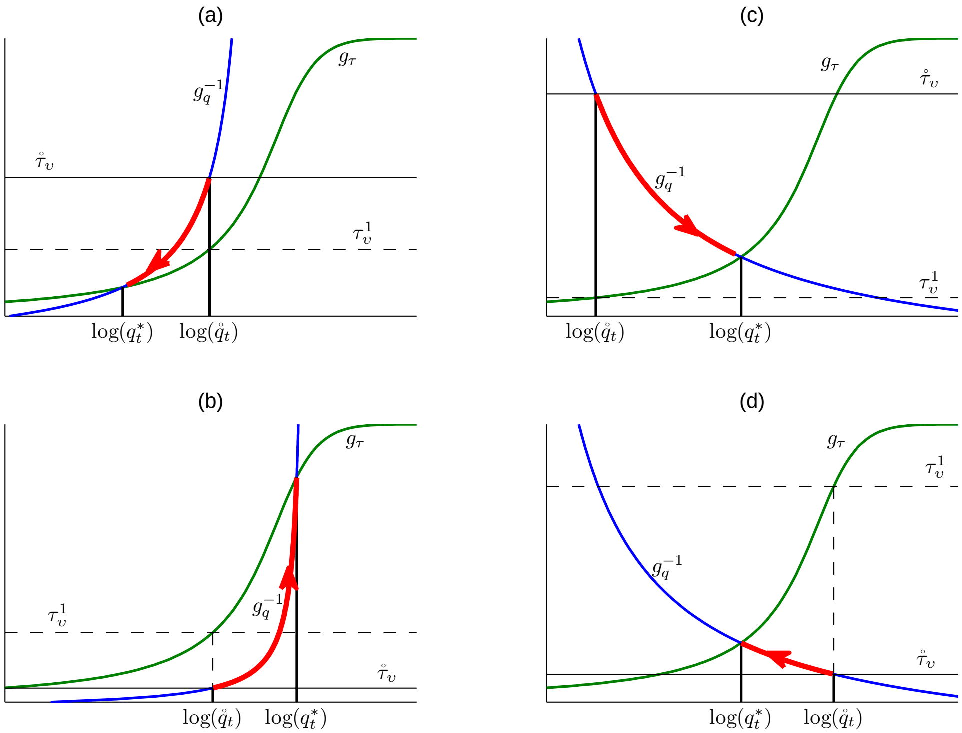

Figure 2. Location of equilibrium asset price relative to the counterfactual price.

To determine the sign of $\{\tau_v^1 - \mathring{\tau}_v\}$, we apply the transformation τ_v^{-1}. Since τ_v is decreasing, the term is positive exactly if $\theta_t + v_t - \log(\mathring{q}_t) < -\gamma_0$, or if $\theta_t + \eta_t > \delta_2(\theta_t + v_t)$, where $\delta_2 : \mathbb{R} \to \mathbb{R}$ is defined by

$$\delta_2(x) = (\tau_\xi + \tau_\eta)^{-1}(-\check{b}_{t-1} + \gamma_1^{-1}(\check{\pi}_{t-1} + (1-\gamma_1)\mathring{\tau}_v)x).$$

Note that δ_2 is decreasing, if and only if $\gamma_1 > \check{\pi}_{t-1}/\mathring{\tau}_v + 1$. This reflects the case, where asset prices have a strong enough impact on τ_v to compensate for any direct impact of $\theta_t + v_t$. By contrast, if γ_1 is sufficiently small, then any effect that a positive realization of $\tilde{\theta}_t + \tilde{v}_t$ has on q_t is dominated by additional labor demand. Thus the economy is effectively more constrained and generates less information as $\theta_t + v_t$ increases, which translates into a positively sloped δ_2.

We are now ready to address the key question in this section. Based on the affine functions δ_1 and δ_2, we can assign each state Ω_t one of the four cases depicted in Figure 2. If we also take into account whether $\mathring{q}_t$ is increased or decreased relative to the no-shock price γ_0, we can therefore determine whether endogenous information amplifies or de-amplifies the impact of shocks in period t. More specifically, when $\mathring{q}_t$ is increased compared to γ_0, then endogenous uncertainty amplifies the impact of Ω_t if $q_t^* > \mathring{q}_t$, and de-amplifies (or, possibly, reverses) the impact if $q_t^* < \mathring{q}_t$. The converse holds true if $\mathring{q}_t$ is decreased compared to the no-shock case. Comparing $\log(\mathring{q}_t)$ with γ_0, we find that $\log(\mathring{q}_t) > \gamma_0$ exactly if $\theta_t + \eta_t > \delta_3(\theta_t + v_t)$, where $\delta_3 : \mathbb{R} \to \mathbb{R}$ is defined by

$$\delta_3(x) = (\tau_\xi + \tau_\eta)^{-1}(-\check{b}_{t-1} - \mathring{\tau}_v x).$$

The state space Ω_t is thus divided into amplification and de-amplification regimes by lines δ_1, δ_2, and δ_3. The following proposition formalizes this result.

PROPOSITION 4: *(a) Suppose that* $\gamma_1 > 1$ *and* $\theta_t + \eta_t > \delta_3(\theta_t + v_t)$, *that is,* $\log(\mathring{q}_t) > \gamma_0$. *Then* $q_t^* > \mathring{q}_t$, *so that the impact of* Ω_t *is amplified if and only if*

$$\delta_1(\theta_t + v_t) < \theta_t + \eta_t < \delta_2(\theta_t + v_t),$$

where the inequalities are reversed for $\gamma_1 < 1$.

(b) Suppose that $\gamma_1 > 1$ *and* $\theta_t + \eta_t < \delta_3(\theta_t + v_t)$, *so that* $\log(\mathring{q}_t) < \gamma_0$. *Then* $q_t^* < \mathring{q}_t$, *so that the impact of* Ω_t *is de-amplified, if and only if*

$$\delta_1(\theta_t + v_t) > \theta_t + \eta_t > \delta_2(\theta_t + v_t),$$

where the inequalities are reversed for $\gamma_1 < 1$.

Figure 3 illustrates the proposition. For all realizations of $\tilde{\theta}_t$, $\tilde{\eta}_t$, and $\tilde{v}_t$ that fall north-east of δ_3, the combined impact on $\mathring{q}_t$ is positive, so that $\log(\mathring{q}_t) > \gamma_0$. For these shocks, the impact of Ω_t compared to the exogenous uncertainty benchmark is de-amplified in region A and amplified in region B. Note that region A is split into two separate areas. The one to the north-west of B corresponds to the case depicted in Panel (d) of Figure 2, and the one to the south-east corresponds to Panel (a). Region B corresponds to either Panel (b) or (c), depending on the value of γ_1. If $\gamma_1 > 1$, δ_2 is steeper than δ_1, and all Ω_t in the area bounded by these two lines are amplified by a decreasing g_q as depicted in Panel (c). If $\gamma_1 < 1$, then δ_2 has a smaller slope than δ_1, and Ω_t is amplified as depicted in Panel (b).

For realizations of $\tilde{\theta}_t$, $\tilde{\eta}_t$, and $\tilde{v}_t$ that fall south-west of δ_3, the combined impact has a negative effect on $\mathring{q}_t$, implying $\log(\mathring{q}_t) < \gamma_0$. In that case, realizations within region C are amplified—corresponding to the cases in Panel (a) in the south-east

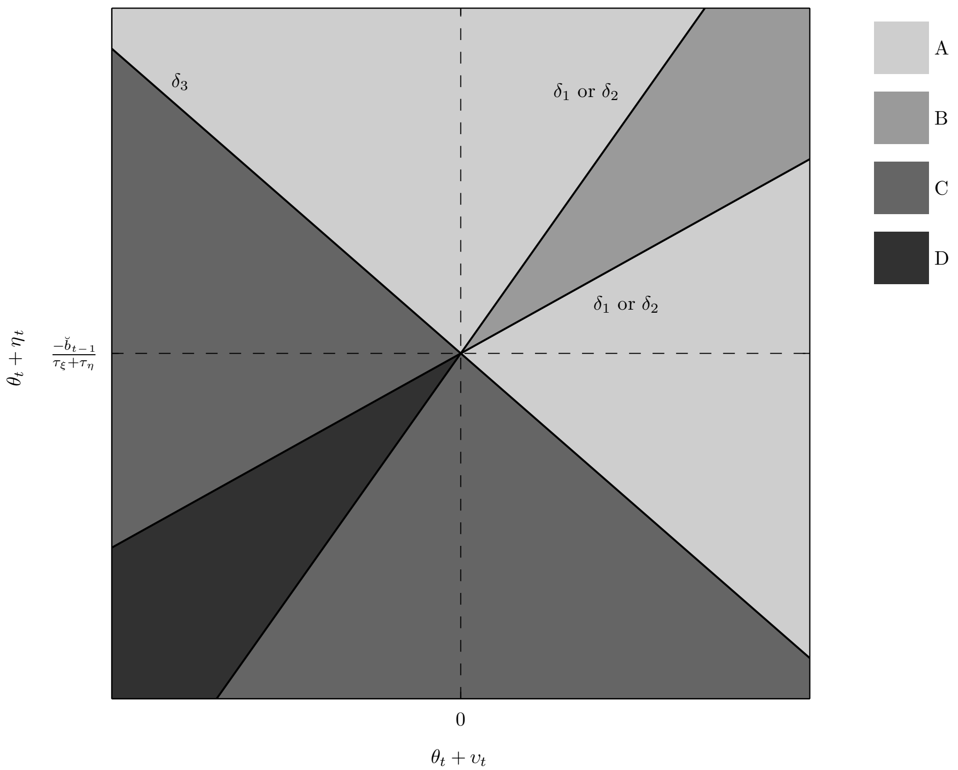

Figure 3. Impact asymmetries. *Note:* Shocks in regions A and B have an overall positive impact on $\mathring{q}_t$, shocks in regions C and D have an overall negative impact. Shocks in regions B and C are endogenously amplified, shocks in regions A and D are endogenously de-amplified.

of region D, and Panel (d) in the north-west of region D. Realizations within region D are de-amplified—corresponding to Panel (b) if $\gamma_1 > 1$, and Panel (c) if $\gamma_1 < 1$.

Macroeconomists are often interested in the special case where the economy is hit by a single shock, shutting down all other stochastic channels through which the economy is impacted. Since δ_1 and δ_2 both have a finite slope, any state in which $|\eta_t|$ is sufficiently large compared to $|\theta_t|$ and $|\upsilon_t|$ is unambiguously amplified for $\eta_t < 0$ and de-amplified for $\eta_t > 0$. In particular, this adverse feedback loop applies to any financial "impulse" shocks; i.e., shocks along the vertical dashed axis through the origin of Figure 3. (See Figure 4 for a schematic illustration of the feedback loop induced by financial shocks.)

Corollary 1: *In the limit as $\theta_t \to 0$, $\upsilon_t \to 0$, and $b_{t-1} \to 0$, financial shocks are*

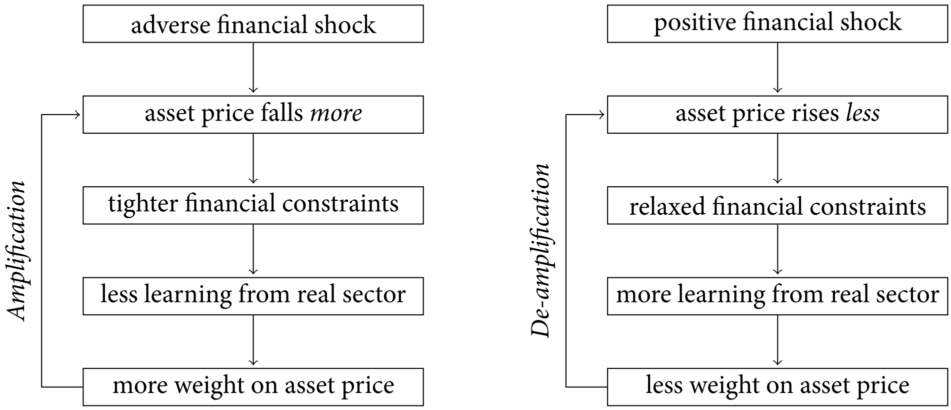

Figure 4. A schematic illustration of the feedback loop in the special case of an isolated financial shock.

amplified if $\eta_t < 0$ *and de-amplified if* $\eta_t > 0$.

Similarly, because b_{t-1} vertically shifts the origin of δ_1, δ_2, and δ_3, any prior pessimism is amplified and any prior optimism is de-amplified along an impulse response path.

COROLLARY 2: *In the limit as* $\theta_t \to 0$, $\eta_t \to 0$, *and* $v_t \to 0$, *prior beliefs are amplified if* $b_{t-1} < 0$ *and de-amplified if* $b_{t-1} > 0$.

The case where the economy is perturbed by a single shock on the labor market (i.e., shocks along the horizontal dashed axis through the origin of Figure 3) is less clear. This is because for large γ_1 the slope of δ_2 becomes negative (but never steeper than the slope of δ_3). Formally, this leads to the following result.

COROLLARY 3: *In the limit as* $\theta_t \to 0$, $\eta_t \to 0$, *and* $b_{t-1} \to 0$, *labor shocks are amplified if* $v_t < 0$ *and de-amplified if* $v_t > 0$ *if and only if* $\gamma_1 < \check{\pi}_{t-1}/\mathring{\tau}_v + 1$. *Otherwise, the converse holds true.*

5.2. *Static asymmetries: Non-proportionality in scale*

Proposition 4 divides the state space into amplifying and de-amplifying regimes. It is silent, however, on how the *degree* of amplification or de-amplification changes within these regimes. We now address this question. In particular, we are interested in how the degree of amplification and de-amplification changes as shocks realize "further away" from the origin, $\mathcal{O}_t \equiv (0, -\breve{b}_{t-1}/(\tau_\xi + \tau_\eta))$, of Figure 3. The following proposition establishes that both the degree of amplification in amplification regimes and the degree of de-amplification in de-amplification regimes monotonically increases as shocks are scaled up relative to $\mathcal{O}_t$.

PROPOSITION 5: *Consider any combination of shocks* $(\theta_t + \upsilon_t, \theta_t + \eta_t) \equiv \mathcal{S}_t + \mathcal{O}_t$. *Then scaling up these shocks to* $a\mathcal{S}_t + \mathcal{O}_t$, $a > 1$, *increases amplification in amplification regimes and decreases amplification in de-amplification regimes. Formally, that is,* $\log(q_t^*) - \log(\mathring{q}_t)$ *decreases in* a *if and only if sign*$\{\theta_t + \eta_t - \delta_1(\theta_t + \upsilon_t)\}$ = *sign*$\{\theta_t + \eta_t - \delta_2(\theta_t + \upsilon_t)\}$.

Proposition 5 states that "scaling up" the combination of shocks that hit the economy at time t, implies that amplification or de-amplification of these shocks is both more pronounced. The reason is that absolute larger shocks lead to larger changes in uncertainty and hence to more pronounced amplification and de-amplification loops, respectively. Figure 5 illustrates these nonlinearities by plotting contours of the degree of amplification $\{\log(q_t^*) - \log(\mathring{q}_t)\} \times \text{sign}(\log(\mathring{q}_t))$. Negative contours (dashed lines) thus correspond to de-amplification regimes, positive contours (solid lines) correspond to amplification regimes. By Proposition 5, these contours are increasing towards the origin in de-amplification regimes, and are decreasing towards the origin in amplification regimes.

5.3. *Dynamic asymmetries: Persistence and non-persistence of beliefs*

We now consider the effect of changes in τ_υ on the persistence of shocks. Because our model abstracts from all intertemporal links other than the fundamental process of $\tilde{\theta}_t$, the only channel through which Ω_t may affect future periods other than through the autocorrelation of $\tilde{\theta}_t$, is through the persistence of public beliefs. For all $s > 0$, we define the persistence of belief b_t onto b_{t+s} as

$$\Lambda_{t,t+s} = \frac{\partial b_{t+s}}{\partial b_t}.$$

In the following, we focus on the specific signal structure underlying our model, but it is worth noting that the arguments generalize to arbitrary quasi-Gaussian signal structures (see the formal proof of Proposition 6). From Lemma 2, we have that

$$b_t = \pi_t^{-1} \times \begin{bmatrix} \tau_\eta & \tau_{\upsilon,t} & \hat{\pi}_{t-1} \end{bmatrix} \times \begin{bmatrix} \theta_t + \eta_t \\ \theta_t + \upsilon_t \\ \rho b_{t-1} \end{bmatrix}. \quad (10)$$

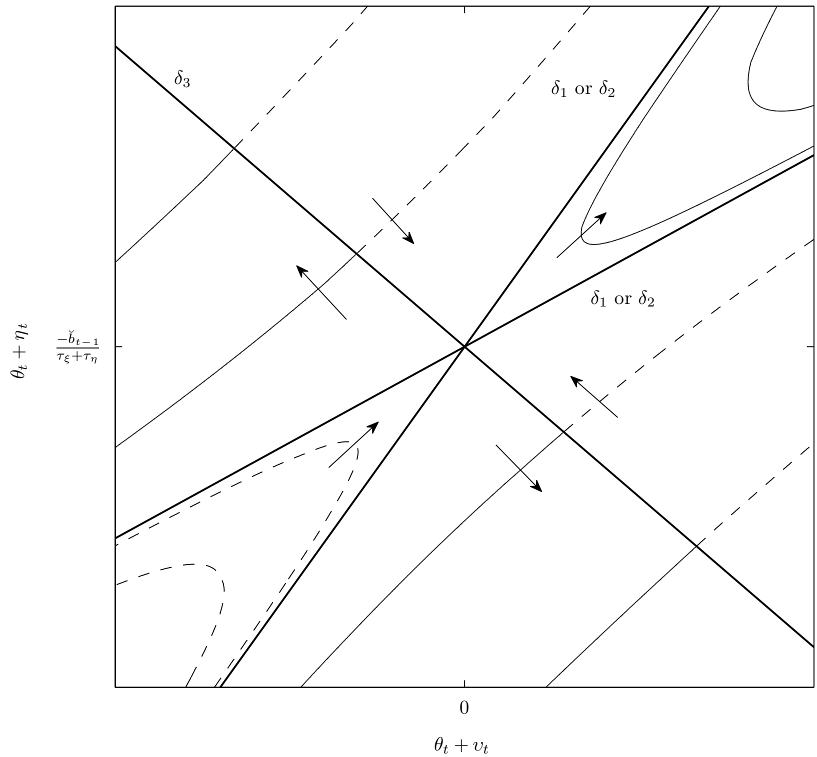

Figure 5. Amplification contours. *Note:* Positive contours (amplification regimes) are plotted solid, negative contours (de-amplification regimes) are dashed. Arrows point into the direction of increasing (more amplifying) contours.

Recursively substituting and differentiating thus yields

$$\Lambda_{t,t+s} = \prod_{q=t+1}^{t+s} \lambda_q,$$

where $\lambda_t \equiv \rho\hat{\pi}_{t-1}/\pi_t$. We are interested in the effect of an increase in uncertainty in period t on the persistence $\Lambda_{t-r,t+s}$ for $r, s \geq 0$. Suppose $\tau_{\upsilon,t}$ changes by a differential $\mathrm{d}\tau_{\upsilon,t}$. Consider the special case where $r = 0$. Then

$$\frac{\mathrm{d}\Lambda_{t,t+s}}{\mathrm{d}\tau_{\upsilon,t}} = \Lambda_{t,t+s} \times \sum_{q=t+1}^{t+s} \left(\frac{1}{\lambda_q} \frac{\partial \lambda_q}{\partial \pi_{q-1}} \prod_{p=t+1}^{q-1} \frac{\partial \pi_p}{\partial \pi_{p-1}} \right).$$

Here, the only term with a nontrivial sign is $\partial\lambda_q/\partial\pi_{q-1}$. However, because $\partial\pi_q/\partial\pi_{q-1} = \partial\hat{\pi}_{q-1}/\partial\pi_{q-1}$, we have

$$\frac{\partial\lambda_q}{\partial\pi_{q-1}} = \frac{\partial\hat{\pi}_{q-1}}{\partial\pi_{q-1}}\frac{\tau_\eta + \tau_{v,q}}{\pi_t^2} > 0,$$

establishing that $\mathrm{d}\Lambda_{t,t+s}/\mathrm{d}\tau_{v,t} > 0$. This reflects that more precise information in period t is unambiguously more relevant for forming *future* beliefs. Contrast this with the case where $s = 0$. Then

$$\frac{\mathrm{d}\Lambda_{t-r,t}}{\mathrm{d}\tau_{v,t}} = \frac{\Lambda_{t-r,t}}{\lambda_t} \times \frac{\partial\lambda_t}{\partial\pi_t} < 0.$$

Now, an increase in period t's information unambiguously decreases the weight on *prior* information. So how do these two opposing effects add up in the general case where $r, s > 0$? The following proposition establishes that the decrease in weight on prior information always dominates all increases in future weights.

PROPOSITION 6: *$\Lambda_{t-r,t+s}$ is decreasing in $\tau_{v,t}$ for all $r, s > 0$.*

Because τ_v is decreased in financial crises, Proposition 6 implies that financial crises are inherently persistent. Intuitively, as the economy receives less news about the current state of the economy, more weight is put on prior information—which during a financial crisis is generally pessimistic. In turn, the asset market continues to constrain the real economy in future periods and, hence, continues to impede learning about $\tilde{\theta}_t$, throwing the economy into a "pessimism trap". In contrast, τ_v increases during financial booms, making them inherently non-persistent. Moreover, as larger shocks have stronger effects on τ_v, persistence and non-persistence are increasing when shocks are "scaled up" as in Proposition 5.

More precisely, from Lemma 2, we have that for $\theta_t = v_t = 0$,

$$\begin{aligned} b_{t+s} &= \pi_t^{-1} \times \begin{bmatrix} \tau_\eta & \hat{\pi}_{t-1} \end{bmatrix} \times \begin{bmatrix} \eta_t \\ \rho b_{t-1} \end{bmatrix} \times \Lambda_{t,t+s} \\ &= (\rho\hat{\pi}_{t-1})^{-1} \times \begin{bmatrix} \tau_\eta & \hat{\pi}_{t-1} \end{bmatrix} \times \begin{bmatrix} \eta_t \\ \rho b_{t-1} \end{bmatrix} \times \Lambda_{t-1,t+s}. \end{aligned}$$

Applying Proposition 6 then yields the formal result.

COROLLARY 4: *For $\theta_t = v_t = 0$, a financial shock η_t is persistent (compared to the fixed-τ_v benchmark) if and only if $\eta_t < -\check{b}_{t-1}/(\tau_\xi + \tau_\eta)$. Otherwise, η_t is non-persistent. Moreover, "scaling up" η_t as in Proposition 5 increases the persistence and non-persistence, respectively.*

5.4. *Summary*

The information-based feedback mechanism underlying our equilibrium creates two types of asymmetries. On the one hand, shocks are either amplified or de-amplified. On the other hand, shocks are also either persistent or non-persistent. In particular, our findings imply that adverse financial shocks are amplified and persistent, while positive financial shocks are de-amplified and non-persistent. Moreover, the underlying mechanisms are highly nonlinear, implying that "scaling up" a shock gives rise to more pronounced (de-)amplification and (non-)persistence, respectively. In consequence, the impact of small shocks is only little amplified and barely persistent, whereas rare adverse shocks virtually destroy the informational capacities of the real sector and thereby induce highly amplified and persistent crashes.

6. Illustration: Impulse responses to financial shocks

In this section, we illustrate our theoretical results using simulated impulse response paths to financial shocks. To highlight the *informational* role of credit constraints, we contrast the impulse responses with counterfactual paths where τ_v is fixed at its steady state level, but credit constraints continue to *constrain* the economy. The only difference between our model and the counterfactual responses is that uncertainty is removed—just as if we were to use the "magic wand" imagined by Blanchard.[31]

6.1. *Impulse responses to financial shocks*

Consider the economy's response to a nonzero realization of noisy asset demand $\tilde{\eta}_t$ and, for simplicity, suppose that the economy is in its steady state prior to the arrival of the shock.[32] From (10),

[31] See the quote in the beginning of the introduction.

[32] As usual, we define the steady state as the situation where true productivity θ_t equals its unconditional expectation and there are no noisy perturbations ($\theta_t = \eta_t = v_t = 0$). Yet, agents are unaware of these realizations and beliefs are formed rationally. Prior expectations are undistorted ($b_{t-1} = 0$) and prior uncertainty is fixed at its stochastic steady state value given the corresponding steady state signal precision (for details see Hamilton, 1994, Ch. 13.5). To streamline the illustration in this section, we also focus on financial noise shocks. For noise that originates in the real sector (i.e., nonzero realizations of $\tilde{v}_t$), similar results hold for $\gamma_1 < \breve{\pi}_{\text{steady state}}/\mathring{\tau}_v + 1$ (see Corollary 3 for details).

$$b_{t+s} = (\tau_\eta/\pi_t)\,\Lambda_{t,t+s} \times \eta_t, \tag{11}$$

where $\{\pi_{t+s}\}$ are recursively defined by $\{\tau_{v,t+s}\}$ and $\{q_{t+s}\}$ solving (9), pinning down all other model variables. In particular, by Lemma 2:

Corollary 5: *If $\eta_t < 0$ (>), then for all $s \geq 0$, b_{t+s} and q_{t+s} are strictly smaller (larger) than their steady state levels in, both, the model and the counterfactual.*

Moreover, by Lemma 1:

Corollary 6: *If $\eta_t < 0$ (>), then for all $s \geq 0$, $\tau_{v,t+s}$ and π_{t+s} are strictly smaller (larger) than their steady state level in the model, but are constant in the counterfactual.*

Because the information-based amplification mechanisms crucially depend on the variability of τ_v, it will be seen that this difference drives a wedge between the model and the counterfactual.

Spillovers to real sector To provide a simple closed form solution for the "real" variables of the model, consider the special case where $v(n_{it}) = n_{it}^2/2$ and $\alpha_1(\chi_t^s) = H^{-1}(\chi_t^s)$. This specification of α_1 ensures that wages will be unperturbed along the impulse response path; i.e., if $v_t = 0$, then $\psi_t = 0$, so that gatherers' labor supply equals farmers' labor demand n_t.[33] From (7), we then have that

$$w_t = \frac{n_t}{1-m} = \left[\frac{q_t \exp(\chi_t^d)}{1-m}\right]^{1/2},$$

along the impulse response path. Because $\chi_t^d = H(\theta_t - \log(q_t))$ is increasing in $\theta_t - \log(q_t)$ with a slope smaller than unity, it follows that w_t and n_t are increasing in q_t. Moreover, field output is given by

$$y_t = m \int_{-\infty}^{\infty} \exp(z)\,(\min\{z, \log(q_t)\} - \log(w_t))\,\mathrm{d}\Phi(\sqrt{\tau_\xi}\,z),$$

along the impulse response path. Substituting w_t, output y_t can be shown to be increasing in q_t, too. That is, tighter financial constraints spill over to the real sector, so that the "real" variables are decreased along the impulse response path, too:

[33] To see this, recall that by Lemma 1, $\tilde{\theta}_t + \tilde{v}_t = \alpha_1(\chi_t^s) + \log(q_t)$. For $v_t = 0$, $\alpha_1 = H^{-1}$ thus implies that $\chi_t^s = H(\theta_t - \log(q_t)) = \chi_t^d$. Hence, $\psi_t = 0$ by (7).

Corollary 7: *If $\eta_t < 0$ (>), then for all $s \geq 0$, w_{t+s}, n_{t+s} and y_{t+s} are strictly smaller (larger) than their steady state levels in, both, the model and the counterfactual.*

Parametrization We set $m = \frac{1}{2}$, implying an equal mass of gatherers and farmers, and set $\rho = 0.98$, $\tau_\epsilon^{-0.5} = \frac{1}{3}$, and, $\tau_\xi^{-0.5} = 2$, corresponding to a persistent and predictable process for the average log-productivity $\tilde{\theta}_t$, with a strong cross-sectional dispersion of $\{\tilde{\theta}_{it}\}$. The standard deviations of market noise, $\tau_\omega^{-0.5}$ and $\tau_\eta^{-0.5}$, are set so that perturbations are high in the financial sector, $\tau_\eta^{-0.5} = 4$, and low in the real sector (i.e., the real sector is the predominant source of information to infer about business conditions). Because $\tau_\omega^{-0.5}$ only matters through $\text{Var}\{\tilde{\chi}_t^d|\chi_t^s\}$, we directly set $\text{Var}\{\tilde{\chi}_t^d|\chi_t^s\} = 4$, avoiding the need to specify $\Psi_{\tilde{\theta}_t,\tilde{q}_t}$. Finally, we use γ_0 and γ_1 to specify the fraction of firms that is constrained in the steady state and the relative amplitude of asset price fluctuations compared to productivity. We set γ_0, so that approximately 2.5 percent of firms are constrained in the steady state ($\gamma_0 = 4$). To emphasize the theoretical results, we pronounce the importance of financial fluctuations by setting γ_1 to 125, implying a relative amplitude that is about two to three times as high as its empirical counterpart.[34]

6.2. *Amplification and persistence of financial crises*

We first illustrate our results on the persistence of financial shocks. From Corollaries 1, 2, 4, and Proposition 5 it follows that:

Corollary 8: *For all $\eta_t \neq 0$ and $s \geq 0$, $\bar{b}_{t+s}$, q_{t+s}, w_{t+s}, n_{t+s}, and y_{t+s} are strictly smaller in the model economy than in the counterfactual. I.e., financial crises are more persistent than in the counterfactual, whereas financial booms are less persistent than in the counterfactual.*

In Figure 6, we plot the responses of the asset price (normalized to $\log(q_{t+s}/q_{\text{steady state}})$), employment n_{t+s}, output y_{t+s}, the fraction of constrained farmers in the economy (i.e., farmers with $n_{it+s} = q_{t+s}/w_{t+s}$), and the endogenous signal precision $\tau_{v,t+s}$ to an adverse (left column) and positive (right column) realization of $\tilde{\eta}_t$. In both columns, we consider a rare tail shock with a magnitude

[34] With more conservative parameters choices for γ_1, we need unrealistically large financial shocks in order to see a notable amplification. In Section 7.3, we illustrate how small levels of risk aversion provide further amplification that substitutes for high values of γ_1, yielding similar responses for realistic values of γ_1.

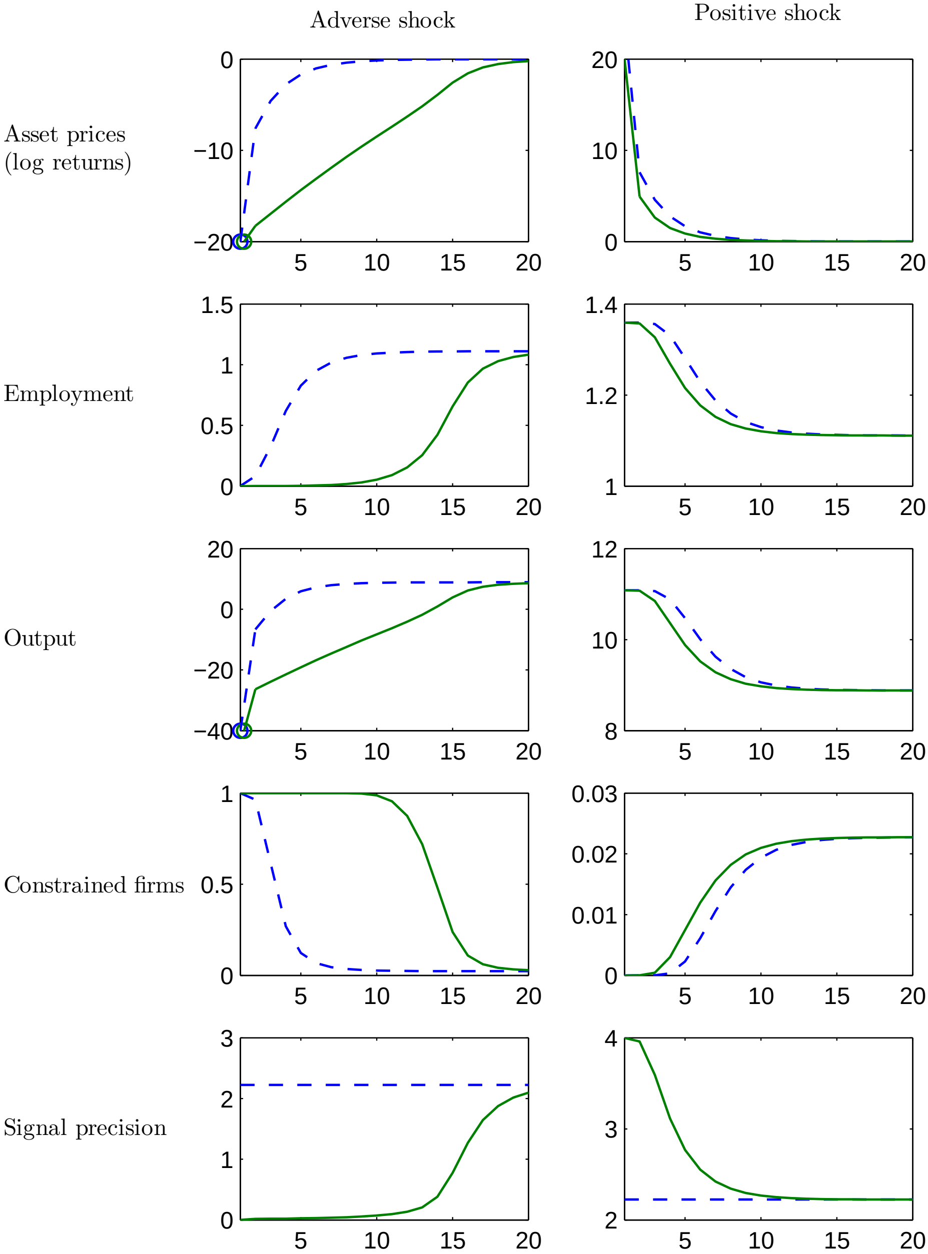

Figure 6. Impulse responses to financial shocks. *Notes:* Solid lines are model responses to shocks. Dashed lines are counterfactual responses in the exogenous uncertainty benchmark. Circles indicate that to retain readability the responses are truncated at the boundary of the graph.

of 2.5 standard errors. With a quarterly interpretation of time, these shocks each correspond to events that occur roughly once every 40 years. As we emphasize in Section 6.4, because of the nonlinearity of our model, the large shock size considered here is crucial for the effects established in Corollary 8 to be economically significant. In all plots, the solid lines correspond to the responses in the model economy and the dashed lines correspond to the fixed-τ_v counterfactual.[35]

Consider first the case of financial crisis in the left column. Qualitatively, asset prices (row 1), employment (row 2), and output (row 3) all decline after a negative financial shock. The causal link between these responses are the financial constraints, implying that an increased fraction of farmers is financially constrained (row 4) after an initial decrease in the asset price, leading to the decline in employment and output. Note that—as known from the financial frictions literature—the contagion of the real sector is present in the counterfactual economy as well. We call this the *constraining* effect of credit constraints.

In addition to this constraining effect, our theoretical results suggest a novel *informational* effect of credit constraints. This effect results from the variations in the signal precision as seen in row 5. By fixing the precision in the counterfactual economy, the difference between our model and the counterfactual exactly amounts to this novel informational effect. As can be seen, this informational effect of credit constraints virtually shuts down the informational capacities of the real sector. This throws the economy in a "pessimism trap" that induces an amplified and highly persistent response to the considered shock. In contrast, removing uncertainty in the counterfactual, agents quickly learn about the noisy character of the crisis, and the crisis largely goes away as has been suggested by Olivier Blanchard.

[35] Note that asset prices and output in the first and third row are truncated at ±20 and 30, respectively, omitting the initial impact of the shock. The initial impact on $\log(q_t)$ amounts to -98 (49) in the model and -63 (63) in the counterfactual in the bust (boom) case. Note that the impact for the endogenous uncertainty model is higher than in the exogenous uncertainty benchmark, reflecting the initial amplification established in Corollary 1. Underlying the strong initial impact in, both, the model and the counterfactual is a dual role of financial shocks. First, a negative realization of noise traders' demand has a direct impact on the asset price in period t and, hence, also a direct impact on economic constraints of the real sector. Second, the perturbation of asset prices also plays an informational role, effecting the beliefs of agents in the economy. Because, the second effect persists over time, while the first effect only applies to period t, this leads to a "discontinuity" between initial impact and the response of the economy starting in period $t+1$. Here, we choose to omit the initial impact to make the graphs more readable.

6.3. De-amplification and non-persistence of financial booms

In contrast, after a positive financial shock (depicted in the right column of Figure 6), the model's response is less persistent than the counterfactual. This is because in financial booms, the real sector aggregates more information, so that agents learn faster about the bullish character of a non-fundamental boom (recall Proposition 6). Importantly, this implies that financial booms have necessarily smaller spillover effects on the real sector than financial crises.[36] (Note the small scale of the y-axis in the right column of Figure 6, when comparing the spillover effects to the left column).

6.4. Convexity of crises

In fact, by Proposition 5 and Corollary 4, the asymmetry between booms and busts generalizes to a non-linearity which, in particular, implies a general convexity of financial crises. The more negative a financial shock, the more information is destroyed and, hence, the more amplified and persistent is the resulting crisis. To illustrate this convexity, consider the half-life of average beliefs in our economy; i.e., the time s it takes until agents in the economy are half as pessimistic as they were at the time the shock hit the economy. In the counterfactual—as in any standard model—shocks are scale-independent, so that the half-life measure is independent of the shock size. That is, scaling up an initial shock leads to a proportional response along the whole response path, so that the *relative* realizations of beliefs across time remain unchanged. In contrast to this, the non-linearity of our model implies that larger shocks have an disproportionately severe effect, resulting in a convexity of financial crises.

COROLLARY 9: *Suppose the economy is in its steady state and let $T_{1/2}(\eta_t)$ denote the half-life of average beliefs in the economy along an impulse response path to a financial shock η_t; i.e., the time s it takes, such that $\bar{b}_{t+r} \leq \bar{b}_t/2$ for all $r \geq s$. Then it holds that $T_{1/2}$ is (weakly) decreasing in η_t.*

[36] Nonlinearities in output and employment are responsible for parts of the asymmetry between the real variables during a financial boom compared to a financial bust. However, the responses for asset prices are necessarily perfectly symmetric in the counterfactual. This implies that the model's responses during a financial boom are bounded above by the responses of the counterfactual that are symmetric to the responses of the counterfactual during a crises. This then further implies that additionally to nonlinearities in output and employment, spillovers caused by the financial market are also necessarily less pronounced during a financial boom than during a bust.

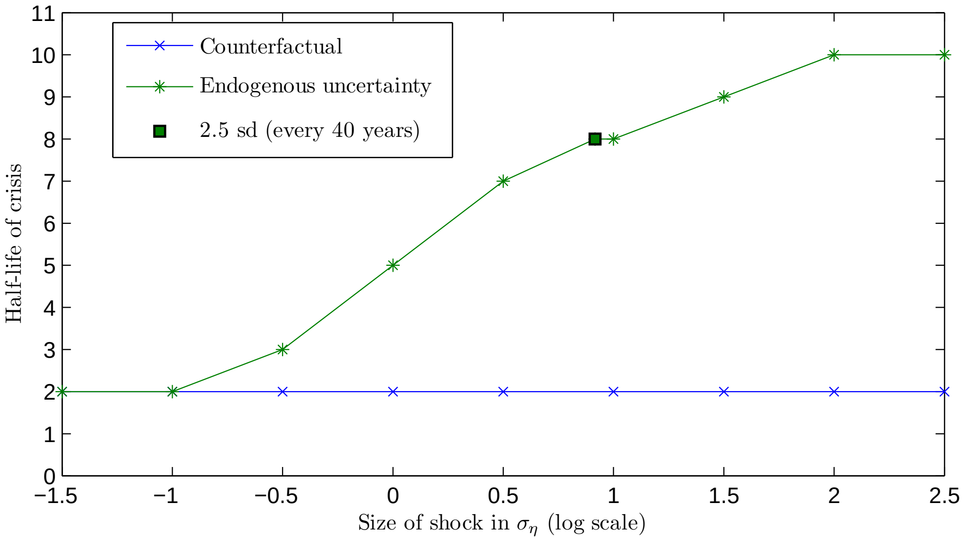

Figure 7. Half-life of adverse financial shocks.

Note that because of the discrete nature of the half-life, it is necessarily locally constant almost everywhere, so that $T_{1/2}$ is only *weakly* decreasing. (Still, from our more general earlier results we know that crises are strictly more persistent as shocks get larger.)

In Figure 7, we plot the half-life in our model economy for negative financial shocks of different sizes. For our baseline simulations with a 2.5 standard error shock, the half-life is 9 quarters, while, in the conterfactual economy, any financial shock leads to a half-life of only 2 quarters.

From Proposition 5 it follows that small financial shocks, which are more frequent, have half-life periods that are similar to the exogenous uncertainty counterfactual, so that for small shocks our model behaves similar to models that only reflect the *constraining* role of credit constraints. However, in the event of a rare negative shock, the feedback loop and pessimism trap studied in Section 5 drive a significant wedge between the predictions of our model and those of the counterfactual.

7. Further empirical predictions

In this section, we point out some further implications of the endogeneity of the information structure. In particular, we show that common proxies of uncertainty, such as the dispersion of beliefs (as can be, for instance, measured by the survey of professional forecasters), risk premia and volatility of financial markets all increase in crisis times.

7.1. Dispersion of beliefs

One "measure of uncertainty" that is commonly used in the empirical literature is the diversity of beliefs in the economy. From Lemma 2 we know that the cross-section of beliefs is normally distributed around $\bar{b}_t$ with standard deviation $\sigma_t = \sqrt{\tau_\xi}/\bar{\pi}_t$. Here, an individual agent's uncertainty $1/\bar{\pi}_t = 1/(\pi_t + \tau_\xi)$ co-moves with the economy's (public) uncertainty $1/\pi_t$. In particular, when the economy is caught in a pessimism trap, agents increasingly refer to their own, private signals, creating a large dispersion in beliefs. On the other hand, during financial booms, public information becomes more valuable compared to private information, which reduces the dispersion of beliefs. In sum, opinions are *aligned* in booms and *dispersed* in crises.

COROLLARY 10: *If $\eta_t < 0$ (>), then for all $s \geq 0$, the cross-sectional dispersion of beliefs $\sigma_{t+s} = \sqrt{\tau_\xi}/(\pi_{t+s} + \tau_\xi)$ is strictly larger (smaller) than its steady state level in the model, but is constant in the counterfactual.*

7.2. Stochastic volatility

Another commonly used measure for uncertainty is the volatility of asset prices. During credit crises, high uncertainty induces volatile asset market behavior, as seen from the perspective of an outside observer. The reason is that, as the real sector becomes less informative, agents place more weight on signals from the financial market. This means that the asset price is more exposed to financial market noise and thus subject to larger conditional volatility.[37]

[37] For technical reasons, we here focus on the conditional volatility of asset prices; i.e., the volatility of asset prices that is induced by noise originating in the financial market. Computing the unconditional volatility would require us to integrate out the distribution of noise in the labor market. While our quasi-Gaussian transformation is aimed at making the updating problem of agents tractable, there is no analytical solution available for computing the actual

Corollary 11: *$Var\{\log(\tilde{q}_t)|\Omega_t \setminus \eta_t\}$ is decreasing in π_t. Hence, if $\eta_t < 0$ (>), then for all $s \geq 0$, the volatility $Var\{\log(\tilde{q}_t)|\Omega_t \setminus \eta_t\}$ is strictly larger (smaller) than its steady state level in the model, but is constant in the counterfactual.*

7.3. Risk premium

Yet another natural consequence of high uncertainty in financial crises are large risk premia.

Although agents in our model are risk-neutral with respect to $\exp(\tilde{\theta}_t)$, we can simulate any risk-attitude towards $\tilde{\theta}_t$ by shifting the realization of noisy asset demand shocks $\tilde{\eta}_t$ by a constant μ. In the baseline model, we use this feature to clear the model from any risk effects on the asset price, so that the price only reflects the first moment of agents' expectations (see also Footnote 18). Here, we apply this feature to obtain an equilibrium asset price which behaves as if agents exhibited risk preferences.

More specifically, suppose that we set μ to $\mu' + r/\tau_\xi$, where μ' denotes our baseline choice of μ that induces risk-neutral behavior, and $r > 0$. Then, the equilibrium asset price q_t reflects a risk-averse attitude towards $\tilde{\theta}_t$. In particular, one can show that in this case, all previous results hold exactly, except that q_t as defined by (8) is replaced by q_t^r, which relates to q_t as follows:

$$R_t^{-1} \equiv \frac{q_t^r}{q_t} = \exp\{-r\gamma_1\pi_t^{-1}\},$$

where R_t is the risk premium associated with each unit of the Lucas tree that is traded in t. Given this specification, Corollary 6 immediately implies that the risk premium is increased along the impulse response path.

Corollary 12: *Suppose $r > 0$ and $\eta_t < 0$ (>). Then for all $s \geq 0$, the risk premium R_t is strictly larger (smaller) than its steady state level in the model, but is constant in the counterfactual.*

Figure 8 re-plots the economy's response to an adverse financial shock of 2.5 standard errors for $r = \frac{1}{2}$. In row 1, we plot asset prices (again, normalized as log returns) and the corresponding risk premium (in logs). The response of the risk premium is hump-shaped. This is because the loss of information that results from tighter constraints on the real sector (see third row) slowly increases the posterior uncertainty. Accordingly, log assets returns in the first

---

distribution of labor market noise (see also the discussion around Lemma 1).

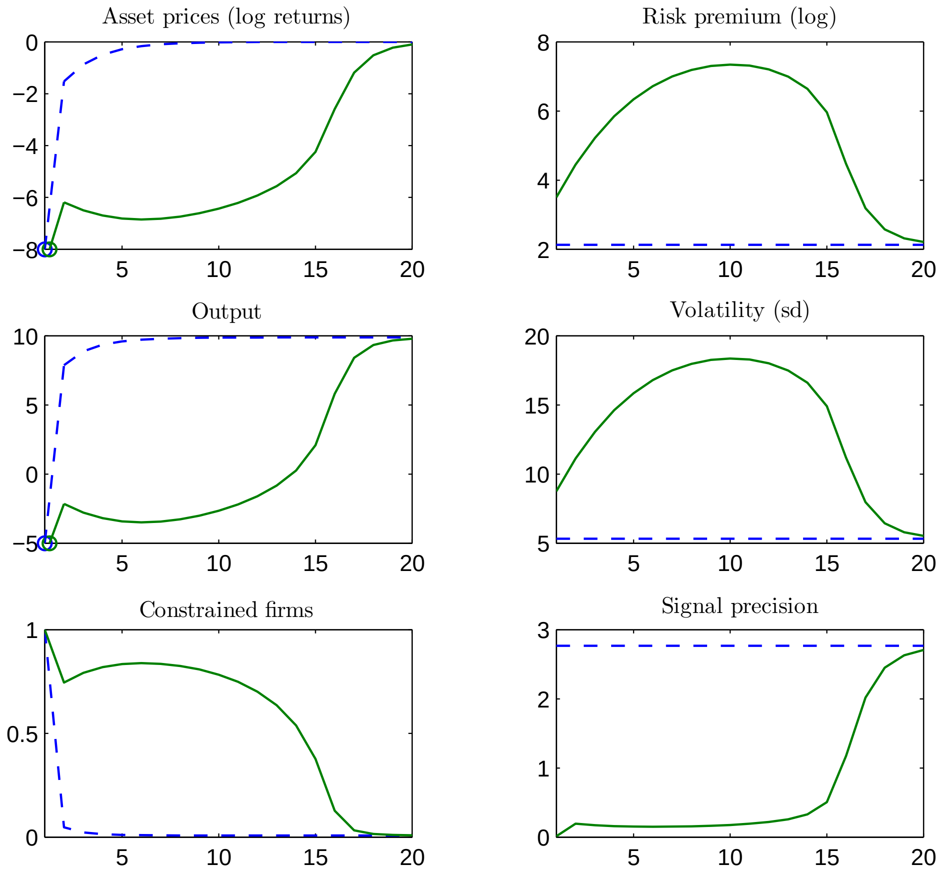

Figure 8. Volatility and risk premium response to an adverse financial shock.

periods after impact mainly reflect the pessimism that is induced through learning. However, while in the case without risk aversion, prices monotonically increase as pessimism ebbs away, prices now further reflect the general uncertainty that characterizes a credit crisis in our model. Accordingly, as uncertainty increases along the crisis path, asset prices are more and more repressed by this increased uncertainty.

This additional downward pressure of prices tends to tighten financial constraints on the real sector even further and, therefore, introduces additional amplification. Note that in our numerical simulation this additional amplification is somewhat obscured by a change in the model parameters.[38]

[38] Because of the additional amplification due to risk-aversion, responses in the baseline

8. Concluding remarks

This chapter of my dissertation proposes a novel mechanism that restricts the ability of agents to learn from the real sector during financial crises. Incorporating this idea into a dynamic macroeconomic model with a financial sector, we show that the transmission of financial shocks is inherently asymmetric and nonlinear. While fluctuations on the financial market have only little impact on the economy during "normal times", unusually adverse shocks destroy the informational capacities of the economy and therefore lead to disproportionately severe and persistent crises.

At a methodological level, we show that a combination of informational and financial frictions gives rise to a nonlinear signal structure, which explains why learning is less efficient in crisis times. Specifically, we establish that learning from "concave" signals leads to higher posterior uncertainty whenever the signal realizes in "flatter" regions. In a general theorem, this is shown to hold for a large class of information structures and to hold independently of the specific properties of our model. Equipped with these results, we then further demonstrate how learning from nonlinear signals can be incorporated within an analytically tractable conjugate Gaussian framework.

Going beyond our main results, our model also provides a number of further predictions that are accessible to an empirical verification. In particular, we show that both uncertainty and also common empirical proxies of uncertainty (such as the dispersion of beliefs, the volatility of asset prices, and risk premia) increase during financial crises.

While these predictions are in line with conventional wisdom and stylized facts, a systematic empirical analysis on the causal links is an important direction for future research. Another promising road is to directly examine how the informational capacity of the economy changes across different states of the world. In particular, applying the empirical methods recently developed by Coibion and Gorodnichenko (2012), a natural investigation suggested by our model is to examine the impact of credit conditions on the persistence of beliefs.

---

parametrization lead to a crisis that spans more than a century. Here, we therefore reduce the value of γ_1 to 30, implying a relative amplitude of asset price fluctuations compared to productivity that is roughly in line with its empirical counterpart. We also set $\gamma_0 = 7$ in order to target again a fraction of 2.5 percent of farmers who are constrained in the steady state.

A. Mathematical appendix

A.1. Proof of Proposition 1

As $\bar{q} \to \infty$, $x_{it} = \exp(\theta_{it})/w_t$ and hence $\tilde{\chi}_t^d = \log\{\int_{-\infty}^{\infty} \exp(z)\, \mathrm{d}\Phi(\sqrt{\tau_\xi}(z - \tilde{\theta}_t))\} - \log(\bar{q}) = \tilde{\theta}_t + 1/(2\tau_\xi) - \log(\bar{q})$. Therefore, $\mathrm{Var}\{\tilde{\theta}_t|w_t\} = \mathrm{Var}\{\tilde{\chi}_t^d|w_t\} = \mathrm{Var}\{\tilde{\chi}_t^d|\chi_t^s\}$, where the last equality follows since $\tilde{\chi}_t^s$ is a monotone transformation of $\tilde{w}_t$. □

A.2. Proof of Theorem 1

We separate the proof into two steps. Lemma 3 establishes that $\mathrm{Var}\{\tilde{\theta}|s\} = \alpha_2(s)^2\,\mathrm{Var}\{\tilde{X}\}$ is increasing in s if we are in the case where $\tilde{\theta}|s = \alpha_1(s) + \alpha_2(s)\tilde{X}$ and where f is concave (the case where f is convex follows analogue). Lemma 5 establishes the corresponding results for the case where $f(\tilde{\theta})|s$ is ordered by the MLRP.

A.2.1. Location-scale distributions under concave transformations

Lemma 3: *Let $g : \mathbb{R} \to \mathbb{R}$ be a differentiable, (strictly) increasing, and (strictly) concave function and $\tilde{X}$ a square-integrable, non-degenerate random variable with mean zero over $\mathbb{R}$. Then, for any positive number $v > 0$, there exists a unique, differentiable function $\alpha_{21} : \mathbb{R} \to \mathbb{R}_{++}$ such that*

$$Var\{g(\alpha_1 + \sqrt{\alpha_{21}(\alpha_1)}\tilde{X})\} = v \qquad \forall \alpha_1 \in \mathbb{R}.$$

Moreover, α_{21} is (strictly) increasing. When $g = H$, α_{21} has limits $\lim_{\alpha_1 \to -\infty} \alpha_{21}(\alpha_1) = (Var\{\tilde{X}\})^{-1}v$ *and* $\lim_{\alpha_1 \to \infty} \alpha_{21}(\alpha_1) = \infty$.

Proof. We only show the "strict" version of the lemma. Define $G(\alpha_1, \alpha_2) = \mathrm{Var}\{g(\alpha_1 + \sqrt{\alpha_2}\tilde{X}\} - v$. G is clearly differentiable in both arguments. We start by establishing that $G_{\alpha_1} < 0$ and $G_{\alpha_2} > 0$.

Suppose $\check{\alpha}_1 < \hat{\alpha}_1 \in \mathbb{R}$. Define $\check{g}(x) = g(\check{\alpha}_1 + x\sqrt{\alpha_2})$ and $\hat{g}(x)$ analogously. The function $\check{g}$ induces the density $F_{\check{g}(X)}(y) = F_X(\check{g}^{-1}(y))$ of $\check{g}(\tilde{X})$. Similarly we find $F_{\hat{g}(X)}(y)$. This means, $F_{\check{g}(X)}(y) = F_{\hat{g}(X)}(k(y))$, where it is straightforward to check that $k(y) \equiv \hat{g}(\check{g}^{-1}(y))$ is a differentiable contraction mapping with $k(y_0) = \mathrm{E}\{\hat{g}(X)\}$ for some y_0. In particular, k satisfies the condition in Lemma 4 and thus $G(\check{\alpha}_1, \alpha_2) > G(\hat{\alpha}_1, \alpha_2)$. This proves $G_{\alpha_1} < 0$.

For $G_{\alpha_2} > 0$, take $\check{\alpha}_2 < \hat{\alpha}_2 \in \mathbb{R}$ and define $\check{g}(x) = g(\alpha_1 + x\sqrt{\check{\alpha}_2})$ and $\hat{g}(x)$ analogously. Proceeding as before, we find that $F_{\hat{g}(X)}(y) = F_{\check{g}(X)}(k(y))$ with $k(y) = g(\alpha g^{-1}(y))$, $\alpha = \sqrt{\check{\alpha}_2/\hat{\alpha}_2} < 1$. Here, k need not be a contraction mapping in general. Still, for $y \le y^*$ with $k(y^*) = y^*$, $k'(y) \le \alpha < 1$ and $k(y) \le y$ for $y > y^*$. Also, there exists a (unique) y_0 such that $k(y_0) = \mathrm{E}\{\check{g}(X)\} < g(0) = k(y^*)$. Using these two properties, the following inequalities hold,

$$|k(y) - k(y_0)| = |k(y) - k(y^*)| + |k(y^*) - k(y_0)| \le |y - y^*| + \alpha|y^* - y_0| < |y - y_0|$$

if $y > y^*$, and

$$|k(y) - k(y_0)| \le \alpha|y - y_0| \le |y - y_0|$$

if $y \le y^*$, ensuring that Lemma 4 is applicable. Consequently, $G(\alpha_1, \hat{\alpha}_2) > G(\alpha_1, \check{\alpha}_2)$ and $G_{\alpha_2} > 0$.

Now turn to the existence of α_{21}. For large α_2, $G(\alpha_{1,}\alpha_2)$ goes to infinity as can be seen by the inequality

$$\mathrm{Var}\{g(\alpha_1 + \sqrt{\alpha_2}\tilde{X}\} \ge \mathrm{Var}\{g'(\alpha_1)\sqrt{\alpha_2}\min\{\tilde{X}, 0\}\} = \alpha_2 \cdot const.$$

Also, $G(\alpha_1, 0) = -v < 0$, that is, by continuity of G and its monotonicity in α_2, a unique $\alpha_2 \equiv \alpha_{21}(\alpha_1) \in \mathbb{R}_{++}$ must exist, with $G(\alpha_1, \alpha_{21}(\alpha_1)) = 0$. By the implicit function theorem, α_{21} is continuous and differentiable with $\alpha_{21}'(\alpha_1) = -G_{\alpha_1}/G_{\alpha_2} > 0$. When $g = H$, it becomes almost linear for large negative values of α_1, i.e.

$$\begin{aligned} v = \lim_{\alpha_1 \to -\infty} \mathrm{Var}\{H(\alpha_1 - \sqrt{\alpha_{21}(\alpha_1)}\tilde{X})\} &= \lim_{\alpha_1 \to -\infty} \mathrm{Var}\{\alpha_1 - H(\alpha_1 - \sqrt{\alpha_{21}(\alpha_1)}\tilde{X})\} \\ &= \mathrm{Var}\{\sqrt{\lim_{\alpha_1 \to -\infty} \alpha_{21}(\alpha_1)}\tilde{X}\} \\ &= \lim_{\alpha_1 \to -\infty} \alpha_{21}(\alpha_1)\mathrm{Var}\tilde{X} \end{aligned}$$

and thus $\lim_{\alpha_1 \to -\infty} \alpha_{21}(\alpha_1) = (\mathrm{Var}\tilde{X})^{-1}v$. Similarly, since $\lim_{\alpha_1 \to \infty} G(\alpha_1, \alpha_2) = -v$ for any level of α_2, $\lim_{\alpha_1 \to \infty} \alpha_{21}(\alpha_1) = \infty$. ◇

The main rationale behind the previous proof was to compare the variances of a random variable $\tilde{X}$ (in our case, this was $\check{g}(\tilde{X})$ or $\hat{g}(\tilde{X})$) and its transform $k(\tilde{X})$ with k satisfying some regularity conditions. When these conditions are merely that k be a contraction mapping, the result is obvious and most textbooks (see for example,) mention it. In our case, however, a (much) more general result is needed since the second set of regularity conditions we derive in Lemma 3 clearly allow for functions with slopes larger than one.

LEMMA 4: *Let $\tilde{X}, \tilde{Y}$ be two random variables over an interval $I \subseteq \mathbb{R}$ with cumulative distributions functions F_X, F_Y, and $k : I \to \mathbb{R}$ be an increasing function such that $F_Y(x) = F_X(k(x))$ and $|k(x) - \mathrm{E}\{\tilde{X}\})| \leq |x - x_0|$ for all $x \in I$ and a given $x_0 \in I$. Then, $Var\{\tilde{X}\} \leq Var\{\tilde{Y}\}$. If $|k(x) - \mathrm{E}\{\tilde{X}\}| < |x - x_0|$ somewhere in the support of $\tilde{X}$ the inequality is strict.*

Proof. We only show the "strict" version of the lemma. Without loss of generality we may assume $\mathrm{E}\{\tilde{X}\} = 0$, $I = \mathbb{R}$ (proof is analogous for any other interval) and $x_0 = 0$ ($\mathrm{Var}\{\tilde{Y}\}$ is invariant under shifts $x \mapsto x + x_0$). First note, that we can restrict our attention to functions $k = k_+$ with $k_+(x) = x$ for all $x \leq 0$. This is the case since any k with $k(0) = 0$ can be split up into two functions, $k_-$and k_+, which are just the identity on the positive (negative) side of zero and k on the other. We recover the result for general k by constructing an intermediate random variable $\tilde{Y}_+$with cdf $F_{Y_+}(x) = F_X(k_+(x))$. The intermediate random variable $\tilde{Y}$ has then mean $\mathrm{E}\{\tilde{Y}\} \geq 0 = \mathrm{E}\{\tilde{X}\}$ and variance $\mathrm{Var}\{\tilde{Y}\} \geq \mathrm{Var}\{\tilde{X}\}$. Now, $F_Y(x) = F_X(k(x)) = F_X(k_+(k_-(x))) = F_{Y_+}(k_-(x))$, and thus $F_{-Y}(x) = F_{-Y_+}(-k_-(-x))$. The function $-k_-(-\cdot)$ is the identity on the negative side of zero and the mean of $-\tilde{Y}$ is also negative. We can use the "reduced" result shown below and get $\mathrm{Var}\{\tilde{Y}\} = \mathrm{Var}\{-\tilde{Y}\} \geq \mathrm{Var}\{-\tilde{Y}_+\} = \mathrm{Var}\{\tilde{Y}_+\} \geq \mathrm{Var}\{\tilde{X}\}$.

Second, it is sufficient to prove the result for functions k_+ which are not only the identity for negative values of x but also above a certain (possibly large) level $M > 0$. This is without loss of generality by a standard limit argument.

Third, even simpler functions may be used to accomplish the proof, namely functions $k_{z,c}$ of the form

$$k_{z,c}(x) = \begin{cases} x & x > z + c \\ z & z < x \leq z + c \\ x & x \leq z \end{cases} \qquad z > 0, c > 0.$$

These functions only slightly differ from the identity but still they are very effective in that they are the building blocks of more general functions k_+. More precisely, any $k_+$with an upper bound of $M > 0$ can be decomposed as follows,

$$k_+ = \lim_{n\to\infty} k_{k_+(Nh),Nh-k_+(Nh)} \circ \ldots \circ k_{k_+(2h),2h-k_+(2h)} \circ k_{k_+(h),h-k_+(h)}, \qquad (12)$$

where $h \equiv M/N$, $N = 2^n$, and where the limit holds under the sup-norm. Hence fix $z > 0$ and consider the random variable $\tilde{Y}_c$ given by cdf $F_{Y_c}(x) = F_X(k_{z,c}(x))$. We would like to show that $g(c) \equiv \mathrm{Var}\{\tilde{Y}_c\} - \mathrm{Var}\{\tilde{X}\}$ is strictly larger than zero for $c > 0$. Clearly $g(0) = 0$. Straightforward computation shows that

$$g'(c) = 2(z+c)(\Delta - \Delta^2) + 2\Delta \int_z^{z+c} x \,\mathrm{d}F_X(x) \geq 0, \tag{13}$$

where $\Delta = F_X(z+c) - F_X(z) \geq 0$. Given our assumption that $|k(x) - \mathrm{E}\{\tilde{X}\}| < |x - x_0|$ somewhere in the support of $\tilde{X}$, the approximation in (12) will always involve functions $k_{z,c}$ such that $\Delta > 0$ and (13) is strict.[39] Therefore, $\mathrm{Var}\{\tilde{X}\} < \mathrm{Var}\{\tilde{Y}\}$. ◇

This concludes the proof of the theorem for the case where $\tilde{\theta}|s$ belongs to a location-scale family of distributions. □

A.2.2. MLRP distributions under concave transformations

LEMMA 5: *Let $I \subseteq \mathbb{R}$ be a nonempty (and possibly unbounded) real interval, let X_1, X_2 be two random variables over I which exhibit the monotone likelihood ratio property, i.e. f_2/f_1 is increasing with f_i being the (possibly degenerate) density of X_i, and let g be an increasing continuous function defined on I.*

1. *$Var\{g(X_1)\} \leq Var\{g(X_2)\}$ if g is convex and $Var\{X_2\} \leq Var\{X_1\} < \infty$.*
2. *$Var\{g(X_1)\} \geq Var\{g(X_2)\}$ if g is concave and $Var\{X_1\} \leq Var\{X_2\} < \infty$.*

In both cases, the inequality is strict whenever f_2/f_1 is strictly increasing and g is strictly convex or concave somewhere in the support of f_2.

Proof. We restrict ourselves to $I = [a, b]$ and distributions of X_1, X_2 with finite support. It is straightforward to generalize the result first to continuous distributions and then to arbitrary, possibly unbounded intervals. Moreover, it is sufficient to prove the result solely for increasing convex functions g since all other cases can be reduced to this case by flipping g either vertically or horizontally. We will prove the result for functions g of the form

$$g_{z,c}(x) = x + (c-1)(x-z)\,\mathbb{1}_{\{x \geq z\}} \quad \forall z \in \mathbb{R}, c \geq 1.$$

[39] Because $\mathrm{Var}\{\tilde{Y}\}$ is monotonic in k—the closer k is to the identity, the closer $\mathrm{Var}\{\tilde{Y}\}$ is to $\mathrm{Var}\{\tilde{X}\}$—the "strictness" does not vanish in the limit of (12). It rather becomes larger with every increase in n.

The general case follows by iteration,

$$g(x) = \lim_{N\to\infty} \left(g_{g(z_N),g'(z_N)/g'(z_{N-1})} \circ \ldots \circ g_{g(z_1),g'(z_1)/g'(z_0)} \right) \times \\ \times \left(g'(z_0)(x - z_0) + g(z_0) \right),$$

where $z_i = a + i(b-a)/N$. To have MLRP well-defined for pairs of discrete distributions, assume X_1 and X_2 share a common support $\{x_1 < x_2 < \ldots < x_N\}$, $N \in \mathbb{N}$, and assign probability weights (p_i) and (q_i) to the respective nodes. MLRP then translates to

$$0 \le \frac{q_1}{p_1} \le \frac{q_2}{p_2} \le \ldots \le \frac{q_N}{p_N} \le \infty.$$

Now define $H : \mathbb{R}^N \to \mathbb{R}$ by

$$H(y_1, \ldots, y_N) = \sum (q_i - p_i) y_i^2 - \left(\sum q_i y_i\right)^2 + \left(\sum p_i y_i\right)^2.$$

Note that $H((x_i)) = \mathrm{Var}\{X_2\} - \mathrm{Var}\{X_1\} \ge 0$. By regarding the first two derivatives with respect to c we will now show that

$$h(c) \equiv H((g_{z,(c+1)}(x_i))_i) \ge 0 \quad \forall z \in \mathbb{R}, c \ge 0.$$

Assume $z \in [x_{j-1}, x_j)$. Then, $h(c)$ is a quadratic polynomial in $c \ge 0$ with the following two derivatives at $c = 0$,

$$\begin{aligned} h'(0) &= 2\sum_{i=j}^{N} q_i (x_i - z)(x_i - \mu_2) - 2\sum_{i=j}^{N} p_i (x_i - z)(x_i - \mu_1) \\ &\equiv 2B_z - 2A_z \end{aligned} \tag{14}$$

and

$$\begin{aligned} h''(0) = 2\left[\sum_{i=j}^{N} q_i (x_i - z)^2 - \left(\sum_{i=j}^{N} q_i (x_i - z)\right)^2\right] \\ - 2\left[\sum_{i=j}^{N} p_i (x_i - z)^2 - \left(\sum_{i=j}^{N} p_i (x_i - z)\right)^2\right] \\ \equiv 2D_z - 2C_z. \end{aligned} \tag{15}$$

If we can show that both expressions are nonnegative, we are done since $h(0) = \mathrm{Var}\{X_2\} - \mathrm{Var}\{X_1\} \ge 0$ by assumption.

1ST DERIVATIVE. First note that B_z and A_z are continuous and piecewise linear in $z \in [x_1, x_N]$ with $B_{x_1} - A_{x_1} \geq 0$ and $B_{x_N} - A_{x_N} = 0$. We will show that $B_z - A_z$ is single-peaked or quasi-concave on $[x_1, x_N]$, meaning it is impossible for $B_z - A_z$ to first drop below 0 and then rise again to 0 at the right boundary x_N. More precisely, we claim that if the derivative $B'_z - A'_z$ is nonpositive at some point z, it stays nonpositive thereafter, preventing u-shaped behavior.

Suppose $z \in [x_{j-1}, x_j)$ and $B'_z \leq A'_z$, i.e.

$$-B'_z = \sum_{i \geq j} q_i(x_i - \mu_2) \geq \sum_{i \geq j} p_i(x_i - \mu_1) = -A'_z. \tag{16}$$

It is sufficient to show that $B'_z \leq A'_z$ also in the next interval $[x_j, x_{j+1})$. The rest follows by iteration. We distinguish between three cases. First, assume $q_j(x_j - \mu_2) \leq p_j(x_j - \mu_1)$. Then, by omitting the terms with $i = j$ on both sides of (16) we only increase the inequality and trivially get that $-B'_z \geq -A'_z$ for all $z \in [x_j, x_{j+1})$. Note that we are automatically in the first case whenever $\mu_2 \geq x_j \geq \mu_1$. Second, assume $q_j(x_j - \mu_2) > p_j(x_j - \mu_1)$ and $x_j > \mu_2 \geq \mu_1$. Thus, using MLRP, we have for all $i > j$,

$$\frac{q_i}{p_i} \geq \frac{q_j}{p_j} \geq \frac{x_j - \mu_1}{x_j - \mu_2} > \frac{x_i - \mu_1}{x_i - \mu_2},$$

where the last inequality follows since $x \mapsto (x - \mu_1)/(x - \mu_2)$ is decreasing when $\mu_2 > \mu_1$ (which follows from MLRP). By $x_i > x_j > \mu_2$ this is easily rearranged to $q_i(x_i - \mu_2) > p_i(x_i - \mu_1)$ for all $i > j$. After summing over all $i > j$ we obtain for $z \in [x_j, x_{j+1})$,

$$-B'_z = \sum_{i > j} q_i(x_i - \mu_2) > \sum_{i > j} p_i(x_i - \mu_1) = -A'_z.$$

Third, assume $q_j(x_j - \mu_2) > p_j(x_j - \mu_1)$ and $\mu_2 \geq \mu_1 > x_j$. In a fashion similar to before we find that for all $i < j$

$$\frac{q_i}{p_i} \leq \frac{q_j}{p_j} \leq \frac{\mu_1 - x_j}{\mu_2 - x_j} < \frac{\mu_1 - x_i}{\mu_2 - x_i},$$

and so $q_i(\mu_2 - x_i) < p_i(\mu_1 - x_i)$ for all $i < j$. Using $\sum_{i=1}^N q_i(x_i - \mu_2) = 0$ we see that for $z \in [x_{j-1}, x_j)$,

$$-B'_z = \sum_{i \geq j} q_i(x_i - \mu_2) = \sum_{i < j} q_i(\mu_2 - x_i) < \sum_{i < j} p_i(\mu_1 - x_i) = -A'_z,$$

contradicting our assumption. The third case is therefore not possible given the assumption. This concludes the proof that $h'(0) \geq 0$ for all possible z.

2ND DERIVATIVE. Again, note that C_z and D_z are both continuous in $z \in [x_1, x_N]$. However, other than before, these two functions are no longer piecewise linear but piecewise quadratic and therefore require a more subtle treatment. First, note that $D_z - C_z$ is quasi-concave on each sub-interval $[x_{j-1}, x_j]$. Suppose this did not hold. This means, the constant second derivative $D''_z - C''_z$ must be positive,

$$D''_z/2 = F_2(x_{j-1})(1 - F_2(x_{j-1})) > F_1(x_{j-1})(1 - F_1(x_{j-1})) = C''_z/2, \tag{17}$$

where F_i denotes the cumulative distribution function of X_i. At the same time, the first derivatives $D'_z - C'_z$ at the left and right boundaries must be negative and positive, respectively. Let us regard the right boundary. A positive first derivative implies,

$$\lim_{z \nearrow x_j} D'_z/2 = -F_2(x_{j-1}) \sum_{i \geq j} q_i(x_i - x_j) > -F_1(x_{j-1}) \sum_{i \geq j} p_i(x_i - x_j) = \lim_{z \nearrow x_j} C'_z/2. \tag{18}$$

It is a well-known feature of MLRP that the corresponding conditional distributions $(X_i - x_j)|(X_i \geq x_j)$ satisfy MLRP again. Necessarily, the conditional means must be ordered again,[40]

$$(1 - F_2(x_{j-1}))^{-1} \sum_{i \geq j} q_i(x_i - x_j) \geq (1 - F_1(x_{j-1}))^{-1} \sum_{i \geq j} p_i(x_i - x_j).$$

Now, we multiply this nonnegative inequality with (17) to obtain,

$$F_2(x_{j-1}) \sum_{i \geq j} q_i(x_i - x_j) \geq F_1(x_{j-1}) \sum_{i \geq j} p_i(x_i - x_j),$$

contradicting (18) and establishing quasi-concavity of $D_z - C_z$ on each interval $[x_{j-1}, x_j)$.

The quasi-concavity on the intervals allows us to restrict our attention to z's that lie on the interval boundaries, i.e. it is sufficient to prove $D_j \geq C_j$ where, with slight abuse of notation, we write D_j for D_{x_j} and similarly for C_j. Still, $D_1 \geq C_1$ and $D_N = C_N$.

We will now show the claim $D_j \geq C_j$ by induction over N. It is trivial for $N = 2$. In the following, suppose it holds for $N - 1$ points in the support of X_1 and X_2. To show it for N points, assume the contrary is true, namely there exists a j such that $C_j > D_j$ while at the same time $\text{Var}\{X_2\} \geq \text{Var}\{X_1\}$. We will try to

[40] Note that $1 - F_2(x_{j-1})$ is precisely $\sum_{i \geq j} q_i$ and the same holds for F_1.

increase $\text{Var}\{X_1\}$ relative to $\text{Var}\{X_2\}$ as much as possible but eventually see that it is not possible that the former exceeds the latter.

Define $\delta = x_2 - x_1$ and rewrite

$$C_1 = \text{Var}\{X_1\} = \delta^2(p_1 - p_1^2) + 2\delta p_1 s_1 + C_1^0,$$

where $s_1 = \sum_{i>3} p_i(x_i - x_2)$. C_1^0 is the variance of X_1 when we collapse x_1 and x_2 by reducing their distance δ to zero. Similar results emerge for $D_1 = \text{Var}\{X_2\}$, s_2 and D_1^0. By the induction hypothesis we must have $C_1^0 > D_1^0$, otherwise D_j would have to be larger than C_j. Note that s_1 and s_2 are independent of the actual levels of p_1 and p_2. From Lemma 6 we can infer that $p_1(1 - p_1) \geq q_1(1 - q_1)$ (otherwise we already have our contradiction $D_j \geq C_j$), thus for C_1 to be possibly larger than D_1, we would certainly need $s_2 > s_1$.

Now, let us study how these expressions change if we shift mass from x_2 to x_1. Since p_2 does not enter C_1 directly,[41] a mass increase towards x_1 by one percent yields

$$p_1 \frac{\partial C_1}{\partial p_1} = C_1 - C_1^0 - \delta^2 p_1^2 \geq 0 \tag{19}$$

$$q_1 \frac{\partial D_1}{\partial q_1} = D_1 - D_1^0 - \delta^2 q_1^2. \tag{20}$$

Here, the positivity of the bottom derivative relies on the fact that $q_1 \leq 1/2$ (which is true if $p_1 - p_1^2 \geq q_1 - q_1^2$ and $p_1 \geq q_1$) and therefore $\delta^2 q_1(1 - 2q_1) \geq 0$. In virtue of the positivity, we increase q_1 to make D_1 as large as possible without violating MLRP, i.e. we set

$$q_1 = p_1\chi \equiv p_1 \frac{1 - \sum_{i>2} q_i}{1 - \sum_{i>2} p_i}.$$

The ratio χ here guarantees that $p_1/q_1 = p_2/q_2$. In other words, we used the positivity of (19) to reduce the two degrees of freedom of the two mass shifts to one. Precisely by percentage shifts up or down, we can now control how much weight *both* distributions lay on x_1, thereby keeping our equated MLRP condition $p_1/q_1 = p_2/q_2$ intact. The overall effect of these remaining percentage shifts on $D_1 - C_1$ is a comparison of (19) and (20).

First, assume $p_1 \partial C_1/\partial p_1 \geq q_1 \partial D_1/\partial q_1$. Then, adding the two inequalities $C_1^0 > D_1^0$ and $\delta^2 p_1^2 \geq \delta^2 q_1^2$ to this yields $C_1 > D_1$, a contradiction to our assumption

[41] Notice that C_1^0 only the depends on the sum $p_1 + p_2$ and is invariant under mass shifts from one to the other.

$\mathrm{Var}\{X_2\} \geq \mathrm{Var}\{X_1\}$. Now, suppose $p_1 \partial C_1/\partial p_1 < q_1 \partial D_1/\partial q_1$. This means, we increase $D_1 - C_1$ by moving more weight from x_2 to x_1 until there is nothing left at x_2, i.e. $p_2 = q_1 = 0$. This is equivalent to omitting x_2, so, again, we know by the induction hypothesis that $D_1 \geq C_1$, another contradiction. In sum, we have shown that $D_j \geq C_j$ for all j whenever $\mathrm{Var}\{X_2\} \geq \mathrm{Var}\{X_1\}$, completing the proof. ◇

Lemma 6: *Let $(p_i)_{1\leq i\leq N}$, $(q_i)_{1\leq i\leq N}$, $(C_j)_{2\leq j\leq N}$, and $(D_j)_{2\leq j\leq N}$ be specified as above.*
If $p_1 - p_1^2 \leq q_1 - q_1^2$ then $D_j \geq C_j$ for all $j \geq 2$ and $N \geq 2$.

Proof. Note that, since $p_1 \geq q_1$ by MLRP, p_1 must be larger than 1/2. In particular, for any $\ell \geq 1$,

$$F_1(x_\ell)(1 - F_1(x_\ell)) \leq F_2(x_\ell)(1 - F_2(x_\ell)). \tag{21}$$

If $F_2(x_\ell) \geq 1/2$, this immediately holds since $F_1(x_\ell) \geq F_2(x_\ell)$ by MLRP so both $F_2(x_\ell)$ and $F_2(x_\ell)$ are in the decreasing region of $x \mapsto x(1-x)$. Now suppose $F_2(x_\ell) < 1/2$. We know that $1/2 \geq F_2(x_\ell) \geq q_1 \geq 1 - p_1$ and all three are in the increasing region, i.e. $F_2(x_\ell)(1-F_2(x_\ell)) \geq p_1(1-p_1) \geq F_1(x_\ell)(1-F_1(x_\ell))$, where the last inequality follows from the fact that $F_1(x_\ell) \geq p_1 \geq 1/2$, so both are in the decreasing region. This establishes (21).

The proof itself works by induction over N. The result is immediate if $N = 2$. Assume it holds for distributions with a support of $N-1$ points. We define $\delta = x_N - x_{N-1}$ and rewrite

$$C_j = \delta^2 p_N(1 - p_N) + 2\delta p_N r_1 + C_j^0,$$

where $r_1 = \left(\sum_{j<i<N} p_i(x_{N-1} - x_i) + (x_{N-1} - x_j)F_1(x_j)\right) \geq 0$. C_j^0 denotes the $N-1$ points analog of C_j where we collapse points x_N and x_{N-1} by reducing their distance δ to zero. Similar results emerge for D_j, r_2 and D_j^0. By the induction hypothesis we must have $D_j^0 \geq C_j^0$. Note that r_1 and r_2 are independent of the actual levels of p_N and p_{N-1}. From (21) we can infer that $p_N(1-p_N) \leq q_N(1-q_N)$, thus if C_j was bigger than D_j, we would certainly need $r_1 > r_2$.

Now, let us study how these expressions change if we shift mass from x_{N-1} to x_N. Since p_{N-1} does not enter C_j directly,[42] a mass increase towards x_N by one

[42] Notice that C_j^0 only the depends on the sum $p_{N-1} + p_N$ and is invariant under mass shifts from one to the other.

percent yields

$$p_N \frac{\partial C_j}{\partial p_N} = C_j - C_j^0 - \delta^2 p_N^2 \geq 0 \tag{22}$$

$$q_N \frac{\partial D_j}{\partial q_N} = D_j - D_j^0 - \delta^2 q_N^2. \tag{23}$$

Here, the positivity of the upper derivative relies on the fact that $p_N \leq 1 - p_1 \leq 1/2$ and therefore $\delta^2 p_N(1 - 2p_N) \geq 0$. In virtue of the positivity, we increase p_N to make C_j as large as possible without violating MLRP, i.e. we set

$$p_N = q_N \chi \equiv q_N \frac{1 - \sum_{i<N-1} p_i}{1 - \sum_{i<N-1} q_i}.$$

The ratio χ here guarantees that $p_N/q_N = p_{N-1}/q_{N-1}$. In other words, we used the positivity of (22) to reduce the two degrees of freedom of the two mass shifts to one. Precisely by percentage shifts up or down, we can now control how much weight *both* distributions lay on x_N, thereby keeping our equated MLRP condition $p_N/q_N = p_{N-1}/q_{N-1}$ intact. The overall effect of these remaining percentage shifts on $D_j - C_j$ is a comparison of (22) and (23).

First, assume $p_N \partial C_j/\partial p_N \leq q_N \partial D_j/\partial q_N$. Then, adding the two inequalities $C_j^0 \leq D_j^0$ and $\delta^2 p_N^2 \leq \delta^2 q_N^2$ yields $C_j \leq D_j$ and we are done. Now, suppose $p_N \partial C_j/\partial p_N > q_N \partial D_j/\partial q_N$. This means, we increase $C_j - D_j$ by moving more weight from x_{N-1} to x_N until there is nothing left at x_{N-1}, i.e. $p_{N-1} = q_{N-1} = 0$. This is equivalent to omitting x_{N-1}, so, again, we know by the induction hypothesis that $D_j \geq C_j$. In sum, we have shown that there is no way in which C_j could exceed D_j, for any $j \geq 2$ and any $N \geq 2$. ◇

This concludes the proof of the theorem for the general MLRP case. □

A.3. Proof of Proposition 2

It is straightforward to show that H is increasing concave. Hence, given Properties 1 and 2, we can apply Theorem 1. In case (i) of Property 2, the claim then follows directly from the theorem. In case (ii), the theorem yields that $\mathrm{Var}\{\tilde{\chi}_t^m | \chi_t^s\}$ is increasing in χ_t^s, implying that $\mathrm{Var}\{\tilde{\theta}_t | \chi_t^s\} = \mathrm{Var}\{\tilde{\theta}_t - \log(\bar{q}) | \chi_t^s\} = \mathrm{Var}\{\tilde{\chi}_t^m | \chi_t^s\}$ is also increasing in χ_t^s. □

A.4. Proof of Lemma 1

First, note that $\tilde{\chi}_t^m = \tilde{\theta}_t - \log(q_t)$, so that given a flat prior[43] over $\tilde{\theta}_t$, the posterior belief $\tilde{\theta}_t | \chi_t^s$ is normally distributed around $\alpha_1(\chi_t^s) + \log(\bar{q})$ with variance $\alpha_2(\chi_t^s)^2$ by Property 2 (ii). This is exactly the same posterior belief a Bayesian updater would hold after observing a Gaussian signal $\tilde{\theta}_t + \tilde{v}_t$ with realization $\alpha_1(\chi_t^s) + \log(\bar{q})$ and variance $\tau_v^{-1} = \alpha_2(\chi_t^s)^2 \equiv \alpha_{21}(\alpha_1(\chi_t^s)) = \alpha_{21}(\theta_t + v_t - \log(\bar{q}))$. Note that τ_v is increasing and has the desired limiting properties by Lemma 3. Thus, an observer with a flat prior updates his information given χ_t^s *as if* the signal he receives is Gaussian with a constant variance that happens to be τ_v^{-1}. The crucial step is now to show that given this informational equivalence holds for a flat prior distribution of $\tilde{\theta}_t$, it continues to hold for *any* normal prior over $\tilde{\theta}_t$.

Suppose an observer holds a normal prior over $\tilde{\theta}_t$ as given by a pdf $p(\theta) = \phi_{\theta_0, \tau_0^{-1}}(\theta)$ and receives some signal s with pdf $q(s|\theta)$ such that he would have updated to an s-dependent normal posterior $p_0(\theta|s) = \phi_{\mu(s), \tau(s)^{-1}}(\theta)$ had he held a flat prior $p_0(\theta) = 1$ over $\tilde{\theta}_t$. This means,

$$p_0(\theta|s) = \frac{q(s|\theta)}{\int q(s|z)\,\mathrm{d}z}.$$

Therefore, the updated posterior pdf $p(\theta|s)$ given a normal prior can be written as

$$p(\theta|s) = \frac{q(s|\theta)p(\theta)}{\int q(s|z)p(z)\,\mathrm{d}z} = \frac{p_0(\theta|s)p(\theta)}{\int p_0(z|s)p(z)\,\mathrm{d}z},$$

which is just a normal pdf with mean $(\tau_0 + \tau(s))^{-1}(\tau_0\theta_0 + \tau(s)\mu(s))$ and variance $(\tau_0 + \tau(s))^{-1}$. This is exactly the posterior distribution a Bayesian updater infers from observing the realization $\mu(s)$ of a Gaussian signal $\tilde{\theta}_t + \tilde{v}_t$ where $\tau_v = \tau(s)$. □

A.5. Proof of Lemma 2

To verify the fixed point note that given the law of motion (8), $\tilde{q}_t$ is informationally equivalent to a signal

$$\tilde{\theta}_t + \delta_\eta \tilde{\eta}_t + \delta_v \tilde{v}_t,$$

[43] We use the flat prior merely for simplicity. The argument goes through for any normal prior with a variance larger than some (constant) upper bound.

where $\delta_\eta = \frac{\tau_\xi+\tau_\eta}{\tau_\xi+\tau_\eta+\tau_v}$ and $\delta_v = \frac{\tau_v}{\tau_\xi+\tau_\eta+\tau_v}$. Straightforward application of Bayes rule yields

$$b_{it} = \bar{\pi}_t^{-1} \times \begin{bmatrix} \tau_\xi & \tau_\eta & \tau_v & \hat{\pi}_{t-1} \end{bmatrix} \times \begin{bmatrix} \theta_t + \xi_{it} \\ \theta_t + \eta_t \\ \theta_t + v_t \\ \rho b_{t-1} \end{bmatrix} \tag{24}$$

when solving the inference problem including the private signal $\tilde{s}_{it}$, and yields b_t as stated in the lemma when considering only the publicly observable history of prices. Aggregating over i, substituting into the equilibrium price as pinned down by the marginal trader

$$q_t = \mathrm{E}\{\tilde{A}_t | b_{it}, \bar{\pi}_t\} = \exp\{\gamma_0 + \gamma_1(\sigma_t\sqrt{\tau_\xi}\,(\tilde{\eta}_t - \mu) + \bar{b}_t + \gamma_1/(2\bar{\pi}_t))\},$$

noting that (24) implies a cross-sectional variation $\sigma_t^2 = \mathrm{Var}\{b_{it}\} = \tau_\xi/\bar{\pi}_t^2$, and using $\mu = \gamma_1/(2\tau_\xi)$ verifies that the mapping (8) is indeed a fixed point.

Uniqueness follows from following the same steps above, but leaving δ_η and δ_v unspecified. Solving the resulting system of equations yields two solutions. The first one being the one stated in the lemma and the second one being $\delta_\eta = \delta_v = 0$. Note that the second solution implies that rational beliefs and, hence, market prices are invariant to the realization of noisy asset demand $\tilde{\eta}_t$. Therefore, $\delta_\eta = \delta_v = 0$ clearly violates market clearing for almost all realizations of $\tilde{\eta}_t$, implying uniqueness of the first solution. □

A.6. Proof of Proposition 3

Please note that the notation in this proof is inconsistent with the notation in the remainder of this chapter. We are sorry for any confusion arising from that.

A.6.1. Existence

To show the existence of a solution to the fixed point problem (9), we first derive the inverse of g_q. We used g_q to describe the functional form of the log-linear equilibrium asset price q_t, defined in equation (8). It can be rewritten in terms of τ_v, $r \equiv \log(q_t)|_{\tau_v \to 0}$, and $s \equiv \gamma_1(\theta_t + v_t) + \gamma_0$,

$$\log(q_t) = g_q(\tau_v, \Omega_t) = (1 - \frac{\tau_v}{\bar{\pi}_t})r + \frac{\tau_v}{\bar{\pi}_t}s.$$

Dropping the second argument Ω_t for simplicity from now on, the inverse reads

$$g_q^{-1}(\log(q_t)) = (\bar{\pi}_t|_{\tau_v\to 0})\,\frac{\log(q_t)-r}{s-\log(q_t)}. \tag{25}$$

Evidently, g_q^{-1} is only defined for $\log(q_t)$ between r and s whenever $r \neq s$. Then, its range is found to be $[0,\infty)$, i.e. $g_\tau(\log(q_t)) - g_q^{-1}(\log(q_t))$ attains $g_\tau(r) > 0$ for $\log(q_t) = r$ and converges to $-\infty$ for $\log(q_t) \to s$. By the intermediate value theorem, there exists a value for $\log(q_t)$ such that $g_\tau(\log(q_t)) - g_q^{-1}(\log(q_t)) = 0$. For $r = s$, g_q^{-1} is not well-defined, so we cannot study the problem in terms of (9). Instead, we consider the usual form of the fixed point equation, $g_q(g_\tau(\log(q_t))) = \log(q_t)$. Obviously, since $g_q = r$, $\log(q_t) = r$ is the unique fixed point if $r = s$. □

A.6.2. Uniqueness

Let $\delta = s - r$. For $\delta < 0$, g_q^{-1} is strictly decreasing and $g_\tau - g_q^{-1}$ is strictly increasing. Thus, a unique fixed point exists if $\delta < 0$. The case $\delta = 0$ is discussed above. Given these considerations, we see that there must be a non-empty set $\Xi \subset \mathbb{R}_+ \times \mathbb{R}^4$ characterizing all parameter constellations Ω_t that lead to unique equilibria. Bear in mind that r and s are just combinations of different components of Ω_t that we use to describe the influence of Ω_t on $g_q(\log(q_t))$. Since the set $\{\Omega_t \mid \delta \leq 0\}$ entirely lies in Ξ, we now derive bounds for the case $\delta > 0$.

Bound 1 In virtue of (25), the fixed-point equation (9) is equivalent to

$$(s - \mathring{s} - z)\Big(1 + \frac{\tau_v(-z)}{\bar{\pi}_t|_{\tau_v\to 0}}\Big) - s = -r, \tag{26}$$

where $\mathring{s} \equiv (s-\gamma_0)/\gamma_1$ and $z \equiv \log(q_t) - \mathring{s} \in (r-\mathring{s}, s-\mathring{s})$. This has a unique solution for any value of r if the left hand side is strictly decreasing for all $z \in (r,s)$, i.e.

$$s - \mathring{s} < f(s-\mathring{s}) \equiv \inf_{z\in(-\infty,s-\mathring{s})} \frac{\bar{\pi}_t|_{\tau_v\to 0} + \tau_v(-z)}{-\tau_v'(-z)} + z. \tag{27}$$

It is evident that the function f is increasing. Apart from that, we show in Lemma 7 below that there exists a unique bound s^* such that below s^*, $f(s) > s$ while $f(s^*) = s^*$. Note that this also implies that $f(s) > -\infty$ for any s. Using s^* the condition for uniqueness independent of r becomes $s - \mathring{s} < s^*$, or, split according to the signs of $\gamma_1 - 1$,

$$\begin{aligned} \gamma_1(\theta_t + v_t) + \gamma_0 = s < M, &\quad \text{for } \gamma_1 > 1 \\ \gamma_1(\theta_t + v_t) + \gamma_0 = s > M, &\quad \text{for } \gamma_1 < 1 \\ M_0 > 0, &\quad \text{for } \gamma_1 = 1, \end{aligned}$$

where $M = (\gamma_1 s^* - \gamma_0)/ \mid \gamma_1 - 1 \mid$ and $M_0 = \gamma_1 s^* - \gamma_0$. Interestingly, the uniqueness condition for $\gamma_1 = 1$ is independent of s, i.e. whenever $M_0 > 0$, *any* equilibrium is unique. On the other hand, for $M_0 < 0$, we always find *some* values for r such that the equilibrium is not unique, irrespective of s. Linearly redefining M establishes the desired bounds on $\theta_t + v_t$.

Bound 2 Similar to (26), the fixed-point equation (9) can be rewritten to

$$\delta \frac{\bar{\pi}_t|_{\tau_v \to 0}}{\tau_v(-z) + \bar{\pi}_t|_{\tau_v \to 0}} + z = s - \mathring{s}, \tag{28}$$

with $\delta, z, s, \mathring{s}$ as above. Now suppose $\gamma_1 \neq 1$. To find an upper bound δ^* for $s - r$ that establishes uniqueness independent of s, we set δ^* to the largest level need to have that the left hand side of (28) is increasing, i.e.

$$\delta^* \equiv \inf_{z \in \mathbb{R}} \frac{(\tau_v(-z) + \bar{\pi}_t|_{\tau_v \to 0})^2}{-\tau_v'(-z)\, \bar{\pi}_t|_{\tau_v \to 0}} > 0.$$

It now holds that, whenever $\delta = \gamma_1(\theta_t + v_t) + \gamma_0 - \log(q_t)|_{\tau_v \to 0} < \delta^*$, the left hand side of (28) is strictly increasing and hence has a unique solution. The bound δ^* is by construction the largest with this property. The inequality stated in the proposition can be derived by a linear transformation.

Let us regard the special case $\gamma_1 = 1$. Here, we can do better than δ^*. Since $s - \mathring{s}$ is constant at γ_0/γ_1, to ensure uniqueness δ must only be small enough for (28) to have a unique solution if the right hand side equals γ_0/γ_1, i.e.

$$\delta < \delta^{**} \equiv \sup\{\delta \mid (28) \text{ has a unique solution for } s - \mathring{s} = \gamma_0/\gamma_1\}.$$

Clearly, $\delta^{**} \geq \delta^* > 0$.

Lemma 7: *The following two statements hold:*

1. *As* $x \to \infty$, $x\, \tau_v'(x) \to 0$.
2. *Let* f *be as defined in* (27). *Then,* $\lim_{s \to -\infty} f(s) - s > 0$ *while* $\lim_{s \to \infty} f(s) - s < 0$.

Proof. We first show part 1. Recall that $\tau_v(x) = 1/\sigma_v(x)^2$ is defined by the following implicit equation,

$$\mathrm{Var}\{H(\sigma_v(x)\tilde{X} + x)\} = \sigma_c^2, \tag{29}$$

where $\tilde{X}$ is standard normally distributed and we denote $\mathrm{Var}\{\tilde{\chi}^d \mid \chi^s\}$ by σ_c^2. Note that H is similar to a standard kink function $\mathring{H} = \min\{\cdot, 0\}$ in that they are almost equal outside the area around the kink where H is smooth while $\mathring{H}$ is not. Now, for x large enough, the probability mass of the distribution of $\sigma_v(x)\tilde{X} + x$ that is assigned to the area around the kink becomes arbitrarily small. Therefore, for large x, the solution σ_v of (29) behaves exactly like the solution $\mathring{\sigma}_v$ to the implicit equation

$$\mathrm{Var}\{\mathring{H}(\mathring{\sigma}_v(x)\tilde{X} + x)\} = \sigma_c^2. \tag{30}$$

In contrast to (29), we can analytically compute the variance on the left hand side of (30),

$$\mathrm{Var}\{\mathring{H}(\mathring{\sigma}_v(x)\tilde{X} + x)\} = \mathring{\sigma}_v(x)^2 F(x/\mathring{\sigma}_v(x)),$$

where $F(\beta) = 1 - \Phi(\beta) - \phi(\beta)^2 + \beta^2(\Phi(\beta) - \Phi(\beta)^2) - \beta\phi(\beta)(2\Phi(\beta) - 1)$. Substituting this in (30), the implicit definition for $\tau_v(x)$ reads,

$$F\left(\sqrt{\tau_v(x)\, x^2}\right) = \sigma_c^2 \tau_v(x).$$

Differentiating this with respect to x yields,

$$\frac{F'}{2}(\tau_v(x))^{-1/2}(x\, \tau_v'(x) + 2\tau_v(x)) = \sigma_c^2 \tau_v'(x). \tag{31}$$

Regard the signs on both sides of (31): $F'(\beta) = -2(\phi(\beta) - \beta(1 - \Phi(\beta)))\Phi(\beta) < 0$ due to the standard bound on the tails of Φ, $1 - \Phi(\beta) < |\beta|^{-1}\phi(\beta)$, and $\tau_v'(x) < 0$ as we saw in Proposition 2. This immediately implies that $x\, \tau_v'(x) + 2\tau_v(x)$ must be positive, or in other words,

$$0 > x\, \tau_v'(x) > -2\tau_v(x),$$

which establishes part 1 of the lemma for we know $\tau_v(x)$ converges to 0 as x tends to infinity (see Lemma 1).

In (27), we defined the function f such that

$$f(-s) = \inf_{z > -s} \frac{\bar{\pi}_t|_{\tau_v \to 0} + \tau_v(z)}{-\tau_v'(z)} - z.$$

The first claimed property $\lim_{s\to-\infty} f(s) - s = \lim_{s\to\infty} f(-s) + s > 0$ means that there exists an $\bar{s} > 0$ such that for any $s > \bar{s}, z \geq s$,

$$s > \frac{\bar{\pi}_t|_{\tau_v\to 0} + \tau_v(z)}{\tau_v'(z)} + z \iff -(z-s)\tau_v'(z) < \bar{\pi}_t|_{\tau_v\to 0} + \tau_v(-z).$$

But the left hand side of the second equation is smaller than $-z\,\tau_v'(z)$ —which we know tends to 0 from part 1—and therefore has to be smaller than $\bar{\pi}_t|_{\tau_v\to 0}$ for large values of z.

The second property $\lim_{s\to\infty} f(s) - s < 0$ is equivalent to finding a $\bar{s} > 0$ such that for any $s > \bar{s}$ there exists a $z \leq s$ for which

$$\frac{\bar{\pi}_t|_{\tau_v\to 0} + \tau_v(-z)}{-\tau_v'(-z)} + z < s \iff \bar{\pi}_t|_{\tau_v\to 0} + \tau_v(-z) - z\,\tau_v'(-z) < -s\,\tau_v'(-z).$$

Trivially, for $z = 0$ the right hand side of the second equation diverges to ∞ for large values of s while the rest remains constant. Thus we can just pick a $\bar{s} > 0$ that is large enough. This establishes the second property of part 2. ◇

This concludes the proof of the proposition. □

A.7. *Proof of Proposition 4*

For the most part, the proof follows from the discussion in the main body of the chapter. It remains to be shown that $\text{sign}\{\tau_v^1 - \mathring{\tau}_v\} = \text{sign}\{\tau_v^* - \mathring{\tau}_v\}$. Given the definition of $\mathring{\tau}_v$, τ_v^1, and τ_v^*, and given that g_τ is increasing (see Lemma 1), proving the claim is equivalent to showing that

$$\log(\mathring{q}_t) > -\gamma_0 \iff \log(q_t^*) > -\gamma_0. \tag{32}$$

We distinguish two cases. First, consider the case where g_q is decreasing. Then we have that

$$\log(\mathring{q}_t) = g_q(g_\tau(-\gamma_0)) > -\gamma_0 \iff g_\tau(-\gamma_0) < g_q^{-1}(-\gamma_0).$$

But given that $g_q' < 0$ and $g_\tau' > 0$, it follows that g_q^{-1} and g_τ intersect to the right of $-\gamma_0$ if and only if $g_q^{-1}(-\gamma_0) > g_\tau(-\gamma_0)$. Hence, (32) holds at the unique fixed point q_t^*.

Now consider the case where g_q is increasing. Then

$$\log(\mathring{q}_t) = g_q(g_\tau(-\gamma_0)) > -\gamma_0 \iff g_\tau(-\gamma_0) > g_q^{-1}(-\gamma_0).$$

From Lemma 2, it follows that $\lim_{x \to \gamma_0 + \gamma_1(\theta_t + v_t)} g_q^{-1}(x) = \infty$, while by Lemma 1 we have that $g_\tau(x)$ is finite for all x. Hence, whenever $g_\tau(-\gamma_0) > g_q^{-1}(-\gamma_0)$, there necessarily exists an intersection between g_q^{-1} and g_τ to the right of $-\gamma_0$. Further, since by Lemma 2 $\lim_{x \to -\infty} g_\tau(x) = 0$, while $\lim_{x \to -\infty} g_q^{-1}(x) < 0$ (since $g_q(0)$ is finite by Lemma 2), we also have that there necessarily exists an intersection between g_q^{-1} and g_τ to the left of $-\gamma_0$ whenever $g_\tau(-\gamma_0) > g_q^{-1}(-\gamma_0)$. Hence, there always exists at least one fixed point q_t^* such that (32) holds. In the case considered in the main body of the text where there always exists a unique equilibrium, this concludes the proof. Moreover, when there are multiple equilibria, in a given state Ω_t, our analysis continues to apply to each equilibrium that satisfies (32) (for further details, see Appendix B.4). □

A.8. Proof of Proposition 5

Let $(\theta_t + v_t, \theta_t + \eta_t) = a\mathcal{S}_t + \mathcal{O}_t \equiv a \cdot (x, y) + (0, z)$, with $z \equiv -\breve{b}_{t-1}/(\tau_\xi + \tau_\eta)$. Then, differentiating $\Delta \equiv \log(q_t^*) - \log(\mathring{q}_t) = g_q(\tau_v^*, \Omega_t) - g_q(\mathring{\tau}_v, \Omega_t)$ with respect to a yields

$$\frac{\mathrm{d}\Delta}{\mathrm{d}a} = \frac{\partial \Delta}{\partial a} + \frac{\partial g_q(\tau_v^*, \Omega_t)}{\partial \tau_v^*} \frac{\mathrm{d}\tau_v^*}{\mathrm{d}a}. \tag{33}$$

Consider the first term first. From Lemma 2, we have that

$$g_q(\tau_v^*, \Omega_t) = a\gamma_1(y(\tau_\xi + \tau_\eta) + x\tau_v^*)/\bar{\pi}^*,$$

and analogous for $\mathring{\tau}_v$. Substituting in Δ, differentiating, and rearranging, we get

$$\frac{\partial \Delta}{\partial a} = \frac{\gamma_1}{\bar{\pi}^* \mathring{\bar{\pi}}(\tau_\xi + \tau_\eta)} \times \frac{1}{\mathring{\tau}_v - \tau_v^*} \times [y + z - \delta_1(x)].$$

From the definitions of x, y, and z, $y + z - \delta_1(x)$ is positive if and only if $\theta_t + \eta_t > \delta_1(\theta_t + v_t)$. Moreover, $\mathring{\tau}_v - \tau_v^*$ is positive if and only if $\theta_t + \eta_t < \delta_2(\theta_t + v_t)$. Thus $\partial\Delta/\partial a < 0$ if and only if sign$\{\theta_t + \eta_t - \delta_1(\theta_t + v_t)\}$=sign$\{\theta_t + \eta_t - \delta_2(\theta_t + v_t)\}$.

Consider now the second term of (33). Substituting $\log(q_t^*) = g_q(\tau_v^*, \Omega_t)$ into (9) and implicit differentiating yields

$$\frac{\mathrm{d}\tau_v^*}{\mathrm{d}a} = \frac{x - \frac{\partial g_q}{\partial a}}{\frac{\partial g_q}{\partial \tau_v} - \frac{\partial g_\tau^{-1}}{\partial \tau_v}}. \tag{34}$$

As illustrated in Figure 2, $\frac{\partial g_q}{\partial \tau_v} < \frac{\partial g_\tau^{-1}}{\partial \tau_v}$ at the fixed point, so that the denominator is necessarily negative. Again, substituting for $\partial g_q(\tau_v^*, \Omega_t)/\partial a$, the numerator simplifies to

$$x - \frac{\partial g_q}{\partial a} = -\frac{\gamma_1(\tau_\xi + \tau_\eta)}{\bar{\pi}^*} \times [y + z - \delta_2^*(x)],$$

where

$$\delta_2^*(x) = (\tau_\xi + \tau_\eta)^{-1}(-\breve{b}_{t-1} + \gamma_1^{-1}(\breve{\pi}_{t-1} + (1-\gamma_1)\tau_v^*)x)$$

is defined such that $\theta_t + \eta_t < \delta_2^*(\theta_t + v_t)$ if and only if $\mathring{\tau}_v - \tau_v^* > 0$. Hence, since it also holds that $\mathring{\tau}_v - \tau_v^* > 0$ if and only if $\mathring{\tau}_v - \tau_v^1 > 0$, we have that from the definitions of x, y, and z, $-[y + z - \delta_2^*(x)]$—and, hence, the numerator of (34)—is positive if and only if $\theta_t + \eta_t < \delta_2(\theta_t + v_t)$. Taken together, we thus have that $\mathrm{d}\tau_v^*/\mathrm{d}a$ is negative if and only if $\theta_t + \eta_t < \delta_2(\theta_t + v_t)$. Moreover, in Section 5.1, we show that $\partial g_q/\partial \tau_v > 0$ if and only if $\theta_t + \eta_t < \delta_1(\theta_t + v_t)$. Hence, like the first term, the second term of (33) is negative if and only if sign$\{\theta_t+\eta_t\text{-}\delta_1(\theta_t+v_t)\}$=sign$\{\theta_t+\eta_t\text{-}\delta_2(\theta_t+v_t)\}$, completing the proof. □

A.9. Proof of Proposition 6

We consider a generic information structure given by

$$s_t = H'\theta_t + \psi_t, \qquad \psi_t \sim N(0, \Psi_t),$$

where H' is a $m \times 1$-vector and Ψ_t is a positive-semidefinite, symmetric $m \times m$ matrix. Note that s_t is informationally equivalent to $\bar{s}_t = B_t H'\theta_t + \bar{\psi}_t$ with $\bar{\psi}_t \sim N(0, B_t\Psi_t B_t')$ for all invertible B_t which match the number of rows in H'. In particular, we can choose B_t, such that $B_t H' = (1, \ldots, 1)'$ and $\bar{\Psi}_t \equiv B_t \Psi_t B_t'$ is diagonal.[44]

Accordingly, suppose without loss of generality that

$$s_t = (1, \ldots, 1)'\theta_t + \psi_t, \qquad \psi_t \sim N(0, \mathrm{diag}(\tau_t)^{-1}),$$

where τ_t is a $m \times 1$-vector of strictly positive signal precisions. Then posterior beliefs at time t are given by

$$b_t = \frac{1}{\pi_t}\sum_{i=0}^{m} \tau_{t,i}s_{t,i} \qquad \text{and} \qquad \pi_t = \sum_{i=0}^{m} \tau_{t,i},$$

[44] To see this, let L_t be the lower Cholesky factor of Ψ_t. Then for all diagonal A_t, we have that $A_t L_t^{-1}\Psi_t L_t'^{-1}A_t'$ is diagonal. Hence, setting $A_t = \mathrm{diag}(L_t^{-1}H')^{-1}$ and defining $B_t \equiv A_t L_t^{-1}$ yields the desired result.

where the subscript i denotes the i-th element of vectors τ_t and s_t with the convention that $\tau_{t,0} = \hat{\pi}_{t-1} \equiv (\rho^2\tau_\epsilon + \pi_{t-1})^{-1}\tau_\epsilon\pi_{t-1}$ and $s_{t,0} = \rho b_{t-1}$. Accordingly, $\lambda_t = \rho\hat{\pi}_{t-1}/\pi_t$, matching exactly the definition in the special case discussed in the main body of the text.

Without loss of generality, consider a generic change of $\tau_{t,1}$ by a differential $\mathrm{d}\tau_{t,1}$. Then for $r, s > 0$,

$$\frac{\mathrm{d}\Lambda_{t-r,t+s}}{\mathrm{d}\tau_{t,1}} = \Lambda_{t-r,t+s} \times \left\{ \frac{1}{\lambda_t}\frac{\partial\lambda_t}{\partial\pi_t} + \sum_{q=t+1}^{t+s} \left(\frac{1}{\lambda_q}\frac{\partial\lambda_q}{\partial\pi_{q-1}} \prod_{p=t+1}^{q-1} \frac{\partial\pi_p}{\partial\pi_{p-1}} \right) \right\},$$

or, after computing the individual terms and dividing by $\Lambda_{t-r,t+s} > 0$,

$$-\frac{1}{\pi_t} + \sum_{q=t+1}^{t+s} \left(\frac{\hat{\pi}_{q-1}}{\pi_{q-1}^2\pi_q} \left(\rho^2 \sum_{i=1}^{m} \tau_{q,i} \right) \prod_{p=t+1}^{q-1} \left(\frac{\rho\hat{\pi}_{p-1}}{\pi_{p-1}} \right)^2 \right). \tag{35}$$

To show that (35) is negative, we proceed in two steps.

Step 1 We claim that (35) is maximized by setting $\tau_{o,1} \to \infty$ for all $o > t$. We prove this claim by proceeding recursively. For $o = s$, the term is obviously increasing in $\tau_{s,1}$ since

$$\frac{\partial}{\partial\tau_{s,1}} \left\{ \frac{\sum_{i=1}^m \tau_{s,i}}{\pi_s} \right\} = \frac{\partial}{\partial\tau_{s,1}} \left\{ \frac{\sum_{i=1}^m \tau_{s,i}}{\sum_{i=0}^m \tau_{s,i}} \right\} > 0.$$

Hence, suppose that $\tau_{o,1} \to \infty$ for all $o > n$. Then, $\hat{\pi}_o \to \tau_\epsilon$ and $\pi_o \to \infty$. Thus, differentiating (35) with respect to $\tau_{n,1}$, $n > t$, simplifies to

$$\begin{aligned}
\frac{\partial}{\partial\tau_{n,1}} & \left\{ \sum_{q=n}^{n+1} \left(\frac{\hat{\pi}_{q-1}}{\pi_{q-1}^2\pi_q} \left(\rho^2 \sum_{i=1}^{m} \tau_{q,i} \right) \prod_{p=t+1}^{q-1} \left(\frac{\rho\hat{\pi}_{p-1}}{\pi_{p-1}} \right)^2 \right) \right\} \\
&= \frac{\partial}{\partial\pi_n} \left\{ \frac{\pi_n - \hat{\pi}_{n-1}}{\hat{\pi}_{n-1}\pi_n} + \frac{\rho^2\hat{\pi}_n}{\pi_n^2} \right\} \times \prod_{p=t+1}^{n} \left(\frac{\rho\hat{\pi}_{p-1}}{\pi_{p-1}} \right)^2 \\
&= \left(\frac{\pi_n - \rho^2\hat{\pi}_n}{\pi_n^2} \right)^2 \times \prod_{p=t+1}^{n} \left(\frac{\rho\hat{\pi}_{p-1}}{\pi_{p-1}} \right)^2 > 0,
\end{aligned}$$

verifying the claim.

Step 2 By step 1, it is sufficient to show that (35) is negative if $\tau_{o,1} \to \infty$ for all $o > t$. Accordingly, (35) simplifies to

$$-\frac{\pi_t - \rho^2 \hat{\pi}_t}{\pi_t^2} = -\frac{\hat{\pi}_t}{\tau_\epsilon \pi_t} < 0,$$

completing the proof. □

B. Supplementary material

B.1. Pooled asset market

First, we prove that in any equilibrium the market clearing price must be the same in both occurrences of the asset market. Suppose the contrary holds and the market clearing price in the first asset market is larger than the price in the second market.[45] Then, given that no informational gains are possible between the two market instances, all gatherers find it optimal to sell their asset (tree) in the first market. Yet, total asset demand is always smaller than $1 - m$ in the first market, a contradiction. Suppose now the price in the first market is smaller than the price in the second market. Consequently, all agents find it optimal not to sell assets in the first market. Again, this is incompatible with market clearing since total asset demand is always strictly positive. Thus, in any equilibrium, the two occurrences of the asset market share the same market price.

To show the equivalence of the two separated markets to one pooled market we need to show that any equilibrium in the separated markets is also an equilibrium in the pooled market and vice versa. Consider an equilibrium in the separate markets. From above we know that there is a single market clearing price equating supply and demand in both markets. This price must also be an equilibrium price in the pooled market for it obviously equates total supply and total demand. Vice versa, suppose a price is an equilibrium price in the pooled market. We construct an equilibrium candidate for the separate markets by letting all gatherers who trade in the pooled equilibrium trade in the first market and all farmers who trade in the pooled equilibrium trade in the second market. Indeed, this is an equilibrium with the pooled price since the fraction of trading agents must be

[45] Note that the price in the second market can be forecasted by the time the first market operates because (i) no exogenous information realizes between the two markets, and (ii) all information aggregated by the second market price is already aggregated by the first.

the same across farmers and gatherers and nobody has an incentive to change his marketplace.

The last argument also implies that there is an equilibrium where the relevant collateral constraint for all farmers is given by $n_{it} \leq q_t/w_t$. Since selling assets on the first market and ending up constrained is dominated for all farmers by waiting and selling assets on the second market, we conclude that every separate asset market equilibrium can be represented by a pooled asset market that operates parallel to the labor market in which $n_{it} \leq q_t/w_t$ is exogenously imposed on all farmers.

B.2. Endogenous learning with alternative financial frictions

For demonstrating how endogenous learning interacts with other financial frictions, we consider four toy models. To be consistent with the structure of the model, all frictions are sited at the firm-level. However, one could also shift constraints to a separate financial sector, which then constrains the real sector depending on the state of the economy. In the following, our strategy is to set up simple versions of these alternative frictions and solve these model fragments up to a point where Theorem 1 is applicable.

Cash-in-advance constraints Consider a continuum of entrepreneurs with an investment opportunity that for an initial investment of k_i pays

$$F(\tilde{A}_i, k_i) = \tilde{A}_i \log(k_i) + \tilde{A}_i \gamma,$$

where $\log(\tilde{A}_i) \sim \mathcal{N}(\theta, 1/\tau_\xi)$. Let p denote the price per unit of investment and assume that investments have to be paid in advance using cash. For the purpose of raising cash, entrepreneurs may sell claims on the investment return on a financial market. For simplicity, assume that each entrepreneur is exogenously endowed with γ units and that claims can only be written on the return $\tilde{A}_i\gamma$ of these units (e.g., because k_i is unobservable or noncontractable). The timing is as follows:

1. Entrepreneurs choose to sell claims on any fraction $x_i \in [0,1]$ of γ on the financial market.

2. The financial market operates and yields an equilibrium price q per claim. (Each share entitles its owner to claim the return of γ units).

3. Entrepreneurs learn the realization of $\tilde{A}_i$ and decide how much to invest, subject to the cash-in-advance constraint $k_i \leq (q/p)\, x_i$.

For any reasonable specification of the financial market, it should be clear that whenever agents on the financial market and entrepreneurs start out with a common prior, then entrepreneurs optimally set $x_i = 1$. Then given p, each entrepreneur optimally sets $k_i = \min\{A_i, q\}/p$, so that aggregated demand k^d resembles (6). Without specifying the details of the capital supply side, assume that there exists a noisy pricing function $\tilde{p} = f(\tilde{k}^d, \tilde{\psi})$ that clears the market and which is increasing in both arguments ($\tilde{\psi}$ being some random variable). Given these assumption, there trivially exists a transform of $\tilde{\psi}$ which gives rise to a concave signal structure that, given the appropriate assumptions on the random variable $\tilde{\psi}$, is isomorphic to the one resulting from our baseline model.

Skin-in-the-game constraints A common generalization of the above cash-in-advance approach is to allow entrepreneurs to give out claims on profits, but assume that in order to provide the right incentives, entrepreneurs must have some "skin-in-the-game" that exogenously restricts the maximal number of shares that can be issued to $x_i \leq \bar{x} < 1$. Keeping the timing identical to our cash-in-advance setup, the difference is now that claims on the financial market are defined on expected entrepreneurs profits:

$$\mathrm{E}\{\Pi(\tilde{A}_i, x_i, \tilde{p}, q)|\mathcal{I}_j\} \equiv \mathrm{E}\{F(\tilde{A}_i, k_i^*(\tilde{A}_i, \tilde{p})) - k_i^*(\tilde{A}_i, \tilde{p})\tilde{p} + x_i q|\mathcal{I}_j\}.$$

Assuming risk-neutrality on the financial market (and some bounds on traders' asset demands that, as in our baseline setup, ensure the existence of a market clearing price), the equilibrium price will be given by the marginal trader m's expectation

$$\begin{aligned} q &= \mathrm{E}\{\Pi(\tilde{A}_i, x_i, \tilde{p}, q)|\mathcal{I}_m\} \\ &= (1 - x_i)^{-1}\mathrm{E}\{F(\tilde{A}_i, k_i^*(\tilde{A}_i, \tilde{p})) - k_i^*(\tilde{A}_i, \tilde{p})\tilde{p}|\mathcal{I}_m\}. \end{aligned}$$

Suppose again that prior information ensures that entrepreneurs optimally set $x_i = \bar{x}$. For any $\bar{x} < 1$, we thus have that q amounts to a finite number that for well-behaved F, $\tilde{A}_i$ and $\tilde{p}$,[46] is increasing in $\mathrm{E}(\tilde{A}_i|\mathcal{I}_m)$. Based on the previous cash-in-advance setting, our results can therefore also be extended to such more general skin-in-the-game settings.

[46] Here, we implicitly assume that as long as the "skin-in-the-game" constraint is fulfilled, entrepreneurs choose $k_i^*(\theta_i) = \arg\max\{F(\theta, k_i) - k_i p\}$. Then, we in particular require that F is increasing in θ_i and that $F - kp$ is concavely increasing in k and has an interior solution.

Costly state verification We argue on an intuitive level. As firms' internal funding decreases, standard auditing models imply that the markup over the risk-free rate increases, implying that firms invest less and are less responsive to the price of output (see, e.g., Carlstrom and Fuerst, 1997). It is straightforward to extent such frameworks to the case where the return to investments also depends on an unobserved (to the financial market) state. To fix ideas, consider the case where $F(k_i, A_i) = A_i k_i$ with $\log(A_i) \sim \mathcal{N}(\theta, \sigma^2)$. With such log-normal specification, the production function transforms to the "standard" setting where $F(k_i, \theta, \omega_i) = p\omega_i k_i$ with $p = \exp(\theta + \sigma^2/2)$ and $\omega_i \sim \mathcal{N}(-\sigma^2/2, \sigma^2)$. That is, one can absorb θ into the output price. Assuming that firms are matched to lenders with zero bargaining power, and further assuming that lenders learn θ upon matching and that lenders finance themselves through an exogenous financial market, this is equivalent to the setting at the core of Carlstrom and Fuerst (1997). In particular, higher effective interest rates imply that firms' investment choices respond less to θ as firms have less internal funds n_i for the purpose of financing k_i; i.e., as $(k_i - n_i)$ increases. Aggregating over k_i gives rise to a concave relation between k and θ. Accordingly, based on this study's analysis the ability of any outside observer (like the financial market), who observes a noisy signal of (aggregate) k, is impeded during financial crises.

Adverse selection We argue on an intuitive level. Suppose there are two types of firms that differ in their probability of defaulting. Then for standard adverse selection setups, good firms are crowded out of the market in crisis times. But if good firms are more likely to succeed, then they will also be more respondent to any change in fundamentals that affect profits in the non-default state. This reduces overall responsiveness to the fundamental among market-financed firms during financial crises, so that based on this study's analysis the information aggregation becomes less efficient.

B.3. Distributional assumptions

To clarify the conditions under which Theorem 1 is applicable to our model, we state them in terms of properties of the posterior distribution $\tilde{\chi}_t^d|\chi_t^s$. Here we illustrate how these properties of $\tilde{\chi}_t^d|\chi_t^s$ can be mapped into assumptions on the distribution of $\tilde{\Psi}_t$.

From $\tilde{\chi}_t^s|\chi_t^s$ to $\tilde{\Psi}_t$ Consider an arbitrary posterior density function $P(\chi_t^d|\chi_t^s)$ given a flat prior over $\tilde{\chi}_t^d$. Trivially, we can choose $P(\chi_t^d|\chi_t^s)$ to satisfy Properties 1 and 2 (an example is the quasi-Gaussian posterior we use throughout most of the chapter). By Bayes' law, reverse engineering gives the corresponding conditions on $P(\chi_t^s|\chi_t^d)$:

$$P(\chi_t^d|\chi_t^s) = \frac{P(\chi_t^s|\chi_t^d)}{\int P(\chi_t^s|\hat{\chi}^d)\,\mathrm{d}\hat{\chi}^d}.$$

Rearranging yields

$$P(\chi_t^s|\chi_t^d) = g(\chi_t^s)P(\chi_t^d|\chi_t^s), \tag{36}$$

where g is indeterminate (i.e., arbitrary).

That is, any $P(\chi_t^s|\chi_t^d)$ that is consistent with (36) implements the chosen posterior $P(\chi_t^d|\chi_t^s)$. In particular note that because of the indeterminacy of g, there are infinite many *improper* conditional distributions $P(\chi_t^s|\chi_t^d)$ that are consistent with our assumptions on $P(\chi_t^d|\chi_t^s)$. Lets for now not worry whether there exists any g which guarantees the existence of a *proper* signal structure (but see below). Then in order to transform $P(\chi_t^s|\chi_t^d)$ into a distribution of $\tilde{\psi}_t$, recall that

$$\tilde{\chi}_t^s = \tilde{\chi}_t^d + \tilde{\psi}_t.$$

Then, from (36),

$$P(\psi_t|\chi_t^d) = P(\chi_t^s - \chi_t^d|\chi_t^d),$$

yielding the following CDF for $\tilde{\psi}_t|\chi_t^d$

$$\begin{aligned} P(\psi_t \leq z|\chi_t^d) &= \int_{-\infty}^{z+\chi_t^d} P(\chi^s|\chi_t^d)\,\mathrm{d}\chi^s \\ &= \int_{-\infty}^{z+\chi_t^d} g(\chi^s)P(\chi_t^d|\chi^s))\,\mathrm{d}\chi^s. \end{aligned} \tag{37}$$

Condition (37) defines the distributional assumptions on $\tilde{\psi}_t$ that lead to the posterior distribution $P(\chi_t^d|\chi_t^s)$. It can be seen that any consistent distribution of $\tilde{\psi}_t$ necessarily varies with the state of the world χ_t^d. But since $\tilde{\psi}_t$ is explicitly allowed to be dependent on $\tilde{\theta}_t$ and $\tilde{q}_t$ (and thus on $\tilde{\chi}_t^d = H(\tilde{\theta}_t - \log(\tilde{q}_t))$), the above distribution is well in line with our model specifications.[47] Moreover, by setting $z \to \infty$, we see that the distribution of $\tilde{\psi}_t$ is also proper whenever $P(\chi_t^s|\chi_t^d)$ is proper.

[47] Because $\tilde{w}_t$ and $\tilde{n}_t^d$ can both be written as functions of $\tilde{\theta}_t$ and $\tilde{q}_t$ and $\tilde{\Psi}_t$, we can also transform $\tilde{\psi}_t$ back to $\tilde{\Psi}_t = (\exp(\tilde{\psi}_t) - 1)\tilde{n}_t^s$.

Existence of a proper signal structure For a class of conditional distributions to be proper, we need that there exists some function g, such that:

$$\int P(\chi^s|\chi_t^d)\,\mathrm{d}\chi^s = 1,$$

or

$$\int g(\chi^s)P(\chi_t^d|\chi^s)\,\mathrm{d}\chi^s = 1 \tag{38}$$

for all $\chi_t^d \in \mathrm{supp}(\chi_t^d)$. This problem turns out to be quite challenging (it is equivalent to solving a Fredholm integral equation of kind 1). For the special case where Property 2b holds with $\tilde{X} \sim \mathcal{N}(0,1)$ (the baseline setup in this chapter), we can verify the existence numerically. More specifically, we use an algorithm that ensures that with a probability arbitrary close to 1, the economy realizes in a state such that $P(\chi_t^s|\chi_t^d)$ integrates arbitrary close to 1.[48] Based on our algorithm, we conjecture that—even if there does not exist an exact solution—there always exists a signal structure which is in this sense "almost" proper.

In summary, we conclude that there always exists an improper signal structure in line with Properties 1 and 2. The question whether there also exists a proper signal structure is analytically unclear. However, even if there does not exist a proper signal structure, then for the quasi-Gaussian case covered in most of this chapter, our numerical algorithm suggests that there always exists an "almost proper" one. This would then suggest that there also exists an almost identical proper signal structure, which does not satisfy our assumptions, but for which the results in this chapter nevertheless describe an arbitrary accurate solution. Or, alternatively, that there exists an almost identical proper signal structure, which does not satisfy our assumptions, but for which the results in this chapter describe the exact solution given that agents make arbitrary small errors by erroneously holding Gaussian beliefs.

[48] For any $\epsilon, \delta > 0$, we first define a set $A \subset \mathrm{supp}(\tilde{\chi}_t^d)$, such that $Pr(\chi_t^d \in A) > 1 - \epsilon$. Given this set, we then ensure that

$$\left| \int g(\chi^s)P(\chi_t^d|\chi^s)\,\mathrm{d}\chi^s - 1 \right| < \delta$$

for all $\chi_t^d \in A$. Given that this condition does not need to hold for $\chi_t^d \in \mathrm{supp}(\chi_t^d) \setminus A$, we have infinite many degrees of freedom in the tails of $P(\chi_t^s|\chi_t^d)$, which allow us to design g such that (38) holds for all $\chi_t^d \in A$ with arbitrary precision.

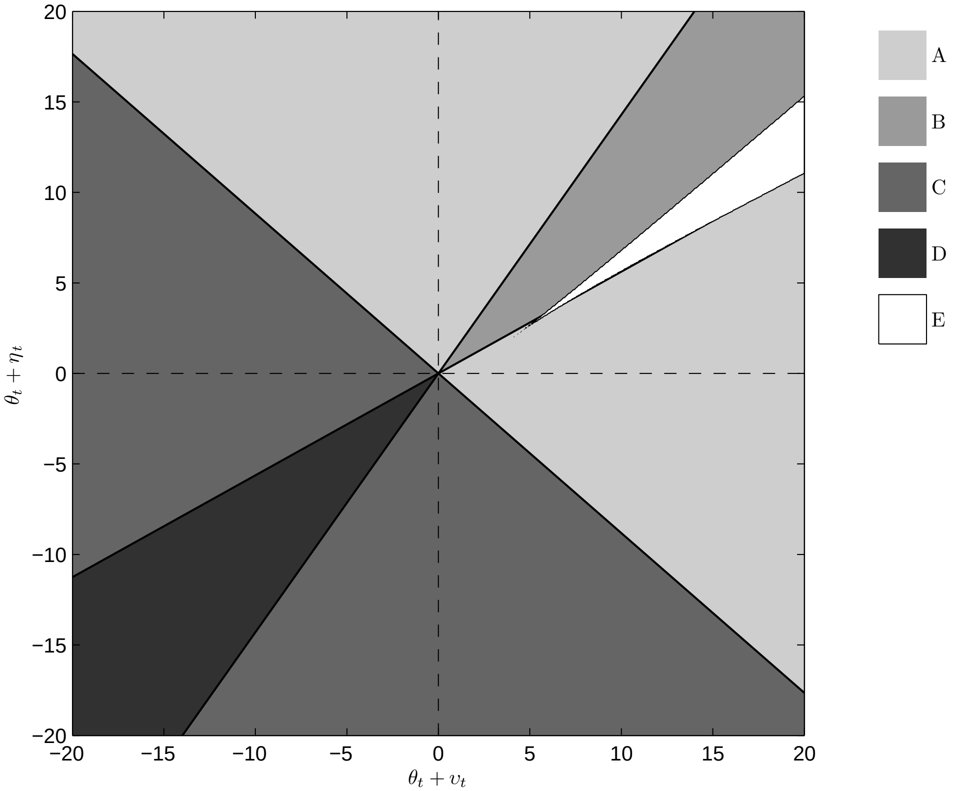

Figure 9. Sunspot regimes. *Note:* Shocks in regions A and B have an overall positive impact on $\mathring{q}_t$, shocks in regions C and D have an overall negative impact. Shocks in regions B and C are endogenously amplified, shocks in regions A and D are endogenously de-amplified. Region E defines the set $\bar{\Xi}$ of sunspot regimes.

B.4. Sunspot regimes

By Proposition 3, a necessary condition for sunspot regimes is that $\theta_t + v_t$ is sufficiently large (small) if $\gamma_1 > 1$ ($\gamma_1 < 1$) in absolute terms and also relative to $\theta_t + \eta_t$ (specifically, sunspots require that $\theta_t + \eta_t < \delta_1(\theta_t + v_t)$). If both of these conditions hold, then the economy could potentially (but not necessarily) be in a sunspot state ($\Omega_t \in \bar{\Xi}$). Figure 9 plots the set $\bar{\Xi}$ of sunspot regimes for the parameter set underlying Figure 3. In the figure, region E defines the set $\bar{\Xi}$. There are three equilibria in the interior of this set and two equilibria at the boundary. It can be shown that at any boundary of $\bar{\Xi}$, one of these equilibria is the continuation of the unique equilibrium outside Ξ. Moreover, by the proof of Proposition 4, there always exists at least one equilibrium within $\bar{\Xi}$ for which our analysis in Section 5.1 applies; i.e., Propositions 4 and 5 continue to hold for this equilibrium. For instance, in Figure 9, there always exists on equilibrium in the intersection of regions A and E in which $\mathring{q}_t > q_t^*$. Similarly, there always exists one equilibrium in the intersection of regions B and E in which $q_t^* > \mathring{q}_t$.[49] For these equilibria, all results in this study apply without any adjustment. Among the other equilibria, our results also continue to hold, but require an adjustment of the conditions that define the respective cases. For instance, for some equilibria in the intersection of regions A and E it holds that $q_t^* > \mathring{q}_t$. Accordingly, these equilibria are described by our characterization of region B rather than the one for region A. Accounting for that, all further results continue to hold.

References

Adam, K. (2007), "Optimal monetary policy with imperfect common knowledge," Journal of Monetary Economics, 54:267–301.

Albagli, E. (2011), "Amplification of Uncertainty in Illiquid Markets," mimeo.

Albagli, E., C. Hellwig, and A. Tsyvinski (2011), "A Theory of Asset Pricing Based on Heterogeneous Information," mimeo.

Amador, M. and P.-O. Weill (2008), "Learning from prices: Public communication and welfare," mimeo.

Amato, J. D. and H. S. Shin (2006), "Imperfect common knowledge and the information value of prices," Economic Theory, 27:213–241.

Angeletos, G.-M. and J. La'O (2012a), "Optimal Monetary Policy with Informational Frictions," mimeo.

——— (2012b), "Sentiments," mimeo.

Angeletos, G.-M., G. Lorenzoni, and A. Pavan (2010), "Beauty Contests and Irrational Exuberance: A Neoclassical Approach," mimeo.

Angeletos, G.-M. and A. Pavan (2004), "Transparency of information and coordination in economies with investment complementarities," The American Economic Review, 94:91–98.

——— (2007), "Efficient use of information and social value of information," Econometrica, 75:1103–1142.

——— (2009), "Policy with dispersed information," Journal of the European Economic Association, 7:11–60.

[49] Since $\mathring{q}_t$ is uniquely pinned down by Ω_t, $\mathring{q}_t > -\gamma_0$ holds for all equilibria in E.

Bachmann, R. and C. Bayer (2009), "Firm-Specific Productivity Risk over the Business Cycle: Facts and Aggregate Implications," mimeo.

Bachmann, R. and G. Moscarini (2011), "Business Cycles and Endogenous Uncertainty," mimeo.

Bernanke, B. and M. Gertler (1989), "Agency Costs , Net Worth , and Business Fluctuations," The American Economic Review, 79:14–31.

Bernanke, B., M. Gertler, and S. Gilchrist (1996), "The Financial Accelerator and the Flight to Quality," The Review of Economics and Statistics, 78:1–15.

——— (1999), "The financial accelerator in a quantitative business cycle framework," in Handbook of Macroeconomics, ed. by J. B. Taylor and M. Woodford, Amsterdam, New York and Oxford: North-Holland, vol. 1, 1341–1393.

Bigio, S. (2011), "Endogenous Liquidity and the Business Cycle," mimeo.

——— (2012), "Financial Risk Capacity," mimeo.

Bloom, N. (2009), "The Impact of Uncertainty Shocks," Econometrica, 77:623–685.

Bloom, N., M. Floetotto, N. Jaimovich, I. Saporta-Eksten, and S. Terry (2012), "Really Uncertain Business Cycles," mimeo.

Boissay, F., F. Collard, and F. Smets (2012), "Booms and Systemic Banking Crises," mimeo.

Brunnermeier, M. K. and Y. Sannikov (2012), "A macroeconomic model with a financial sector," mimeo.

Carlstrom, C. T. and T. S. Fuerst (1997), "Agency costs, net worth, and business fluctuations: A computable general equilibrium analysis," The American Economic Review, 87:893–910.

Christiano, L., R. Motto, and M. Rostagno (2009), "Financial Factors in Economic Fluctuations," mimeo.

Coibion, O. and Y. Gorodnichenko (2012), "Information Rigidity and the Expectations Formation Process : A Simple Framework and New Facts," mimeo.

Gertler, M. and P. Karadi (2011), "A model of unconventional monetary policy," Journal of Monetary Economics, 58:17–34.

Gertler, M. and N. Kiyotaki (2011), "Financial intermediation and credit policy in business cycle analysis," in Handbook of Monetary Economics, Vol. 3A, ed. by B. M. Friedman and M. Woodford, Amsterdam, New York and Oxford: North-Holland.

Gilchrist, S., J. W. Sim, and E. Zakrajšek (2010), "Uncertainty, Financial Frictions, and Investment Dynamics," mimeo.

Goldstein, I., E. Ozdenoren, and K. Yuan (2011), "Trading Frenzies and Their Impact on Real Investment," mimeo.

Hamilton, J. D. (1994), Time Series Analysis, Princeton University Press.

Hart, O. and J. Moore (1994), "A Theory of Debt Based on the Inalienability of Human Capital," The Quarterly Journal of Economics, 109:841–879.

——— (1998), "Default and Renegotiation: A Dynamic Model of Debt," The Quarterly Journal of Economics, 113:1–41.

Hassan, T. A. and T. M. Mertens (2011), "The social cost of near-rational investment," mimeo.

Hellwig, C. and L. L. Veldkamp (2009), "Knowing What Others Know: Coordination Motives in Information Acquisition," The Review of Economic Studies, 76:223–251.

Hellwig, M. F. (1980), "On the aggregation of information in competitive markets," Journal of Economic Theory, 22:477–498.

Iacoviello, M. (2005), "House prices, borrowing constraints, and monetary policy in the business cycle," The American Economic Review, 95:739–764.

Kiyotaki, N. and J. Moore (1997), "Credit cycles," Journal of Political Economy, 105:211–248.

——— (2008), "Liquidity, Business Cycles, and Monetary Policy," mimeo.

Krishnamurthy, A. (2003), "Collateral constraints and the amplification mechanism," Journal of Economic Theory, 111:277–292.

Kurlat, P. (2010), "Lemons, Market Shutdowns and Learning," mimeo.

La'O, J. (2010), "Collateral Constraints and Noisy Fluctuations," mimeo.

Lorenzoni, G. (2009), "A Theory of Demand Shocks," The American Economic Review, 99:2050–2084.

——— (2010), "Optimal Monetary Policy with Uncertain Fundamentals and Dispersed Information," The Review of Economic Studies, 77:305–338.

Lucas, R. E. (1978), "Asset prices in an exchange economy," Econometrica, 46:1429–1445.

Maćkowiak, B. and M. Wiederholt (2009), "Optimal Sticky Prices under Rational Inattention," The American Economic Review, 99:769–803.

——— (2010), "Business cycle dynamics under rational inattention," mimeo.

Mertens, T. M. (2011), "Volatile Stock Markets: Equilibrium Computation and Policy Analysis," mimeo.

Milgrom, P. R. (1981), "Good news and bad news: Representation theorems and applications," The Bell Journal of Economics, 12:380–391.

Morris, S. and H. S. Shin (2002), "Social value of public information," The American Economic Review, 92:1521–1534.

——— (2006), "Inertia of forward-looking expectations," The American economic review, 96:152–157.

Ordoñez, G. L. (2010), "Larger crises, slower recoveries: the asymmetric effects of financial frictions," mimeo.

Sim, J. W. (2008), "Uncertainty driven business cycle," mimeo.

Sims, C. A. (2003), "Implications of rational inattention," Journal of Monetary Economics, 50:665–690.

——— (2006), "Rational inattention: Beyond the linear-quadratic case," The American economic review, 96:158–163.

Townsend, R. M. (1979), "Optimal contracts and competitive markets with costly state verification," Journal of Economic Theory, 21:265–293.

——— (1983), "Forecasting the forecasts of others," The Journal of Political Economy, 91:546–588.

Van Nieuwerburgh, S. and L. L. Veldkamp (2006), "Learning asymmetries in real business cycles," Journal of Monetary Economics, 53:753–772.

Veldkamp, L. L. (2005), "Slow boom, sudden crash," Journal of Economic Theory, 124:230–257.

Woodford, M. (2003), "Imperfect common knowledge and the effects of monetary policy," in Knowledge, Information and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps, ed. by P. Aghion, R. Frydman, J. Stiglitz, and M. Woodford, Princeton: Princeton University Press, 25–58.

——— (2009), "Information-constrained state-dependent pricing," Journal of Monetary Economics, 56:100–124.

Yuan, K. (2005), "Asymmetric Price Movements and Borrowing Constraints: A Rational Expectations Equilibrium Model of Crises, Contagion, and Confusion," The Journal of Finance, 60:379–412.

Chapter 2

EMERGENCE AND PERSISTENCE OF EXTREME POLITICAL SYSTEMS — TRANSITION DYNAMICS IN AN UNRESTRICTED POLITY SPACE*

Contents

* This chapter is based on joint work with Lukas Buchheim. We would like to thank Toke Aidt, Matthias Doepke, Larry Christiano, Georgy Egorov, Florian Englmaier, Matthias Fahn, Florian Herold, Monika Piazzesi, Klaus M. Schmidt, Martin Schneider, Daniel Sturm, Uwe Sunde, Mirko Wiederholt, and Fabrizio Zilibotti for many helpful comments and discussions. We are also grateful for comments by seminar participants at Munich and Northwestern Universities, and the 2012 Silvaplana Workshop in Political Economy. A significant amount of the work leading to this chapter was done while I was visiting Northwestern in 2010. I thank the Economics and the Management and Strategy departments for their hospitality.

Abstract

This chapter of my dissertation introduces a dynamic model of political transitions, in which the outcomes of reforms and revolts are determined endogenously from a set of potential political systems that ranges continuously from single-man dictatorships to full-scale democracies. We find that while revolts result in autocracies, political reforms always enfranchise the majority of the population. Moreover, we show that democracies are intrinsically stable, leading to long episodes without political change. In contrast, autocracies are subject to frequent regime changes. Yet, our findings suggest that autocratic systems are persistent over time as they are frequently overthrown by small groups of insurgents, resulting in political systems similar to their predecessors. Taken together, our results imply that the long-run distribution of political systems is double hump-shaped with mass concentrated on the extremes. The model's predictions are consistent with descriptive statistics from cross-country data.

Keywords

Emergence of extreme political systems, endogenous political transitions, stability of political systems, unrestricted polity space.

JEL Classification: D74, D78, N40, P16.

1. Introduction

There is a growing economics literature exploring causes and circumstances of political transitions. This recent literature describes how political reforms and revolts can be supported within a rational-agents framework and how political systems are changed by these transition mechanisms. Even though many aspects of political transitions are dynamic in nature, these studies have largely abstracted from dynamic issues and focused on isolated transition events. This chapter of my dissertation takes a step towards filling this gap, by placing the dynamic process that describes the evolution of political systems at the core of the analysis. To this end, we endogenize outcomes of political transitions to a continuum of *a priori* attainable political systems and ensure the co-existence of reforms and revolts along the equilibrium path, allowing us to focus on the endogenously arising dynamic properties of political transitions.

More specifically, we introduce a dynamic framework where the space of political systems ranges continuously from single-man dictatorships to full-scale democracies. Actual political systems are determined endogenously and result from political transitions that either can be initiated from within a regime (i.e., reforms) or can be enforced from outside (revolts); the likelihood of transitions is thereby determined endogenously. Within this framework, we address the following key questions. Which types of political systems arise from reforms, and which arise from revolts? Similarly, through which of these transition mechanisms are particular systems such as democracies most likely to emerge? And how frequently is either type of transition observed depending on the political system in place?

Model overview Our modeling approach aims to resemble the key mechanisms behind political transitions explored in the literature, but generalizes them in order to ensure the co-existence of reforms and revolts along the equilibrium path and to endogenize their outcomes.

To endogenize political systems that emerge after revolts, we dispense with the simplifying approach of a representative "political outsider". Instead we consider an economy in which agents that are excluded from political power are heterogeneously adapted to the current regime. As is standard in the literature, political outsiders can attempt to acquire political power by supporting a subversive attempt against the regime. Prospects of subverting depend on, first, an unobserved ability of the regime to withstand such an attempt and, second, the total mass of outsiders supporting it. For deciding whether or not to support a revolt, agents weight these prospects against their individual adaptation utility to the current regime. As a consequence, a coordination game similar to the literature on global games endogenously determines the regime type after a successful revolt.[1]

Reforms are modeled as in the seminal paper by Acemoglu and Robinson (2000b) in that members of the current regime ("political insiders") may conduct preemptive reforms in order to alleviate the threat from a revolt. However, we generalize their original approach by permitting political insiders to enfranchise an arbitrary fraction of the population, allowing for a continuum of *a priori* unspecified political systems to emerge from these reforms.

Finally, we assume that while insiders are perfectly informed about their ability to withstand a revolt, outsiders are strictly less informed about the prospects of subverting. As a consequence, conducting reforms will be endogenously associated with being intrinsically weak, which in equilibrium helps outsiders to coordinate their actions. This effectively increases the costs of reforming and provides an incentive for weak regimes to take tough stance rather than to negotiate on moderate reforms. Because, in equilibrium, excessive repression translates into a substantial risk to be overthrown, asymmetric information, crucially, ensures the co-existence of reforms and revolts along the equilibrium path and allows us to jointly analyze these two transition mechanisms in our model.

Results Our first set of findings characterizes the political systems that endogenously arise in equilibrium. We show that while revolts result in autocracies where a minority of the population forms the ruling class, political reforms enfranchise the majority of the population and establish democratic political systems. Intermediate types of political regimes, by contrast, do not arise along the equilibrium path, so that political systems tend to be extreme.

Furthermore, this first set of results implies that democracies are only established from within regimes, giving theoretical support to a long-standing view in political science according to which members of former autocracies are key actors in the establishment of democracies (Rustow, 1970; O'Donnell and Schmitter, 1973; Huntington, 1991). Or, as Karl (1990, p. 8) puts it: "no stable political democracy [in South America] has resulted from regime transitions in which mass

[1] Although outsiders in our model share the same amount of information, we use heterogeneous opportunity costs to ensure that subverting and not subverting is always a dominant strategy for some outsiders in our model. Iterated elimination of (interim) dominated strategies then gives rise to a unique outcome of this coordination game. This is essentially the same mechanism that determines equilibria in global games.

actors have gained control, even momentarily, over traditional ruling classes".

Our second set of results concerns the stability and persistence of political systems. From our analysis it follows that democratic regimes are intrinsically stable, characterized by long episodes without political change. In contrast, autocracies are subject to frequent regime changes—either via revolts or reforms. This is in line with the empirical literature on regime stability, which observes that democratic political systems are significantly more stable than autocratic ones (Przeworski, 2000; Gates et al., 2006; Magaloni and Kricheli, 2010).[2]

Nevertheless, our findings suggest that despite their instability, autocratic systems are persistent over time. This is because even though single autocratic regimes are relatively short lived, political change is frequently initiated by a small group of insurgents, resulting in autocracies very similar to their predecessors. Interestingly, this reasoning further implies that revolts tend to be serially correlated over time as they go along with a selection into politically instable regimes, leading to periods of political instability.

In combination, our results imply that the long-run distribution of political systems is double hump-shaped with mass concentrated on extreme political systems. Our model thus provides a foundation to the empirically observed distribution of political systems since World War I, plotted in Figure 1.[3] Taking a look at the underlying dataset (for details, see Section 5), we also find similar support for the findings outlined above.

Related literature So far, the literature on political transitions has primarily focused on developing arguments for why autocratic regimes may conduct democratic reforms. Bourguignon and Verdier (2000), Lizzeri and Persico (2004), and Llavador and Oxoby (2005) argue that reforms are reflective of situations where autocratic decision makers are better off in a democratized political system than under the status quo. A number of other studies are based on the idea of preemptive reforms introduced by Acemoglu and Robinson (2000b) (e.g.,

[2] From these results it follows that the mode of transition—peaceful reforms or violent revolts—is important for the characteristics of the resulting regimes. For transitions to democracy, a similar point has been highlighted by Cervellati et al. (2007, 2011), who show that consensual transitions foster civil liberties and property rights provision in contrast to violent transitions.

[3] The underlying data is taken from the Polity IV Project (for details, see Section 5). It has been disputed whether intermediate scored regimes on this index should nevertheless be classified as either democratic or autocratic due to nonlinearities in the index (Cheibub et al., 2010). Note, however, that this is to say that different measurements would only lead to more mass on the extremes, not altering the basic conclusion drawn for our purposes.

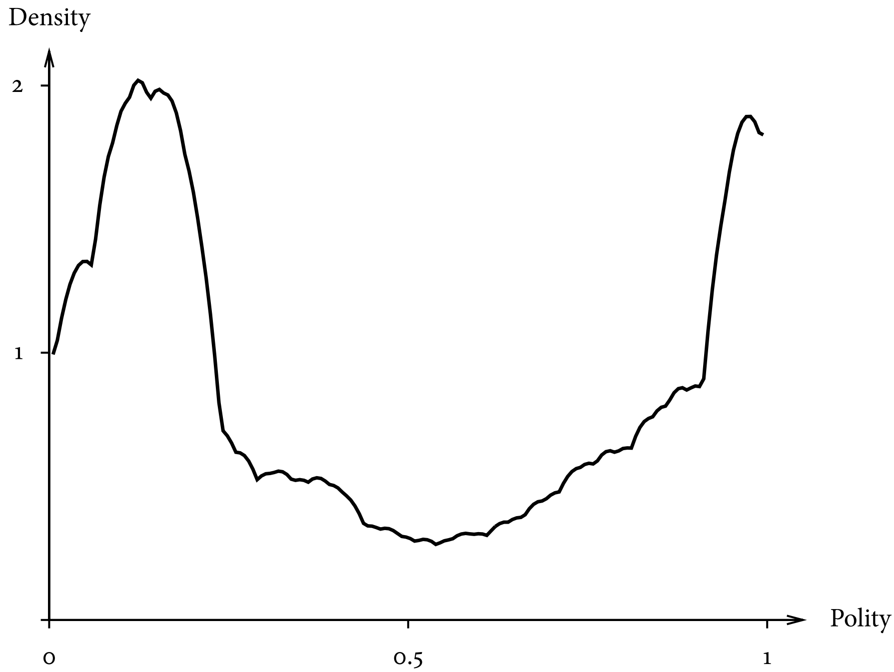

Figure 1. Distribution of political systems since World War I. Political systems range from extremely autocratic (0) to extremely democratic (1). Units of observation are country-days.

Conley and Temini, 2001; Boix, 2003). These papers share with ours the basic logic behind reforms; i.e., autocratic regimes may use political reforms to credibly commit to redistribution and to reduce revolutionary pressure.[4]

In contrast to these papers, the emphasis of this analysis is on the dynamics of political transitions, including but not restricted to democratization. In this respect, this chapter relates more closely to Acemoglu and Robinson (2001) and Acemoglu et al. (2010), who consider settings where preemptive reforms co-exist with coups along the equilibrium path, and to Ellis and Fender (2011), who consider preemptive reforms that co-exist with mass revolutions. In particular, Ellis and Fender choose a similar approach in studying how autocracies may strategically manipulate the degree of subversive coordination in the presence of asymmetric information. In their model, outsiders sequentially choose whether or not to support a subversive attempt, which succeeds only if it is unanimously

[4] See Aidt and Jensen (2012) and Przeworski (2009) for empirical studies suggesting that subversive threats are indeed the driving force behind democratization.

supported. They find that asymmetric information provides an incentive to refrain from stabilizing reforms despite the presence of revolutionary pressure (see also Acemoglu and Robinson, 2000a; and for information manipulation in global games, see Angeletos et al., 2006 and Edmond, 2011).

However, all of these papers have in common that they exogenously restrict the set of political systems that result from transitions. In contrast, our approach of an unrestricted space of political systems leaves the outcomes of reforms and revolts unspecified. This is central to our analysis, allowing us to endogenously derive the properties of these transition mechanisms and to analyze their implications for the stability and persistence of political systems.

We also relate to Justman and Gradstein (1999), Jack and Lagunoff (2006), and Gradstein (2007), who study the incentives of political regimes to conduct democratic reforms in frameworks in which—as in our approach—continuous extensions of the franchise are possible. Similar to the literature discussed above, these authors provide conditions under which (possibly gradual) extensions of the franchise are to be expected. In contrast to our work, however, they do not allow for political change to be initiated from political outsiders (via revolts), preventing them from analyzing transition dynamics in their generality that follows from the interplay between reforms and revolts, which is at the core of our contribution.

Outline The remainder of the chapter is organized as follows. Section 2 introduces the model economy. In Section 3, we characterize the equilibrium and illustrate the strategic considerations determining political transitions. The law of motion of the dynamic economy and our main predictions are derived in Section 4. In Section 5, we present some empirical evidence, and Section 6 concludes.

2. The model

We consider an infinite horizon economy with a continuum of two-period lived agents. Each generation has a mass equal to 1. At time t, fraction λ_t of the population has the power to implement political decisions, whereas the remaining agents are excluded from political power. We refer to these two groups as (political) "insiders" and "outsiders".

When born, the distribution of political power among the young is inherited from their parent generation; that is, λ_t agents are born as insiders, while $1 - \lambda_t$

agents are born as outsiders. However, agents who are born as outsiders can attempt to overthrow the current regime and thereby acquire political power. To this end, outsiders choose individually and simultaneously whether or not to participate in a revolt.[5] Because we will assume that all political change takes effect at the beginning of the next period, only young outsiders have an interest in participating in a revolt. Accordingly, we denote young outsider i's choice by $\phi_{it} \in \{0,1\}$ and use the aggregated mass of supporters, $s_t = \int \phi_{it}\,\mathrm{d}i$, to refer to the size of the resulting revolt.

The probability that a revolt is successful is given by

$$p(\theta_t, s_t) = \theta_t h(s_t), \tag{1}$$

where $\theta_t \in \Theta$ is a random state of the world that reflects the vulnerability of the current regime or their ability to put down a revolt, and h is an increasing and twice differentiable function, $h : [0,1] \to [0,1]$, with $h(0) = 0$. That is, the threat of a revolt to the current regime is increasing in the mass of its supporters and in the vulnerability of the regime. When a revolt has no supporters ($s_t = 0$) or the regime is not vulnerable ($\theta_t = 0$), it fails with certainty.

The purpose of θ_t in our model is to introduce asymmetric information between insiders and outsiders that, as will become clear below, explains the prevalence of revolts along the equilibrium path. Formally we have that the state θ_t is uniformly distributed on $\Theta = [0,1]$, is i.i.d. from one period to the next, and is revealed to insiders at the beginning of each period. Outsiders only know the prior distribution of θ_t.

After they learn θ_t, insiders may try to alleviate the threat of revolt by conducting reforms. We follow Acemoglu and Robinson (2000b) by modeling these reforms as an extension of the franchise to outsiders, which is effective in credibly preventing them from supporting a revolt.[6] However, since our model is aimed at endogenizing the political system λ_t, we generalize this mechanism by allowing insiders to continuously extend the regime by any fraction, $x_t - \lambda_t$, of young outsiders, where $x_t \in [\lambda_t, 1]$ is the reformed political system.[7] Because

[5] For notational convenience, we abstract from the possibility of insiders participating in a revolt. In Appendix A.1, however, we show that this is without loss of generality, since it is never optimal for insiders to support a revolt against fellow members of the regime.

[6] As argued in Footnote 5 and shown in Appendix A.1, it is indeed individually rational for enfranchised outsiders to not support a revolt.

[7] Note that by assuming $x_t \in [\lambda_t, 1]$, we are ruling out reforms that withdraw political power once it has been granted. This is in line with the idea that granting someone the status of an

preferences of insiders will be perfectly aligned, there is no need to specify the decision making process leading to x_t in detail.

Given the (aggregated) policy choices s_t and x_t, and conditional on the outcome of a revolt, the political system evolves as follows:

$$\lambda_{t+1} = \begin{cases} s_t & \text{if the regime is overthrown, and} \\ x_t & \text{otherwise.} \end{cases} \tag{2}$$

When a revolt fails (indicated by $\eta_t = 0$), reforms take effect and the old regime stays in power. The resulting political system in $t+1$ is then given by x_t. In the complementary case, when a revolt succeeds ($\eta_t = 1$), those who have participated will form the new regime. Accordingly, after a successful revolt, the fraction of insiders at $t+1$ is equal to s_t. Note that this specification prevents non-revolting outsiders from reaping the benefits from overthrowing a regime; there are no gains from free-riding in our model.[8]

To complete the description of our model, we still have to specify how payoffs are distributed across the different groups of agents at t. As for outsiders, we assume that they receive a constant per period payoff of γ_{it} which is privately assigned to each agent at birth and is drawn from a uniform distribution on $[0,1]$. We interpret this heterogeneity of outsiders as different degrees of economical or ideological adaptation to a regime, determining their propensity to revolt.

In contrast, insiders enjoy per period payoffs $u(\lambda_t)$, where u is twice differentiable, $u' < 0$, and $u(1)$ is normalized to unity. One should think of $u(\cdot)$ as a reduced form function that captures the various benefits of having political power (e.g., from extracting a common resource stock, implementing preferred policies, etc.). One important feature of u is that it is decreasing in the current regime size and, hence, extending the regime is costly for insiders (e.g., because resources have to be shared, or preferences about policies become less aligned). Another thing to note is that $u(\lambda_t) \geq \gamma_{it}$ for all λ_t and γ_{it}; that is, being part of the regime is always desirable. In the case of full democracy ($\lambda_t = 1$) all citizens

---

insider is a credible and irreversible commitment in the logic of Acemoglu and Robinson (2000b).

8 The theoretical possibility for free-riding arises since we depart from the common assumption of treating the opposition as a single player in order to endogenize the political system resulting from a revolt. However, as long as there are some private benefits that provide incentives for outsiders to support a revolt, the working of this model is unaffected by (moderate) incentives to free-ride. Entirely abstracting from the collective action problem is merely a model simplification.

are insiders and enjoy utility normalized to the one of a best-adapted outsider (i.e., $u(1) = 1$).

To simplify the analysis, we assume that members of an overthrown regime and participants in a failed revolt are worst-adapted to the new regime. Formally, $\gamma_{it} = 0$, resulting in zero payoff.

For the upcoming analysis it will be convenient to define the expected utility of agents that are born at time t, which is given as follows:

$$V^I(\theta_t, \lambda_t, s_t, x_t) = u(\lambda_t) + [1 - p(\theta_t, s_t)] \times u(x_t), \tag{3}$$

$$V^O(\theta_t, \gamma_{it}, s_t, \phi_{it}) = \gamma_{it} + \phi_{it}\, p(\theta_t, s_t) \times u(s_t) + (1 - \phi_{it}) \times \gamma_{it}, \tag{4}$$

where superscript I and O denote agents that are born as insiders and outsiders, respectively. In both equations, the first term corresponds to the first period payoff (unaffected by the policy choices of the young agent's generation), while the other terms correspond to second period payoffs. (Since agents do not face an intertemporal tradeoff, we do not need to define a discount rate here).

The timing of events within one period can be summarized as follows:

1. The state of the world θ_t is revealed to insiders.

2. Insiders may extend political power to a fraction $x_t \in [\lambda_t, 1]$ of the population.

3. Outsiders individually and simultaneously decide whether or not to participate in a revolt.

4. Transitions according to (1) and (2) take place, period $t + 1$ starts with the birth of a new generation, and payoffs determined by λ_{t+1} are realized.

In what follows, we characterize the set of perfect Bayesian equilibria that satisfy the trembling-hand criterion (due to Selten, 1975); that is, perfect Bayesian equilibria that are the limit of some sequence of perturbed games in which strategy profiles are constrained to embody "small" mistakes.[9] To increase the predictive

[9] Here, the concept of trembling-hand perfection rules out "instable" equilibria, in which $s_t = 0$, but iteratively best-responding to a (perceived) second-order perturbation of s_t would lead to a different equilibrium with a first-order change in s_t. For details see the proof of Proposition 1. Except for these instabilities, the set of trembling-hand perfect equilibria coincides with the set of perfect Bayesian equilibria in our model. An alternative approach to rule out these instabilities would be to restrict attention to equilibria which are the limit to a sequence of economies with a finite number of outsiders, where each agent's decision has non-zero weight on s_t.

power of our model, we thereby limit attention to equilibria that are consistent with the D1 criterion introduced by Cho and Kreps (1987), a standard refinement for signaling games. The D1 criterion restricts outsiders to believe that whenever they observe a reform x' that is not conducted in equilibrium, the reform has been implemented by a regime with vulnerability θ', for which a deviation to x' would be most attractive.[10]

Anticipating our results, we simplify our notation as follows. First, outsiders' beliefs regarding the regime's vulnerability will be uniquely determined in our setup. We therefore denote the commonly held belief by $\hat{\theta}_t$, dropping the index i. Second, there are no nondegenerate mixed strategy equilibria in our game. Accordingly, we restrict the notation in the main text to pure strategies and introduce mixed strategies only to define the perturbations required by trembling-hand perfection.

This leads to the following definition of equilibrium for our economy.

Definition: Given a history $\delta = \{\lambda_0\} \cup \{\{\phi_{i\tau} : i \in [0,1]\}, \theta_\tau, x_\tau, \eta_\tau\}_{\tau=0}^{t-1}$, an equilibrium in this economy consists of policy mappings $x_\delta : (\theta_t, \lambda_t) \mapsto x_t$ and $\{(\phi_{i\delta} : (\hat{\theta}_t, x_t) \mapsto \phi_{it}) : i \in [0,1]\}$, and beliefs $\hat{\theta}_\delta(\lambda_t, x_t) \mapsto \hat{\theta}_t$, such that for all possible histories δ:

a. Reforms x_δ maximize insider's utility (3), given states (θ_t, λ_t), beliefs $\hat{\theta}_\delta$, and perturbed policy mappings $\{\omega_{i\delta}^k : i \in [0,1]\}$ for all values of k;

b. Each outsider's policy choice $\phi_{i\delta}$ maximizes (4), given perturbed policy mappings σ_δ^k, $\{\omega_{j\delta}^k : j \in [0,1] \setminus i\}$, and corresponding beliefs $\hat{\theta}_\delta^k$ for all values of k;

c. Beliefs $\hat{\theta}_\delta = \lim_{k\to\infty} \hat{\theta}_\delta^k(x_t)$, where $\hat{\theta}_\delta^k$ are obtained using Bayes rule given σ_δ^k; and $\hat{\theta}_\delta$ satisfies the D1 criterion;

d. States (λ_t, η_t) are consistent with (1) and (2);

[10] Formally, let $\bar{V}^I(\theta', \lambda_t)$ be the insiders' payoff in a candidate equilibrium when the regime has a vulnerability θ'. Then the D1 criterion restricts beliefs to the state θ' that maximizes $D_{\theta',x'} = \{\hat{\theta} : V^I(\theta', \lambda_t, s(\hat{\theta}, x'), x') \geq \bar{V}^I(\theta', \lambda_t)\}$, where $s(\hat{\theta}, x')$ is the mass of outsiders supporting a revolt, given the beliefs $\hat{\theta}$ and reform x'. $D_{\theta',x'}$ is maximal here, if there is no θ'', such that $D_{\theta',x'}$ is a proper subset of $D_{\theta'',x'}$. That is, beliefs are attributed to the state in which a deviation to x' is attractive for the largest set of possible inferences about the regime's vulnerability (implying that the regime gains most by deviating).

e. The perturbed policy mappings $\{\{\omega^k_{i\delta} : i \in [0,1]\}, \sigma^k_\delta\}^\infty_{k=0}$ are sequences of completely mixed strategy profiles converging to profiles that place all mass on $\{\phi_{i\delta} : i \in [0,1]\}$ and x_δ, respectively.

3. Political equilibrium

In this section, we characterize the political equilibrium in the model economy. Our analysis will be simplified considerably by the overlapping generations structure of our model, which gives rise to a sequence of "generation games" between young insiders and young outsiders. Since the distribution of political power at time t captures all payoff-relevant information of the history up to t, the only link between generations is λ_t. We can therefore characterize the set of equilibria in our model by characterizing the equilibria of the generation games as a function of λ_t. All other elements of the history up to time t may affect the equilibrium at t only by selecting between multiple equilibria of the generation game.

The generation game consists of two stages that determine the political system at $t+1$. First, outsiders have to choose whether or not to support a revolt. Because the likelihood that a revolt succeeds depends on the total mass of its supporters, outsiders face a coordination problem in their decision to revolt. Second, prior to this coordination problem, insiders decide on the degree to which political power is extended to outsiders. On the one hand this will decrease revolutionary pressure along the extensive margin by contracting the pool of potential insurgents. However, extending the regime may also contain information about the regime's vulnerability. As a result, reforms may increase revolutionary pressure along the intensive margin by increasing coordination among outsiders who are not subject to reforms. Insiders' policy choices will therefore be governed by signaling considerations.

We proceed by backward induction in solving for the equilibrium, beginning with the outsiders' coordination problem.

3.1. *Stage 2: Coordination among outsiders*

Consider the outsiders' coordination problem at time t. For any given belief, $(\hat{\theta}_t, \hat{s}_t) \in \Theta \times [0,1]$, individual rationality requires all outsiders to choose a ϕ_{it}

that maximizes their expected utility $\mathrm{E}_t\{V^O(\cdot)\}$.[11] At time t, outsider i with adaptation utility γ_{it} will therefore participate in a revolt if and only if

$$\gamma_{it} \leq p(\hat{\theta}_t, \hat{s}_t)\, u(\hat{s}_t) \equiv \bar{\gamma}(\hat{s}_t). \tag{5}$$

Here $\bar{\gamma}(\hat{s}_t)$ is the expected benefit of participating in a revolt that is supported by a mass of $\hat{s}_t$ outsiders. Since $\bar{\gamma}(\hat{s}_t)$ is independent of γ_{it}, it follows that in any equilibrium the set of outsiders who support a revolt at t is given by the agents who are least adapted to the current regime. Suppose for the time being that $\bar{\gamma}(\hat{s}_t) \leq 1$. Then, $\bar{\gamma}(\hat{s}_t)$ defines the fraction of young outsiders that participates in a revolt, and, therefore, the size of a revolt, s_t, that would follow from $\bar{\gamma}(\hat{s}_t)$ is given by

$$f(\hat{s}_t) \equiv (1 - x_t)\, \bar{\gamma}(\hat{s}_t). \tag{6}$$

Further note that in any equilibrium it must hold that $s_t = \hat{s}_t$. Therefore, as long as $\bar{\gamma}(\hat{s}_t) \leq 1$, the share of outsiders that support a revolt at t has to be a fixed point to (6). To guarantee that this is always the case and to further ensure that a well-behaved fixed point exists, we impose the following assumption.

Assumption 1: *For* $\psi(s) \equiv h(s) \cdot u(s)$,

a. $\psi' \geq 0$ *and* $\psi'' \leq 0$;

b. $\lim_{s \to 0} \psi'(s) = \infty$.

Intuitively, Assumption 1 states that participating in a revolt becomes more attractive if the total share of supporters grows. This requires that the positive effect of an additional supporter on the success probability outweighs the negative effect of being in a slightly larger regime after a successful revolt. Put differently, Assumption 1 states that the participation choices of outsiders are strategic complements. To ensure existence, we further require that the strategic complementarity is sufficiently strong when a revolt is smallest, and is decreasing as it grows larger.

Using Assumption 1, the above discussion leads to the following proposition.

[11] Note that by our specification of p, V^O is linear in θ_t, and thus $\mathrm{E}_t\{V^O(\theta_t, \cdot)\} = V^O(\hat{\theta}_t, \cdot)$, where $\hat{\theta}_t \equiv \mathrm{E}_t\{\theta_t\}$. That is, the expected value of θ_t, given the posterior distribution of θ_t (outsiders' beliefs), is a sufficient statistic for computing V^O. Henceforth we define $\hat{\theta}_t$ accordingly, disregarding any higher moments of outsiders' beliefs.

Proposition 1: *In any equilibrium, the mass of outsiders supporting a revolt at time t is uniquely characterized by a time-invariant function, $s : (\hat{\theta}_t, x_t) \mapsto s_t$, which satisfies $s(0, \cdot) = s(\cdot, 1) = 0$, increases in $\hat{\theta}_t$, and decreases in x_t.*

All formal proofs are in the appendix. Proposition 1 establishes the already discussed tradeoff of conducting reforms: On the one hand, reforms reduce support for a revolt along the extensive margin. In the limit, as regimes reform to a full-scaled democracy, any subversive threat is completely dissolved. On the other hand, if reforms signal that the regime is vulnerable, they may backfire by increasing support along the intensive margin.

3.2. Stage 1: Policy choices of insiders

We now turn to the insiders' decision problem. Since more vulnerable regimes have higher incentives to reform than less vulnerable ones, conducting reforms will shift beliefs towards being vulnerable and, therefore, indeed stipulate coordination among outsiders who are unaffected by reforms. This generates the tradeoff established in Proposition 1, which is the main driving force behind the following result.

Proposition 2: *In any equilibrium, policy choices of insiders and beliefs of outsiders are uniquely characterized by time-invariant functions $x : (\theta_t, \lambda_t) \mapsto x_t$ and $\hat{\theta} : (\lambda_t, x_t) \mapsto \hat{\theta}_t$, such that*

$$x(\theta_t, \lambda_t) = \begin{cases} \lambda_t & \text{if } \theta_t < \bar{\theta}(\lambda_t) \\ \xi(\theta_t) & \text{if } \theta_t \geq \bar{\theta}(\lambda_t), \end{cases}$$

and

$$\hat{\theta}(\lambda_t, x_t) = \begin{cases} \bar{\theta}(\lambda_t)/2 & \text{if } x_t = \lambda_t \\ \bar{\theta}(\lambda_t) & \text{if } \lambda_t < x_t < \xi(\bar{\theta}(\lambda_t)) \\ \xi^{-1}(x_t) & \text{if } \xi(\bar{\theta}(\lambda_t)) \leq x_t \leq \xi(1) \\ 1 & \text{if } x_t > \xi(1), \end{cases}$$

where ξ is a unique increasing function with $\xi(\theta_t) > \lambda_t + \mu$, and $\bar{\theta}(\lambda_t) > 0$ for all λ_t and some $\mu > 0$.

Proposition 2 defines insiders' policy choices for generation t as a function of (θ_t, λ_t). Because the logic behind these choices is the same for all values of λ_t,

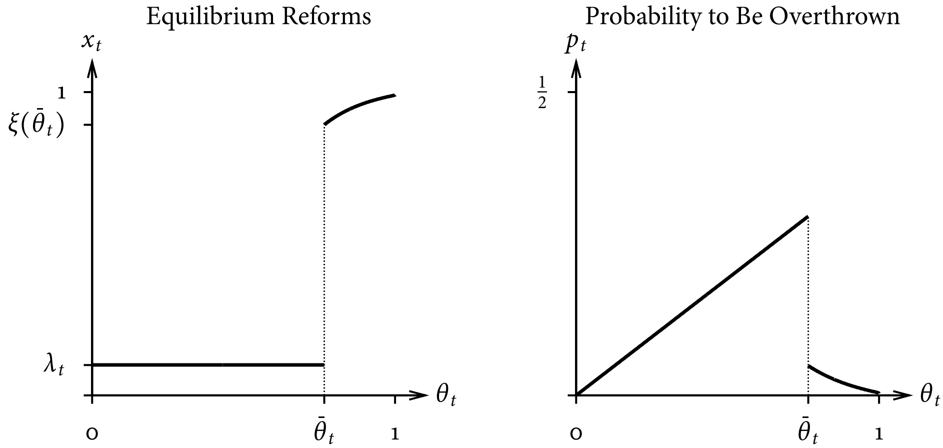

Figure 2. Equilibrium reforms and implied probability to be overthrown.

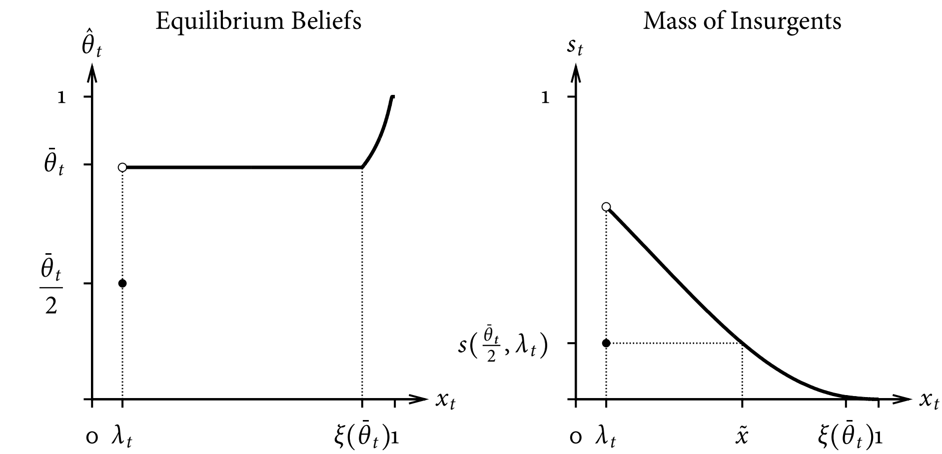

Figure 3. Equilibrium beliefs and implied mass of insurgents.

we can discuss the underlying intuition keeping λ_t fixed. Accordingly, in Figure 2 we plot reform choices (left panel) and the implied probability to be overthrown (right panel), sliced along a given λ_t plane. It can be seen that whenever a regime is less vulnerable than $\bar{\theta}(\lambda_t)$, insiders prefer to not conduct any reforms (i.e., $x_t = \lambda_t$), leading to a substantial threat for regimes with θ_t close to $\bar{\theta}(\lambda_t)$. Only if $\theta_t \geq \bar{\theta}(\lambda_t)$, reforms will be conducted ($x_t = \xi(\theta_t)$), which in equilibrium effectively mitigate the threat to be overthrown, ruling out marginal reforms where $\xi(\theta_t) \to \lambda_t$.

To see why marginal reforms are not effective in reducing revolutionary pressure consider Figure 3. Here we plot equilibrium beliefs (left panel) and the corresponding mass of insurgents (right panel) as functions of x_t. If the political system is left unchanged by insiders, outsiders only learn the average state $\bar{\theta}(\lambda_t)/2$ of all regimes that pool on $x_t = \lambda_t$ in equilibrium. On the other hand, every extension of the regime—how small it may be—leads to a non-marginal change in outsiders' beliefs from $\hat{\theta}_t = \bar{\theta}(\lambda_t)/2$ to $\hat{\theta}_t \geq \bar{\theta}(\lambda_t)$ and, hence, results in a non-marginal increase in revolutionary pressure along the intensive margin. It follows that there exists some $\tilde{x}(\lambda_t)$, such that for all $x_t < \tilde{x}(\lambda_t)$ the increase of pressure along the intensive margin dominates the decrease along the extensive margin. Thus, reforms smaller than $\tilde{x}(\lambda_t)$ will backfire and *increase* the mass of insurgents (as seen in the right panel of Figure 3), explaining why effective reforms have to be non-marginal.

Furthermore, optimality of reforms requires that the benefit of reducing pressure compensates for insiders' disliking of sharing power. Because $\tilde{x}(\lambda_t) - \lambda_t > 0$, it follows that $u(\tilde{x}(\lambda_t)) - u(\lambda_t) < 0$. Moreover, any reform marginally increasing the regime beyond $\tilde{x}(\lambda_t)$ leads only to a marginal increase in the likelihood to stay in power. Hence, there exists a non-empty interval, given by $[\tilde{x}(\lambda_t), \xi(\bar{\theta}(\lambda_t))]$, in which reforms are effective, yet insiders prefer to gamble for their political survival in order to hold on to the benefits of not sharing power in case they survive. This explains the substantial threat for regimes with θ_t close to $\bar{\theta}(\lambda_t)$, as seen in the right panel of Figure 2.[12]

3.3. Existence and uniqueness of equilibrium

Propositions 1 and 2 uniquely pin down the policy choices in every state, which in return determine the evolution of political systems. We conclude that there is no scope for multiple equilibria in our model economy; if there exists an equilibrium it is unique. Verifying the existence then permits us to reach the following result.

PROPOSITION 3: *There exists an equilibrium, in which for all histories δ, policy mappings x_δ and $\{\phi_{i\delta} : i \in [0,1]\}$, as well as beliefs $\hat{\theta}_\delta$ correspond to the time-invariant mappings given by Propositions 1 and 2. Furthermore, for any given initial political system λ_0, this equilibrium is unique.*

[12] More precisely, gambling for survival increases the likelihood to be overthrown in two ways. First, since on the margin it is more vulnerable regimes that join the pool at $x_t = \lambda_t$, these regimes obviously face a high threat by not conducting reforms. Second, since these regimes also shift the pooling belief towards more vulnerable, the threat further increases for regimes of all vulnerabilities in the pool.

4. Transition dynamics

In the preceding section, we have established that in the unique equilibrium, policy mappings are time-invariant, implying that (λ_t, θ_t) is a sufficient statistic for characterizing the transition dynamics of the political system from time t to $t+1$. Integrating out θ_t, political systems in our equilibrium follow a Markov process where the probability that $\lambda_{t+1} \in \Lambda$ is given by

$$Q(\lambda_t, \Lambda) = \rho^S(\lambda_t) \times Q^S(\lambda_t, \Lambda) + \rho^R(\lambda_t) \times Q^R(\lambda_t, \Lambda) + \{1 - \rho^I(\lambda_t) - \rho^R(\lambda_t)\} \times \mathbb{1}_{\lambda_t \in \Lambda}. \quad (7)$$

Here ρ^S and ρ^R denote the probabilities that a transition occurs via subversive attempts (i.e., revolts) and reforms, respectively; Q^S and Q^R are conditional transition functions; and $\mathbb{1}$ is an indicator function equal to unity whenever $\lambda_t \in \Lambda$.[13] Accordingly, the first term in (7) defines the probability that state $\lambda_{t+1} \in \Lambda$ emerges through a revolt, the second term defines the probability that $\lambda_{t+1} \in \Lambda$ emerges from a reform, and the third term refers to the event of no transition. Decomposing the law of motion into these conditional channels, we are now ready to state our main predictions.

4.1. *Political systems after transition*

By (7), political systems that arise after transitions are summarized by Q^S and Q^R. Our first result states that political systems that emerge after reforms differ fundamentally from those that emerge from revolts. The following proposition states the formal result.

Proposition 4: *For all states λ_t,*

$$Q^R(\lambda_t, (\tfrac{1}{2}, 1]) = 1 \qquad \textit{and} \qquad Q^S(\lambda_t, (0, \tfrac{1}{2})) = 1;$$

[13] Formally, we have that

$$\begin{aligned} \rho^S(\lambda_t) &= \textstyle\int_0^1 \grave{p}(\theta)\, \mathrm{d}\theta \\ \rho^R(\lambda_t) &= \textstyle\int_{\bar{\theta}(\lambda_t)}^1 \{1 - \grave{p}(\theta)\}\, \mathrm{d}\theta \\ Q^S(\lambda_t, \Lambda) &= \{\rho^S(\lambda_t)\}^{-1} \textstyle\int_{\theta:\grave{s}(\theta) \in \Lambda} \grave{p}(\theta)\, \mathrm{d}\theta \\ Q^R(\lambda_t, \Lambda) &= \{\rho^R(\lambda_t)\}^{-1} \textstyle\int_{\theta:\grave{x}(\theta) \in \Lambda \setminus \lambda_t} \{1 - \grave{p}(\theta)\}\, \mathrm{d}\theta, \end{aligned}$$

where $\grave{x}(\theta) \equiv x(\lambda_t, \theta)$, $\grave{s}(\theta) \equiv s(\hat{\theta}(\lambda_t, \grave{x}(\theta)), \grave{x}(\theta))$, and $\grave{p}(\theta) \equiv p(\theta, \grave{s}(\theta))$.

i.e., reforms lead to majority regimes with $\lambda_{t+1} > \frac{1}{2}$ *and revolts lead to minority regimes with* $\lambda_{t+1} < \frac{1}{2}$.

The first part of Proposition 4 states that any reform leads to a democratic system, in which the majority of citizens holds political power. The intuition for this result mirrors the one for Proposition 2. Because conducting reforms will be associated with being intrinsically weak, coordination is increased along the intensive margin. For the benefits along the extensive margin to justify these costs, reforms therefore have to be far-reaching, leading to the enfranchising of the majority of the population.

In contrast, the second part of Proposition 4 establishes that successful revolts always lead to minority regimes, in which a small elite rules over a majority of political outsiders. Underlying this result is that in equilibrium subversive attempts are conducted by only a small group of insurgents. Mass revolutions on the other hand are off-equilibrium. To see what drives the result, first note that rationality of reforms implies that revolts are largest when regimes abstain from reforms and choose to repress the population. However, because abstaining from reforms is optimal, both, in times when regimes are strong and when they hide their weakness through taking tough stance, uncertainty about a regime's weakness is largest from the perspective of outsiders exactly when a regime abstains from reforms. Accordingly, prospects of revolting are only moderate and only those with large gains from winning political power (i.e., outsiders who are least adapted to the current regime) will find it rational to take the risk of revolting.

An interesting implication of Proposition 4 is that democratic regimes arise if and only if it is optimal for the regime to enfranchise former political outsiders. The commonly made assumption in the previous literature that democracies are established by means of reforms conducted by the elites is thus an endogenous outcome of our model. The other channel through which democracies hypothetically could be established are mass revolutions. Their severe threat, however, is always mitigated by rational regimes, such that mass revolts are events off the equilibrium path. This observation gives support to a long-standing view in political science according to which members of former autocracies are key actors in the establishment of democracies, which is based on, e.g., the observation of Karl (1990, p. 8) that no stable South American democracy has been the result of mass revolutions (see also Rustow, 1970; O'Donnell and Schmitter, 1973; Huntington, 1991).

Finally, note that from Proposition 4 it follows that there is a (possibly quite large) open interval $\bar{\Lambda}$ around 1/2, such that $Q(\lambda_t, \bar{\Lambda}) = 0$ for all λ_t. That is, there is a range of intermediate regimes that are completely off the equilibrium path, suggesting a long-run distribution with mass only on the extremes. In a parametric example below, we will see that this is indeed the case.

4.2. Probabilities of transition

The next proposition describes how the likelihood of either type of political transition depends on the political system λ_t.

PROPOSITION 5: *For all* $\lambda_t > \bar{\lambda}$, $\partial\rho^S/\partial\lambda_t < 0$ *and* $\partial\rho^R/\partial\lambda_t \leq 0$; *and for all* $\lambda_t < \underline{\lambda}$, $\partial\rho^S/\partial\lambda_t < 0$ *and* $\partial\rho^R/\partial\lambda_t > 0$ *if* $\lim_{\lambda\to 0} \partial u/\partial\lambda < \underline{u}$, *and* $\partial\rho^S/\partial\lambda_t > 0$ *and* $\partial\rho^R/\partial\lambda_t < 0$ *if* $\lim_{\lambda\to 0} \partial u/\partial\lambda > \bar{u}$, *and some* $(\bar{\lambda}, \underline{\lambda}, \bar{u}, \underline{u}) \in [0,1)^2 \times \mathbb{R}^2_-$, *whereas* $\bar{\lambda} \geq \underline{\lambda} > 0$ *if* $\bar{\theta}(0) < 1$.

From Proposition 5 it follows that as regimes become more democratic, they eventually become more stable. This is generally true for political systems in which no reforms are conducted; and further holds for sufficiently democratic regimes ($\lambda_t > \bar{\lambda}$). For autocratic systems, in contrast, the properties of the likelihood of political change depend on the exact specification of u. Still, Proposition 5 suggests that ρ^R and ρ^S are hump-shaped when marginal reforms for autocratic regimes are very costly or rather cheap, respectively. Otherwise, the likelihood for either type of transition tends to be decreasing as the political system becomes more democratic.

4.3. A parametric example

To illustrate the dynamics implied by Propositions 4 and 5 and to further study the implications of the model in the long-run, we now introduce a parametrized version of our model economy. We choose the following functional forms,

$$h(s_t) = s_t^\alpha$$

and

$$u(\lambda_t) = -\exp(\beta_1\lambda_t) + \beta_0.$$

Here one may think of β_0 as a common resource stock or some other type of private benefits, which decline at an exponential rate β_1 as power is shared with more agents. To pin down the free parameters, we further assume that $\psi'(1) = 0$; i.e., the strategic effect of an additional outsider supporting a revolt becomes negligible when revolts are supported by the full population. Together with our assumptions on u and h, this pins down α and β_0 in terms of β_1, which is restricted to approximately satisfy $\beta_1 \in (0, 0.56)$.[14]

Intuitively, β_1 measures the costs of enfranchising political outsiders. In practice, these costs are expected to be high if members of the regime have access to a large pool of resources, or if there is a large degree of economic and political inequality.[15] Thus, when β_1 is close to its upper bound, extending the franchise is costly and the incentives to gamble for survival are strong. Consequently, for large β_1, one should expect to observe revolts frequently in equilibrium. On the other hand, if β_1 is low, conducting reforms is cheap and one should expect political insiders to quickly reform to a fully integrated society.

To give an overview of the transition dynamics, Figure 4 displays a simulated time series of the model economy for different values of β_1 and for 500 periods each. For each time path, we plot the political system, λ_t, at time t and indicate the dates where transitions occur via revolts (marked by Δ) and reforms (marked by $\times$). It can be seen that low costs of reforms in Setting 1 ($\beta_1 = 0.35$) result in immediate democratic reforms and the absence of successful subversive attempts. As the costs of reforms are increasing in Setting 2 ($\beta_1 = 0.40$) and Setting 3 ($\beta_1 = 0.45$), successful revolts become more frequent and are followed by periods of frequent regime changes, where autocracies succeed each other. In contrast, democratic reforms give rise to long periods of political stability.

Polarization Although Figure 4 is the result of a random simulation, it captures many essential transition dynamics that arise in our model. First, in line with Proposition 4, it can be seen that transitions lead to a polarization of regimes; i.e., revolts lead to autocratic regimes, whereas reforms result in fairly inclusive democracies. A more complete picture is provided by Figure 5, which displays the distribution of political systems that emerge from each transition mechanism for $\beta_1 = 0.4$.[16] From the left panel, it becomes apparent that approximately two

[14] The implied values for the other two parameters are $\alpha = \beta_1 \exp(\beta_1)$ and $\beta_0 = \exp(\beta_1) + 1$, restricting $\beta_1 \in (0, \exp(-\beta_1)) \approx (0, 0.56)$.

[15] In particular, note that $u(\lambda) = \exp(\beta_1) - \exp(\beta_1 \lambda) + 1$ is increasing in β_1 for all λ, so that also the inequality between insiders and the average outsider, $\int (u(\lambda) - \gamma)\, \mathrm{d}\gamma$, is increasing in β_1 for all λ.

[16] For computing the distributions, originating political systems are weighted by their long-run distribution Ψ; e.g., the distribution of political systems after reforms is given by $\mathrm{pdf}(\lambda_{t+1}) = \int_0^1 Q^R(\lambda_t, \lambda_{t+1})\, \mathrm{d}\Psi(\lambda_t)$. While the long-run distribution itself varies considerably with β_1 (see also Figure 8), the conditional distributions displayed in Figure 5 remain largely unaffected by changes in β_1.

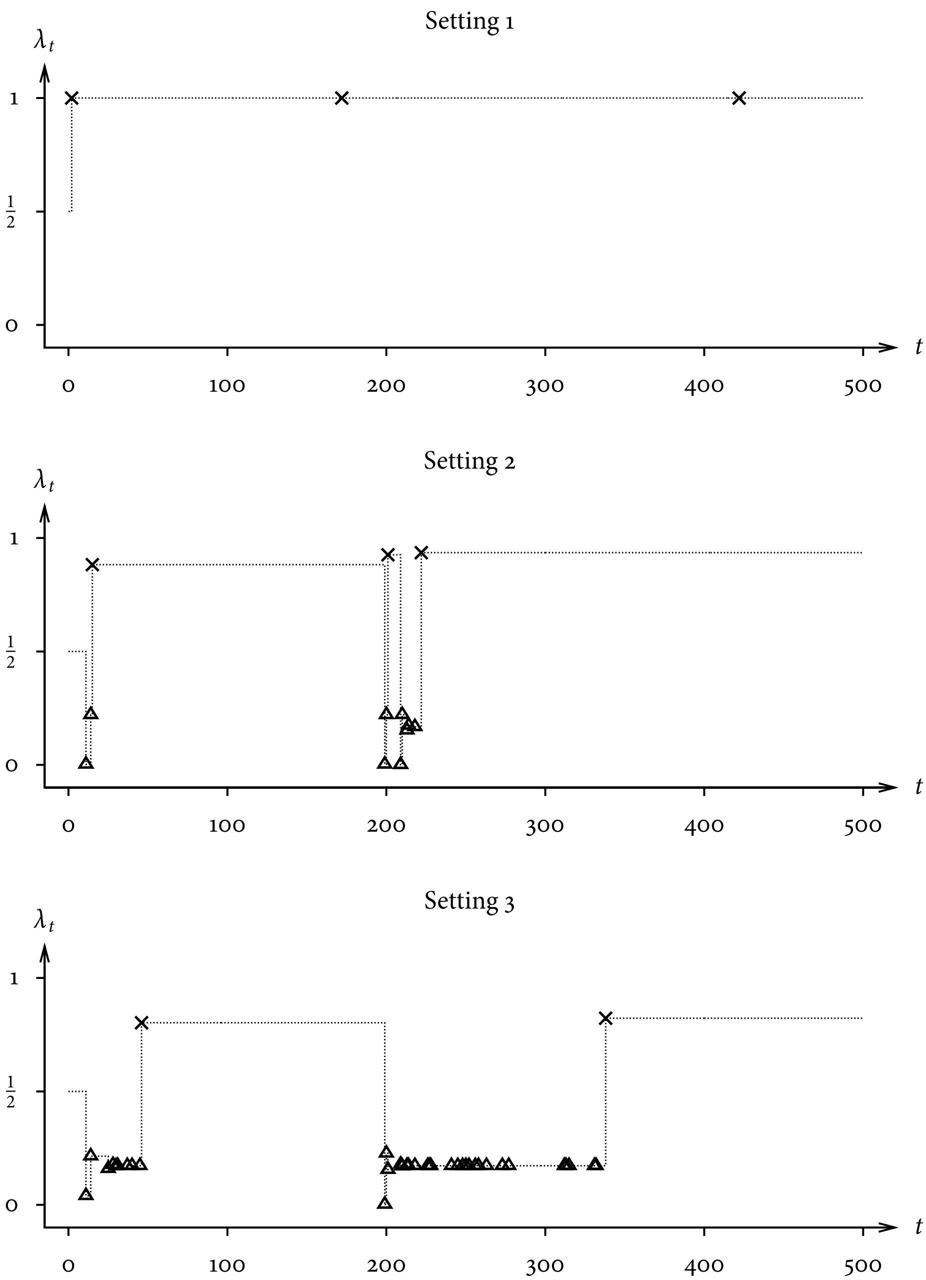

Figure 4. Simulated time series of the model economy. *Notes:* Reforms are marked by "×", successful revolts are marked by "△".

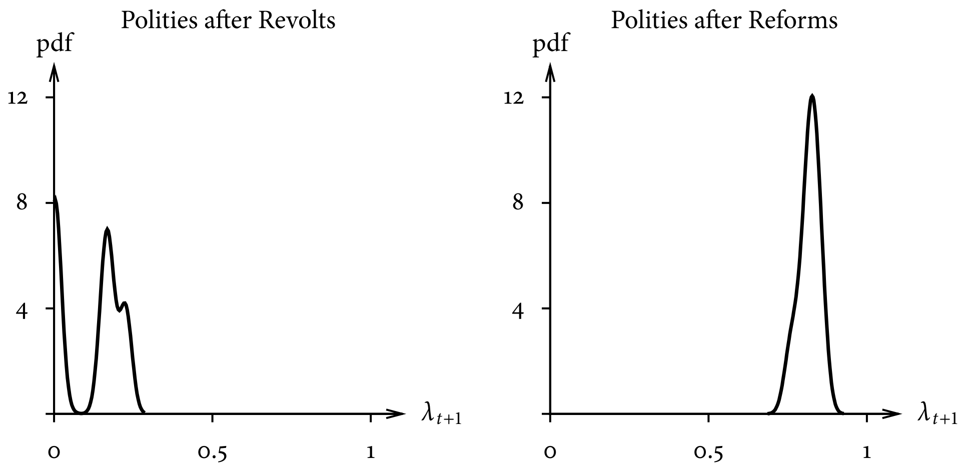

Figure 5. Distribution of political systems after revolts and reforms.

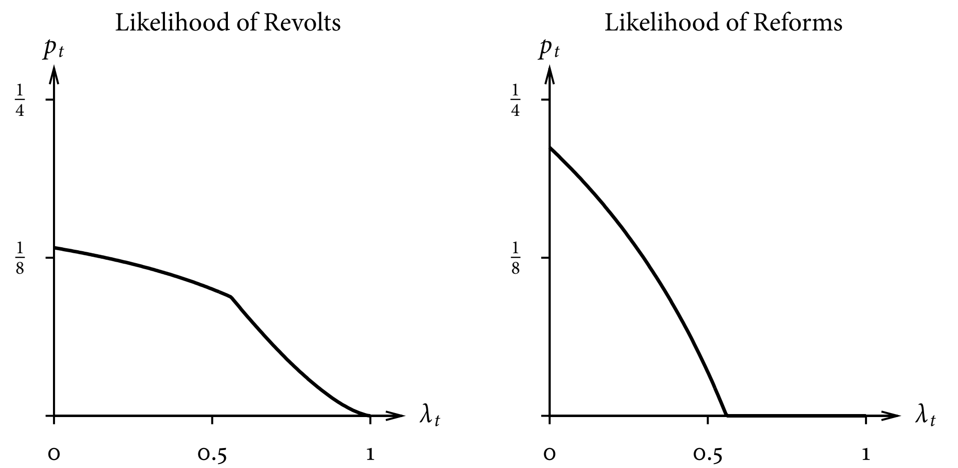

Figure 6. Likelihood of revolts and reforms.

different types of autocracies emerge after revolts: dictatorships, corresponding to regimes that emerge after revolts against democracies, and autocracies which emerge after succeeding other autocracies. From the right panel of Figure 5, it becomes apparent that reforms lead to democratic political systems where political power is shared among the majority of the population. Furthermore, it can be seen that a large set of political systems around 1/2 is neither emerging from reforms, nor from revolts.

Stability The second observation that can be drawn from the simulations in Figure 4 concerns the stability of political regimes. In line with Proposition 5, it is evident that democracies are characterized by long episodes without political change. In contrast, autocracies are subject to frequent regime changes. The underlying transition probabilities are depicted in Figure 6. Here we plot the likelihood of political transitions via revolts (left panel) and reforms (right panel) as a function of λ_t. It can be seen that both relations are decreasing in λ_t, such that autocracies are more likely than democracies to experience transitions of either type.

Turbulent and peaceful times Another interesting observation suggested by the simulations in Figure 4 is that revolts tend to be serially correlated over time. Underlying this observation is a statistical selection into autocratic regimes after successful revolts, seen in Figure 5. Because succeeding autocracies are frequently overthrown themselves, seen in Figure 6, the serial correlation follows. A direct assessment of this effect is provided in Figure 7, which plots the likelihood of a revolt at time $t + s$ conditional on a successful revolt at time t (represented by the downward sloping solid line).

The converse is true for reforms, which by Propositions 4 and 5 lead to democratic regimes, for which further political change is unlikely. Our model predicts, therefore, that via selection into particular political systems, revolts lead to "turbulent" times, while reforms lead to "peaceful" periods.

Persistence A side effect of the considerations in the preceding paragraph is that despite their instability, autocratic systems are persistent over time. That is, while individual autocracies are relatively short-lived, they are frequently overthrown by small groups of insurgents, resulting in autocracies very similar to their predecessors. Settings 2 and 3 of our simulations in Figure 4 illustrate this implication further.

Long-run distribution Taken together, polarization to extreme regimes and the persistence of these suggests that the long-run distribution of political systems is polarized as well. In Figure 8, we plot the invariant distribution of political systems for different values of β_1. It can be seen that the distributions are double hump-shaped, with most mass concentrated on extreme political systems. Whether political systems are mostly democratic or autocratic depends on the costs of reform as given by β_1. For low values of these costs (Settings 1 and 2), reforms are commonly used to mitigate most subversive threats, revolts are unlikely, and mass is mainly concentrated on democratic systems. If the costs

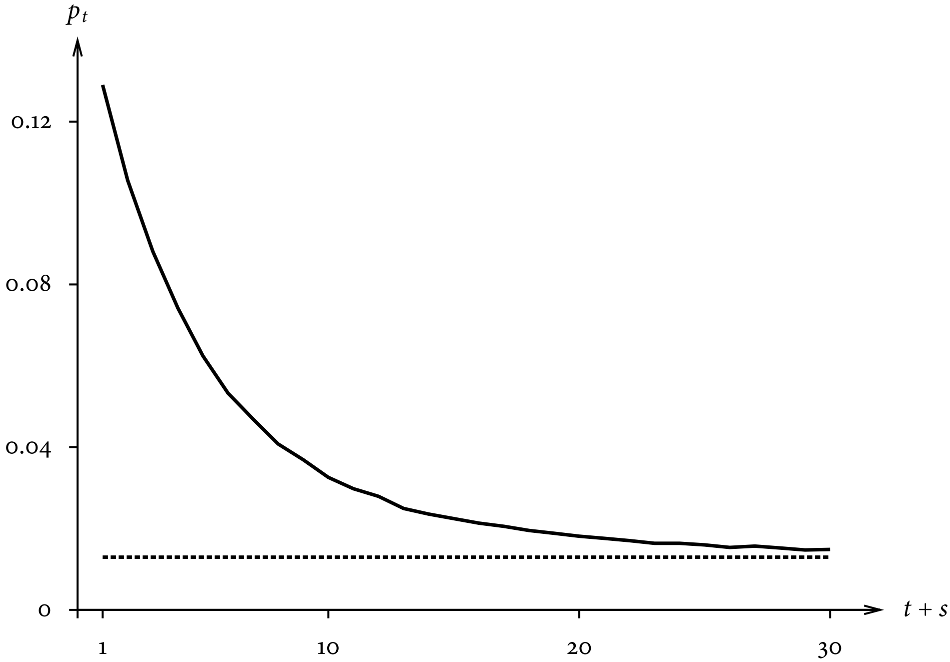

Figure 7. Likelihood of a successful revolt at time $t + s$ conditional on a revolt s periods before (solid) and unconditional likelihood (dashed).

of conducting reforms are high (Settings 3 and 4), less reforms are conducted, revolts are more frequent, and most mass is concentrated on autocratic political systems.

5. A look at the data

Our model predicts a number of properties about political transitions that are in principle accessible to an empirical investigation. In this section, we take an exploratory look at data that combines information on political transitions and political systems to evaluate the model's predictions. While we are able to demonstrate that our predictions are consistent with descriptive statistics from the data, we make no claims of capturing causal relations, which would be beyond the scope of this exercise.

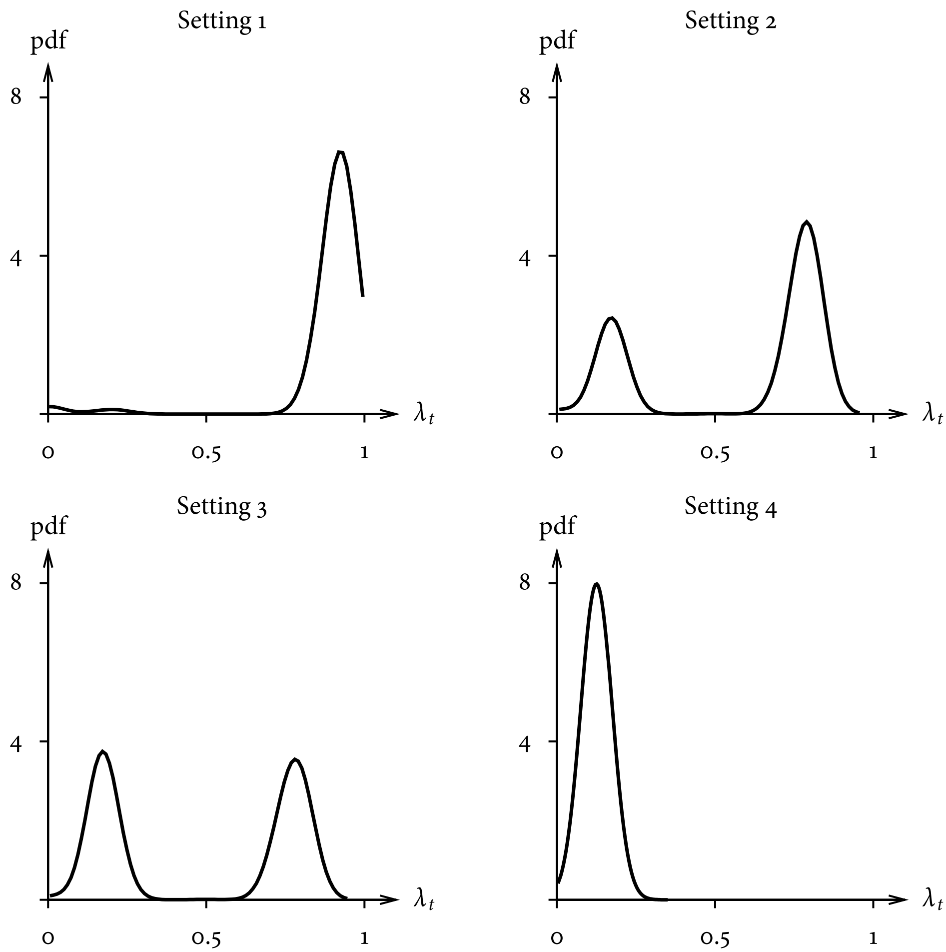

Figure 8. Invariant distribution of political systems.

5.1. Data construction

As a measure for the model's political system, we use the *polity* variable, scaled to $[0, 1]$, from the Polity IV Project (Marshall and Jaggers, 2002), which ranks political regimes on a 21 point scale between autocratic and democratic. In order to examine the model's predictions, we combine this dataset with data on political transitions.

To classify successful revolts, we use the Archigos Dataset of Political Leaders (Goemans et al., 2009). The dataset is available for the time period between 1919 and 2004, such that we limit attention to political systems and transition in these years. We record a successful revolt if a leader is irregularly removed from office due to domestic popular protest, rebel groups, or military actors (defined by Archigos' *exitcodes* 2, 4 and 6), and if at the same time the leader's successor takes office in irregular manner (defined by an *entrycode* 1). Furthermore, we take a revolt to be causal for a change in the political system if a change in the political system is recorded in the Polity IV database within a two week window of the revolt.

Finally, we use the dataset on the Chronology of Constitutional Events from the Comparative Constitution Project (Elkins et al., 2010) to classify reforms. We define reforms by a constitutional change (*evnttype* equal to *new*, *reinstated*, or *amendment*) accompanied by a positive change in the political system (as indicated by the variable *durable* from the Polity IV Project) which is not matched to a revolt or another irregular regime change from the Achigos Dataset.

The resulting dataset is a daily panel on the country level, which covers 175 countries and records 251 revolts and 97 reforms.

5.2. *Empirical properties of political systems and transitions*

Overview Table 1 summarizes the resulting dataset. Panel A displays average political systems and annualized empirical likelihoods for a transition of either type. It can be seen that on average, revolts are observed with a frequency of 2.8 percent per year and country, and reforms are observed with a frequency of 1.1 percent. On average, this corresponds to a transition every 25 years per country.

The mean polity is given by 0.49—almost exactly the midpoint of the polity scale. As can be seen in the second column, however, the standard deviation of political systems is quite large. The reason for this becomes clear in light of Figure 1, which displays the distribution of political systems in our dataset: Only a minority of regimes are located in the middle of the polity scale. Instead, in line with our predictions, most mass is concentrated on extreme political systems. More precisely, 44 percent of all regimes are rather autocratic with a polity index of 0.25 and below, while 38 percent of all regimes are rather democratic with an index value of 0.75 and above.

Our model identifies two reasons for why the distribution of political systems is extreme: Polarization via the transition mechanism and persistence of extreme political systems.

Table 1. Descriptive Statistics

	Mean	Standard Deviation	Observations
	A. Regimes		
Political systems	0.493	0.376	3 289 400
Annual likelihood of a revolt			
Unconditional	0.028		3 289 400
If polity ≤ 0.25	0.030		1 452 533
If polity ≥ 0.75	0.012		1 238 720
Annual likelihood of a reform			
Unconditional	0.011		3 289 400
If polity ≤ 0.25	0.018		1 452 533
If polity ≥ 0.75	0.001		1 238 720
	B. Transitions		
Resulting political systems			
After revolts	0.316	0.235	251
After reforms	0.672	0.242	97

Notes.— Units of observation in Panel A are country-days. Units of observation in Panel B are transitions.

Polarization To examine whether regimes are polarized via political transitions, consider Panel B of Table 1, which displays the mean polity index for regimes emerging after revolts and reforms, respectively. As predicted by Proposition 4, revolts on average lead to autocratic regimes with a polity index of 0.32, while reforms lead to rather democratic political systems with a mean polity index equal to 0.67.

Further insight can be gained from the conditional distribution of political systems emerging after either type of transition. Figure 9 displays these distributions. From the left panel it is obvious that indeed the majority of political systems that emerge after revolts is autocratic. In contrast, the evidence about political reforms is less clear. On the one hand, the right panel of Figure 9 suggests that the majority of systems that are established through reforms are democratic. On the other hand, it also can be seen that, in contrast to the model's predictions, a significant number of reforms lead to regimes that are less democratic.

However, while some reforms are less democratic than predicted, Figure 9 still suggests that the majority of democratic regimes are established via reforms, consistent with Proposition 4.

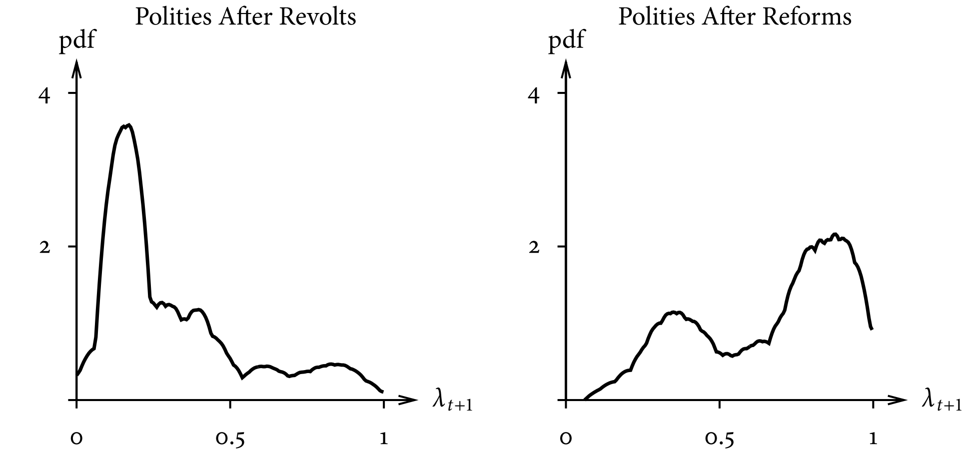

Figure 9. Empirical distribution of political systems after revolts and reforms.

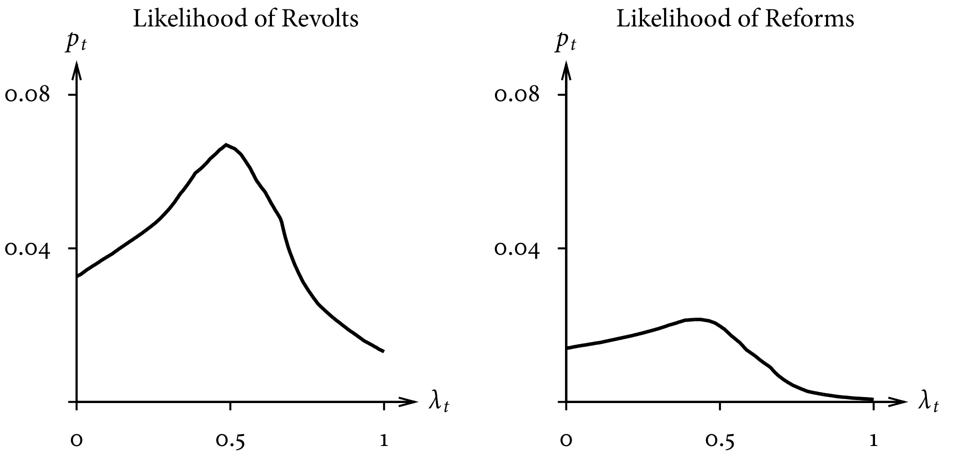

Figure 10. Empirical likelihood of revolts and reforms.

Stability and persistence To examine the stability of political systems, consider Figure 10. Here we plot the empirical likelihood functions for revolts and reforms, derived from a local polynomial estimation. Both likelihoods are hump-shaped in the polity index, with regimes in the middle of the scale being most likely to be overthrown. Nevertheless, as can be seen in Panel A of Table 1, autocracies with a polity index of 0.25 or below are more than twice as likely to fall to a revolt than democratic regimes with an index value of 0.75 and above.

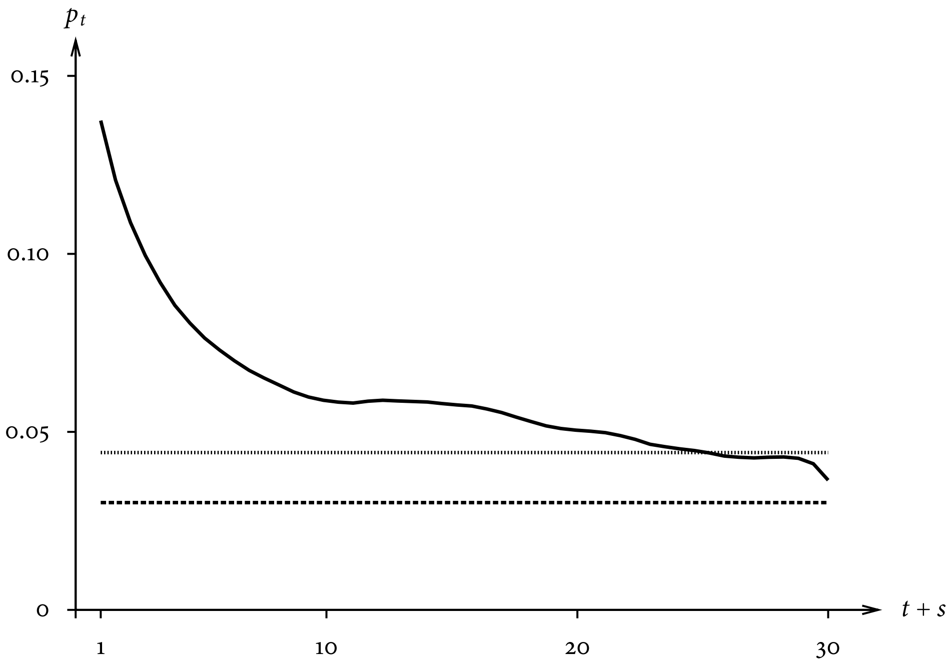

Figure 11. Empirical likelihood of a revolt at date $t + s$ conditional on a revolt s years before (solid) and unconditional likelihood for all countries (dashed) and countries with at least one transition (dotted).

Moreover, autocratic regimes are about 18 times more likely to conduct reforms than democracies. Overall, autocracies survive for an average of about 21 years, while democracies survive for an average of about 79 years. Hence, while in contrast to Proposition 5 full-scale democracies face a nonzero probability to be overthrown, they are nevertheless considerably more stable than all other regime types, confirming the qualitative predictions made by the model.

According to our model, even though autocracies are more instable than democracies, a serial correlation between revolts results in a persistence of autocratic political systems. The descriptive statistics reported above already suggest that the statistical selection mechanism underlying the persistence in our model might also be at work in the data. That is, we have seen that revolts are likely to result in autocracies, which are themselves likely to be overthrown again (see the left panels of Figures 9 and 10). As can be seen in Figure 11, the suggested correlation is indeed present in the data. The solid line in Figure 11 reflects the likelihood of observing a revolt at date $t + s$ conditional on that there was a successful revolt s years before. This likelihood is considerably larger than the unconditional likelihood of revolts across all countries (dashed line) and also compared to the unconditional likelihood in countries with at least one observed transition (dotted line). Compared to the latter benchmark, the difference is statistically significant at the 5 percent level for $s \leq 15$.

Summary In summary, the moments and correlations predicted by our model are consistent with the corresponding empirical moments and correlation. As predicted by the model, transitions lead to a polarization of political regimes, giving rise to autocracies after revolts and democracies after political reforms. While democracies are found to be empirically stable, autocracies are found to be short-lived. Yet, consistent with the model, a statistical selection gives rise to autocorrelation of successful subversions, explaining persistence of autocracies in the long-run. Consistently, as predicted by the model, the overall empirical distribution has mass mainly concentrated on extreme political systems.

6. Concluding remarks

This is the first study, which explores the dynamic properties of political transitions in a general framework allowing for endogenous outcomes of reforms and revolts. Our results suggest that transitions to democracy occur peacefully via reforms under participation of the former ruling elites. In contrast, violent transitions are driven by a small groups of insurgents and thus always lead to autocratic political regimes. Furthermore, democratic political systems face only a small opposition and are, hence, inherently stable, while autocratic regimes are short-lived due to the significant threat of revolts and the resulting strong incentives to conduct reforms.

These predictions are derived from a model in which the threat of revolt is the driving force of political change. We enrich the pioneering work of Acemoglu and Robinson (2000b) by allowing for arbitrary reforms conducted by the elite and endogenous formation of revolts through coordination of outsiders. While the predictions from this model fit descriptive statistics on political transitions quite well, our work points out promising avenues for future research. In particular, one simplifying assumption of our model is that the vulnerability of the incumbent regime is independently drawn anew in each period. Relaxing this assumption by allowing for serial correlation of the incumbents' strength would allow outsiders to learn about the prospects of revolting over time. A model, in which such endogenous learning is possible, could thus foster our understanding of the dynamic processes which ultimately lead to transition events (revolts or reforms).

Another interesting question regards the existence of mass movements. While from Figure 9 one can see that the majority of regimes that emerge after successful revolts is indeed autocratic in the data, there is also a nonzero mass of democratic regimes emerging from revolts. In our model mass revolutions and, hence, violent transitions to democracy are events off the equilibrium path. Therefore, only strategic mistakes could trigger mass revolts within our framework. For example, the elite may erroneously signal weakness by making small concessions, or outsiders may rally because of a commonly held belief that the regime is weak (for example due to information cascades as in Kuran, 1989 or Lohmann, 1994). While it seems plausible that costly mass revolutions are the result of strategic mistakes, there thus remains the challenge to find a rational explanation for the emergence of mass revolutions when the regime has the power to counteract them via reforms.

A. Mathematical appendix

A.1. Insiders never subvert, outsiders always join the regime

Insiders' choice set includes $x_t \in [\lambda_t, 1]$. It thus holds that $(1 - p(\cdot, x_t))u(x_t) \geq (1 - p(\cdot, 1))u(1) = u(1) \geq \psi(1) \geq \psi(s_t) \geq \hat{\theta}_t \psi(s_t)$, where the first inequality follows from revealed preferences, the second inequality follows from $h(\cdot) \in [0,1]$, the third inequality follows from ψ increasing, and the last inequality follows from $\theta_t \in [0,1]$. Hence, it is not attractive for any individual insider to support a revolt against his own regime. As for outsiders we need to differentiate two cases. First, outsiders that are targeted by a reform and would otherwise support a revolt prefer to join the regime using exactly the same argument as above. Second, outsiders that are targeted by a reform and would otherwise not support a revolt prefer to join the regime since again by revealed preferences it holds that $(1 - p(\cdot, x_t))u(x_t) \geq (1 - p(\cdot, 1))u(1) = u(1) \geq \gamma_{it}$ for all i and t. □

A.2. Proof of Proposition 1

We first establish that any solution to the outsiders' coordination problem is a fixed point to equation (6). From our discussion in the main body of the chapter it is clear that this is the case if and only if $\bar{\gamma}(\hat{s}_t) \leq 1$ for all $\hat{s}_t$. From

Assumption 1 it follows that $\bar{\gamma}$ is increasing in $\hat{s}_t$, and therefore $\bar{\gamma}(\hat{s}_t) \leq 1$ holds if $\bar{\gamma}(1) = p(\hat{\theta}_t, 1)\, u(1) \leq 1$. Since $u(1) = 1$ and $p(\cdot) \in [0,1]$ this is indeed the case.

Hence, consider any fixed point to (6). Since $f(0) = 0$ for all $(\hat{\theta}_t, x_t) \in \Theta \times [0,1]$, there always exists a fixed point at $\hat{s}_t = 0$. Whether or not $\hat{s}_t = 0$ is consistent with the concept of trembling-hand perfection, and whether or not other fixed points exist, depends on the values of $\hat{\theta}_t$ and x_t. We have to distinguish two cases.

First, if $\hat{\theta}_t = 0$ or $x_t = 1$, then $f(\hat{s}_t) = 0$ for all $\hat{s}_t$, and therefore $\hat{s}_t = 0$ is obviously the only fixed point to (6). To establish that $\hat{s}_t = 0$ is also trembling-hand perfect, it suffices to show that for all i, $\phi_{it} = 0$ is a best response to some sequence of totally mixed strategy profiles $\{\omega_{jt}^k : j \in [0,1] \setminus i\}_{k=0}^{\infty}$ that converges to the equilibrium profile where all i play $\phi_{it} = 0$ with probability 1. Since for $\hat{\theta}_t = 0$ and $x_t = 1$ playing $\phi_{it} = 0$ is a (weakly) dominant strategy, this is trivially true.

Second, consider the case where $\hat{\theta}_t \neq 0$ and $x_t \neq 1$. In this case the fixed point at $\hat{s}_t = 0$ is not trembling-hand perfect. To see this let $z^k = \min_i\{\omega_{it}^k(1)\}$ denote the minimum probability with which any agent i plays $\phi_{it} = 0$ in the kth element of sequence ω_{it}^k. The requirement of trembling-hand perfection that $\{\omega_{it}^k\}$ is totally mixed for all i and k implies that $z^k > 0$ for all k. Hence, $s_t^k = (1-x_t) \int_i \omega_{it}^k(1)\, \mathrm{d}i \geq (1-x_t)\, z^k > 0$. However, from $h(0) = 0$ in combination with Assumption 1(b) it follows that for any $s_t^k > 0$, $\bar{\gamma}(s_t^k) = \hat{\theta}_t \psi(s_t^k) > 0$ and, hence, a strictly positive fraction of outsiders strictly prefers to choose $\phi_{it} = 1$ in response to $\{\omega_{jt}^k : j \in [0,1]\}$. We conclude that $\hat{s}_t = 0$ can not be supported in any trembling-hand perfect equilibrium if $\hat{\theta}_t \neq 0$ and $x_t \neq 1$.

Having ruled out $\hat{s}_t = 0$ as a solution to the coordination problem for $\hat{\theta}_t \neq 0$ and $x_t \neq 1$, we now show that there is a unique $\hat{s}_t > 0$ solving (6) for $\hat{\theta}_t \neq 0$ and $x_t \neq 1$, which is also consistent with the concept of trembling-hand perfection. From $\bar{\gamma} \in [0,1]$ it follows that f is bounded by its support, $[0, 1-x_t]$. Moreover, by Assumption 1 we have that $\lim_{\hat{s}\to 0} \psi'(\hat{s}) = \infty$, implying that $\lim_{\hat{s}\to 0} f'(\hat{s}) = \infty$. Hence, there exists a $\tilde{s} > 0$, such that $f(\tilde{s}) > \tilde{s}$. Together with continuity of ψ (and thus of f), it follows that there exists a strictly positive fixed point to (6), which by concavity of ψ (and thus of f) is unique on $(0,1]$.

Let $s_t^* = f(s_t^*)$ denote this fixed point. It remains to be shown that s_t^* is consistent with the concept of trembling-hand perfection. To show this, consider the following sequences $\omega_{it}^k(1) = 1 - \varepsilon^k$ for all $i \in \{j : \gamma_{jt} \leq \bar{\gamma}(s_t^*)\}$ and $\omega_{it}^k(1) = \frac{\bar{\gamma}(s_t^*)}{1-\bar{\gamma}(s_t^*)}\varepsilon^k$ for all $i \in \{j : \gamma_{jt} > \bar{\gamma}(s_t^*)\}$, with some $\{\varepsilon^k\}_{k=0}^{\infty}$ such that $\lim_{k\to\infty} \varepsilon^k = 0$.

Then, by construction,

$$\begin{aligned} s_t^k &= (1 - x_t)\left((1 - \varepsilon^k)\,\bar{\gamma}(s_t^*) + \frac{\bar{\gamma}(s_t^*)}{1 - \bar{\gamma}(s_t^*)}\,\varepsilon^k(1 - \bar{\gamma}(s_t^*))\right) \\ &= (1 - x_t)\,\bar{\gamma}(s_t^*) \\ &= f(s_t^*), \end{aligned}$$

and hence $\{\phi_{it} : i \in [0,1]\}$ being mutually best responses implies that $\{\phi_{it} : i \in [0,1]\}$ are best responses to $\{\omega_{it}^k : i \in [0,1]\}$ for all values of k.

The above arguments establish that s_t is uniquely determined by a (time-invariant) function $s : (\hat{\theta}_t, x_t) \to s_t$. It remains to be shown that $\partial s/\partial\hat{\theta}_t \geq 0$ and $\partial s/\partial x_t \leq 0$. Given that s_t is a fixed point to (6), we have that

$$\pi(s_t, x_t) \equiv s_t - (1 - x_t)\,\hat{\theta}_t\,\psi(s_t) = 0.$$

Implicit differentiation implies that

$$\frac{\partial s_t}{\partial x_t} = -\hat{\theta}_t\,\psi(s_t) \times \left(\frac{\partial \pi_t}{\partial s_t}\right)^{-1}$$

and

$$\frac{\partial s_t}{\partial \hat{\theta}_t} = (1 - x_t)\,\psi(s_t) \times \left(\frac{\partial \pi_t}{\partial s_t}\right)^{-1},$$

where

$$\frac{\partial \pi_t}{\partial s_t} = -(1 - x_t)\,\frac{\partial \bar{\gamma}}{\partial s_t} + 1.$$

Since ψ is bounded by $\psi(1) = 1$, (6) implies that $\lim_{\hat{\theta}_t \to 0} s_t^* = \lim_{x_t \to 1} s_t^* = 0$, and therefore the case where $\hat{\theta}_t = 0$ or $x_t = 1$ is a limiting case of $\hat{\theta} \neq 0$ and $x_t \neq 1$. From the implicit function theorem it then follows that s is differentiable on its whole support. Moreover, the previous arguments imply that $f(\tilde{s}) > \tilde{s}$ for all $\tilde{s} < s_t^*$ and $f(\tilde{s}) < \tilde{s}$ for all $\tilde{s} > s_t^*$, implying that $f'(s_t^*) < 1$ or, equivalently, $\partial\bar{\gamma}/\partial s_t < (1 - x_t)^{-1}$ at s_t^*. Thus $\partial\pi_t/\partial s_t > 0$ for all $(\hat{\theta}_t, x_t) \in \Theta \times [0,1]$, which yields the desired results.

Finally, while we focused on pure strategies when proving the results above, it is easy to see that the proposition generalizes to mixed strategies. By the law of large numbers, any mixed strategy equilibrium beliefs about s are of zero variance and, hence, the arguments above apply, implying that all outsiders, except a zero mass i with $\gamma_i = \bar{\gamma}(s_t^*)$, strictly prefer $\phi_i = 0$ or $\phi_i = 1$. We conclude that there is no scope for (nondegenerate) mixed best responses. □

A.3. Proof of Proposition 2

The proof proceeds by a series of lemmas. To simplify notation, in what follows we drop λ_t as an argument of x and $\hat{\theta}$ where no confusion arises. Furthermore, we use $\tilde{V}^I(\theta_t, \hat{\theta}_t, x_t) = (1 - \theta_t h(s_t))\, u(x_t)$ to denote insider's indirect utility (up to a constant $u(\lambda_t)$), as follows from $s_t = s(\hat{\theta}_t, x_t)$ given Proposition 1.

LEMMA 1: *x is weakly increasing in θ_t.*

Proof. Suppose to the contrary that $x(\theta'') < x(\theta')$ for $\theta' < \theta''$. Let $x' \equiv x(\theta')$, $x'' \equiv x(\theta'')$, $u' \equiv u(x')$, $u'' \equiv u(x'')$, $h' \equiv h(s(\hat{\theta}(x'), x'))$, and $h'' \equiv h(s(\hat{\theta}(x''), x''))$. Optimality of x' then requires that $\tilde{V}^I(\theta', \hat{\theta}(x''), x'') \leq \tilde{V}^I(\theta', \hat{\theta}(x'), x')$, implying $u'h' - u''h'' \leq (u' - u'')/\theta' < (u' - u'')/\theta''$, where the last inequality follows from $\theta' < \theta''$ and $u' < u''$. Hence, $\tilde{V}^I(\theta', \hat{\theta}(x''), x'') \leq \tilde{V}^I(\theta', \hat{\theta}(x'), x')$ implies that $\tilde{V}^I(\theta'', \hat{\theta}(x''), x'') < \tilde{V}^I(\theta'', \hat{\theta}(x'), x')$, contradicting optimality of x'' for θ''. ◇

LEMMA 2: *Suppose x is discontinuous at θ', and define $x^- \equiv \lim_{\varepsilon\uparrow 0} x(\theta' + \varepsilon)$ and $x^+ \equiv \lim_{\varepsilon\downarrow 0} x(\theta' + \varepsilon)$. Then for any $x' \in (x^-, x^+)$, the only beliefs consistent with the D1 criterion are $\hat{\theta}(x') = \theta'$.*

Proof. Let $\theta'' > \theta'$, and let $x'' \equiv x(\theta'')$. Optimality of x'' then requires that $\tilde{V}^I(\theta'', \hat{\theta}(x''), x'') \geq \tilde{V}^I(\theta'', \hat{\theta}(x^+), x^+)$ and, thus for any $\tilde{\theta}$,

$$\tilde{V}^I(\theta'', \tilde{\theta}, x') \geq \tilde{V}^I(\theta'', \hat{\theta}(x''), x'') \quad \text{implies that} \\ \tilde{V}^I(\theta'', \tilde{\theta}, x') \geq \tilde{V}^I(\theta'', \hat{\theta}(x^+), x^+)\,.$$

Moreover, arguing as in the proof of Lemma 1,

$$\tilde{V}^I(\theta'', \tilde{\theta}, x') \geq \tilde{V}^I(\theta'', \hat{\theta}(x^+), x^+) \quad \text{implies that} \\ \tilde{V}^I(\theta', \tilde{\theta}, x') > \tilde{V}^I(\theta', \hat{\theta}(x^+), x^+)\,.$$

Hence, if $\tilde{V}^I(\theta'', \tilde{\theta}, x') \geq \tilde{V}^I(\theta'', \hat{\theta}(x^+), x^+) = \bar{V}^I(\theta'')$, then $\tilde{V}^I(\theta', \tilde{\theta}, x') > \tilde{V}^I(\theta', \hat{\theta}(x^+), x^+) = \bar{V}^I(\theta')$. Therefore, $D_{\theta'',x'}$ is a proper subset of $D_{\theta',x'}$ if $\theta'' > \theta'$. (For the definition of $D_{\theta,x}$, see Footnote 10.) A similar argument establishes that $D_{\theta'',x'}$ is a proper subset of $D_{\theta',x'}$ if $\theta'' < \theta'$ and, thus, the D1 criterion requires that $\hat{\theta}(x') = \theta'$ for all $x' \in (x^-, x^+)$. ◇

LEMMA 3: *There exists $\bar{\theta}(\lambda_t) > 0$, such that $x(\theta_t, \lambda_t) = \lambda_t$ for all $\theta_t < \bar{\theta}(\lambda_t)$. Moreover, $x(\theta'') > x(\theta') > \lambda_t + \mu$ for all $\theta'' > \theta' \geq \bar{\theta}(\lambda_t)$ and some $\mu > 0$.*

Proof. First, consider the existence of a connected pool at $x_t = \lambda_t$. Because for $\theta_t = 0$, $x_t = \lambda_t$ dominates all $x_t > \lambda_t$, we have that $x(0) = \lambda_t$. It follows that there exists a pool at $x_t = \lambda_t$, because otherwise $\hat{\theta}(\lambda_t) = 0$ and, therefore, $p(\cdot, s(\hat{\theta}(\lambda_t), \lambda_t)) = 0$, contradicting optimality of $x(\theta) > \lambda_t$ for all $\theta > 0$. Moreover, by Lemma 1, x is increasing, implying that any pool must be connected. This proves the first part of the claim.

Now consider $x(\theta'') > x(\theta')$ for all $\theta'' > \theta' \geq \bar{\theta}(\lambda_t)$ and suppose to the contrary that $x(\theta'') \leq x(\theta')$ for some $\theta'' > \theta'$. Since x is increasing, it follows that $x(\theta) = x^+$ for all $\theta \in [\theta', \theta'']$ and some $x^+ > \lambda_t$. W.l.o.g. assume that θ' is the lowest state in this pool. Then Bayesian updating implies that $\theta^+ \equiv \hat{\theta}(x^+) \geq (\theta' + \theta'')/2 > \theta'$ and, therefore, $\tilde{V}^I(\theta', \theta^-, x^+) > \tilde{V}^I(\theta', \theta^+, x^+)$ for all $\theta^- \leq \theta'$. Hence, because θ' prefers x^+ over $x(\theta^-)$, it must be that $x(\theta^-) \neq x^+$ for all $\theta^- \leq \theta'$ and, hence, $x(\theta^-) < x^+$ by Lemma 1. Accordingly, let $x^- \equiv \max_{\theta^- \leq \theta'} x(\theta^-)$. Then from continuity of $\tilde{V}^I$ and $\theta^+ > \theta'$ it follows that there exists an off-equilibrium reform $x' \in (x^-, x^+)$ with $\tilde{V}^I(\theta', \theta', x') > \tilde{V}^I(\theta', \theta^+, x^+)$. Hence, to prevent θ' from choosing x' it must be that $\hat{\theta}(x') > \theta'$. However, from Lemma 2 we have that $\hat{\theta}(x') = \theta'$, a contradiction.

Finally, to see why there must be a jump-discontinuity at $\bar{\theta}(\lambda_t)$ note that $\tilde{V}^I(\bar{\theta}(\lambda_t), \bar{\theta}(\lambda_t)/2, \lambda_t) = \tilde{V}^I(\bar{\theta}(\lambda_t), \bar{\theta}(\lambda_t), x(\bar{\theta}(\lambda_t)))$; otherwise, there necessarily exists a θ in the neighborhood of $\bar{\theta}(\lambda_t)$ with a profitable deviation to either λ_t or $x(\bar{\theta}(\lambda_t))$. From the continuity of $\tilde{V}^I$ and the non-marginal change in beliefs from $\bar{\theta}(\lambda_t)/2$ to $\bar{\theta}(\lambda_t)$ it follows that $x(\bar{\theta}(\lambda_t)) > \lambda_t + \mu$ for all λ_t and some $\mu > 0$. ◇

LEMMA 4: *x is continuous and differentiable in θ_t on $[\bar{\theta}(\lambda_t), 1]$.*

Proof. Consider continuity first and suppose to the contrary that x has a discontinuity at $\theta' \in (\bar{\theta}(\lambda_t), 1)$. By Lemma 1, x is monotonically increasing in θ_t. Hence, because x is defined on an interval, it follows that for any discontinuity θ', $x^- \equiv \lim_{\varepsilon \uparrow 0} x(\theta')$ and $x^+ \equiv \lim_{\varepsilon \downarrow 0} x(\theta')$ exist, and that x is differentiable on $(\theta' - \varepsilon, \theta')$ and $(\theta', \theta' + \varepsilon)$ for some $\varepsilon > 0$. Moreover, from Lemmas 2 and 3 it follows that in equilibrium $\hat{\theta}(x') = \theta'$ for all $x' \in [x^-, x^+]$. Hence, $\tilde{V}^I(\theta', \theta', x^-) = \tilde{V}^I(\theta', \theta', x^+)$, since otherwise there necessarily exists a θ in the neighborhood of θ' with a profitable deviation to either x^- or x^+. Accordingly, optimality of $x(\theta')$ requires $\tilde{V}^I(\theta', \theta', x') \leq \tilde{V}^I(\theta', \theta', x^-)$ and, thus, $\tilde{V}^I(\theta', \theta', x^-)$ must be weakly decreasing in x. Therefore, $\partial \tilde{V}^I / \partial \hat{\theta}_t < 0$ and $\lim_{\varepsilon' \downarrow 0} \partial \hat{\theta}(x^- - \varepsilon')/\partial x_t > 0$ (following from Lemma 3) imply that $\lim_{\varepsilon' \downarrow 0} \partial \tilde{V}^I(\theta', \hat{\theta}(x^- - \varepsilon'), x^- - \varepsilon')/\partial x_t < 0$. Hence, a profitable deviation to $x^- - \varepsilon'$ exists for some $\varepsilon' > 0$, contradicting optimality of $x(\theta')$.

We establish differentiability by applying the proof strategy for Proposition 2 in Mailath (1987). Let $g(\theta, \hat{\theta}, x) \equiv \tilde{V}^I(\theta, \hat{\theta}, x) - \tilde{V}^I(\theta, \theta', x(\theta'))$, for a given $\theta' > \bar{\theta}(\lambda_t)$, and let $\theta'' > \theta'$. Then, optimality of $x(\theta')$ implies $g(\theta', \theta'', x(\theta'')) \leq 0$, and optimality of $x(\theta'')$ implies that $g(\theta'', \theta'', x(\theta'')) \geq 0$. Letting $a = (\alpha\theta' + (1-\alpha)\theta'', \theta'', x(\theta''))$, for some $\alpha \in [0,1]$ this implies

$$0 \geq g(\theta', \theta'', x(\theta'')) \geq -g_\theta(\theta', \theta'', x(\theta''))(\theta'' - \theta') - \tfrac{1}{2} g_{\theta\theta}(a)(\theta'' - \theta')^2,$$

where the second inequality follows from first-order Taylor expanding $g(\theta'', \theta'', x(\theta''))$ around $(\theta', \theta'', x(\theta''))$ and rearranging the expanded terms using the latter optimality condition. Expanding further $g(\theta', \theta'', x(\theta''))$ around $(\theta', \theta', x(\theta'))$, using the mean value theorem on $g_\theta(\theta', \theta'', x(\theta''))$, and noting that $g(\theta', \theta', x(\theta')) = g_\theta(\theta', \theta', x(\theta')) = 0$, these inequalities can be written as

$$\begin{aligned}
0 \geq g_{\hat{\theta}}(\theta', \theta', x(\theta')) &+ \frac{x(\theta'') - x(\theta')}{\theta'' - \theta'} \times [g_x(\theta', \theta', x(\theta')) \\
&+ \tfrac{1}{2} g_{xx}(b(\beta))(x(\theta'') - x(\theta')) + g_{\hat{\theta}x}(b(\beta))(\theta'' - \theta')] \\
&\qquad + \tfrac{1}{2} g_{\hat{\theta}\hat{\theta}}(b(\beta))(\theta'' - \theta') \\
\geq -[g_{\theta\hat{\theta}}(b(\beta')) &+ \tfrac{1}{2} g_{\theta\theta}(a)](\theta'' - \theta') - g_{\theta x}(b(\beta'))(x(\theta'') - x(\theta')),
\end{aligned}$$

for $b(\beta) = (\theta', \beta\theta' + (1-\beta)\theta'', \beta x(\theta') + (1-\beta)x(\theta''))$ and some $\beta, \beta' \in [0,1]$. Because $\tilde{V}^I$ is twice differentiable, all the derivatives of g are finite. Moreover, continuity of x implies that $x(\theta'') \to x(\theta')$ as $\theta'' \to \theta'$ and, therefore, for $\theta'' \to \theta'$,

$$0 \geq g_{\hat{\theta}}(\theta', \theta', x(\theta')) + \lim_{\theta'' \to \theta'} \frac{x(\theta'') - x(\theta')}{\theta'' - \theta'} \, g_x(\theta', \theta', x(\theta')) \geq 0.$$

By Lemma 3, x and, hence, $\hat{\theta}$ are strictly increasing for all $\theta \geq \bar{\theta}(\lambda_t)$. Arguing similarly as we did to show continuity, optimality of x, therefore, requires that $g_x = \partial\tilde{V}^I/\partial x_t \neq 0$ and, hence, the limit of $(x(\theta'') - x(\theta'))/(\theta'' - \theta')$ is well defined, yielding

$$\frac{\mathrm{d}x}{\mathrm{d}\theta_t} = -\frac{\partial\tilde{V}^I/\partial\hat{\theta}_t}{\partial\tilde{V}^I/\partial x_t}. \quad \diamond \tag{8}$$

Lemma 5: *$x(\theta_t, \lambda_t) = \xi(\theta_t)$ for all $\theta_t > \bar{\theta}(\lambda_t)$, where ξ is unique and $\partial\xi/\partial\theta_t > 0$.*

Proof. From Lemma 4 we have that ξ is differentiable, and by Lemma 3, $\partial\xi/\partial\theta_t > 0$. We thus only need to show that ξ is unique. By the proof to Lemma 4, $\mathrm{d}x/\mathrm{d}\theta_t$

is pinned down by the partial differential equation (8), which must hold for all $x_t \geq x(\bar{\theta}(\lambda_t))$. Moreover, whenever $\bar{\theta}(\lambda_t) < 1$, in equilibrium $\hat{\theta}(x(1)) = 1$ and, therefore, it obviously must hold that $x(1, \lambda_t) = \arg\max_{x_t} \tilde{V}^I(1, 1, x_t)$, providing a boundary condition for (8). Because $\tilde{V}^I$ is independent of λ_t, it follows that $x(\theta_t, \lambda_t)$ is uniquely characterized by a function, i.e., $\xi : \theta_t \mapsto x_t$, for all $\theta_t \geq \bar{\theta}(\lambda_t)$. $\diamond$

LEMMA 6: *$\bar{\theta}(\lambda_t)$ is unique.*

Proof. Suppose to the contrary that $\bar{\theta}(\lambda_t)$ is not unique. Then there exist $\bar{\theta}'' > \bar{\theta}'$, defining two distinct equilibria for a given λ_t. By Lemma 5, there is a unique $\xi(\theta)$ characterizing reforms outside the pool for both equilibria. Optimality for type $\theta \in (\bar{\theta}', \bar{\theta}'')$ then requires $\tilde{V}^I(\theta, \theta, \xi(\theta)) \geq \tilde{V}^I(\theta, \bar{\theta}'/2, \lambda_t)$ in the equilibrium defined by $\bar{\theta}'$, and $\tilde{V}^I(\theta, \theta, \xi(\theta)) \leq \tilde{V}^I(\theta, \bar{\theta}''/2, \lambda_t)$ in the equilibrium defined by $\bar{\theta}''$. However, $\tilde{V}^I(\theta, \bar{\theta}'/2, \lambda_t) > \tilde{V}^I(\theta, \bar{\theta}''/2, \lambda_t)$, a contradiction. $\diamond$

This establishes uniqueness of $x(\theta_t, \lambda_t)$, with all properties given by Lemmas 3 and 5, and the corresponding beliefs $\hat{\theta}(\lambda_t, x_t)$ following from Lemma 2 and Bayesian updating. Again, for the purpose of clarity we have established this proposition by focusing on pure strategy equilibria. In the following we outline how the proof generalizes to mixed strategy equilibria; a detailed version of these steps can be attained from the authors on request.

Replicating the proof of Lemma 1, it is trivial to show that if $\tilde{V}^I(\theta', \hat{\theta}(x'), x') = \tilde{V}^I(\theta', \hat{\theta}(x''), x'')$, then $\tilde{V}^I(\theta'', \hat{\theta}(x'), x') < \tilde{V}^I(\theta'', \hat{\theta}(x''), x'')$ for all $\theta' < \theta''$ and $x' < x''$. It follows that (i) supports, $\mathcal{X}(\theta)$, are non-overlapping, and (ii) $\min \mathcal{X}(\theta'') \geq \max \mathcal{X}(\theta')$. Moreover, noting that $\tilde{x}(\theta) \equiv \max \mathcal{X}(\theta)$ has a jump-discontinuity if and only if type θ mixes in a nondegenerate way, (ii) further implies that there can be only finitely many types that mix on the closed interval $[0,1]$. Lemmas 2, 3, and 4 then apply with minor changes, ruling out any jumps of $\tilde{x}$ on $[\bar{\theta}(\lambda_t), 1]$. This leads to the conclusion that at most a mass zero of types (i.e., $\theta_t = \bar{\theta}(\lambda_t)$) could possibly mix in any equilibrium (with no impact on $\hat{\theta}$) and, thus, there is no need to consider any nondegenerate mixed strategies. $\square$

A.4. Proof of Proposition 3

From the discussion in the main body of this chapter it is clear that the equilibrium is uniquely pinned down by the time-invariant mappings given by Propositions 1 and 2 if it exists. We are thus left to show existence, which requires us to verify that the equilibrium mappings are consistent with the D1 and trembling-hand criterion. The first is a direct implication from the proof of Proposition 2 where we apply Lemma 2 to restrict off-equilibrium beliefs, such that $\hat{\theta}$ is necessarily consistent with the D1 criterion.

To show consistency with the concept of trembling-hand perfection, we need to show that $\{\phi_i : i \in [0,1]\}$ and x are best responses to a sequence of completely mixed strategy profiles $\{\{\omega_i^k : i \in [0,1]\}, \sigma^k\}_{k=0}^{\infty}$ that converge to a profile that places all mass on $\{\phi_i : i \in [0,1]\}$ and x, respectively.

Accordingly, for $\phi_i(\hat{\theta}^k(\cdot, x_t), x_t)$ to be a best-response to x_t and the perturbed strategy profile $\{\omega_i^k : i \in [0,1]\}$ for the marginal outsider i with $\gamma_i = \bar{\gamma}(s_t)$, we need that $\hat{\theta}^k(\cdot, x_t)\,\psi(s_t^k(x_t)) = \bar{\gamma}(s_t)$, requiring any change in beliefs along the perturbation path to be offset by trembles of outsiders $j \neq i$. Because for $x \in [\xi(1), 1]$, $\hat{\theta}(\cdot, x) = 1$ can never be sustained in a completely mixed equilibrium with a continuum of types, this implies that we need to adjust for $\hat{\theta}^k(\cdot, x) < \hat{\theta}(\cdot, x)$ by introducing asymmetric trembles, leading to $s^k(x) > s(\hat{\theta}(\cdot, x), x)$. Hence, let $s^k(x(1)) = s(x(1)) + \varepsilon_k$ for some $\{\varepsilon^k\}_{k=0}^{\infty}$ such that $\lim_{k\to\infty} \varepsilon^k = 0$ and $\varepsilon^k \in (0, \bar{\varepsilon})$ for all k.

A necessary (and for $\theta \in (\bar{\theta}(\cdot), 1)$ sufficient) condition for $x \in [\xi(\bar{\theta}(\cdot)), \xi(1)]$ to be optimal against s^k is that $s^k(x)$ satisfies the inverse differential equation (8) for $x(\cdot, \theta)$ fixed,

$$\frac{\mathrm{d}s^k}{\mathrm{d}x} = -\frac{\partial V^I/\partial x}{\partial V^I/\partial s}\bigg|_{s=s^k}, \tag{9}$$

which in combination with $s^k(x(1))$ pins down $s^k(x)$ for all $x \in [\xi(\bar{\theta}), \xi(1)]$. Note that $s^k(x(1)) > s(\cdot, x(1))$ implies that $s^k(x) > s(\cdot, x)$ for all $x \in [\xi(\bar{\theta}), \xi(1)]$ since the indifference condition (8) is unique. Moreover, since optimality of x requires that $\bar{\theta}$ is necessarily indifferent between λ_t and $\xi(\bar{\theta})$, $s^k(\xi(\bar{\theta}))$ pins down $s^k(\lambda_t) > s(\cdot, \lambda_t)$.

For off-equilibrium $x \in (\lambda, \xi(\bar{\theta})) \cup (\xi(1), 1]$ we are free to assign any $s^k(x)$ that (1) assures optimality of x, and (2) converges to $s(\cdot, x)$. As to (1), we can for instance set $s^k(x) = s(\bar{\theta}, x) + s^k(\xi(\bar{\theta})) - s(\cdot, \xi(\bar{\theta}))$ for $x \in (\lambda, \xi(\bar{\theta}))$ (which is continuous around $\xi(\bar{\theta})$ and has slope $\mathrm{d}s(\bar{\theta}, x)/\mathrm{d}x \geq \mathrm{d}s^k(\xi(\bar{\theta}))/\mathrm{d}x$, so that by (9) no type has an incentive to deviate), and $s^k(x) = s(\cdot, x) + \varepsilon^k f^k(x)$ for $x \in (\xi(1), 1]$ with some $f^k : [\xi(1), 1] \to \mathbb{R}_+$ such that $\mathrm{d}f^k(\xi(1))/\mathrm{d}x = \{\mathrm{d}s^k(\xi(1))/\mathrm{d}x - \mathrm{d}s(\cdot, \xi(1))/\mathrm{d}x\}/\varepsilon^k$ and f^k sufficiently convex for V^I to be concave on $[\xi(1), 1]$, so that $\xi(1)$ is the global optimum for $\theta = 1$.

Note that these definitions imply that $s^k(x) \downarrow s(\hat{\theta}(\cdot, x), x)$ for all x and, hence, $\hat{\theta}^k(\cdot, x) \uparrow \hat{\theta}(\cdot, x)$ for all x as implied by the indifference condition of the marginal outsider, $\hat{\theta}^k(x) = \bar{y}(s(\cdot,x))/\psi(s^k(x)) \in (0, \hat{\theta}(\cdot,x))$. By construction, these sequences assure optimality of $\{\phi_i : i \in [0,1]\}$ and x along the perturbation path. To conclude the proof it therefore suffices to show the existence of $\{\{\omega_i^k : i \in [0,1]\}, \sigma^k\}_{k=0}^{\infty}$ yielding $\{s^k, \hat{\theta}^k\}_{k=0}^{\infty}$.

Consider $\{s^k\}_{k=0}^{\infty}$ first. Define $\tilde{\varepsilon}$ such that $\max_x s^k(x) < 1 - \lambda$ for $\varepsilon^k = \tilde{\varepsilon}$ and suppose that $\bar{\varepsilon} \leq \tilde{\varepsilon}$.[17] Then any s^k can be sustained by setting

$$\omega_i^k(1)(x) = \begin{cases} 1 - \varepsilon^k & \text{for all } i : y_i \leq \bar{y}(s(\hat{\theta}(\cdot,x),x)) \\ c^k(x)\varepsilon^k & \text{for all } i : y_i > \bar{y}(s(\hat{\theta}(\cdot,x),x)), \end{cases}$$

with $c^k(x) = \{s^k(x) - (1-\varepsilon^k)s(\cdot,x)\}/\{(1-x)(1-\bar{y}(x))\varepsilon^k\}$. Note that ω_i^k is completely mixed if $\bar{\varepsilon} < 1$ and $\varepsilon^k c^k(x) \in (0,1) \iff c^k(x) \in (0, 1/\varepsilon^k) \iff s^k(x) + \varepsilon^k s(\cdot,x) < 1 - x$. From $s^k(x) > s(\cdot,x)$ we have that $c^k(x) > 0$ and because $s^k \to s$, using the same arguments as in Footnote 17, there exists some $\hat{\varepsilon}$ such that $c^k(x) < 1/\varepsilon^k$ holds for all $\bar{\varepsilon} \leq \hat{\varepsilon}$.

Finally, consider $\{\hat{\theta}^k\}_{k=0}^{\infty}$. It is straightforward to verify by Bayes rule that any $\hat{\theta}^k$ with $\hat{\theta}^k(x) > 0$ for all x can be sustained by setting

$$\sigma^k(x)(\theta,\cdot) = \begin{cases} \varepsilon^k & \text{if } \theta > \hat{\theta}^k(x) \text{ and } (x > \lambda_t \text{ or } \theta > \bar{\theta}_t) \\ d^k(x)\varepsilon^k & \text{if } \theta < \hat{\theta}^k(x) \text{ and } x > \lambda_t \\ 1 - R^k(\theta) & \text{if } \theta \geq \bar{\theta}(\lambda_t) \text{ and } x = \xi(\theta) \\ T^k & \text{if } \theta \leq \hat{\theta}^k(\lambda_t) \text{ and } x = \lambda_t \\ Z^k & \text{if } \theta \in (\hat{\theta}^k(\lambda_t), \bar{\theta}(\lambda_t)) \text{ and } x = \lambda_t, \end{cases}$$

with $d^k(x) = (1 - \hat{\theta}^k(x))^2/\hat{\theta}^k(x)^2$, $R^k(\theta) = \int_{\theta > \hat{\theta}(x)} \varepsilon^k \, \mathrm{d}x + \int_{\theta < \hat{\theta}(x)} d^k(x)\varepsilon^k \, \mathrm{d}x$, $T^k = \inf_{\theta < \bar{\theta}(\lambda_t)}(1 - R^k(\theta))$, and $Z^k = \{T^k \hat{\theta}^k(x)^2 + \varepsilon^k[2(1 - \bar{\theta}(\lambda_t))\hat{\theta}^k(\lambda_t) - 1 + \bar{\theta}(\lambda_t)^2]\}/\{\bar{\theta}(\lambda_t) - \hat{\theta}^k(\lambda_t)\}^2$. With a slight abuse of notation, in the definition of σ^k, R^k , T^k and Z^k denote probabilities, while ε^k are understood to be probability densities. Note that σ^k is completely mixed if $T^k, R^k(\theta) \in (0,1)$ and $Z^k \in (0, R^k(\theta))$ for all θ. This is obviously true for some $\breve{\varepsilon}$, such that $\bar{\varepsilon} < \breve{\varepsilon}$. Finally, note that the above definition is incomplete in the sense that $R^k(\theta) + T^k < 1$ or $R^k(\theta) + Z^k < 1$ for some types $\theta < \bar{\theta}(\lambda_t)$. In these cases the remaining probability

[17] To see that $\tilde{\varepsilon}$ exists, note that $s(\hat{\theta}(\cdot,x),x) < 1 - x \leq 1 - \lambda_t$ since otherwise $\bar{y}_t = 1$, which requires $\bar{\theta}_t = 1$ and $s_t = 1$, contradicting that s is strictly decreasing in x. Convergence of s^k to s then implies that one can always find some $\tilde{\varepsilon}$ that is sufficiently small.

mass can be distributed (almost) arbitrary over atoms on $(\lambda_t, 1]$ without impact on the resulting beliefs.[18]

We conclude the proof by setting $\bar{\varepsilon} = \min\{1, \tilde{\varepsilon}, \hat{\varepsilon}, \breve{\varepsilon}\}$. □

A.5. *Proof of Proposition 4*

Consider $Q^R(\lambda_t, (\frac{1}{2}, 1]) = 1$ first. By Proposition 2, for any reform $x_t > \lambda_t$, $x_t = \xi(\theta_t)$, with ξ increasing. To show the claim, it thus suffices to show that $\tilde{x} \equiv \xi(\tilde{\theta}) > 1/2$ for $\tilde{\theta} = \min_\lambda \bar{\theta}(\lambda)$. Also, define $\tilde{\lambda} = \arg\min_\lambda \bar{\theta}(\lambda)$. Then, optimality of $\tilde{x}$ implies $s^* \equiv s(\tilde{\theta}/2, \tilde{\lambda}) > s(\tilde{\theta}, \tilde{x}) \equiv s^{**}$. Using (6),

$$s^* = (\tilde{\theta}/2)(1-\tilde{\lambda})\,\psi(s^*) \equiv w^*\psi(s^*), \tag{10}$$

$$s^{**} = \tilde{\theta}(1-\tilde{x})\,\psi(s^{**}) \equiv w^{**}\psi(s^{**}). \tag{11}$$

Note that, in analogue to the proof of Proposition 1, for a general $w_t \equiv \hat{\theta}_t(1-x_t)$ it holds that

$$\frac{\partial s_t}{\partial w_t} = -\psi(s_t)\left(\frac{\partial \pi_t}{\partial s_t}\right)^{-1} > 0.$$

Hence, $s^* > s^{**}$ implies $w^* > w^{**}$, or $(\tilde{\theta}/2)(1-\tilde{\lambda}) > \tilde{\theta}(1-\tilde{x})$. Rearranging, then proves the claim,

$$\tilde{x} > 1 - \frac{1-\tilde{\lambda}}{2} \geq \frac{1}{2}.$$

Now consider $Q^S(\lambda_t, (0, \frac{1}{2})) = 1$. Again, optimality of x_t implies that $s(\hat{\theta}(\lambda_t, x), x)$ is decreasing in x. Hence, for all λ_t,

$$s(\hat{\theta}(\lambda_t, x_t), x_t) \leq s(\bar{\theta}(\lambda_t)/2, \lambda_t) \leq s(1/2, 0),$$

where the last inequality follows since s is increasing in its first and decreasing in its second argument. Hence, it suffices to show that $s(1/2, 0) < 1/2$.

Let $s^* \equiv s(1,0) \leq 1$ and let $s^{**} \equiv s(1/2, 0)$. From (6), $s^* = \psi(s^*)$ and $s^{**} = \psi(s^{**})/2$. Moreover, by Proposition 1, $s^* > s^{**}$. Hence, since ψ is strictly increasing,

$$s^{**} = \frac{\psi(s^{**})}{2} = \frac{\psi(\psi(s^{**})/2)}{2} < \frac{\psi(\psi(s^*)/2)}{2} = \frac{\psi(s^*/2)}{2} < \frac{\psi(s^*)}{2} = \frac{s^*}{2} \leq \frac{1}{2}.$$

□

[18] For instance, we can dispose of the atomic waste without any hazard by having each type θ place the remaining probability mass on $x = \lambda_t + \theta(1-\lambda_t)/\bar{\theta}(\lambda_t)$.

A.6. Proof of Proposition 5

From Footnote 13,

$$\rho^S(\lambda_t) = \int_0^{\bar{\theta}(\lambda_t)} \theta h\left(s\left(\bar{\theta}(\lambda_t)/2, \lambda_t\right)\right) \mathrm{d}\theta + \int_{\bar{\theta}(\lambda_t)}^1 \theta h\left(s\left(\theta, x(\theta)\right)\right) \mathrm{d}\theta, \quad (12)$$

and

$$\rho^R(\lambda_t) = \int_{\bar{\theta}(\lambda_t)}^1 \left(1 - \theta h\left(s\left(\theta, x(\theta)\right)\right)\right) \mathrm{d}\theta. \quad (13)$$

Also, note that $\bar{\theta}(\lambda_t) \in (0,1]$ is implicitly defined as the solution to

$$F(\bar{\theta}, \lambda_t) \equiv \tilde{V}^I(\bar{\theta}, \bar{\theta}/2, \lambda_t) - \tilde{V}^I(\bar{\theta}, \bar{\theta}, \xi(\bar{\theta})) = 0, \quad (14)$$

if an interior solution exists. Otherwise, for λ_t there is a corner solution $\bar{\theta}(\lambda_t) = 1$, which implies $\tilde{V}^I(1, 1/2, \lambda_t) > \tilde{V}^I(1, 1, \xi(1))$.

First, consider $\lambda_t > \bar{\lambda}$. Suppose that there exists $\bar{\lambda}$, such that for all $\lambda_t \in (\bar{\lambda}, 1]$, $\bar{\theta}(\lambda_t)$ is a corner solution. Then clearly for all $\lambda_t > \bar{\lambda}$, $\partial\bar{\theta}(\lambda_t)/\partial\lambda_t = 0$, such that $\partial\rho^S(\lambda_t)/\partial\lambda_t = \partial h(s(1/2, \lambda_t))/\partial\lambda_t < 0$, by Proposition 1. Furthermore, $\partial\rho^R(\lambda_t)/\partial\lambda_t = 0$. Otherwise, if there exists no $\bar{\lambda}$, such that for all $\lambda_t \in (\bar{\lambda}, 1]$, $\bar{\theta}(\lambda_t)$ is a corner solution, then there necessarily exists a λ^*, such that for $\lambda_t \in (\lambda^*, 1]$, $\bar{\theta}(\lambda_t)$ is an interior solution. But then, because $\rho^S(1) = \rho^R(1) = 0$, continuity of ρ^S and ρ^R implies that $\partial\rho^S(\lambda_t)/\partial\lambda_t < 0$ and $\partial\rho^R(\lambda_t)/\partial\lambda_t < 0$ for all $\lambda_t > \bar{\lambda}$ and some $\bar{\lambda} < 1$.

Now consider $\lambda_t < \underline{\lambda}$ and $\bar{\theta}(0) < 1$. Then, F differentiable implies that $\bar{\theta}(\lambda_t)$ has an interior solution and is differentiable for all $\lambda_t \in [0, \lambda^*)$ for some $\lambda^* > 0$. Implicit differentiation of F, substituting for $x'(\bar{\theta})$ from (8), and using $F(\bar{\theta}, \lambda_t) = 0$ yields

$$\frac{\partial\bar{\theta}(\lambda)}{\partial\lambda} = \frac{-\bar{\theta} h_1^p s_2^p u^p + (1 - p^p) u_1^p}{\frac{\bar{\theta}}{2} h_1^p s_1^p u^p + \frac{u^p - u^s}{\bar{\theta}}}, \quad (15)$$

where subscript i denotes the derivative with respect to the ith argument, and superscripts p and s denote that the function is evaluated at the pooling or separating values, respectively (where $\hat{\theta}^p = \frac{\bar{\theta}}{2}$, $x^p = \lambda$ and $\hat{\theta}^s = \bar{\theta}$, $x^s = x(\bar{\theta})$).

Using this, the signs of $\partial\rho^S/\partial\lambda_t$ and $\partial\rho^I/\partial\lambda_t$ are given by

$$\operatorname{sign}\left\{\frac{\partial\rho^S(\lambda_t)}{\partial\lambda_t}\right\} = \operatorname{sign}\left\{u^P\left(\frac{(p^P-p^S)(1-2p^S)}{1-p^S}\right)+ \right.$$
$$\left. + (1-p^P)u_1^P\left((1-\lambda_t)-\frac{2(p^P-p^S)}{\bar{\theta}h_1^P s_2^P}\right)\right\} \quad (16)$$

and

$$\operatorname{sign}\left\{\frac{\partial\rho^R(\lambda_t)}{\partial\lambda_t}\right\} = \operatorname{sign}\left\{-\frac{\partial\bar{\theta}(\lambda_t)}{\partial\lambda_t}(1-p^S)\right\}, \quad (17)$$

where we have used that $(1-p^P)u^P = (1-p^S)u^S$ from (14) and $s_1^P/(-s_2^P) = 2(1-\lambda_t)/\bar{\theta}$ by the proof of Proposition 1.

Evaluated at $\lambda_t = 0$, all terms except u_1 in (16) are strictly positive.[19] Thus, $\partial\rho^S(0)/\partial\lambda_t$ is weakly positive if and only if for $\lambda_t = 0$ it holds that

$$u_1^P \geq -u^P\left(\frac{(p^P-p^S)(1-2p^S)}{1-p^S}\right)\left[(1-p^P)\left((1-\lambda_t)-\frac{2(p^P-p^S)}{\bar{\theta}h_1^P s_2^P}\right)\right]^{-1}. \quad (18)$$

Likewise, note that the sign of $\partial\rho^R/\partial\lambda_t$ is the opposite sign of $\partial\bar{\theta}(\lambda_t)/\partial\lambda_t$. Hence, because all terms except u_1 in (15) are strictly positive, $\partial\rho^R/\partial\lambda_t$ is weakly negative if and only if

$$u_1^P \geq \bar{\theta}h_1^p s_2^p u^p(1-p^P)^{-1}. \quad (19)$$

Let u' and u'' be the values of the right hand sides of (18) and (19) when evaluated at $\lambda_t = 0$. Then, from our discussion above it follows, that $\partial\rho^S(0)/\partial\lambda_t > 0$ and $\partial\rho^R(0)/\partial\lambda_t < 0$ if $u_1(0) > \bar{u} \equiv \max\{u', u''\}$. The converse—that is, $\partial\rho^S(0)/\partial\lambda_t < 0$ and $\partial\rho^R(0)/\partial\lambda_t > 0$—holds true, if $u_1(0) < \underline{u} \equiv \min\{u', u''\}$. Differentiability of ρ^S and ρ^R around 0 thus establishes the claim for all $\lambda_t \in [0, \underline{\lambda}]$ for some $\underline{\lambda} > 0$. □

[19] Note that $p^S = \bar{\theta}h(s^S) < 1/2$ for $\lambda_t = 0$ is not obvious. To see that this is indeed the case, assume to the contrary $p^S > 1/2$ implying $p^P = \bar{\theta}/2\,h(s^P) > 1/2$. By Proposition 4, $s^P = \bar{\theta}/2\,h(s^P)u(s^P) = p^P/2\,u(s^P) < 1/2$ and hence $u(s^P) < 1/p^P < 2$ by $p^P > 1/2$. Furthermore, optimality of $\bar{\xi} \equiv \xi(\bar{\theta})$ requires $(1-p^S)u(\bar{\xi}) \geq 1$, since an indirect utility of 1 is always attainable by setting $x = 1$. This implies $u(\bar{\xi}) \geq 2$ by $p^S > 1/2$. Thus, $p^S > 1/2$ implies $u(s^P) < 2 \leq u(\bar{\xi})$ for $\lambda_t = 0$. However, by Proposition 4, $s^P < 1/2 < \bar{\xi}$ such that $u(s^P) > u(\bar{\xi})$, a contradiction.

References

Acemoglu, D. and J. A. Robinson (2000a), "Democratization or Repression?" European Economic Review, 44:683–693.

——— (2000b), "Why Did the West Extend the Franchise? Democracy, Inequality, and Growth in Historical Perspective," Quarterly Journal of Economics, 115:1167–1199.

——— (2001), "A Theory of Political Transitions," American Economic Review, 91:938–963.

Acemoglu, D., D. Ticchi, and A. Vindigni (2010), "A Theory of Military Dictatorships," American Economic Journal: Macroeconomics, 2:1–42.

Aidt, T. S. and P. S. Jensen (2012), "Workers of the World, Unite! Franchise Extensions and the Threat of Revolution in Europe, 1820-1938," Cambridge Working Paper in Economics.

Angeletos, G.-M., C. Hellwig, and A. Pavan (2006), "Signaling in a Global Game: Coordination and Policy Traps," Journal of Political Economy, 114:452–484.

Boix, C. (2003), Democracy and Redistribution, Cambridge and UK: Cambridge University Press.

Bourguignon, F. and T. Verdier (2000), "Oligarchy, Democracy, Inequality and Growth," Journal of Development Economics, 62:285–313.

Cervellati, M., P. Fortunato, and U. Sunde (2007), "Consensual and Conflictual Democratization, Rule of Law, and Development," CEPR Discussion Paper, 6328.

——— (2011), "Democratization and Civil Liberties: The Role of Violence during the Transition," IZA Discussion Paper, 5555.

Cheibub, J. A., J. Gandhi, and J. R. Vreeland (2010), "Democracy and Dictatorship Revisited," Public Choice, 143:67–101.

Cho, I.-K. and D. M. Kreps (1987), "Signaling Games and Stable Equilibria," Quarterly Journal of Economics, 102:179–221.

Conley, J. P. and A. Temini (2001), "Endogenous Enfranchisement when Groups' Preferences Conflict," Journal of Political Economy, 79–102.

Edmond, C. (2011), "Information Manipulation, Coordination, and Regime Change," NBER Working Paper, 17395.

Elkins, Z., T. Ginsburg, and J. Melton (2010), "Chronology of Constitutional Events, Version 1.1," .

Ellis, C. J. and J. Fender (2011), "Information Cascades and Revolutionary Regime Transitions," Economic Journal, 121:763–792.

Gates, S., H. Hegre, M. P. Jones, and H. Strand (2006), "Institutional Inconsistency

and Political Instability: Polity Duration, 1800–2000," American Journal of Political Science, 50:893–908.
Goemans, H. E., K. S. Gleditsch, and G. Chiozza (2009), "Introducing Archigos: A Dataset of Political Leaders," Journal of Peace Research, 46:269–283.
Gradstein, M. (2007), "Inequality, Democracy and the Protection of Property Rights," Economic Journal, 117:252–269.
Huntington, S. P. (1991), The Third Wave: Democratization in the Late Twentieth Century, Norman (OK): University of Oklahoma Press.
Jack, W. and R. Lagunoff (2006), "Dynamic Enfranchisement," Journal of Public Economics, 90:551–572.
Justman, M. and M. Gradstein (1999), "The Industrial Revolution, Political Transition, and the Subsequent Decline in Inequality in 19th-Century Britain," Explorations in Economic History, 36:109–127.
Karl, T. L. (1990), "Dilemmas of Democratization in Latin America," Comparative Politics, 23:1–21.
Kuran, T. (1989), "Sparks and Prairie Fires: A Theory of Unanticipated Political Revolution," Public Choice, 61:41–74.
Lizzeri, A. and N. Persico (2004), "Why did the Elites Extend the Suffrage? Democracy and the Scope of Government, with an Application to Britain's 'Age of Reform,'" Quarterly Journal of Economics, 119:707–765.
Llavador, H. and R. J. Oxoby (2005), "Partisan Competition, Growth, and the Franchise," Quarterly Journal of Economics, 120:1155–1189.
Lohmann, S. (1994), "The Dynamics of Informational Cascades: The Monday Demonstrations in Leipzig, East Germany, 1989–91," World Politics, 47:42–101.
Magaloni, B. and R. Kricheli (2010), "Political Order and One-Party Rule," Annual Review of Political Science, 13:123–143.
Mailath, G. J. (1987), "Incentive Compatibility in Signaling Games with a Continuum of Types," Econometrica, 55:1349–1365.
Marshall, M. and K. Jaggers (2002), "Polity IV Project: Political Regime Characteristics and Transitions, 1800-2002," .
O'Donnell, G. and P. C. Schmitter (1973), Transitions from Authoritarian Rule: Tentative Conclusions about Uncertain Democracies, Baltimore (MD): John Hopkins University Press.
Przeworski, A. (2000), Democracy and Development: Political Institutions and Material Well-Being in the World, 1950-1990, Cambridge (UK): Cambridge University Press.
——— (2009), "Conquered or Granted? A History of Suffrage Extensions," British Journal of Political Science, 39:291.

Rustow, D. A. (1970), "Transitions to Democracy: Toward a Dynamic Model," Comparative Politics, 2:337–363.

Selten, R. (1975), "Reexamination of the Perfectness Concept for Equilibrium Points in Extensive Games," International Journal of Game Theory, 4:25–55.

Chapter 3

OPTIMAL DELEGATED SEARCH WITH ADVERSE SELECTION AND MORAL HAZARD*

Contents

* I am extremely grateful to Klaus M. Schmidt for his guidance throughout this project. I also thank Guy Arie, Niko Matouschek, and Marco Ottaviani for useful comments and suggestions.

Abstract

This chapter of my dissertation analyzes optimal search when it is delegated to an agent. There are two informational asymmetries. First, adverse selection: *ex ante*, prospects of search are privately known by the agent. Second, moral hazard: search itself cannot be observed by the principal. In this environment, the principal's problem is to bring the agent to reveal the optimal search policy and, simultaneously, to induce him to actually search according to the revealed policy. We show that the solution to this problem uses a screening menu, which is exclusively comprised of simple bonus contracts. Search policies are almost surely inefficient; either search is terminated prematurely, or it is completely undirected. In contrast, if either of the two informational asymmetries is resolved, the first-best outcome can be supported in equilibrium.

Keywords

Adverse selection, bonus contracts, delegating search, hidden action, optimal search.

JEL Classification: D82, D83, D86, C72.

1. Introduction

Searching is an important element of many types of agency relationships. Recruiting agencies are hired to search for job candidates. Real estate agents are contracted to search for prospective buyers or, alternatively, to search for attractive houses. Insurance brokers are often hired to find new clients. At a more general level, many more agency relationships require agents to be "original" rather than performing routine tasks and can be thought of as search agencies. This includes all relationships where agents are expected to "think" and search for "ideas" in order to provide a solution for a given problem; e.g., research centers exploring new product designs, advocates searching for good defense strategies, and business consultancies searching for promising business plans.

This chapter of my dissertation analyzes optimal searching when it is delegated to an agent. For this purpose, we consider the canonical search model introduced by McCall (1970), in which a single agent sequentially samples "solutions" from a time-invariant distribution. The framework gives rise to an optimal stopping rule that determines at which point the agent stops sampling new solutions in order to utilize the best available one.

Our study only deviates from this setting by assuming that the payoffs from adopting a solution are not realized by the agent who operates the searching, but from another agent—i.e., the "principal". There are two informational asymmetries that govern the relationship. First, motivated by the agent's role as an expert in the aforementioned examples, the agent has an informational advantage over the principal in assessing the prospects of searching. Specifically, we assume that payoffs x are sampled from a time-invariant, but state-dependent distribution $F(x|\theta)$, upon which θ is private information of the agent. Second, search itself cannot be observed (or verified) by the principal, a natural assumption given the soft and unverifiable nature of finding qualified job candidates, serious buyers, or good "solutions". In this environment, the principal's problem is to bring the agent to reveal the optimal search policy (which depends on θ) and, simultaneously, to induce him to actually search according to the revealed policy.

Using this model we study how delegation affects the searching process, and what contractual arrangements a profit maximizing principal would offer to optimally delegate search to the agent. A key aspect of the analysis is the dealing with the simultaneous presence of adverse selection and moral hazard and showing how the interaction of these asymmetries affects the optimal contract design.

Not surprisingly, we find that searching is almost surely inefficient. Search is either stopped too early (compared to the efficient solution) or is completely undirected (leading the agent to unconditionally adopt the first solution he finds). These results lend theoretical support to recent empirical evidence provided by Rutherford et al. (2005) and Levitt and Syverson (2008) that the search effort by real estate agents is inefficiently low.

The combination of the two informational asymmetries is crucial for the inefficiency. If either adverse selection or moral hazard is shut down, then the efficient benchmark can be implemented as a rational equilibrium outcome. Each of the two asymmetries acts hereby as an "catalyst" for the other one: If search is unobservable but the prospects of searching are known to the principal, then the efficient benchmark could be achieved by the agent "buying" all prospective benefits from searching. Once θ is hidden from the principal, this resolution is undermined by the price being exposed to adverse selection. Conversely, if the moral hazard is dissolved, then the principal could simply reimburse the agent for his search efforts and thereby solve the adverse selection problem, a strategy which is undermined by moral hazard forcing him to also provide search incentives to the agent.

Regarding the optimal contract design, we show that search is optimally delegated through the use of a screening menu, which is exclusively comprised of simple bonus contracts. These contracts pay a fixed bonus to the agent when an *ex ante* specified target is reached, and nothing otherwise. All other information about the realized solution is optimally ignored. Underlying this result is the adverse selection problem which causes any contracts that are more sensible to the realized payoffs to increase the costs of bringing the agent to reveal the prospects of search. On the other hand, moral hazard precludes the usage of even simpler "fixed wage" agreements, since then the agent would not be induced to provide any search effort.

Having in mind the general interpretation of search agencies as problem-solving specialist-agents, the optimality of bonus contracts provides a novel explanation for the observed popularity of this simple compensation scheme.[1] This finding is related to a small but important literature that shows how in environments that are more complex than baseline moral hazard models (e.g., Holmstrom, 1979) optimal compensation schemes can be simpler than the baseline analysis would suggest. In particular Townsend (1979), Holmstrom and Milgrom (1987), and Innes (1990) have shown in their seminal contributions

[1] See, for instance, Moynahan (1980) for a documentation of the widespread usage of bonus contracts in many industries, and Oyer (2000) for specific evidence on the frequent usage of bonus schemes for salesmen in the food manufacturing industry.

that linear contracts (Holmstrom and Milgrom) and standard debt contracts (Townsend and Innes) are optimal. More recently Herweg et al. (2010) have used the idea that agents have nonstandard preferences that make them averse to losses compared to an expectations-based reference point in order to explain the usage of bonus schemes.[2] Our approach deviates from these studies in that we use a combination of adverse selection and moral hazard to argue that contracts are simple because simplicity helps solving the adverse selection problem.

For surveys of the literature on optimal search, see, for example, Mortensen (1986) and Rogerson et al. (2005). So far, this literature has primarily focused on single agent decision problems. Two important exceptions are Lewis and Ottaviani (2008) and Lewis (2012). In these related contributions, Lewis and Ottaviani analyze delegated search problems that consist of multiple stages over which a principal repeatedly interacts with the same agent. In contrast to this study, these papers focus in informational asymmetries that emerge within a repeated agency relationship that differs considerably from the standard search model. While both studies also find that search is inefficient, the contracts that are optimally used in these dynamic relationships are considerably more complex than the simple contracts which we find to be optimal in our setting.

The remainder of this chapter is organized as follows. Section 2 presents the model. Section 3 provides the first-best benchmark and shows how it can be implemented as an equilibrium outcome if either adverse selection or moral hazard is shut down. Section 4 analyzes the solution to our model with both informational asymmetries. Section 5 concludes. All proofs are deferred to the Appendix.

2. A simple model of delegated search

In this section, we set up a simple search model, in which search is operated by an agent and payoffs from search are realized by a principal. Both the principal and the agent are risk-neutral. The "problem" of the principal is of generic nature. In order to solve the problem, the principal hires an agent. The job of the agent

[2] It is also known that the combination of risk-neutrality, limited liability and moral hazard makes bonus schemes the unique optimal contract (see, e.g., Park, 1995; Kim, 1997; Oyer, 2000 and Demougin and Fluet, 1998. As demonstrated by Jewitt et al. (2008), this finding breaks, however, down if agents are risk-averse to only the slightest degree. In contrast, even though agents are risk averse in our setting, too, our results do not rely on the preferences of the agent.

is to first sample a selection of different solutions, and then to select a specific one. Importantly, different solutions are also differently valued by the principal.

We denote the value of a solution by x, $x \in X$, where $X = [0, B]$. The agent can sample solutions at constant costs c from a distribution $F(x|\theta)$, which depends on an exogenously given state of the world θ. We assume that F is twice continuously differentiable in x and θ. The state of the world is randomly selected before the principal contracts the agent, and has a support equal to $[\underline{\theta}, \bar{\theta}]$, denoted by Θ. The prior cumulative distribution function of θ is common knowledge, is denoted by P, and has a differentiable density p such that $p(\theta) > 0$ for all $\theta \in \Theta$. Each time the agent samples a new solution, he can either adopt that solution, or continue to search for other solutions. The outside option from not adopting any solution and not contracting is normalized to zero for both the principal and the agent. We restrict attention to the case where, in the absence of informational asymmetries, solving the problem is profitable in all states of the world, that is $\mathrm{E}\{x|\theta\} \geq c$ for all $\theta \in \Theta$.

To complete the model, we need to specify the information that is available to the principal and the agent. We impose the following two informational asymmetries.

Assumption 1 (Adverse Selection): *The state of the world θ is privately revealed to the agent before he contracts with the principal. The principal knows the set of potential states Θ and their distribution $P(\theta)$.*

Assumption 2 (Hidden Action): *Search by the agent and the sampled selection of solutions cannot be observed by the principal. However, the value of the adopted solution is observable and verifiable.*

Assumptions 1 and 2 define the two uncertainties which the principal faces in the main model. Assumption 1 states that the principal does not know the state of the world. This implies that the principal relies on the agent to select the optimal search policy, since optimal search generally depends on the state of the world. Assumption 2 adds an additional dimension of uncertainty by assuming that search by the agent cannot be observed. The principal learns the value of the solution only after the agent adopts it. This implies that the principal has to rely on the contractual arrangements to ensure that the agent not only reveals the optimal search policy but also searches according to this policy.

3. Preliminary analysis

Before proceeding to the main analysis, we briefly describe the first-best benchmark, and analyze the model's solution when either Assumption 1 or Assumption 2 is relaxed.

Suppose that the principal is both able to observe the state of the world and to monitor search by the agent. Then, in the first best, search by the agent maximizes the (joint) surplus of search. Therefore first-best search policies are identical to the optimal search policies in the standard search model. We skip the derivation and simply state the solution in the following observation. For details, see for instance McCall (1970).

OBSERVATION: *In the first best the agent searches as long as for all previously sampled solutions it holds that $x \leq \bar{x}^{FB}(\theta)$. Otherwise he stops search and adopts the last-sampled solution. The first-best stopping rule is given by a function $\bar{x}^{FB}$: $\Theta \to X$, which is defined pointwise, such that $\bar{x}^{FB}(\theta)$ for a given state θ is pinned down by the following condition:*

$$c = \int_{\bar{x}^{FB}(\theta)}^{B} (x' - \bar{x}^{FB}(\theta))\, \mathrm{d}F(x'|\theta). \tag{1}$$

In the first best, an agent who knows the world to be in state θ, searches for better solutions until he finds one of at least a value of $\bar{x}^{FB}(\theta)$. The optimal "stopping rule" $\bar{x}^{FB}(\theta)$ equates the marginal expected benefits of finding a better solution than $\bar{x}^{FB}(\theta)$ with the marginal costs of searching c.

To better understand the mechanics of our model consider now a situation where the principal is able to observe and verify the selection of solutions which the agent has sampled, but faces uncertainty from not knowing the true state of the world. This resembles a situation where assumption 1 holds, but 2 does not. In this case, the first-best outcome, in which the agent pursues the first-best optimal search policies in every state of the world, can be implemented by using a simple contractual arrangement. Essentially all we have to do is compensate the agent for his search costs independently from the search policy he pursues, and then he finds it (weakly) optimal to search according to the first best. We state the precise result in the following.

PROPOSITION 1: *Suppose that Assumption 1 holds, but that the principal is able to observe and verify the selection of solutions which the agent has sampled. Then the first-best search policies can be implemented by paying a transfer $T(N)$ to the agent after he adopts a solution, where $T(N) = N\,c$, and N is the number of solutions in the final sample.*

Proof Sketch. Since for the above described contract the agent breaks even independently of his search behavior, there exists a first-best equilibrium, where the agent accepts the contract, pursues first-best search policies, and adopts the solution that yields the highest value. This is trivially true as none of these choices is payoff-relevant under the considered contract. Moreover, it is also trivially true that the principal has no incentive to deviate from the first-best equilibrium by offering another contract. □

The point here is that by offering a contract that fully compensates the agent for his search costs, the principal can provide a contract, in which the agent's private knowledge about the state of the world is not payoff-relevant to him. As a result, the agent is willing to reveal the state of the world without any explicit incentives. Critical to this contract is that the principal is able to verify the sampled selection of solutions, allowing him to assess the *actual* costs of the agent. This is not possible anymore once we introduce Assumption 2, preventing the principal from differentiating bad luck while searching from a fundamentally bad distribution. In this sense Assumption 2 *catalyzes* Assumption 1 by rendering the agent's private information necessarily payoff-relevant for any non-constant contract.

Note, however, that also in the case where Assumption 2 holds but Assumption 1 does not, there exists again a simple contract which implements the first-best search policies. In this case it is sufficient that the agent is the residual claimant, as it then will be in his own interest to pursue first-best search policies.

Proposition 2: *Assume 2. Suppose that the principal learns the state of the world prior to contracting the agent. Then first-best search policies can be implemented by paying a transfer $T(x)$ to the agent after he adopts a solution, where $T(x) = -\bar{x}^{FB}(\theta) + x$, and $\bar{x}^{FB}(\theta)$ is the first-best stopping rule in state θ.*

Proof Sketch. Since the agent effectively becomes the residual claimant under the described contract, search obviously is efficient. The only question is, whether both the principal and the agent would agree to the price $\bar{x}^{FB}(\theta)$ that the agent pays to become the residual claimant. To see that this is indeed the case, note that by condition (1), the first-best expected surplus, $[\bar{F}(\bar{x}^{FB}(\theta)|\theta)]^{-1} \times$ $\times(\int_{\bar{x}^{FB}(\theta)}^{B} x' \,\mathrm{d}F(x'|\theta) - c)$, is equal to $\bar{x}^{FB}(\theta)$. Hence the principal reaps all the surplus, and therefore happily proposes this contract, which the agent accepts in equilibrium since he breaks even. □

As before, this contract is not feasible anymore as soon as we introduce both informational asymmetries simultaneously. The reason is that with θ unknown the price $\bar{x}^{FB}(\theta)$ will be subject to adverse selection, as originally pointed out by Akerlof (1970). In this sense, adverse selection unleashes the moral hazard problem, much alike risk aversion and limited liability do in other moral hazard setups.

We conclude that whenever the principal is either able to perfectly monitor search by the agent, or is equally well informed about the state of nature, delegated search comes without any efficiency loss and is identical to search by a single agent. Only when the agent has some informational advantage about the state of world (e.g., because he is an expert) *and* search cannot be perfectly monitored, delegated search may differ from the standard search model. The remainder of the chapter analyzes how to optimally deal with such a situation.

4. Optimal delegation of search

The problem of the principal when, both, search and the state of the world is unobservable is to bring the agent to reveal his information on the state of the world θ, and simultaneously induce him to actually search according to the search policies that the principal finds optimal given θ. An important feature of this problem is that search has to be self-enforcing given the contractual arrangements. That is, given Assumption 2, a contract is simply a function $T : [0, B] \to \mathbb{R}$, which specifies, for every solution x, a transfer from the principal to the agent. For any contract T, searching is then determined by the search policy which optimizes the agent's payoff given the state of the world θ.

Taking into account these search policies, the principal's objective is to maximize her expected payoffs. By the revelation principle, a solution to this problem may be obtained via a direct revelation mechanism in which the agent truthfully reports the state of the world, and for each state θ is assigned a contract T_θ. The principal's problem is then to find the optimal set of contracts $\{T_\theta\}_{\theta\in\Theta}$.

We approach this problem as follows. Since the search environment of the agent is designed by the principal through the choice of $\{T_\theta\}_{\theta\in\Theta}$, we first characterize the solution to the search problem of the agent for an arbitrary contract T. With the solution to this problem at hand, we then turn to the optimization problem of the principal and obtain some defining properties of the optimal menu. In particular, we establish the optimality of bonus contracts. After simplifying the problem accordingly, we then compute the optimal menu and derive the equilibrium search policies.

4.1. Search problem of the agent

Once the agent has chosen a contract from the menu offered to him, sequential rationality requires that he pursues the search policy which is then optimal for him. Since the agent is effectively facing a search problem over the transfers specified by the chosen contract, equilibrium search is characterized by the solution to this search problem. Because the underlying distribution $F(x|\theta)$ is in terms of solutions x rather than in terms of $T_{\tilde{\theta}}(x)$, we introduce an indicator function that indicates whether for a particular x, $T_{\tilde{\theta}}(x)$ is smaller or larger than the optimal stopping rule of the agent. Otherwise, search by the agent resembles the results of standard search theory.

LEMMA 1: *An agent with distribution θ and contract $T_{\tilde{\theta}}$ searches as long as for all previously sampled solutions it holds that $T_{\tilde{\theta}}(x) \leq \bar{T}_{\tilde{\theta}}(\theta)$. Otherwise he stops search and adopts the last-sampled solution. Let $\psi_{\tilde{\theta}} : X \times \mathbb{R} \to \{0, 1\}$ be an indicator function, such that*

$$\psi_{\tilde{\theta}}(x, \bar{T}_{\tilde{\theta}}(\theta)) = \begin{cases} 0 & \text{if } T_{\tilde{\theta}}(x) \leq \bar{T}_{\tilde{\theta}}(\theta), \text{ and} \\ 1 & \text{if } T_{\tilde{\theta}}(x) > \bar{T}_{\tilde{\theta}}(\theta). \end{cases}$$

Then the stopping rule is given by function $\bar{T}_{\tilde{\theta}} : \Theta \to X$, which is defined pointwise, such that $\bar{T}_{\tilde{\theta}}(\theta)$ is pinned down by

$$c = \int \Big(\big\{ T_{\tilde{\theta}}(x') - \bar{T}_{\tilde{\theta}}(\theta) \big\} \cdot \psi_{\tilde{\theta}}(x', \bar{T}_{\tilde{\theta}}(\theta)) \Big) \, \mathrm{d}F(x'|\theta) \tag{2}$$

whenever $\int T_{\tilde{\theta}}(x') \, \mathrm{d}F(x'|\theta) \geq c$. Otherwise the agent does not search at all.

Here the optimal stopping rule equates marginal costs of searching c and the marginal expected benefits from finding a solution x' which yields a higher transfer $T(x')$ as given by the right hand side of equation (2).

Because the search policy of the agent simplifies to a stopping rule in terms of payoffs $T_{\tilde{\theta}}(x)$ rather than in terms of the value of the underlying solution x, the solution to (2) does not necessarily map back into a unique solution to the principal's problem. To ensure that T_θ is invertible, we impose the following assumption.

ASSUMPTION 3: *Contracts are monotonically increasing, i.e. $T_\theta(x') \leq T_\theta(x'')$ for all $(x', x'', \theta) \in \{X^2 \times \Theta \,|\, x' \leq x''\}$.*

It is well known that this assumption can be rationalized by free disposal. That is, given free disposal, the agent can guarantee himself a payoff of $T^*_\theta(x) \equiv \max_{x' \in [0,x]}\{T_\theta(x')\}$. Hence, w.l.o.g., one could replace T_θ by $\hat{T}_\theta$, which for all x, pays $\hat{T}_\theta(x) = T^*_\theta(x)$. It can easily be verified that $\hat{T}_\theta$ is indeed increasing in x. Intuitively, Assumption 3 thus requires that the agent can freely downscale any realized solution.

Given that all contracts $\{T_\theta\}_{\theta\in\Theta}$ are increasing in x, we can reformulate the optimal search policy of the agent in terms of the value of the underlying solution. Then the stopping rule of the agent in terms of x is given by,

$$\bar{x}(T_{\tilde{\theta}}, \theta) = \max_x \left\{x : T_{\tilde{\theta}}(x) \le \bar{T}_{\tilde{\theta}}(\theta)\right\} . \tag{3}$$

Although (3) uniquely identifies $\bar{x}(T_{\tilde{\theta}}, \theta)$, it will be convenient to formulate the search policies directly in terms $T_{\tilde{\theta}}$. Because it may be optimal (and indeed will be in equilibrium) to offer a contract that is discontinues at $\bar{x}(T_{\tilde{\theta}}, \theta)$, equilibrium search may be given by a corner solution. The following characterization accounts for that.

Proposition 3: *Suppose Assumptions 2 and 3 hold. Let $\mathcal{M}$ be the space of monotonically increasing functions $X \to \mathbb{R}$. Then search by the agent is completely summarized by function $\bar{x} : \mathcal{M} \times \Theta \to X$, which specifies, for a contract $T_{\tilde{\theta}} \in \mathcal{M}$ and a state of the world $\theta \in \Theta$, a number $\bar{x}(T_{\tilde{\theta}}, \theta)$, such that the agent searches as long as for all previously sampled solutions it holds that $x \le \bar{x}(T_{\tilde{\theta}}, \theta)$. Otherwise he stops search and adopts the last-sampled solution. The stopping rule is given by function $\bar{x}$, which is defined pointwise by the following inequalities.*

$$c \le \int_{\hat{x}}^{B} \left(T_{\tilde{\theta}}(x') - T_{\tilde{\theta}}(\hat{x})\right) \mathrm{d}F(x'|\theta) \quad \textit{for all } \hat{x} \le \bar{x}(T_{\tilde{\theta}}, \theta) \tag{4a}$$

$$c > \int_{\hat{x}}^{B} \left(T_{\tilde{\theta}}(x') - T_{\tilde{\theta}}(\hat{x})\right) \mathrm{d}F(x'|\theta) \quad \textit{for all } \hat{x} > \bar{x}(T_{\tilde{\theta}}, \theta) . \tag{4b}$$

Proposition 3 characterizes the search policies of the agent given Assumption 2. Accordingly, (4a) and (4b) can be interpreted as *hidden action* constraints to the principal's problem.

4.2. *Optimality of bonus contracts*

We are now ready to characterize the optimal menu. The optimal menu of contracts $\{T_\theta\}_{\theta\in\Theta}$ is given by the solution to the following maximization problem:

$$\max_{\{T_\theta\}_{\theta\in\Theta}} \left\{ \int_{\theta\in\Theta} \int_{\bar{x}(T_\theta,\theta)}^{B} \left(\frac{x' - T_\theta(x')}{\bar{F}(\bar{x}(T_\theta,\theta)|\theta)} \right) \mathrm{d}F(x'|\theta)\,\mathrm{d}P(\theta) \right\}$$

subject to the constraints,

$$\frac{1}{\bar{F}(\bar{x}(T_\theta,\theta)|\theta)} \left[\int_{\bar{x}(T_\theta,\theta)}^{B} T_\theta(x')\,\mathrm{d}F(x'|\theta) - c \right] \geq 0 \qquad (IR_\theta)$$

$$\frac{1}{\bar{F}(\bar{x}(T_\theta,\theta)|\theta)} \left[\int_{\bar{x}(T_\theta,\theta)}^{B} T_\theta(x')\,\mathrm{d}F(x'|\theta) - c \right] \geq \frac{1}{\bar{F}(\bar{x}(T_{\tilde{\theta}},\theta)|\theta)} \left[\int_{\bar{x}(T_{\tilde{\theta}},\theta)}^{B} T_{\tilde{\theta}}(x')\,\mathrm{d}F(x'|\theta) - c \right] \qquad (IC_{\theta,\tilde{\theta}})$$

for all $(\theta,\tilde{\theta}) \in \Theta^2$, where $\bar{x}(T_{\tilde{\theta}},\theta)$ is characterized by

$$c \leq \int_{\hat{x}}^{B} \left(T_{\tilde{\theta}}(x') - T_{\tilde{\theta}}(\hat{x}) \right) \mathrm{d}F(x'|\theta) \quad \text{for all } \hat{x} \leq \bar{x}(T_{\tilde{\theta}},\theta) \qquad (SP^-_{\theta,\tilde{\theta}})$$

$$c > \int_{\hat{x}}^{B} \left(T_{\tilde{\theta}}(x') - T_{\tilde{\theta}}(\hat{x}) \right) \mathrm{d}F(x'|\theta) \quad \text{for all } \hat{x} > \bar{x}(T_{\tilde{\theta}},\theta)\,. \qquad (SP^+_{\theta,\tilde{\theta}})$$

The objective of the principal here is to maximize her expected payoff subject to three kind of constraints. First, constraints (IR_θ) require that it must be individually rational for the agent in state θ to accept contract T_θ, rather then choosing his outside option. Second, constraints $(IC_{\theta,\tilde{\theta}})$ require that it must be optimal for the agent in state θ to truthfully reveal the state to the principal by choosing T_θ from the menu of all contracts $\{T_{\tilde{\theta}}\}_{\tilde{\theta}\in\Theta}$. These constraints stem from the principal not knowing the state of the world. The third set of constraints is due to the additional uncertainty from not observing search by the agent, as analyzed in the previous subsection. Accordingly, $(SP^-_{\theta,\tilde{\theta}})$ and $(SP^+_{\theta,\tilde{\theta}})$ pin down the agent's search policies, $\bar{x}(T_{\tilde{\theta}},\theta)$, as given by Proposition 3.

The hidden action constraints also suggests that for an arbitrary contract T_θ, the search policy differs with the distribution $\tilde{\theta}$. In particular, off-equilibrium search policies of an agent with distribution $\tilde{\theta}$ who chooses contract T_θ are not bound to be the same to the ones this contract T_θ implements for an agent with distribution θ. A crucial question is therefore, what action should a contract which is designed for state θ implement for the agent who knows the world to be in state $\tilde{\theta}$? An approximate answer here is that the principal wants to design a contract T_θ which implements an action for all states $\tilde{\theta}$ other than θ that makes it as unattractive as possible for an agent in these states to choose contract T_θ. As we will formally see below, this goal can be achieved by a menu of bonus contracts.

To analyze the model, we impose the following assumptions on the distribution $F(x|\theta)$. Let $\bar{F} \equiv 1 - F$, and let $H \equiv \partial \bar{F}^{-1}/\partial x$. Then:

ASSUMPTION 4: $\partial H/\partial\theta \leq 0$, *and* $\partial^2 H/\partial\theta^2 \geq 0$.

ASSUMPTION 5: $\partial H/\partial x \leq 0$, *and* $\partial^2 H/\partial x \partial\theta \leq 0$.

The first part of Assumption 4 states that distributions can be ordered according to expression H. Since $1 - F(x|\theta)$ is decreasing in θ, a sufficient condition for H to be decreasing is the commonly used monotone likelihood ratio condition.[3] The intuition of this assumption is that at any point of search, when one continues search, one will do better—in the sense of first-order stochastic dominance—in state θ'' than in state θ', $\theta' < \theta''$. The role of the first part of Assumption 4 in this chapter is that it guarantees in a stochastical sense what is commonly referred to as the single crossing property.[4] The second part of Assumption 4 strengthens this, such that H is convexly increasing in the state of the world. Intuitively this requires that the benefit of being in a higher state of the world than θ is decreasing as the state θ becomes better.

Assumption 5 is of more technical nature, ensuring that the objective function of the principal is concave.

[3] Another condition which is less strict than the monotone likelihood ratio condition, and which is also sufficient to guarantee the first part of Assumption 4 is sometimes referred to as the monotone hazard ratio condition.

[4] More precisely, Assumption 4 implies that the agent's indifference curves between expected transfers and different stopping rules cross only once over different states. To see this, let $T^e_\theta \equiv \mathrm{E}\{T(x)\,|\,x \geq \bar{x}, \theta\}$ denote the expected transfers to the agent in state θ with a given contract T, and let $u_\theta(T^e_\theta, \bar{x}) \equiv T^e_\theta - c/\bar{F}(\bar{x}|\theta)$ denote the expected utility of the agent when pursuing stopping rule $\bar{x}$. Then the single crossing property holds, if for any $(\theta, \theta') \in \{\Theta^2|\theta > \theta'\}$,

$$-\frac{\partial u_\theta/\partial\bar{x}}{\partial u_\theta/\partial T^e_\theta} \leq -\frac{\partial u_{\theta'}/\partial\bar{x}}{\partial u_{\theta'}/\partial T^e_{\theta'}},$$

which is equivalent to $H(x|\theta) \leq H(x|\theta')$, as given by the first part of Assumption 4.

To keep the results clear, we also impose the following two assumptions which guarantee that the principal's maximization problem has an interior solution that can be characterized by first order conditions.

Assumption 6: $\frac{\mathrm{d}}{\mathrm{d}\theta}\left(\frac{p(\theta)}{1-P(\theta)}\right) \geq 0$.

Assumption 7: $\frac{\partial^2}{\partial\bar{x}\partial\theta}\ \mathrm{E}\{x|x \geq \bar{x}, \theta\} \geq 0$.

Assumption 6 is a standard assumption in adverse selection problems, which states that the conditional density to be in state θ', given that the state of the world is $\theta'' \geq \theta'$ is increasing in θ'. In other words, for higher states it becomes less likely, that the world is in a even better state. A sufficient condition for Assumption 6 to hold is that the likelihood $p(\theta)$ is decreasing in θ.

Assumption 7 states that the marginal benefit of search is increasing in θ. As mentioned before, together with Assumptions 4 and 6 this assumption guarantees that the optimal solution is an interior one. In many cases this assumption will, however, not be necessary at all. Whenever $p(\theta)/(1 - P(\theta))$ is increasing at a sufficiently high rate, or whenever H is sufficiently convex, we can completely drop this assumption. For details see the proof of Proposition 5.

In the following, we now argue that the principal optimally designs a menu of contracts which exclusively utilizes a particular simple form of contracts, namely bonus contracts as defined in the following.

Definition: Let τ be a nonrandom constant. Then a contract T is called a *bonus contract* when it is of the following form:

$$T(x) = \begin{cases} 0 & \text{if } x < \bar{x} \\ \tau & \text{if } x \geq \bar{x}\,. \end{cases}$$

Our argumentation establishes that a menu of such contracts implements any set of equilibrium search policies in a weakly optimal way and, therefore, any equilibrium can be implemented by a menu of such contracts. First, we investigate the implications of a menu that consists exclusively of bonus contracts. The following lemma asserts that in this case the optimal menu resembles many of the characteristics of a standard adverse selection problem. In particular, we have that (IR_θ), $(IC_{\theta,\tilde{\theta}})$, $(SP^-_{\theta,\tilde{\theta}})$ and $(SP^+_{\theta,\tilde{\theta}})$ in the principal's maximization problem can be replaced by (a), (b) and (c) below.

Lemma 2: *Suppose Assumption 1–4 hold. Let $\{T_\theta\}_{\theta\in\Theta}$ be a menu of bonus contracts, let $\bar{x}(\theta)$ be the search targets implemented by contract T_θ, and let $\tau(\theta)$ be* *the transfer paid to the agent when $x > \bar{x}(\theta)$. Then if $\{T_\theta\}_{\theta\in\Theta}$ is a solution to the principal's optimization program, then it is also a solution to a program where the principal maximizes her payoff subject to the following constraints:*

(a) $\bar{x}(\theta)$ *is nondecreasing in* θ,

(b) $U(\underline{\theta}) = 0$, *and*

(c) $\frac{\mathrm{d}U}{\mathrm{d}\theta} = -\frac{\partial}{\partial\theta}\left(\frac{c}{\bar{F}(\bar{x}(\theta)|\theta)}\right)$,

where $U(\theta) \equiv u(\bar{x}(\theta), \tau(\theta), \theta)$ *denotes the agent's indirect utility in state* θ, *i.e. given contract* $T_\theta = (\bar{x}(\theta), \tau(\theta))$.

Given these results, which hold for any menu of bonus contracts, we can now show that a menu of bonus contracts is at least as good as any other menu, allowing us to restrict attention to such a menu.

Proposition 4: *Suppose Assumptions 1–7 hold. Then to any equilibrium in the game defined in Section 2, there corresponds an associated menu of bonus contracts* $\{T_\theta\}_{\theta\in\Theta}$, *which implements the same actions and is (weakly) preferred by the principal.*

The main argument in the proof is that, for an arbitrary set of actions that are to be implemented, the utility of the agent in a menu of bonus contracts that implements these actions is a lower bound for any other contract, resulting in a weak optimality of bonus contracts. The intuition behind the result is that an agent in state θ'' will profit from choosing a contract designed for state $\theta' < \theta''$, because he faces a "better" distribution $F(x|\theta'')$ than the agent in state θ' (in the sense of Assumption 4). This stochastic advantage translates into an expected utility advantage via two channels. First, expected costs of pursuing a particular search policy are lower in state θ'' than in state θ'. Second, given that contracts are monotonically increasing, an increase in θ shifts probability mass to those solutions x where payments to the agent are weakly higher. Since for any $\bar{x}(T_{\theta'}, \theta')$ this effect is increasing in the slope of $T_{\theta'}$, a bonus contract minimizes the information rents paid in state θ''.

The usage of bonus contracts also answers the previously posed question, which actions a contract should implement after every possible deviation in the contracting stage. Specifically, as shown in the proof to Lemma 2, bonus contracts imply that $\bar{x}(T_\theta, \tilde{\theta}) = \bar{x}(T_\theta, \theta)$ in all relevant cases.[5]

[5] Here, for any given contract T_θ, "relevant" are all states $\tilde{\theta}$ in which it is necessary and sufficient to prevent them from choosing contract T_θ.

Corollary: *Let $\{T_\theta\}_{\theta\in\Theta}$ be a solution to the principal's optimization program. Then without loss of generality we may assume that $\bar{x}(T_\theta, \tilde{\theta}) = \bar{x}(T_\theta, \theta) \equiv \bar{x}(\theta)$ for all $(\theta, \tilde{\theta}) \in \Theta^2$.*

In other words, we may assume without loss of generality that any given contract T_θ implements the same stopping rule in all states of the world $\tilde{\theta}$. Accordingly, we henceforth suppress the second argument of $\bar{x}(T_\theta, \theta')$ and denote with $\bar{x}(\theta)$ the search policy which is implemented by contract T_θ in all states.

Furthermore, in combination with Lemma 2, Proposition 4 also pins down $\{T_\theta\}_{\theta\in\Theta}$ as a function of the search policies $\{\bar{x}(\theta)\}_{\theta\in\Theta}$ which are implemented.

Corollary: *The second-best optimum can be achieved by a menu of bonus contracts of the following form. For all $\theta \in \Theta$,*

$$T_\theta(x) = \begin{cases} 0 & \text{if } x < \bar{x}(\theta) \\ [\bar{F}(\bar{x}(\theta)|\theta)]^{-1}c + U(\theta) & \text{if } x \geq \bar{x}(\theta)\,, \end{cases}$$

where

$$U(\theta) = \int_{\underline{\theta}}^{\theta} -\frac{\partial}{\partial\tilde{\theta}}\left(\frac{c}{\bar{F}(\bar{x}(\tilde{\theta})|\tilde{\theta})}\right) \mathrm{d}\tilde{\theta}$$

is the minimal rent which prevents an agent in state θ to deviate to another contract than T_θ.

This last corollary gives us all pieces at hand in order to compute the second-best optimal search policies. Specifically, it reduces the combined problem of bringing the agent to reveal the state θ, while simultaneously inducing him to pursue search policies that the principal finds optimal given θ, to a standard adverse selection problem, which we can solve using standard techniques.

4.3. *Equilibrium search policies*

Using the results from the previous section, we can compute the second-best optimal search policies $\{\bar{x}_\theta\}_{\theta\in\Theta}$ using standard techniques.

Proposition 5: *Suppose Assumptions 1–7 hold. Let $\Phi \subseteq \Theta$ denote a set of states for which "directed" search is implemented. Then for all $\theta \in \Phi$, second best optimal search policies $\{\bar{x}(\theta)\}_{\theta\in\Phi}$ are characterized by,*

$$c + D_\theta(\bar{x}(\theta)) = \int_{\bar{x}(\theta)}^{B} (x' - \bar{x}(\theta))\, \mathrm{d}F(x'|\theta)\,, \tag{5}$$

where function $D_\theta : [0, B] \to \mathbb{R}_+$ is defined as

$$D_\theta(x) = -\frac{1 - P(\theta)}{p(\theta)} \frac{\partial H(x|\theta)/\partial\theta}{H(x|\theta)} c\,. \tag{6}$$

For all $\theta \notin \Phi$, search is "undirected", that is $\bar{x}(\theta) = 0$.

Proposition 5 establishes that delegated search is either determined by equation (5), or is otherwise "undirected" (i.e., "non-sequential"). In the latter case, the agent unconditionally adopts the first sampled solution. Because directed search is efficient in all states, this is clearly inefficient. To evaluate the distortion in the first case where search is directed, we need to compare equation (5) to the first-best stopping rule given by (1). For the first-best case we have seen that the optimal stopping rule equates (fundamental) costs of searching and benefits of further search. While benefits of further search in the delegated search model are the same, costs are now inflated by an additional agency cost term D_θ. Intuitively, D_θ reflects the costs of learning the state θ from the agent. More precisely, D_θ resembles the effect of a marginal increase in $\bar{x}(\theta)$ on the information rents that are to be paid to the agent in all states $\theta' \in \{\theta' \in \Theta : \theta' > \theta\}$. Since D_θ is strictly positive whenever $\bar{x} > 0$ and $\theta < B$, it follows that the stopping rule is generally inefficiently low.

Corollary: *Delegated search is almost surely inefficient. That is, $\bar{x}^{SB}(\theta) < \bar{x}^{FB}(\theta)$ for all $\theta \in [0, B)$.*

The left panel of Figure 1 illustrates the case where search is still sufficiently attractive to be conducted sequentially ($\theta \in \Phi$). Here $\bar{x}^{FB}$ denotes the first-best stopping rule that solves (1), and $\bar{x}^{SB}$ denotes the second-best stopping rule that solves (5). On the other hand, when the delegation costs increase the relevant costs above the benefits of directed search ($\theta \notin \Phi$), it will be optimal for the principal to implement a stopping rule $\bar{x} = 0$, so that search is undirected. That is, the principal prefers the agent to abstain from conditional sampling and instead adopt the first sampled solution—no matter how bad (or good) it turns out to be. This case is illustrated in the right panel of Figure 1.

To complete the description of equilibrium we have to characterize the set $\Phi \subseteq \Theta$, which tells us for which states search is directed in equilibrium. Search will be undirected when either the state of the world in θ is sufficiently bad compared to the better states θ', or when *a priori* it is sufficiently unlikely to be in state θ, such that it is not worth to distort more likely states from an *ex ante* perspective. In these cases it will be optimal to abstain from sequential search in

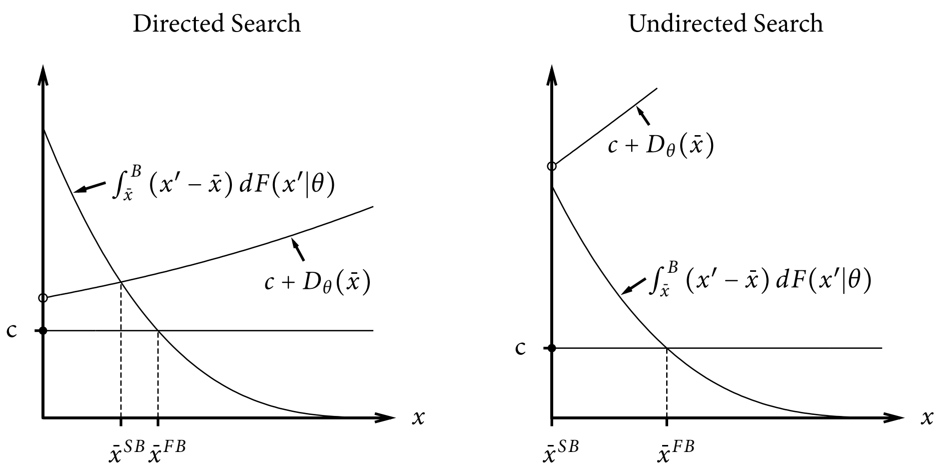

Figure 1. Second-best search policies for $\theta \in \Phi$ (left panel) and $\theta \notin \Phi$ (right panel).

state θ. Obviously, a sufficient condition for search to be directed, is that marginal costs for an infinitesimal $\bar{x}$ are sufficiently low, such that $c + D_\theta(\bar{x}) \leq E(x|\theta)$, or formally,

$$c + \hat{D}_\theta(0) \leq \int x \, \mathrm{d}F(x|\theta), \tag{7}$$

where $\hat{D}_\theta(0) = \lim_{\hat{x} \downarrow 0} D_\theta(\hat{x})$.[6]

However, since the information rent in all states $\theta' > \theta$ is increasing in $x(\theta)$, we may strengthen this results.

Proposition 6: *Suppose Assumptions 1–7 hold. Then $\theta \in \Phi$ if and only if θ fulfills condition* (7). *When $\theta \in \Phi$, search is distorted according to Proposition 5. Otherwise search is undirected, that is $\bar{x}(\theta) = 0$.*

In particular, since $\bar{x}$ is increasing in θ, Proposition 6 implies that Φ has the following "monotonicity" property.

Corollary: *Let $\theta'' > \theta'$. Then it holds that (i) if $\theta' \in \Phi$, then $\theta'' \in \Phi$; and (ii) if $\theta'' \notin \Phi$, then $\theta' \notin \Phi$.*

[6] Because delegation costs are discontinues at $\bar{x} = 0$, where $D_\theta(0) = 0$, we need to consider the limit of $D_\theta(\hat{x})$ as $\hat{x}$ approaches 0 from above. Details can be found in the proof of Proposition 6.

5. Concluding remarks

In this project we study how delegation affects the search for information when a principal is both uninformed about the prospects of search and unable to observe the searching of an agent. We find that searching is almost surely inefficient. Search is either stopped too early (compared with the efficient solution) or is completely undirected (leading the agent to unconditionally adopt the first solution he encounters). These results lend theoretical support to recent empirical evidence provided by Rutherford et al. (2005) and Levitt and Syverson (2008) that the search effort by real estate agents is inefficiently low.

The combination of the two informational asymmetries is thereby crucial for the inefficiency. If either adverse selection or moral hazard is shut down, then the efficient benchmark can be implemented as a rational equilibrium outcome. Each of the two asymmetries thus acts as an "catalyst" for the other one. Furthermore, the combination of moral hazard and adverse selection is also key to our finding that search is optimally delegated via the use of simple bonus contracts. More precisely, adverse selection precludes the principal from using contracts that are more sensible to realized payoffs, since such types of contracts would increase the costs of bringing the agent to reveal the prospects of search. On the other hand, moral hazard precludes the usage of even simpler "fixed wage" agreements, as then the agent would not be induced to provide any search effort.

As with any other framework, our analysis is based on a number of important modeling choices. A key choice is the kinds of informational asymmetries that we consider and, clearly, our results are linked with that choice. In particular, the role of adverse selection as a key source of uncertainty may seem unconventional. However, adverse selection is known to be relevant in many areas, and introspection suggests that it is likely to be relevant in the case of agents specializing on search services. Specifically, we feel that by an accumulation of expertise, specialist-agents are likely to have an informational advantage in assessing the prospects of search from an *ex ante* point of view. We therefore feel justified in studying adverse selection as a key source of uncertainty.

The other key choice in our modeling is the process of searching itself. Here our modeling strategy is aimed at resembling the "standard search model" (e.g., McCall, 1970) as closely as possible, deviating only in introducing asymmetric information. This standard search model has been successfully applied to many situations, motivating the use of this model as the point of departure for an analysis of delegated search. Possible applications include recruiting agencies searching for job candidates, real estate agents searching for buyers and properties, research centers searching for new product designs, and insurance brokers searching for clients.

At a more general level, our model also lends itself to think about agency relationships where the task of the agent tends to be original rather than routine; i.e., when agents are expected to "think" and search for "ideas" in order to come up with a solution for a given problem. Examples along these lines are advocates who have to find a good defense strategy and business consultancies, searching for promising business plans. With such a general interpretation in mind, an important question is whether the findings of our model and, specifically, the optimality of bonus contracts also extent to other specifications of the search process.

In particular, relaxing our assumption that the distribution of search payoffs is exogenously given seems to be a natural extension.[7] Exploring which contractual arrangements are optimal when the agent is able to either *ex ante* or *ex tempore* control the quality of the distribution is thereby a promising direction for future research.

A. Mathematical appendix

A.1. Proof of Lemma 1

Let $V(x|\theta)$ be the indirect utility function of the agent in state θ who has sampled solution x. Then the indirect utility function satisfies the Bellman equation

$$V(x|\theta) = \max\left\{T_{\tilde{\theta}}(x),\ -c + \int V(x'|\theta)\,\mathrm{d}F(x'|\theta)\right\}, \tag{8}$$

where the agent adopts solution x whenever the associated wage $T_{\tilde{\theta}}(x)$ exceeds the expected utility from continuing search. Since this expected utility is independent from x, we have that the agent adopts solution x whenever $T_{\tilde{\theta}}(x) > \bar{T}_{\tilde{\theta}}(\theta)$, where $\bar{T}_{\tilde{\theta}}(\theta) = -c + \int V(x'|\theta)\,\mathrm{d}F(x'|\theta)$. Let $\psi_{\tilde{\theta}} : X \times \mathbb{R} \to \{0, 1\}$ be an indicator, such that

$$\psi_{\tilde{\theta}}(x, \bar{T}_{\tilde{\theta}}(\theta)) = \begin{cases} 0 & \text{if } T_{\tilde{\theta}}(x) \leq \bar{T}_{\tilde{\theta}}(\theta)\text{ , and} \\ 1 & \text{if } T_{\tilde{\theta}}(x) > \bar{T}_{\tilde{\theta}}(\theta)\,. \end{cases} \tag{9}$$

[7] See Lewis and Ottaviani (2008) and Lewis (2012) for some work in that direction.

Then, using (8), we can rewrite $\bar{T}_{\tilde{\theta}}(\theta)$ as follows

$$\bar{T}_{\tilde{\theta}}(\theta) = -c + \int \Big(\big(1 - \psi_{\tilde{\theta}}(x', \bar{T}_{\tilde{\theta}}(\theta))\big)\, \bar{T}_{\tilde{\theta}}(\theta) + \psi_{\tilde{\theta}}(x', \bar{T}_{\tilde{\theta}}(\theta))\, T_{\tilde{\theta}}(x') \Big)\, \mathrm{d}F(x'|\theta), \quad (10)$$

or

$$\bar{T}_{\tilde{\theta}}(\theta) \Big(1 - \int \big(1 - \psi_{\tilde{\theta}}(x', \bar{T}_{\tilde{\theta}}(\theta))\big)\, \mathrm{d}F(x'|\theta)\Big) = -c + \int \psi_{\tilde{\theta}}(x', \bar{T}_{\tilde{\theta}}(\theta))\, T_{\tilde{\theta}}(x')\, \mathrm{d}F(x'|\theta), \quad (11)$$

or

$$c = \int \psi_{\tilde{\theta}}(x', \bar{T}_{\tilde{\theta}}(\theta))\, T_{\tilde{\theta}}(x')\, \mathrm{d}F(x'|\theta) - \bar{T}_{\tilde{\theta}}(\theta) \int \psi_{\tilde{\theta}}(x', \bar{T}_{\tilde{\theta}}(\theta))\, \mathrm{d}F(x'|\theta) \quad (12)$$

$$= \int \psi_{\tilde{\theta}}(x', \bar{T}_{\tilde{\theta}}(\theta)) \Big(T_{\tilde{\theta}}(x') - \bar{T}_{\tilde{\theta}}(\theta)\Big)\, \mathrm{d}F(x'|\theta). \quad (13)$$

Since an increase in $\bar{T}_{\tilde{\theta}}(\theta)$ weakly decreases $\psi_{\tilde{\theta}}(x', \bar{T}_{\tilde{\theta}}(\theta))$, the RHS of (13) is strictly decreasing in $\bar{T}_{\tilde{\theta}}(\theta)$. Thus whenever there exists a solution to (13), it is unique. Moreover, since X is a compact interval, $T^{MAX} \equiv \max_x T_{\tilde{\theta}}(x)$ exists, and therefore for all $\bar{T}_{\tilde{\theta}}(\theta) \geq T^{MAX}$, the RHS of (13) is equal to 0. Thus (13) uniquely characterizes $\bar{T}_{\tilde{\theta}}(\theta)$ whenever $\int T(x)\, \mathrm{d}F(x|\theta) \geq c$. Otherwise, marginal costs of searching do always exceed the marginal benefits, and therefore, the agent trivially abstains from search. □

A.2. Proof of Proposition 3

By construction of $\bar{x}(T_{\tilde{\theta}}, \theta)$, for some $x \in [0, B]$, $T_{\tilde{\theta}}(x) \geq \bar{T}_{\tilde{\theta}}(\theta)$ (see the proof to Lemma 1). Further, whenever search is directed, for some $x \in [0, B]$, $T_{\tilde{\theta}}(x) \geq \bar{T}_{\tilde{\theta}}(\theta)$. Thus for $T_{\tilde{\theta}}$ strictly increasing and continuous, $\bar{x}(T_{\tilde{\theta}}, \theta) \equiv T_{\tilde{\theta}}^{-1}(\bar{T}_{\tilde{\theta}}(\theta)) \in [0, B]$ obviously exists, and is given by (4a) and (4b). To verify the remaining cases, suppose that $\bar{T}_{\tilde{\theta}}(\theta)$ is not attained by $T_{\tilde{\theta}}(x)$ on $[0, B]$. Then from (4a) and (4b), $\bar{x}(T_{\tilde{\theta}}, \theta)$ is assigned to the point of discontinuity where $\lim_{x \uparrow \bar{x}(T_{\tilde{\theta}}, \theta)} T_{\tilde{\theta}}(x) < \bar{T}_{\tilde{\theta}}(\theta)$ and $\lim_{x \downarrow \bar{x}(T_{\tilde{\theta}}, \theta)} T_{\tilde{\theta}}(x) > \bar{T}_{\tilde{\theta}}(\theta)$. So search given by $\bar{x}(T_{\tilde{\theta}}, \theta)$ is identical to search given by $\bar{T}_{\tilde{\theta}}(\theta)$. Finally, suppose that $\bar{T}_{\tilde{\theta}}(\theta)$ is attained on an interval $[\underline{x}, \bar{x}]$. Then from Lemma 1, the agent continues search for all $x \leq \bar{x}$ and stops search for $x > \bar{x}$. Thus $\bar{x}(T_{\tilde{\theta}}, \theta) = \bar{x}$, identical to the rule given by (4a) and (4b). □

A.3. Proof of Lemma 2

To proof this lemma, we first characterize the set of stopping rules $\bar{x}(\theta)$ that are implementable via bonus contracts. Though formally the choice of the stopping rule is taken by the agent, the resulting problem shares the basic logic of a standard mechanism design problem, and can be solved using similar techniques. Here we proceed along the lines of Fudenberg and Tirole (1991, pp. 257–68).

From $(SP^{-}_{\theta,\tilde{\theta}})$ and $(SP^{+}_{\theta,\tilde{\theta}})$ it follows that an agent with bonus contract $T_{\tilde{\theta}} = (\bar{x}(\tilde{\theta}), \tau(\tilde{\theta}))$ chooses a stopping rule

$$\bar{x}(T_{\tilde{\theta}}, \theta) = \begin{cases} 0 & \text{if } \tau(\tilde{\theta}) < [\bar{F}(\bar{x}(\tilde{\theta})|\theta)]^{-1}\, c, \text{ and} \\ \bar{x}(\tilde{\theta}) & \text{if } \tau(\tilde{\theta}) \geq [\bar{F}(\bar{x}(\tilde{\theta})|\theta)]^{-1}\, c. \end{cases} \tag{14}$$

Accordingly, let $u(\theta, \tilde{\theta}) \equiv \max\{0, -[\bar{F}(\bar{x}(\tilde{\theta})|\theta)]^{-1}c + \tau(\tilde{\theta})\}$ denote the agent's indirect utility in state θ when he chooses bonus contract $T_{\tilde{\theta}} = (\bar{x}(\tilde{\theta}), \tau(\tilde{\theta}))$. Note that (IR_{θ}) implies that $\tau(\tilde{\theta}) \geq [\bar{F}(\bar{x}(\tilde{\theta})|\tilde{\theta})]^{-1}\, c$, and thus we can ignore $(SP^{-}_{\theta,\tilde{\theta}})$ and $(SP^{+}_{\theta,\tilde{\theta}})$ if $(IC_{\theta,\tilde{\theta}})$ holds.

Moreover, Assumption 4 implies that $u(\theta, \tilde{\theta}) \geq u(\tilde{\theta}, \tilde{\theta})$ for all $\theta \geq \tilde{\theta}$. Thus, $u(\theta, \tilde{\theta}) = -[\bar{F}(\bar{x}(\tilde{\theta})|\theta)]^{-1}c + \tau(\tilde{\theta})$ for all $\theta \geq \tilde{\theta}$. Therefore, for the agent in state θ to not locally deviate, it must be that the first order condition

$$-H(\bar{x}(\tilde{\theta})|\theta)\frac{\mathrm{d}\bar{x}(\tilde{\theta})}{\mathrm{d}\tilde{\theta}}\, c + \frac{\mathrm{d}\tau(\tilde{\theta})}{\mathrm{d}\tilde{\theta}} = 0 \quad \text{for } \tilde{\theta} = \theta, \tag{15}$$

and the second order condition

$$\frac{\mathrm{d}}{\mathrm{d}\tilde{\theta}}\left(-H(\bar{x}(\tilde{\theta})|\theta)\frac{\mathrm{d}\bar{x}(\tilde{\theta})}{\mathrm{d}\tilde{\theta}}\, c + \frac{\mathrm{d}\tau(\tilde{\theta})}{\mathrm{d}\tilde{\theta}}\right) \leq 0 \quad \text{for } \tilde{\theta} = \theta \tag{16}$$

hold locally at $\tilde{\theta} = \theta$. Moreover, since (15) must hold for all $\theta \in \Theta$, it is an identity in θ, and thus

$$\frac{\mathrm{d}}{\mathrm{d}\tilde{\theta}}\left(-H(\bar{x}(\tilde{\theta})|\theta)\frac{\mathrm{d}\bar{x}(\tilde{\theta})}{\mathrm{d}\tilde{\theta}}\, c + \frac{\mathrm{d}\tau(\tilde{\theta})}{\mathrm{d}\tilde{\theta}}\right) - \frac{\mathrm{d}}{\mathrm{d}\theta}\left(H(\bar{x}(\tilde{\theta})|\theta)\frac{\mathrm{d}\bar{x}(\tilde{\theta})}{\mathrm{d}\tilde{\theta}}\, c\right) = 0 \tag{17}$$

for $\tilde{\theta} = \theta$. Substituting (16) in (17) together with Assumption 4 thus yields $\mathrm{d}\bar{x}/\mathrm{d}\theta \geq 0$. This establishes that any solution to the principal's optimization program implies condition (a).

Next we argue that (15) is also sufficient to prevent the agent from deviating globally. First note that in the case where $\bar{x}(T_{\tilde{\theta}}, \theta) = 0$, $u(\theta, \tilde{\theta}) = 0 \leq u(\theta, \theta)$ and thus we can restrict attention to the case where the agent chooses $\bar{x}(T_{\tilde{\theta}}, \theta) = \bar{x}(\tilde{\theta})$ under contract $T_{\tilde{\theta}}$. Suppose to the contrary that the incentive constraint is violated in at least one state, i.e. $u(\theta, \tilde{\theta}) - u(\theta, \theta) > 0$ for some $(\theta, \tilde{\theta}) \in \Theta^2$, or by the fundamental theorem of calculus,

$$\int_{\theta}^{\tilde{\theta}} \left(-H(\bar{x}(\theta')|\theta) \frac{\mathrm{d}\bar{x}(\theta')}{\mathrm{d}\theta'} c + \frac{\mathrm{d}\tau(\theta')}{\mathrm{d}\theta'} \right) \mathrm{d}\theta' > 0. \tag{18}$$

Suppose $\tilde{\theta} > \theta$. Then, Assumption 4 implies that $H(\bar{x}(\tilde{\theta})|\tilde{\theta}) \leq H(\bar{x}(\tilde{\theta})|\theta)$, and therefore (18) implies

$$\int_{\theta}^{\tilde{\theta}} \left(-H(\bar{x}(\theta')|\theta') \frac{\mathrm{d}\bar{x}(\theta')}{\mathrm{d}\theta'} c + \frac{\mathrm{d}\tau(\theta')}{\mathrm{d}\theta'} \right) \mathrm{d}\theta' > 0 \tag{19}$$

since $\mathrm{d}\bar{x}/\mathrm{d}\theta \geq 0$. However, equation (15) implies that the integrand in (19) is equal to 0 for all θ', contradicting that for any $\tilde{\theta} > \theta$, contract $T_{\tilde{\theta}}$ is preferred over T_{θ}. The same logic establishes a contradiction for the case where $\tilde{\theta} < \theta$.

Now, let $U(\theta) \equiv u(\theta, \theta)$. Then

$$\frac{\mathrm{d}U}{\mathrm{d}\theta} = -H(\bar{x}(\theta)|\theta) \frac{\mathrm{d}\bar{x}(\theta)}{\mathrm{d}\theta} c + \frac{\mathrm{d}\tau(\theta)}{\mathrm{d}\theta} - \frac{\partial}{\partial\theta} \frac{c}{\bar{F}(\bar{x}(\theta)|\theta)} . \tag{20}$$

Thus (c) holds if and only if (15) holds. Therefore, (a) and (c) are both sufficient and necessary for $(IC_{\theta,\tilde{\theta}})$ to hold. Moreover, as shown above, $(IC_{\theta,\tilde{\theta}})$ implies $(SP^-_{\theta,\tilde{\theta}})$ and $(SP^+_{\theta,\tilde{\theta}})$. Thus we are left to show that (a), (b) and (c) are sufficient and (b) is necessary for (IR_{θ}) to hold. Consider sufficiency first. From (c) we have that $\mathrm{d}U/\mathrm{d}\theta$ is increasing in θ (by Assumption 4 implies $\partial\bar{F}(\bar{x}(\theta)|\theta) \geq 0$). Thus $U(\underline{\theta}) = 0$ implies (IR_{θ}) for all $\theta \in \Theta$. That $U(\underline{\theta}) = 0$ is also necessary for a solution to the principal's optimization program follows trivially from the above analysis, since shifting $U(\underline{\theta})$ doesn't affect any other constraints. □

A.4. Proof of Proposition 4

Consider an arbitrary menu of contracts $\{T_{\theta}\}_{\theta\in\Theta}$, and let $u(T, \bar{x}, \theta)$ be the utility of the agent in state θ when he chooses contract T and search policy $\bar{x}$. Then $u(T_{\tilde{\theta}}, \bar{x}(T_{\tilde{\theta}}, \theta), \theta) \geq u(T_{\tilde{\theta}}, \bar{x}(T_{\tilde{\theta}}, \tilde{\theta}), \theta)$, and therefore a necessary condition for $(IC_{\theta,\tilde{\theta}})$ to hold is that

$$U(\theta) \equiv u(T_{\theta}, \bar{x}(T_{\theta}, \theta), \theta) \geq u(T_{\tilde{\theta}}, \bar{x}(T_{\tilde{\theta}}, \tilde{\theta}), \theta) \quad \text{for all } \tilde{\theta} \in \Theta .$$

Hence $\tilde{\theta} = \theta$ maximizes the RHS of the inequality, with $U(\theta)$ also being the value function of $\max_{\tilde{\theta}} u(T_{\tilde{\theta}}, \bar{x}(T_{\tilde{\theta}}, \tilde{\theta}), \theta)$. The envelope theorem implies

$$\frac{\mathrm{d}U}{\mathrm{d}\theta} = \left\{ \frac{\partial}{\partial\theta}\left(\frac{\int_{\bar{x}(T_{\tilde{\theta}},\tilde{\theta})}^{B} T(x')\,\mathrm{d}F(x'|\theta)}{\bar{F}(\bar{x}(T_{\tilde{\theta}},\tilde{\theta})|\theta)}\right) - \frac{\partial}{\partial\theta}\left(\frac{c}{\bar{F}(\bar{x}(T_{\tilde{\theta}},\tilde{\theta})|\theta)}\right)\right\}\Bigg|_{\tilde{\theta}=\theta}, \tag{21}$$

and since by Assumptions 3 and 4 the first term in (21) is positive, we have that

$$\frac{\mathrm{d}U}{\mathrm{d}\theta} \geq -\frac{\partial}{\partial\theta}\left(\frac{c}{\bar{F}(\bar{x}(T_\theta,\theta)|\theta)}\right). \tag{22}$$

Moreover, (IR_θ) implies $U(\underline{\theta}) \geq 0$. Thus for any $\bar{x}$ that is nondecreasing in θ, the agent's utility under a menu of bonus contracts as given by Lemma 2(b) and (c) constitutes a lower bound on the agent's utility under any menu of contracts that implements $\bar{x}$. As long as the restriction on $\bar{x}$ is not binding for bonus contracts, a menu of bonus contracts is therefore optimal. That the restriction that $\bar{x}$ is increasing is indeed not binding is established in the proof of Proposition 5. □

A.5. Proof of Proposition 5

Using $\{T_\theta\}_{\theta\in\Theta}$ as given by the second corollary to Proposition 4, the principal's objective function is

$$\int_{\underline{\theta}}^{\bar{\theta}}\left(\int_{\bar{x}(\theta)}^{B}\frac{x'}{\bar{F}(\bar{x}(\theta)|\theta)}\,dF(x'|\theta) - \frac{c}{\bar{F}(\bar{x}(\theta)|\theta)} + \int_{\underline{\theta}}^{\theta}\frac{\partial}{\partial\tilde{\theta}}\left(\frac{c}{\bar{F}(\bar{x}(\tilde{\theta})|\tilde{\theta})}\right)\mathrm{d}\tilde{\theta}\right)\mathrm{d}P(\theta), \tag{23}$$

or, after an integration by parts,

$$\int_{\underline{\theta}}^{\bar{\theta}}\left(\int_{\bar{x}(\theta)}^{B}\frac{x'}{\bar{F}(\bar{x}(\theta)|\theta)}\,dF(x'|\theta) - \frac{c}{\bar{F}(\bar{x}(\theta)|\theta)} + \frac{1-P(\theta)}{p(\theta)}\frac{\partial}{\partial\tilde{\theta}}\left(\frac{c}{\bar{F}(\bar{x}(\tilde{\theta})|\tilde{\theta})}\right)\right)\mathrm{d}P(\theta). \tag{24}$$

By Lemma 2 the only relevant constraint which we have to take care of is that $\bar{x}(\theta)$ is nondecreasing in θ. Ignoring this constraint for the moment, the maximizer of (24) is given by

$$c + D_\theta(\bar{x}(\theta)) = \int_{\bar{x}(\theta)}^{B} (x' - \bar{x}(\theta))\, \mathrm{d}F(x'|\theta), \tag{25}$$

where function $D_\theta : [0, B] \to \mathbb{R}_+$ is defined by,

$$D_\theta(x) = -\frac{1 - P(\theta)}{p(\theta)} \frac{\partial H(x|\theta)}{\partial \theta} \frac{c}{H(x|\theta)}\,. \tag{26}$$

(From Assumption 5, (24) is concave at any $\bar{x}(\theta)$ that satisfies (25), so (24) is globally quasi-concave, and the second-order condition is satisfied.)

We still have to show two things. First, that $\bar{x}(\theta)$ as given by (25) is indeed nondecreasing. Second, we have to extend our analysis to the possibility that for some $\theta \in \Theta$, a stopping rule of $\bar{x} = 0$ is preferred over an interior solution.

From Assumption 4, the RHS of (25) is increasing in θ, so $\bar{x}(\theta)$ will trivially be nondecreasing whenever $D_\theta(\bar{x}(\theta))$ is nonincreasing in θ. This will be the case whenever $p/(1 - P)$ is increasing at a sufficiently high rate, and/or H is sufficiently convex in θ. Otherwise, $\bar{x}(\theta)$ will still be increasing if the RHS of (25) is increasing at a sufficiently high rate. A sufficient condition for this to be the case is that

$$H(\bar{x}(\hat{\theta})|\theta) \int_{\bar{x}(\hat{\theta})}^{B} (x' - \bar{x}(\hat{\theta}))\, \mathrm{d}F(x'|\theta) \tag{27}$$

is increasing in θ at $\hat{\theta} = \theta$. Assumption 7 ensures that this is always the case. Hence, any interior solution to the principal's program is characterized by equations (25) and (26).

As for $\bar{x}(\theta) = B$ benefits of search (the RHS of (25)) are equal to 0, corner solutions may at most be given by $\bar{x}(\theta) = 0$. Let $\Phi \subseteq \Theta$ denote the set of states for which the principal implements an interior solution. Our previous reasoning implies that if $\theta' \in \Phi$, then for all $\theta'' > \theta'$, $\theta'' \in \Phi$. Thus, for all $\theta \in \Phi$, D_θ does not depend on Φ, and therefore, for all $\theta \in \Phi$, implemented search policies are given by the interior solution characterized above. That $\bar{x}(\theta) = 0$ is optimal for all $\theta \notin \Phi$ follows from $\partial H/\partial\theta = \partial^2 \bar{F}^{-1}/\partial\theta\partial x = 0$ at $x = 0$ which implies $D_\theta(0) = 0$. Per assumption we have that $\mathrm{E}(x|\theta) \geq c$ for all $\theta \in \Theta$, and therefore undirected search is preferred over terminating search whenever $\bar{x}(\theta) \ngtr 0$. □

A.6. Proof of Proposition 6

From Proposition 5 marginal costs of searching are given by $c + D_\theta(\bar{x})$. Differentiating with respect to $\bar{x}$ yields

$$-\frac{1-P}{P}\left(\frac{\partial^2 H}{\partial x \partial \theta}\frac{1}{H} - \frac{\partial H}{\partial \theta}\frac{\partial H}{\partial x}\frac{1}{H^2}\right)c \geq 0\,, \tag{28}$$

by Assumptions 4 and 5. Moreover, marginal benefits are trivially decreasing in $\bar{x}$. Thus a necessary and sufficient condition for $\bar{x}(\theta)$ to be optimally strictly greater than 0, is that for $\hat{x} \downarrow 0$ directed search is beneficial, or formally,

$$\lim_{\hat{x}\downarrow 0}\left\{\int_{\hat{x}}^{B}(x' - \hat{x})\,\mathrm{d}F(x'|\theta) - c - D_\theta(\hat{x})\right\} \geq 0\,. \tag{29}$$

The only term which may not be continuous in the limit is D_θ, so we can write (29) as

$$c + \hat{D}_\theta(0) \leq \int x\,\mathrm{d}F(x|\theta), \tag{30}$$

where $\hat{D}_\theta(0) = \lim_{\hat{x}\downarrow 0} D_\theta(\hat{x})$. □

References

Akerlof, G. A. (1970), "The market for 'lemons': Quality uncertainty and the market mechanism," The quarterly journal of economics, 84:488–500.

Demougin, D. and C. Fluet (1998), "Mechanism sufficient statistic in the risk-neutral agency problem," Journal of Institutional and Theoretical Economics (JITE)/Zeitschrift für die gesamte Staatswissenschaft, 622–639.

Fudenberg, D. and J. Tirole (1991), Game theory, MIT Press.

Herweg, F., D. Muller, and P. Weinschenk (2010), "Binary payment schemes: Moral hazard and loss aversion," The American Economic Review, 100:2451–2477.

Holmstrom, B. (1979), "Moral hazard and observability," The Bell Journal of Economics, 74–91.

Holmstrom, B. and P. Milgrom (1987), "Aggregation and linearity in the provision of intertemporal incentives," Econometrica: Journal of the Econometric Society, 55:303–328.

Innes, R. (1990), “Limited liability and incentive contracting with ex-ante action choices,” Journal of Economic Theory, 52:45–67.

Jewitt, I., O. Kadan, and J. Swinkels (2008), “Moral hazard with bounded payments,” Journal of Economic Theory, 143:59–82.

Kim, S. (1997), “Limited liability and bonus contracts,” Journal of Economics & Management Strategy, 6:899–913.

Levitt, S. D. and C. Syverson (2008), “Market Distortions when Agents Are Better Informed: The Value of Information in Real Estate Transactions,” The Review of Economics and Statistics, 90:599–611.

Lewis, T. R. (2012), “A theory of delegated search for the best alternative,” The RAND Journal of Economics, 43:391–416.

Lewis, T. R. and M. Ottaviani (2008), “Search Agency,” Mimeo.

McCall, J. J. (1970), “Economics of Information and Job Search,” The Quarterly Journal of Economics, 84:113–126.

Mortensen, D. (1986), “Job search and labor market analysis,” Handbook of labor economics, 2:849–919.

Moynahan, J. (1980), Designing an effective sales compensation program, New York: Amacom.

Oyer, P. (2000), “A theory of sales quotas with limited liability and rent sharing,” Journal of labor Economics, 18:405–426.

Park, E. (1995), “Incentive contracting under limited liability,” Journal of Economics & Management Strategy, 4:477–490.

Rogerson, R., R. Shimer, and R. Wright (2005), “Search-theoretic models of the labor market: a survey,” Journal of Economic Literature, 43:959–988.

Rutherford, R. C., T. M. Springer, and A. Yavas (2005), “Conflicts Between Principals and Agents: Evidence from Residential Brokerage,” Journal of Financial Economics, 76:627–665.

Townsend, R. (1979), “Optimal contracts and competitive markets with costly state verification,” Journal of Economic theory, 21:265–293.

Chapter 4

INTERTEMPORAL PRICE DIFFERENTIATION IN THE AIRLINE INDUSTRY*

Contents

* This chapter is based on joint work with Caspar Siegert. We are extremely grateful to Anton Vasilev for invaluable help and technical support throughout the process of collecting our dataset. We also would like to thank Ken Boyer, Meghan Busse, Sarit Markovich, Niko Matouschek, Klaus M. Schmidt, Mauricio Varela, and seminar participants at the University of Munich, Northwestern, and the 2011 IIOC conference for useful comments and discussions.

Abstract

This chapter of my dissertation develops a theory of dynamic pricing in industries with short-term capacity constraints. When supply is fixed in the short term, firms price more aggressively in earlier periods in order to relax competition in the future. This leads to intertemporal price differentiation in competitive environments, even if firms are unable to commit to future prices. Applied to the airline industry, this provides a novel explanation for the rise in ticket prices close to the scheduled departure date. Importantly, when the number of competitors increases, each firm benefits less from altering the competitive environment, causing prices to become intertemporally less dispersed. Using a hand-collected data set of 1.4 million airline ticket prices on 92 intra-European routes, we successfully test our theoretical predictions.

Keywords

Airline industry, capacity constraints, dynamic pricing, price dispersion.

JEL Classification: D43, D92, L11, L93.

1. Introduction

The fact that prices for airline tickets rise as the scheduled departure date approaches belongs to the most well-known regularities in airline pricing. While a number of papers have shown that intertemporal price differentiation may be an equilibrium even in competitive settings, such theories rely on the airlines' ability to commit to future price schedules. Recent technological advances like the arrival of online booking and dynamic pricing have, however, arguably undermined airlines' capacity to commit to price schedules and have thus led to a need to reconsider such explanations. This chapter of my dissertation proposes a simple model that shows that intertemporal price differentiation is an equilibrium in oligopolistic settings, even if airlines can not commit to future prices. Using a novel dataset of the European airline industry, we find that the empirical patterns in airline pricing lend support to our model, while they do not confirm previous theories of airline pricing.

Our model is based on a simple observation. Once an airline has made its capacity choices, it is typically unable to change the number of seats available on a particular flight, causing airlines to be capacity constraint in the short-term. As a result, whenever all flights on a given route are at the verge of being sold out, competition for customers will no longer ensure that airlines price close to their marginal costs, so that airlines effectively have monopoly power over consumers. This possibility of ending up with monopoly power makes it very attractive for airlines to hold back capacity until shortly before departure. On the other hand, the capacity an airline has available on the short term market will also affect the chance that it gains monopoly power in the first place. In order to increase this chance and relax competition in the future, airlines are thus inclined to price aggressively in earlier periods, causing prices to be intertemporally dispersed.

Importantly, the propensity to price aggressively depends largely on the competitive environment. While the cost of foregoing high last-minute returns in favor of low advanced-booking prices is borne by each airline individually, the benefit of reducing aggregate capacity on the late market is enjoyed by all airlines serving a particular route. The lack of last-minute capacity can thus be seen as a public good. Accordingly, as the number of airlines serving a route increases, airlines have less incentives to sell capacity in early periods, and prices on the early market increase relative to last-minute prices. This gives rise to the key prediction of our model that differentiates it from alternative theories. While the standard explanation for intertemporal increases in prices developed by the previous literature (see below) predicts that the intertemporal slope of prices is increasing in the number of competitors, our model predicts that it is decreasing.

Using an extensive hand-collected sample of over 1.4 million ticket prices on 92 intra-European routes, we test this theoretical prediction. Consistent with our model, the amount of intertemporal price differentiation is decreasing in the number of competing airlines on a route. The effects we observe are economically significant and have important policy implications. While on routes served by a single airline prices increase by an average of 1.31 percent with every day that a customer waits to book, this slope is reduced to 1.19 percent on duopoly routes and amounts to only 0.68 percent on routes with six competing airlines. Conversely, our data rejects traditional theories of peak-load pricing such as Dana (1999a) for the markets we study. While the observed patterns may also be caused by price discrimination, based on rudimentary tests we are unable to find direct evidence that airlines price discriminate against late bookers. In particular, we do not find any effect of some measure of customer heterogeneity on a route on pricing behavior. Finally, our results are robust to a variety of different specifications of competition and subsamples.

Related literature In a seminal paper Borenstein and Rose (1994) have empirically shown that price dispersion for airline tickets is substantial and robust to competition, a finding that is inconsistent with the view that price dispersion is the result of a price-discriminating monopolist. Even though the empirical literature finds a variety of dimensions along which price differentiation takes place, Advanced Purchase discounts and high prices for late bookings have moved to the center stage of the theoretical debate. Some of the most prominent explanations of Advanced Purchase discounts in both, monopoly and competitive settings include Gale and Holmes (1992; 1993), Dana (1998; 1999a; 1999b; 2001), and Nocke et al. (2010). In most of these models, price differentiation reflects differences in the cost of capacity. The last seats on a plane sell with a low probability and must hence sell at a high price if the airline is to recoup its marginal investments into capacity. In a competitive setting, this requires firms to commit to price schedules before capacity investments are made, since those investments are sunk afterward. This commitment power however seems to have been undermined in recent years by the widespread use of online booking and the possibility for airlines to change prices at will. The theoretical contribution of this study is to provide a novel explanation for intertemporal price differentation that does not require commitment power.

On the empirical front, this study also contributes to an extensive literature that analyzes the factors determining price dispersion in the airline market (see for instance, Morrison and Winston, 1990, Borenstein and Rose, 1994, Stavins, 2001 or Puller et al., 2008). Most of these papers do, however, not differentiate between different dimensions along which prices differ; e.g., Saturday-night stay-over requirements, time of the day or Advanced Purchase discounts. In particular, due to limitations in the available data, the intertemporal dimension of price differentiation has so far received little attention. In this study, we overcome these limitations by exploiting a novel three-dimensional dataset. Specifically, we construct a panel of markets, where each market consists of all direct flights offered on a particular day and route. For each of these markets, we then record a time series of prices that ranges between 10 weeks and 1 day prior to the date of departure.

Outline The remainder of the chapter proceeds as follows. In the next section, we introduce a simple model of the airline industry, which is aimed at analytical tractability, and yet gives rise to strategic pricing effects due to short-run capacity constraints. In Section 3, we characterize the equilibrium. The key predictions of the model are derived in Section 4, where we also contrast them with the predictions of alternative models of price differentiation. Section 5 introduces the data set, which we use to test our predictions. The baseline empirical analysis is conducted in Section 6; and further robustness specifications are studied in Section 7. Section 8 concludes.

2. A simple model of airline pricing

We consider a symmetric industry, in which airlines compete in prices. Each airline operates on two markets, an early market in period $t = 1$ and a last minute market in period $t = 2$. There are $N \geq 2$ airlines active on each market.

In period 1, airlines have an overall capacity of $\bar{X}_1 = 2$, which is equally distributed amongst airlines such that for all airlines $i \in \{1, \dots, N\}$, first period capacity equals $\bar{x}_1^i = \bar{x}_1 = \bar{X}_1/N$.[1] Given that an airline sells Q_1^i seats on the early market, its capacity on the last minute market is then given by $\bar{x}_2^i = \bar{x}_1^i - Q_1^i$. As long as airlines sell less than their available capacity, marginal costs of selling a ticket are c. In case an airline sells more than its capacity, however, it has to

[1] While the assumption that the installed capacity of an individual airline is smaller on more competitive routes than on less competitive ones appears to be natural, one may also think about an alternative, but formally equivalent setting, in which the size of the market is proportional to N and each airline has a fixed capacity $\bar{x}_1$.

reschedule some passengers and faces an increased marginal cost of $\bar{c} > c$. Below, we will assume $\bar{c}$ to be prohibitively high, so that no flight will be overbooked in equilibrium. That is, rather than assuming capacity constraints to be "strict", we allow for the theoretical possibility of overbooking, but impose that airlines never choose to do so. This assumption goes back to Maggi (1996) and ensures that there always exists a pure strategy equilibrium in prices, which wouldn't be the case if capacity constraints were strict (see Kreps and Scheinkman, 1983).

In both periods, airlines offer differentiated products and are located in $\mathbb{R}^K$ where $K \geq N - 1$. Hence, flights offered by different airlines differ in up to K dimensions. For simplicity, airlines are positioned at the vertices of a regular $N - 1$ simplex, ensuring that no matter which two flights one compares, they always display the same degree of heterogeneity. Consumers are uniformly distributed on the edges of this simplex and receive positive utility if and only if they buy a flight from an airline located on either end of the relevant edge. Otherwise, they receive zero utility. Henceforth, we also refer to these edges as *segments*. The location of every individual customer is unobservable, so that airlines have to set a single price for all customers.

Figure 1 illustrates the location of airlines and customers for $N \in \{2, \ldots, 4\}$.[2] As it can be seen, for $N \in \{2, 3\}$ this setup is equivalent to competition on a Hotelling Street and Salop Circle, respectively. The main difference is that for $N > 3$, every airline still competes directly with all other airlines that are active in the market. While a marginal increase in the price of a Salop oligopolist only affects demand for the two adjacent products, in our setting any such increase directly affects demand for *all* alternative products. In this respect our model is similar to the Logit model of oligopolistic competition. Unlike the more general Logit model, however, competition on a simplex is highly tractable and allows us to derive closed-form analytic results. For a more detailed discussion of competition on a simplex see Thompson et al. (2007).[3]

Because consumers in period t are only interested in buying a single airline ticket, potential demand for airline tickets in t is given by the mass of consumers in the market, denoted by μ_t. Suppose that consumers are distributed across the

[2] Note that airlines' location in Figure 1 are projected into $\mathbb{R}^2$, leading to a distortion in the length of the edges. Per assumption, all the edges are of the same length when viewed in $\mathbb{R}^K$. Furthermore, for $N = 3$ the edges are displayed as curves in order to illustrate the conceptional equivalence to the Salop Circle.

[3] Although the simple and symmetric structure of competition on a simplex helps in streamlining our analysis, the main results are robust to the two major alternatives for oligopolistic competition in prices, Salop and Logit.

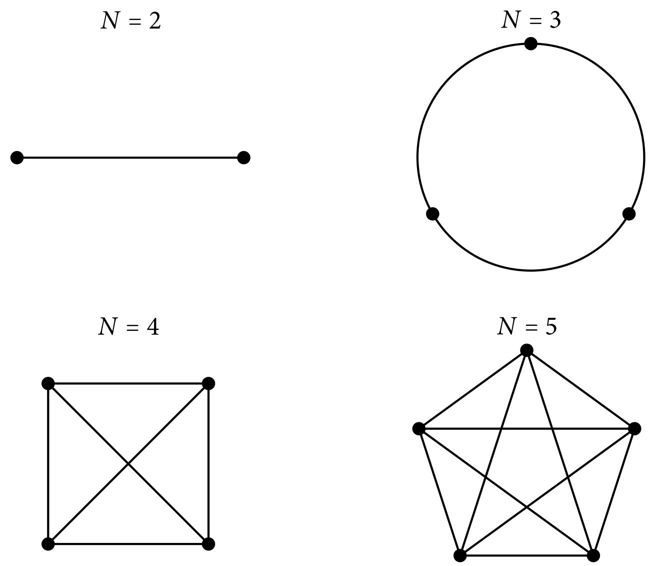

Figure 1. Projection of airlines' location for $N \in \{2, \ldots, 4\}$

simplex with a constant density d that is inversely proportional to the number of edges.[4] Then the length of the edges of the simplex are pinned down to equal μ_t, too. Without changing any of our results, we normalize first period demand μ_1 to unity. Last minute demand, on the other hand, is *ex ante* unknown and is drawn from a uniform distribution over the interval $[0, 2]$.[5] It is worthwhile pointing out that with this specification the length of the edges is independent of N. Accordingly, without capacity constraints, adding additional airlines to any market would not have an effect on prices.

Importantly, we require that from the perspective of each airline the set of its competitors in period 1 is *not* identical to the set of its competitors in period 2. Without this assumption, if airlines choose to serve all customers (i.e., if the market is covered), then price changes by any airline only shift demand towards its competitors, but do not affect overall sales—ruling out the possibility for airlines to affect overall capacity in the last minute market. In fact, to simplify the exposition, we make the extreme assumption that the set of competitors between

[4] That is, $1/d = N(N-1)/2$.

[5] Despite simplifying the exposition, keeping first period demand deterministic has no qualitative effect on any of our results. That is, including demand shock to the first period market neither adds additional insights, nor changes our predictions.

the early and the late market is completely disjoint for all airlines.[6] In an extended version of the model, however, which is available from the authors on request, we show that it is sufficient to allow for an imperfect overlap between the two markets to generate results that are qualitatively identical to the simplified version considered here.[7] Such a situation would arise when some of the early customers consider a different set of alternatives than the one considered by last-minute customers. For instance, customers on the early market might be undetermined with respect to the destination of travel or could consider using other means of transportation, whereas customers on the late market are more determined on buying a ticket for a specific route.

Preferences of a potential customer $\theta \in [0, \mu_t]$ who is located on the edge between airline i and j at time t are given by:

$$U(\theta, p_t^i, p_t^j) = \begin{cases} v - \theta - \alpha p_t^i & \text{if he buys from airline } i \\ v - (\mu_t - \theta) - \alpha p_t^j & \text{if he buys from airline } j \neq i, \end{cases}$$

where $\mu_1 = 1$ and $\mu_2 \sim U[0, 2]$. Note that because consumers in both periods have the same valuation v for a ticket, there is no scope for price discrimination between the two periods. The only way that different prices in $t = 1$ and $t = 2$ can be sustained in equilibrium is through the strategic effects of capacity constraints.

The timing of events can be summarized as follows.

1. Airlines simultaneously choose period 1 prices and serve all consumers willing to buy.

2. Period 2 demand μ_2 is publicly realized and airlines observe the full vector $(\bar{x}_2^1, \bar{x}_2^2, \ldots, \bar{x}_2^N)$ of capacities available to their competitors.

3. Airlines simultaneously choose period 2 prices and serve all consumers willing to buy.

6 To be precise, we assume the existence of N distinct markets in both periods. In any of the N last minute market, there operates exactly one airline from each of the N early markets.

7 In the generalized version of the model, airlines end up competing on another market in period 2 with an arbitrary (and potentially small) probability $p \in (0, 1)$. Hence, there is some positive probability that they face a different set of competitors on the last minute market. While the equilibrium properties remain qualitatively unchanged, the analysis becomes quite cumbersome due to a large number of asymmetric situations which arise off the equilibrium path.

Since airlines are *ex ante* identical, we will concentrate on symmetric subgame perfect equilibria throughout our analysis. Moreover, we will impose the following two parameter restrictions.

Assumption 1: *Airlines do not overbook their flights:*

$$\bar{c} > v/\alpha.$$

Following our discussion above, this assumption ensures that even though it is theoretically possible to , an airline will never voluntarily do so, since the costs are prohibitively high. Clearly, in reality airlines do regularly choose to overbook their flights. Note, however, that in the context of our model customers never choose to cancel a ticket. Accordingly, Assumption 1 merely implies that airlines would not want to overbook a flight if they could be sure that all passengers who bought a ticket will actually turn up at the gate.

Assumption 2: *Airlines sell all available capacity if possible:*

$$v > 4 + c\alpha.$$

This assumption states that the valuation for tickets is sufficiently high, such that airlines don't let any seats go unsold whenever there is sufficient demand. In particular, Assumption 2 implies that if an airline has monopoly power on the last minute market, it decides to sell all of its seat inventory.

3. Equilibrium

We now examine the properties of optimal dynamic pricing. Starting in the next subsection, we first characterize equilibrium prices in the early market, while taking the shadow cost of capacity as given. The value of capacity is then derived subsequently in Subsection 3.2 by solving for the equilibria in the last-minute markets. The analysis of how competition effects the dispersion of prices over time is deferred to Section 4.

3.1. The early market

First note that, since $\mu_1 < \bar{X}_1$, there can be no symmetric equilibrium where airlines are constrained by their capacity in the early market. Accordingly, consider a candidate equilibrium where no airline is constrained by its capacity in $t = 1$

and suppose that airlines set prices such that every customer finds it optimal to buy a ticket. Just like in the case of Salop competition, we can derive the demand for any airline by finding the customer θ who is indifferent between purchasing from airline i and airline j for any $j \neq i$. Solving the indifference condition yields that demand for airline i on this particular segment is given by $q_1^i = (1 - \alpha\{p_1^i - p_1^j\}) \times d/2$. Hence, in a symmetric equilibrium with $p_1^j = p_1$ for all j, total demand per airline is given by $Q_1^i = (1 - \alpha\{p_1^i - p_1\})/N$, after aggregating over all $N-1$ segments of the market in which airline i operates.

In order to characterize equilibrium prices in the early market, we also need to account for how expected profits $\Pi_2^i(\bar{x}_2^i, \bar{x}_2)$ in the late market change as a function of airline i's capacity, given that all other airlines $j \neq i$ are equipped with the symmetric equilibrium capacity $\bar{x}_2^j = \bar{x}_2 = \bar{x}_1 - Q_1^j$. Taking the shadow value of capacity into account, airlines choose their price p_1^i to maximize $Q_1^i(p_1^i, p_1) \times (p_1^i - c) + \Pi_2^i(\bar{x}_2^i, \bar{x}_2)$. It follows that as long as expected future profits Π_2^i are concave in their first argument, airlines' best response functions are given by[8]

$$p_1^i(p_1) = \frac{1}{2}\left[\frac{1}{\alpha} + c + p_1 + \frac{\mathrm{d}\Pi_2^i}{\mathrm{d}\bar{x}_2^i}\right].$$

Hence, in a symmetric equilibrium all airlines set the equilibrium price,

$$p_1^i = p_1 = \frac{1}{\alpha} + c + \frac{\mathrm{d}\Pi_2^i}{\mathrm{d}\bar{x}_2^i},$$

and share the early market,

$$Q_1^i = Q_1 = 1/N.$$

Hence, because $\bar{x}_1 > 1/N$, capacity constraints are indeed not binding in the early market. Also, as long as the value of capacity in period 2 is not too large, firms do set prices such that all customers want to buy a ticket and the period 1 market is covered. We therefore conclude that, as long as Π_2^i satisfies the technical condition discussed above and prices in period 1 are sufficiently low, any symmetric equilibrium in the early market will be uniquely pinned down by the late markets.

[8] To keep the presentation comprehensible, we do not require airlines to set positive prices in the early market. In light of our results, it will become clear that this simplification is without loss of generality and no airline will ever charge negative prices in equilibrium.

LEMMA 1: *Fix any $\Pi_2^i : [0,2]^2 \to \mathbb{R}_+$ and suppose that Π_2^i is concave in its first argument and that $\mathrm{d}\Pi_2^i(\bar{x}_2, \bar{x}_2)/\mathrm{d}\bar{x}_2^i < (v - \alpha c - \frac{3}{2})/\alpha$. Then, in any symmetric equilibrium, airlines split the early market equally and charge a price of*

$$p_1 = \frac{1}{\alpha} + c + \frac{\mathrm{d}\Pi_2^i}{\mathrm{d}\bar{x}_2^i}(\bar{x}_2, \bar{x}_2), \tag{1}$$

where $\bar{x}_2 = 1/N$.

Lemma 1 implies that airlines always enter the late market with the same residual capacity $\bar{x}_2 = 1/N$, irrespective of how valuable capacity is in period 2. Nevertheless, as the marginal value of capacity $\mathrm{d}\Pi_2^i/\mathrm{d}\bar{x}_2^i$ increases, airlines have smaller incentives to underbid each other in the early market, leading to higher equilibrium prices in period 1.

3.2. *The late market*

In order to fix ideas, consider the symmetric equilibrium situation, in which all airlines have the same residual capacity $\bar{x}_2^i = \bar{x}_2$ available at the beginning of period 2. Airlines then end up to be capacity constrained in period 2 whenever the last minute demand μ_2 turns out to exceed the available market capacity $\bar{X}_2 = N\bar{x}_2$. In this case, airlines can set monopoly prices and serve as many customers as they have free seats available, without interfering with any of their competitors. By Assumptions 1 and 2 airlines then find it optimal to sell off exactly their remaining capacity and to extract the full surplus of the marginal customer by charging $p_2(\mu_2) = (v - \bar{X}_2/2)/\alpha$.

On the other hand, in states of the world where market capacity $\bar{X}_2$ exceeds aggregate demand μ_2, airlines can no longer commit to charging monopoly prices and enter into price competition. Because airlines now have an incentive to marginally undercut their competitors in order to sell additional seats, prices are determined by the same logic as in the early market and profits drop discontinuously. In the unique equilibrium of the subgame, each airline sells μ_2/N seats at a price of $p_2(\mu) = \mu_2/\alpha + c$ each.

It follows that in any situation where all firms have the same capacity available expected profits in the late market are given by

$$\Pi_2^i(\bar{x}_2, \bar{x}_2) = \int_0^{\bar{X}_2} \frac{\mu^2}{\alpha N}\,\mathrm{d}(\mu/2) + \int_{\bar{X}_2}^2 \left(v - \alpha c - \frac{\bar{X}_2}{2}\right) \frac{\bar{X}_2}{\alpha N}\,\mathrm{d}(\mu/2), \tag{2}$$

where the first term reflects states where airlines are in competition, and the second term reflects states where industry demand exceeds industry supply and airlines charge monopoly prices.

From Lemma 1, we know that the incentive to shift capacity from the early to the late market is an important determinant of equilibrium prices in the first period. Intuitively, the main benefit from selling an extra seat in period 1 lies in restricting the overall capacity, which increases the probability of ending up with monopoly power in period 2. Standing against this effect is that conditionally on having monopoly power, having an extra seat available in the late market is valuable since it allows airlines to sell this seat at a mark-up.

Formally, this tradeoff corresponds to the incentive for an airline i to unilaterally deviate from a symmetric equilibrium by selling one more seat in period 1. Because this incentive depends on the expected continuation payoff from all reachable subgames *off* the equilibrium path, we can no longer restrict ourselves to symmetric situations in period 2. In Appendix A, we show that (i) the expected payoff $\Pi_2^i(\bar{x}_2^i, \bar{x}_2)$ resulting from the asymmetric equilibria in these subgames is differentiable around $\Pi_2^i(\bar{x}_2^i, \bar{x}_2)|_{\bar{x}_2^i=\bar{x}_2}$ as given by equation (2); and (ii) that $\Pi_2^i(\bar{x}_2^i, \bar{x}_2)$ is globally concave in $\bar{x}_2^i$. Accordingly, we can characterize the incentives to deviate from the symmetric equilibrium by differentiating (2) with respect to its first argument. The following lemma summarizes the discussion and states the relevant conclusion.

LEMMA 2: *In any symmetric equilibrium, Π_2^i is globally concave in its first argument, and marginal returns to an unilateral increase in second period capacity $\bar{x}_2^i$ are given by*

$$\left.\frac{\mathrm{d}\Pi_2^i(\bar{x}_2^i, \bar{x}_2)}{\mathrm{d}\bar{x}_2^i}\right|_{\bar{x}_2^i=\bar{x}_2} = -\left(v - \alpha c - \frac{3\bar{X}_2}{2}\right)\frac{\bar{X}_2}{2\alpha N} + \int_{\bar{X}_2}^{2} \alpha^{-1}\left(v - \alpha c - \bar{X}_2\right) \mathrm{d}(\mu/2). \quad (3)$$

Equation (3) reflects the previously discussed tradeoff. The first term stands for the strategic benefit from *restricting* capacity in order to avoid price competition; the second term defines the expected value of reserving capacity in order to serve last minute customers in high demand states.[9]

[9] There is another case, not reflected in (3), in which demand is so low that supply always exceeds demand. In this case, however, having extra capacity available on the last minute market is worthless, because airlines are already unable to sell their full capacity and any additional seats will go unsold.

3.3. *Equilibrium prices and quantities*

Equipped with Lemmas 1 and 2, we can now state the equilibrium predictions for the full game. In particular, by Lemma 2, $\mathrm{d}\Pi_2^i/\mathrm{d}\bar{x}_2^i$ satisfies the conditions of Lemma 1, so that we can characterize the equilibrium by substituting (3) into (1).

PROPOSITION 1: *There exists a unique symmetric equilibrium in pure strategies. In this equilibrium, prices in the early market are given by*

$$p_1 = \left(\frac{1}{\alpha} + c\right) - \frac{v - \alpha c - 3/2}{2\alpha N} + \frac{v - \alpha c - 1}{2\alpha}, \tag{4}$$

and airlines share the market, selling $1/N$ *tickets each. In the late market, prices are given by*

$$p_2(\mu_2) = \begin{cases} \mu_2/\alpha + c & \text{if } \mu_2 \leq 1 \\ (v - 1/2)/\alpha & \text{if } \mu_2 > 1, \end{cases} \tag{5}$$

and airlines sell $\min\{1, \mu_2\}/N$ *tickets each.*

In equilibrium, first period prices ensure that airlines are indifferent between selling tickets on the early and on the late market. The value of capacity in period 2 is determined by the tradeoff discussed above and is captured by the second and third term in equation (4). That is, as reserving capacity for the period 2 market becomes more attractive, selling tickets on the early market becomes less attractive, which increases the equilibrium price in period 1.

4. **Competition and price dispersion**

We are now ready to investigate the relationship between competition and price dispersion. The next subsection states our main result. Subsection 4.2 compares this result to the literature.

4.1. *Price dispersion due to capacity constraints*

From equation (4), it is obvious that the benefit of reserving capacity for high demand states, represented by the third term, is independent of the number of competitors. This is because these benefits only accrue conditionally on being in a monopoly—in which case airlines are naturally not affected by their competitors.

On the other hand, the strategic incentive to restrict capacity on the late market in order to change the competitive environment is decreasing in the number of competitors N. Intuitively, if the industry is fragmented and there is a large number of competitors, each airline has a small market share and the benefit of avoiding a price war is small. Put differently, a low level of capacity on the late market can be seen as a public good. While the benefit of increasing the probability of monopoly power is enjoyed by all airlines alike, the cost of loosing out on revenues in high demand states is carried by each airline individually. Accordingly, airlines in a more concentrated industry will be more tempted to reduce period 2 capacity than airlines on more competitive routes.

Before proceeding, let us define the expected last-minute price as follows.

Definition: The expected last-minute price is the price that airlines are expected to post on the late market:

$$\mathrm{E}\{p_2\} \equiv \int_0^2 p_2(\mu)\,\mathrm{d}(\mu/2).$$

Note that this is not identical to the average price at which tickets are sold in the market. This is because in the above definition prices are not weighted by the amount of tickets that sell for a particular realization of μ_2. Instead, $\mathrm{E}\{p_2\}$ is the model equivalent to the average offer that an airline makes, which is the variable that we observe in our dataset.

We now turn to our key empirical prediction; namely that prices are expected to be intertemporally less dispersed as the number of competitors increases. By Lemma 1, airlines' capacity at the beginning of the second period $\bar{X}_2$ is independent of N. As established in Proposition 1, this causes prices in the second period to be unaffected by the number of competitors, too. Any effect of N on the difference in expected prices over time ($\Delta \equiv \mathrm{E}\{p_2\} - p_1$) must therefore result from changes in p_1. As already discussed above, airlines' incentive to price aggressively in the early market in order to relax future competition decreases in the number of competitors, causing first-period prices to increase in N. Hence, $\mathrm{d}\Delta/\mathrm{d}N < 0$. Moreover, it is straightforward to verify that expected last-minute prices exceed those on the early market; i.e., $\Delta > 0$, so that the intertemporal *dispersion* of prices $|\Delta|$ decreases in N, too. The following proposition summarizes the discussion.

Proposition 2: *The slope of expected prices over time is strictly decreasing in the number of competitors:*

$$\frac{\mathrm{d}\Delta}{\mathrm{d}N} < 0,$$

where $\Delta = \mathrm{E}\{p_2\} - p_1 > 0$.

As a final note, mind that in our model all prices are weakly increasing in the number of competitors since the attempt to restrict capacity in the late market is self-defeating in equilibrium. However, while our main prediction in Proposition 2 is robust to other competitive frameworks, price levels are unlikely to increase in competition in alternative settings. This observation should therefore be regarded as a peculiarity of our framework, where the early market is always covered and actual capacity in the late market is independent from airlines' pricing choices in the early market.

4.2. Alternative explanations for price dispersion

COMPETITIVE PRICE DISCRIMINATION We have proposed one possible explanation for why prices for airline tickets typically increase over time. Another possible explanation are differences in the price elasticity of consumers. Arguably, business travelers are more likely to make travel plans at short notice than more price elastic leisure travelers. So the price elasticity of customers on the last minute market is lower than the price elasticity of customers on the early market and in an oligopolistic setting, airlines may discriminate against customers that book their flights late. For a discussion of the equilibrium level of price discrimination that we should expect in markets with and without free entry, see e.g. Borenstein (1985); Holmes (1989) and Armstrong and Vickers (2001). For a detailed discussion of third degree price discrimination in competitive settings see Stole (2007).

Unfortunately, the literature on competitive price discrimination does not yield clear-cut testable predictions on how price discrimination depends on the level of competition in an industry. The relationship between competition and price discrimination crucially depends upon the relation between cross-price elasticities and the market elasticity of demand. Depending on these quantities, price discrimination can be either increasing or decreasing in competition. However, if price dispersion is due to discriminatory pricing, we should expect dispersion to increase in the heterogeneity of customers on a route. In our empirical analysis, we try to find evidence in favor of competitive price discrimination by looking at routes where we expect customers to be rather homogeneous. If price discrimination plays a role in the determination of ticket prices, these routes should feature less price dispersion than other itineraries.

STOCHASTIC PEAK-LOAD PRICING A large body of theoretical literature has been trying to explain price dispersion on the market for air travel by models of stochastic peak-load pricing. Under stochastic peak-load pricing, airlines decide on a price schedule *ex ante* and are able to commit to it. Seats that are still available on the last minute market sell with a smaller probability than the average seat that an airline offers. So if investing into capacity is costly, an airline will charge a higher price for the last seat than for the average seat in order to recoup the cost of capacity in expectation. Dana (1999a) shows that if price dispersion is due to peak-load pricing, we should expect intertemporal price differences to be increasing in the number of competitors. The closer an industry moves towards perfect competition, the more differences in costs are going to translate into differences in prices. The expected cost of selling a ticket is given by the cost of increasing capacity by one seat divided by the probability of selling this particular seat. Since this expected cost is highest for seats available at the last minute, these seats sell at a higher price and the price difference to the average seat is strictly increasing in competition.

From a theoretical perspective, recent innovations of internet booking and dynamic pricing systems are likely to have undermined airlines' power to commit to future price schedules. However, the question whether or not a model of peak-load pricing is better suited to explain price dispersion than the one put forward in this study is ultimately an empirical one. While a positive relationship between price dispersion and competition lends support to peak-load pricing, our model predicts an inverse relationship.

5. Description of the dataset

In order to test our theoretical predictions, we use a hand-collected data set of ticket prices on 92 intra-European routes. The routes are randomly selected and cover a variety of different regions, market characteristics, and different numbers of competing airlines. A geographical overview of these routes is given by Figure 2, and a full list of the routes is provided in Table 3 in the Appendix.[10] On each

[10] More specifically, our selection of routes reflects the following criteria. First, we include all routes that connect the largest airports in France, Germany, Italy, Spain, and the UK (in a randomly chosen direction). The remaining routes correspond to 100 randomly chosen routes between European airports with international connections, excluding all routes that are served by Ryanair, a major competitor who does not offer tickets through our data source (see below). Hereby a route is defined as being served by Ryanair if Ryanair serves either the main airport or a secondary airport located less than 35 miles from the city center at both endpoints of the route.

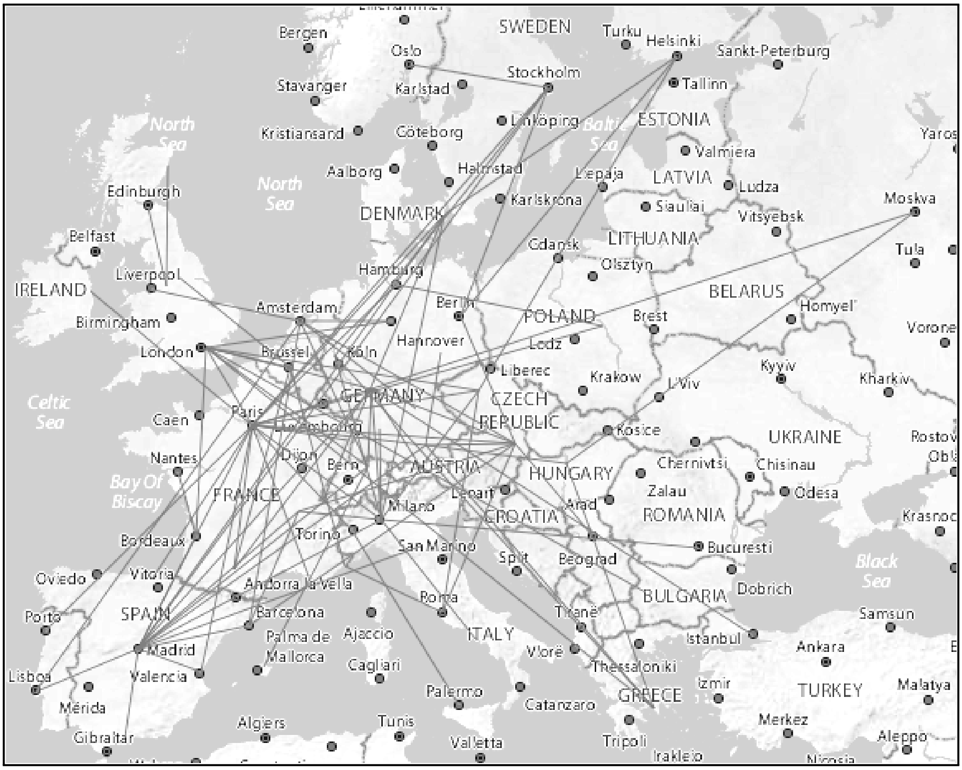

Figure 2. Routes

route, we record prices for all direct flights leaving on Friday and returning on Sunday, as well as all flights leaving on Monday and returning on Thursday in a given week. So for any given route there are two route-date pairs per week in our sample. We refer to these route-date pairs as "markets". For each market, we record flights and prices once a week, starting 10 weeks prior to the departure date. In the last week prior to departure, prices are recorded on a daily basis, giving us a total of up to 17 different prices for each flight. A flight is hereby defined as a roundtrip, which is uniquely characterized by a combination of two individual flight numbers. For example, in our terminology one "flight" on the route Paris–London would be using flight number BA 333 on the outbound leg and flight number BA 334 on the inbound leg.

Our data set runs from October 31, 2010, to March 26, 2011, which covers the complete 2010/2011 European winter flight schedule.[11] This corresponds to 41 distinct flight dates, or 3772 distinct markets (41 flight dates times 92 routes). Each market averages 376 prices that are recorded over the 17 different dates prior to departure, corresponding to an average of 22.1 flights per market. Overall, our data set consists of 1.42 million individual prices (92 routes times 41 flight-dates times 17 recorded prices per flight times on average 22.1 flights per market).[12] Routes are on average 560 miles long and connect Metropolitan Areas with an average of 3.9 million inhabitants.[13] The share of domestic routes in our sample is roughly 13 percent (12 of the 92 routes).

Prices represent offers by a leading website for airline ticket purchases, which accounts for a major share of bookings on the European market. While we cannot rule out that prices may differ to those offered by other online retailers, differences across retailers for intra-European flights are typically small. Moreover, unless the influence of competition on intertemporal *changes* in prices consistently differs across retailers, our empirical analysis is representative. The recorded prices in our sample range from 27 to 2581 Euros, with an average of 409 Euros and a standard deviation of 466.[14]Conditioning on the time remaining until departure, prices increase from an average of approximately 280 Euros ten weeks prior to departure to more than 500 Euros within the last week before departure (see Figure 3 for a more comprehensive summary of the evolution of average prices over the time remaining until departure).

To investigate the impact of competition on the observed pricing dynamics, we measure competition as the number of airlines that compete in a given market.[15]

[11] Flight schedules and routings within Europe are planned on a semiannual basis. Within these periods, individual flight numbers serve as a unique identifier for individual flight characteristics such as the exact route, time of departure and approximate flight length.

[12] Not every flight was offered on every 17 dates prior to departure. Given the total of 1.42 million recorded prices in our sample, this increases the average flights per market, so that 22.1 should be more accurately interpreted as a lower bound on average flights per market.

[13] The large average size of Metropolitan Areas is due to the over-sampling of large airports (see also Footnote 10).

[14] Because of the increased frequency with which we record prices in the last week prior to departure, the average price in our sample does *not* equal the average price at which a ticket is offered over the last 10 weeks prior to departure. Downsampling the last week's observations in each market, the 10 week average in our sample amounts to 364 Euros with a standard deviation of 466.

[15] In 7.9 percent of our sample, the number of airlines offering services on the outbound leg differs from the number of firms offering services on the inbound leg. In these cases, we set competition to the rounded up average.

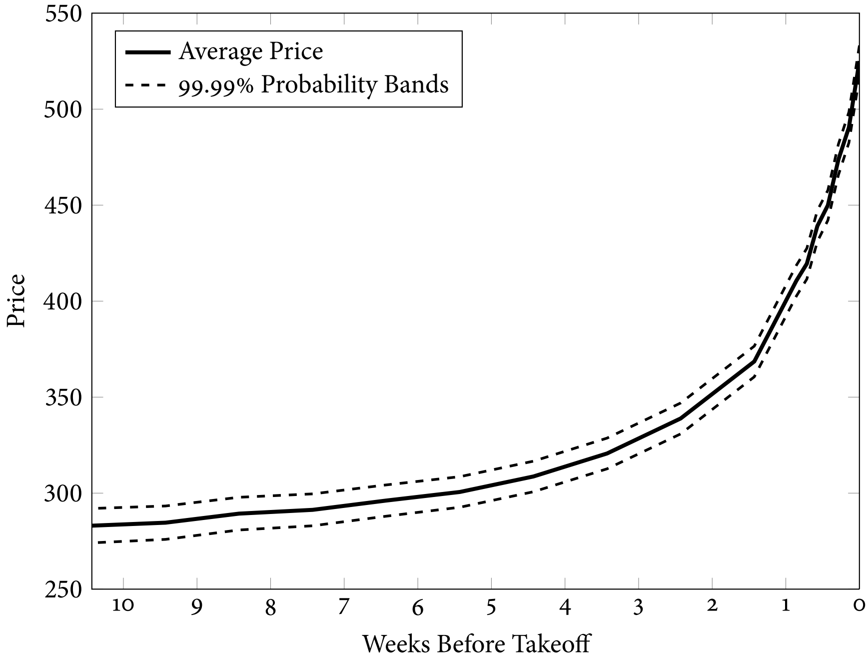

Figure 3. Average prices (in Euros) as a function of time remaining until departure

For these purposes, we treat airlines that are affiliated through cross-holdings as single competitors. More precisely, an airline is matched to an affiliate group if that group owns more than 25% of the airline's equity. Table 1 summarizes the distribution of competitors in the sample.

Table 1. Competition in the sample

	Prices		Markets	
Competing airlines	Frequency	Percent	Frequency	Percent
1	229 218	16.17	905	24.06
2	648 371	45.74	1696	45.08
3	275 680	19.45	656	17.44
4	185 051	13.05	382	10.15
5	68 237	4.81	107	2.84
6	11 078	0.78	16	0.43

6. Empirical specification and results

Let $Price_{ijtd}$ denote the price for a round trip that involves the outgoing itinerary i and the returning itinerary j (both identified by their flight numbers), which is posted at date d, and for which the outgoing flight departs at date t. Further, let $Competition_{ijt}$ denote a vector of dummy variables that covers all competition categories, and let $Daysleft_{td}$ denote the difference between t and d in days. As a baseline, we estimate the following equation:

$$\log(Price_{ijtd}) = (\alpha + \beta Daysleft_{td}) \times Competition_{ijt} + \lambda_i + \mu_j + \nu_t + \xi_d + \varepsilon_{ijtd}, \quad (6)$$

where we treat λ_i, μ_j, ν_t, and ξ_d as fixed effects.[16] Here, α is a vector of competition-specific constants and β is the relevant coefficient-vector on the interaction term $Competition_{ijt} \times Daysleft_{td}$. Note that λ_i and μ_j both nest a complete set of route specific fixed effects since any flight number uniquely pins down the corresponding route. In particular, the specified set of fixed effects absorbs all flight-related effects such as departure time or length of flight; all route characteristics such as connected cities or alternative means of transportation; and all time-related effects such as day of travel and day of price offer.

The impact of competition on the observed pricing dynamics is captured by our estimates of β. Table 2 reports the estimated coefficients. Our estimates for the corresponding standard errors are adjusted for clustering at the market level. All reported coefficients are economically and statistically significant. It can be seen that, consistent with our model, the amount of intertemporal price differentiation is decreasing in the number of competitors. While on monopoly routes prices increase by an average of 1.31 percent with every day that a customer waits to book, this slope is reduced to 1.19 percent on duopoly routes and is reduced to 0.68 percent for routes with 6 competitors.[17]

[16] Because our sampling is weekly for all but the last week before departure, fixed effects for d can not be separately identified from $Daysleft_{dt}$ on a daily level; ξ_d is therefore modeled on a weekly level. Moreover, it is worth pointing out that flight numbers i and j are allocated on a semi-annual basis and identify a particular flight leg flown by a particular airline at a particular time within a weekly schedule. Because i and j remain constant across weeks, we can identify β and control for i and j at the same time.

[17] Note that the (unreported) competition-specific constants are only weakly identified in our sample by variations across markets but within routes. This is because competition typically does not vary within routes for a given flight schedule. Accordingly, the competition-specific constants are statistically not significant.

Table 2. The effects of competition on pricing dynamics

Dependent variable is $\log(Price_{ijtd})$

	Coefficient	Clustered Std. Errors
$(Comp_{ijt} = 1) \times Daysleft_{td}$	-1.31	0.09
$(Comp_{ijt} = 2) \times Daysleft_{td}$	-1.19	0.09
$(Comp_{ijt} = 3) \times Daysleft_{td}$	-1.15	0.09
$(Comp_{ijt} = 4) \times Daysleft_{td}$	-1.07	0.09
$(Comp_{ijt} = 5) \times Daysleft_{td}$	-0.93	0.10
$(Comp_{ijt} = 6) \times Daysleft_{td}$	-0.68	0.12
Observations	1 417 635	
R-squared (adj.)	0.58	

Notes.— The estimation also includes a complete set of (weakly identified) competition-specific constants, and fixed effects for both outgoing i and incoming flights j, the date where prices are recorded d, and the date of departure t. Reported coefficients and standard errors are multiplied by 100. Standard errors are clustered at the market level. All coefficients are significant at the 5×10^{-7} percent level.

Note that even though we include routes that are served by only one airline in our sample, we should not think of these routes as being completely protected from competition. In fact, most itineraries are also offered by airlines that offer indirect flights and that are subject to similar capacity constraints as airlines offering direct flights. Hence, we think of markets served by only one airline as markets with a particularly low number of competitors, rather than a proper "monopoly".

From the discussion in Section 4.2, we conclude that dynamic pricing under capacity constraints appears to better explain intra-European airline pricing than theories based on peak-load pricing (which predict that the absolute slope should be increasing in the number of competitors). With respect to competitive price discrimination, the inconclusiveness of theories whether the intertemporal slope should be increasing or decreasing in competition prevents any final conclusions. However, as discussed above, if price discrimination drives the intertemporal differentiation, then one should expect intertemporal price differences to be less pronounced on routes where there is less heterogeneity in customers' willingness to pay. One big source of heterogeneity is arguably the co-existence of business and leisure travelers. Brueckner et al. (1992) and Goolsbee and Syverson (2008) have argued that a good proxy for the share of leisure travelers on a route are temperature differentials between the destination and origin. In particular, we should expect the share of leisure travelers to be higher, the warmer the destination is in comparison to the origin.[18] Since the share of leisure travelers is in general markedly larger than the share of business travelers, we would hence expect routes with small or negative temperature differentials to have a more heterogenous customer base and to exhibit more intertemporal price differentiation.[19]

In order to provide a rudimentary assessment to which extent discriminatory pricing might be the source of intertemporal price dispersion, we split the data set into three subsamples. The first subsample consists of all 30 routes in our data set where the yearly average temperature in the destination city is more than one degree Celsius above the average temperature in the departing city. The second subsample consists of all 37 routes in the data set where the destination city is more than one degree Celsius colder than the departing city, and the third subsample consists of the remaining 25 routes, linking cities in which the temperature is approximately the same.[20]

For all three subsamples, we again estimate equation (6). Figure 4 reports the estimated coefficients. It can be seen that the relationship between the estimated coefficients and competition is overall increasing in all three subsamples. Moreover, the coefficients are in a similar range as in our baseline estimation and there is no clear correlation between our measure of customer heterogeneity and the amount of intertemporal price dispersion. This holds true for any number of competitors. While only rudimentary in nature, we thus find no strong support for a discrimination based theory of pricing.

7. Robustness specifications

7.1. Nonlinear effect of days before departure

Our baseline regression imposes that prices linearly depend on $Daysleft_{td}$. Clearly, as can be seen in Figure 3, this is not the case. To address this shortcoming, we repeat the estimation, but replace $Daysleft_{td}$ by a vector of dummy variables, which covers all values of $Daysleft_{td}$. The resulting nonlinear relationships are reported in Figure 5. In the figure, we normalize the estimated coefficients such

[18] Recall that all prices in our sample are per round trip, allowing airlines operating on route A–B to systematically discriminate between travelers visiting city A and those visiting city B.

[19] For example, leisure travelers accounted for 60% of departures at the Airport of Frankfurt and for 68.7% in London Heathrow (Fraport, 2011; CAA, 2011).

[20] For a full listing of the three subsamples, see Table 3 in the Appendix.

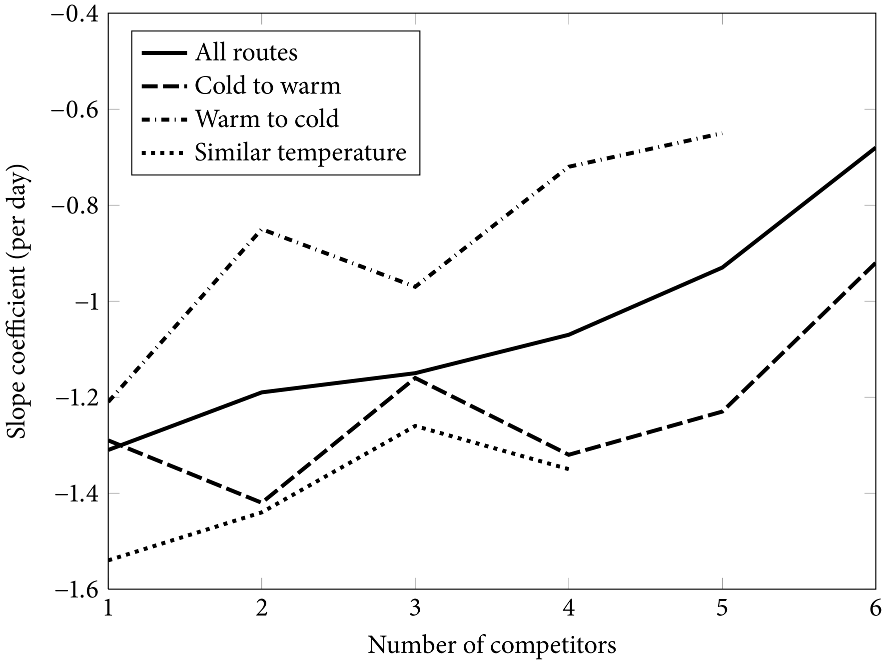

Figure 4. Estimated coefficients for different levels of competition and different subsamples. *Notes:* Reported coefficients are multiplied by 100. All coefficients are significant at the 5×10^{-5} significance level (using standard errors that are clustered at the market level). All estimations control for all four fixed effects specified in the baseline regression and a full set of competition-specific constants.

that prices one day before takeoff are set to zero. Accordingly, the numbers on the y-axis report

$$\log\left(\frac{p_{ijtd}}{p_{ijtd|d=t-1}}\right),$$

which approximately amounts to the early-booking discount in percent of the price charged shortly before takeoff. It can be seen that, although nonlinear, the slopes are monotonically decreasing in the number of competitors. That is, the relative discount for booking a flight in advance is less pronounced on routes that are served by a larger number of competitors. This reinstates the conclusion drawn from our baseline estimation.

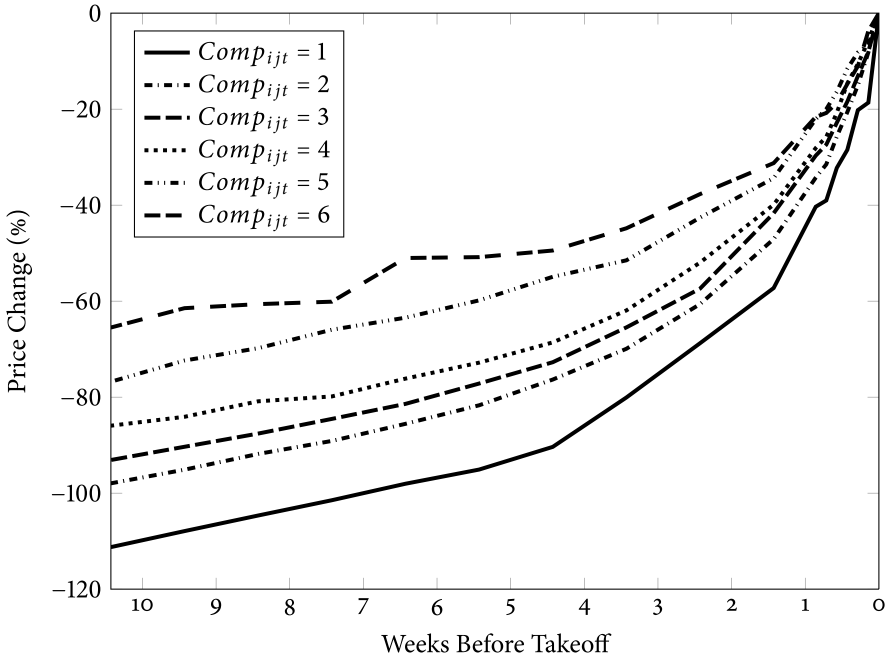

Figure 5. Empirical relationships between $\log(Price_{ijtd})$ and $Daysleft_{td}$ by degree of competition. *Notes:* Coefficients are multiplied by 100 and are normalized, such that prices 1 day before takeoff are set to zero, so that the reported estimates approximately correspond to percentage price changes relative to the price before takeoff. The estimation controls for all four fixed effects specified in the baseline regression and a full set of competition-specific constants.

7.2. Symmetric markets only

Another worry might be that our theoretical exploration is based on the simplifying assumption that all competing airlines are symmetric, while markets in our empirical analysis can be strongly asymmetric. To test whether our results are driven by a correlation between market asymmetries and competition, we repeat our baseline estimation (6) for symmetric markets only. For this purpose, we define a market to be symmetric if the market share (total number of flights offered by a competitor, as defined above, relative to the total number of flights in the market) of the smallest competitor relative to the largest competitor is at least 1/3. The resulting sample is approximately two thirds the size of our baseline sample and has 995 903 observations. As can be seen from Figure 6, the estimated

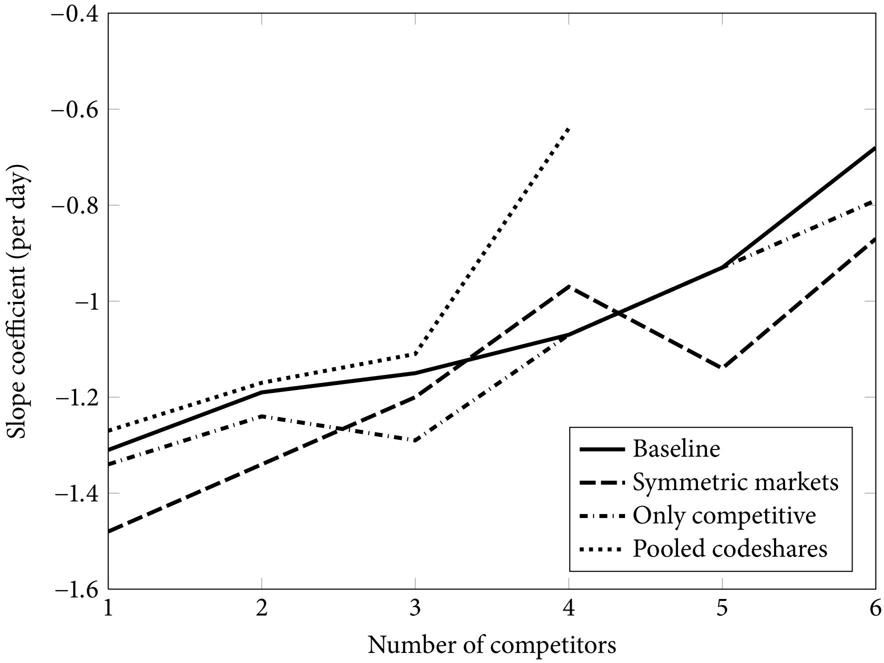

Figure 6. Estimated coefficients for different levels of competition—robustness specifications. *Notes:* Reported coefficients are multiplied by 100. All coefficients are significant at the 5×10^{-7} significance level (using standard errors that are clustered at the market level). All estimations control for all four fixed effects specified in the baseline regression and a full set of competition-specific constants.

coefficients closely resemble those in our baseline estimation.

7.3. Alternative measures for competition

In the baseline estimation, we treat codesharing airlines as competitors. Accordingly, if the same physical connection is marketed under different flight numbers that correspond to different airlines, this increases our measure of competition. The reasoning behind this choice is that in so-called "block space" codeshare agreements, each of the codesharing partner still controls a distinct, *ex ante* fixed amount of seats. In practice, by the pricing agreements between the carrier operating a service and the codesharing partner, the codesharer is usually granted considerable freedom to set prices independently (European Commission, 2007). Accordingly, prices are indeed often observed to differ across different codesharers. Alternatively, one could also define competition as the number of airlines that operate their own services on a particular market. To check the implications of this approach, we re-estimate equation (6), using such an alternative measure for competition.[21] As can be seen from Figure 6, the estimated coefficients closely resemble those in our baseline estimation.

7.4. Correcting for outliers

Since we do not observe transaction data but only posted prices, our dataset includes some prices that are extremely high when compared to comparable fares. Arguably, such offers are the result of mistakes by the respective airline and are never taken up by consumers. To account for such "outliers", we also re-estimate the baseline regression (6) after dropping all observed prices that exceed the mean price offered at any given date in a given market by more than two standard deviations. The competition measure is adjusted accordingly. The resulting sample is about 4 percent smaller than our baseline sample and has 1 360 748 observations. As can be seen from Figure 6, the estimated coefficients again closely resemble those in our baseline estimation.

8. Conclusion

In this study, we introduce a novel argument for why intertemporal price-dispersion can be an equilibrium, even if firms are unable to commit to future prices. In environments with short-term capacity constraints, firms price aggressively in early periods in order to have less capacity available in later periods. Using this strategy, firms thereby increase the probability that aggregate demand exceeds aggregate supply on the last minute market and that they end up with monopoly power over last-minute consumers, yielding them considerable profits whenever there is excess demand. Importantly, this propensity to price aggressively decreases as markets become more competitive, implying that prices differ less across time in more competitive markets.

Using a unique data set, we find a high degree of intertemporal price dispersion. Moreover, on routes with a large number of competing firms, intertemporal price dispersion is less pronounced than on oligopolistic routes. This result is in line

[21] To be consistent with this approach, we also pool all physically identical roundtrips into a single observation, where at each date the pooled roundtrip is assigned the price offered by the cheapest partner.

with the predictions of our capacity-based model. In contrast, the standard explanation used in the literature to explain intertemporal price differentiation—peak-load pricing—is rejected by our findings. Regarding price discrimination, the inconclusive predictions of theories of competitive price discrimination do not allow us to effectively test whether price discrimination is a relevant cause for intertemporal price dispersion. Based on a rudimentary argument, we, however, find no support for discriminatory pricing.

A. Mathematical appendix

Let us examine the subgame on the last minute market more closely. Since we are looking for a symmetric equilibrium in the full game, we can assume that all but one firm have the same capacity at their disposal in period 2. We will denote the deviating firm by the superscript i while all other firms will collectively be referred to as type j. For notational simplicity we will often drop the superscript for firms of type j.

In general, firms set monopoly prices if aggregate demand exceeds aggregate supply and competitive prices otherwise. When firms do not have the same residual level of capacity at the beginning of period 2, we also allow for asymmetric equilibria in the off-equilibrium subgames. In the proof, we will consider the case where $N \geq 3$. The proof for $N = 2$ is very similar. The only difference is that in case $N = 2$ some asymmetric cases can be ruled out. We will derive equilibria that obtain for different realizations of μ_2 first. In a second step, we will then show that for any given μ_2 the equilibrium on the period 2 market is unique. As it will be seen, the resulting continuation value of capacity is identical to the one in equation (2), allowing us to focus on the symmetric case in the main body of the text.[22]

A.1. First case $\bar{x}_2^i \leq \bar{x}_2$

Consider the case where firm i deviates by choosing a smaller period 2 capacity than in the symmetric equilibrium. Furthermore, suppose that for $\mu_2 \leq N\bar{x}_2^i$ no firm ends up capacity constrained in equilibrium and that firms of type j set a

[22] In the following, we present the analysis by focusing on interior solutions, omitting some cumbersome steps that show that also deviations to corner solutions are not profitable in the unique equilibrium. The omitted steps are available from the authors on request.

symmetric price of $p_2^j = p_2$. In this case firms choose prices according to their best response functions

$$
\begin{aligned}
p_2(p_2^i, p_2) &= \frac{1}{2}\left[\frac{\mu_2}{\alpha} + \frac{1}{N-1}p_2^i + \frac{N-2}{N-1}p_2 + c\right] \\
p_2^i(p_2) &= \frac{1}{2}\left[\frac{\mu_2}{\alpha} + p_2 + c\right]
\end{aligned}
$$

and in equilibrium all firms set symmetric prices of $p_2 = p_2^i = \frac{\mu_2}{\alpha} + c$ and share the market equally. So firms are indeed not constrained by their capacity. However, once demand increases beyond $N\bar{x}_2^i$ it can no longer be optimal for firm i to set the same price as everybody else since its capacity constraint becomes binding.

Subcase A Suppose $\mu > N\bar{x}_2^i$ and suppose that in equilibrium firm i is capacity constrained while nobody else is constrained. In this subcase firm i will set a price that ensures that it doesn't face any excess demand. If all customers decide to buy a ticket, this yields the following best response function for firm i:

$$p_2^i = \frac{1}{\alpha}\left[\mu_2 - N\bar{x}_2^i\right] + p_2.$$

All firms other than i set prices according to their unrestricted best response function, which yields the equilibrium prices

$$
\begin{aligned}
p_2 &= \frac{N(\mu_2 - \bar{x}_2^i)}{\alpha(N-1)} + c \\
p_2^i &= \frac{(2N-1)\mu_2 - N^2\bar{x}_2^i}{\alpha(N-1)} + c.
\end{aligned}
$$

We can check that for $v > 4 + \alpha c$ all customers do indeed find it optimal to purchase a ticket and firms other than i are not capacity constrained whenever $\bar{x}_2^i + (N-1)\bar{x}_2 = \hat{\mu}_i \geq \mu_2$. In order to verify that this is an equilibrium, note that firms j set prices according to their unrestricted best response function. Firm i on the other hand doesn't have any incentive to reduce prices due to capacity constraints. Increasing prices is not profitable, either: The marginal return to an increase in prices is given by $-\alpha(p_2^i - c)/N + \bar{x}_2^i$ which is negative.

Subcase B Let us now consider the subcase where $\mu_2 > \hat{\mu}_i$, i.e. airlines are unable to serve all customers without exceeding their joint capacity. We claim that there is an equilibrium in which firms set prices that effectively amount to monopoly prices and sell all of their seats. However, the prices of firms other than i need to account for the fact that capacity is unevenly distributed across the market: Firms j serve all customers located on segments connecting two firms of type j and share demand on those segments symmetrically. Additionally, firms j have to set sufficiently low prices in order to attract customers that are located in the vicinity of firm i but that firm i chooses not to serve due to capacity constraints. Airlines j will never be able to cater to all customers that are not served by firm i, so firm i is able to set monopoly prices.

As long as $N\bar{x}_2 \geq \mu_2$ firms set prices of

$$\begin{aligned} p_2 &= \frac{v}{\alpha} + \frac{(N-2)\mu_2 - (N-1)N\bar{x}_2}{2\alpha} \\ p_2^i &= \frac{v}{\alpha} - \frac{N\bar{x}_2^i}{2\alpha}. \end{aligned}$$

while for $N\bar{x}_2 < \mu_2$ all firms set monopoly prices and we get $p_2 = \frac{v}{\alpha} - \frac{N\bar{x}_2}{2\alpha}$. Firm i does not have any incentive to reduce prices due to capacity constraints. Moreover, the return to increasing prices is strictly negative. Similarly, none of he other firms has any incentive to deviate from this equilibrium due to capacity constraints.[23]

Returns to capacity in case $\bar{x}_2^i \leq \bar{x}_2$ Summing up our results so far, whenever firm i chooses a weakly smaller capacity for the late market than everybody else, its expected profits are given by

$$\begin{aligned} \Pi_2^i(\bar{x}_2^i, \bar{x}_2) = \int_0^{N\bar{x}_2^i} \frac{\mu^2}{N\alpha} f(\mu)\,\mathrm{d}\mu + \int_{N\bar{x}_2^i}^{\hat{\mu}_i} \left[\frac{(2N-1)\mu - N^2\bar{x}_2^i}{\alpha(N-1)}\right] \bar{x}_2^i f(\mu)\,\mathrm{d}\mu \\ + \int_{\hat{\mu}_i}^{2} \left[\frac{1}{\alpha}\left(v - \frac{N\bar{x}_2^i}{2}\right) - c\right] \bar{x}_2^i f(\mu)\,\mathrm{d}\mu. \end{aligned}$$

It is easy to see that the value of an increase in the capacity available on the last minute market is given by

[23] Generally, when checking for deviations in which a firm chooses a larger price it is sufficient to check if the deviation is profitable for $\partial q_2^i / \partial p_2^i = -\alpha/(N(N-1))$. While the responsiveness of demand is strictly larger in absolute terms whenever the marginal customer is indifferent between buying a ticket from firm i or not buying a ticket at all, using $\partial q_2^i / \partial p_2^i = -\alpha/(N(N-1))$ gives us sufficient conditions for prices to be part of an equilibrium.

$$\begin{aligned}\frac{\mathrm{d}\Pi_2^i(\bar{x}_2^i,\bar{x}_2)}{\mathrm{d}\bar{x}_2^i} &= \int_{N\bar{x}_2^i}^{\hat{\mu}_i}\left[\frac{(2N-1)\mu-2N^2\bar{x}_2^i}{\alpha(N-1)}\right]f(\mu)\,\mathrm{d}\mu \\ &\quad -\frac{1}{N\alpha}\left[\nu-(2N-1)\bar{x}_2+\left(\frac{N-2}{2}\right)\bar{x}_2^i-\alpha c\right]N\bar{x}_2^i f(\hat{\mu}_i) \\ &\quad +\int_{\hat{\mu}_i}^{2}\left[\frac{1}{\alpha}\left(\nu-N\bar{x}_2^i\right)-c\right]f(\mu)\,\mathrm{d}\mu. \qquad (7)\end{aligned}$$

Evaluating equation (7) at $\bar{x}_2^i = \bar{x}_2$ gives us expression (3). Moreover, if the first period market is covered, we get $\bar{x}_2 = 1/N$ and using the fact that $f(\mu)$ is constant for all $0 < \mu_2 < 2$ we can check that the second order condition

$$\begin{aligned}\frac{\mathrm{d}^2\Pi_2^i(\bar{x}_2^i,\bar{x}_2)}{\mathrm{d}\left(\bar{x}_2^i\right)^2} &= \frac{N^2\bar{x}_2^i}{\alpha(N-1)}f(N\bar{x}_2^i)-\int_{N\bar{x}_2^i}^{\hat{\mu}_i}\left[\frac{2N^2}{\alpha(N-1)}\right]f(\mu)\,\mathrm{d}\mu \\ &\quad -\int_{\hat{\mu}_i}^{2}\frac{N}{\alpha}f(\mu)\,\mathrm{d}\mu-\frac{1}{\alpha}\left[\nu-(2N-1)\bar{x}_2+(N-2)\bar{x}_2^i-\alpha c\right]f(\hat{\mu}_i) \\ &\quad -\frac{1}{\alpha}\left[\nu-(2N-1)\bar{x}_2+N\bar{x}_2^i+\frac{1}{(N-1)}\bar{x}_2^i-\alpha c\right]f(\hat{\mu}_i) \qquad (8)\end{aligned}$$

is satisfied for all values of $\bar{x}_2^i \in [0,\bar{x}_2]$ since the sum of the first three terms is negative.

Uniqueness Finally, let us show that the pure-strategy equilibria considered above are unique for any given μ_2. First, let's establish that firms with the same capacity must set the same price in period 2. In order to do so, let us assume otherwise and order firms according to their prices:

$$p_2^1 \le p_2^2 \le p_2^3 \le \ldots \le p_2^N,$$

where at least two inequalities are strict. We define τ_N as the profit that firm N would make on a given segment if the competitor were to charge the same price. If all customers were to buy a ticket on that segment we have $\tau_N = \left(p_2^N - c\right)\frac{\mu_2}{N(N-1)}$. But firm N might also charge monopoly prices which would result in some customers dropping out. Similarly, firm 1's profits under symmetry are denoted τ_1. Let us assume that neither firm N nor firm 1 is of type i. If firm 1 does not serve all customers on the segment towards firm N, no segment can be at a corner solution. It is easy to see that if firm N doesn't have an incentive to reduce prices, firm 1 has a strict incentive to increase prices. If however firm 1 serves the segment towards firm N fully, we must have $\Pi_2^1 - \Pi_2^N \leq \tau_1$: Otherwise firm N would have an incentive to imitate firm 1 even if we account for the fact that they would have to share demand on their joint segment. Similarly, if firm 1 has no incentive to imitate firm N we must have $\Pi_2^N - \Pi_2^1 \leq -\tau_N$. However, using the fact that $\Pi_2^1 > (N-1)\tau_1$ and $\Pi_2^N < (N-1)\tau_N$ this yields a contradiction. In case firm N is of type i firm N might be capacity constrained when imitating firm 1. However, in this case we can use similar reasoning for firms 1 and $N-1$. If firm 1 is of type i, firm N might again be capacity constrained when imitating firm 1. In this case firm N could imitate firm 2. If firm 2 does not serve the segment leading to firm N fully, firm N faces corner solutions on weakly less markets than firm 1 and if firm N doesn't have any incentive to reduce prices, firm 1 has an incentive to increase prices. Hence, firms that have the same capacity available must set the same prices in period 2.

In case aggregate supply exceeds aggregate demand it is straightforward to see that the equilibrium is unique. Only the type of firm that has less capacity available can be capacity constrained: Otherwise, the constrained firm would set a smaller price and would have an incentive to increase its price. But if the set of firms that are capacity constrained is uniquely defined and symmetric firms set symmetric prices, the equilibrium derived above must be unique.

Let us now turn to the case where aggregate demand exceeds aggregate supply. Assume that firm i sets a larger price than the one prescribed above. In this case firm i can not sell its full capacity and has an incentive to reduce prices. Instead, assume that firm i charges a lower price. If firm i is not to exceed its capacity, firms j have to set a price strictly below p_2. But if firms j charge a price below p_2, firms face more demand than their joint capacity, so this can not be an equilibrium. Similarly we can also show that for the given price p_2^i the prices charged by all firms of type j are uniquely determined.

A.2. Second case: $\bar{x}_2^i \geq \bar{x}_2$

Let us now turn to the case where firm i chooses a higher capacity than everybody else. If $N\bar{x}_2 \geq \mu_2$ we again have a symmetric equilibrium in which no firm is capacity constrained and all airlines set a price of $p_2 = \frac{\mu_2}{\alpha} + c$.

Subcase A Let us assume that $\mu > N\bar{x}_2$ and in equilibrium, all firms except for firm i are capacity constrained. In this subcase firms j sets prices such that they sell their full capacity while firm i prices according to its familiar best response function. This gives us equilibrium prices of

$$p_2 = \frac{1}{\alpha}\left((2N-1)\mu_2 - 2(N-1)N\bar{x}_2\right) + c$$
$$p_2^i = \frac{1}{\alpha}\left(N\mu_2 - (N-1)N\bar{x}_2\right) + c$$

if all customers decide to buy a ticket. We can check that this is indeed the case whenever $v > 4+\alpha c$. So we know that firm i sells a quantity of $Q_2^i = \mu_2-(N-1)\bar{x}_2$ and firm i is not capacity constrained in equilibrium as long as $\bar{x}_2^i + (N-1)\bar{x}_2 = \hat{\mu}_a \geq \mu_2$. Clearly, firms j have no incentive to lower prices, given that they are capacity constrained. Firm i follows its unrestricted best response function and will also have no incentive to change prices. So we are only left to check that firms j do not find it worthwhile to increase prices: The return to a marginal increase in prices is given by $-\alpha(p_2-c)/N+\bar{x}_2$ which is negative in the candidate equilibrium. Finally, we need to check that firm i does not sell more tickets that there are customers on segments accessible by firm i: $\mu_2 - N\bar{x}_2 \leq 2\mu_2/N$. This condition is always satisfied as long as deviations $\bar{x}_2^i - \bar{x}_2$ are sufficiently small. Moreover, if $\bar{x}_2 = 1/N$ this holds true for any $\bar{x}_2^i$.

Subcase B We turn to the subcase where $\hat{\mu}_a < \mu_2$, so in equilibrium everybody must end up capacity constrained. Our candidate equilibrium looks as follows: Firm i prices as to sell all its capacity but anticipates that the marginal customer will be indifferent between purchasing a ticket from i or its competitors. Firms j on the other hand anticipate that a marginal customer located on segments connecting two firms of type j is indifferent between purchasing from that firm or not buying a ticket at all. This gives us the following equilibrium prices:

$$p_2 = \frac{v}{\alpha} + \frac{2\mu_2 - (N-1)N\bar{x}_2 - N\bar{x}_2^i}{2\alpha(N-2)}$$
$$p_2^i = \frac{v}{\alpha} + \frac{2(N-1)\mu_2 - (N-1)N\bar{x}_2 - (2N-3)N\bar{x}_2^i}{2\alpha(N-2)}$$

Note that this equilibrium requires that $\frac{1}{2}N\left(\bar{x}_2 + \bar{x}_2^i\right) = \hat{\mu}_b \geq \mu_2$ since otherwise, firms j are no longer affected by the prices set by firm i. This implied that we can only end up in this regime if $N \geq 3$. It is easy to check that all customers between firm i and a firm of type j purchase a ticket while some customers between two firms of type j abstain. Moreover, no airline has an incentive to increase prices since prices are close to the monopoly level.

Subcase C Finally, assume that $\hat{\mu}_b < \mu_2$. In this subcase even if all firms sell their full capacity, the marginal customer on each segment is indifferent between buying a ticket from a given firm or not buying any ticket at all and we end up in a standard monopoly setting where airlines set prices of

$$p_2 = \frac{v}{\alpha} - \frac{N\bar{x}_2}{2\alpha}$$
$$p_2^i = \frac{v}{\alpha} - \frac{N\bar{x}_2^i}{2\alpha}.$$

Returns to capacity in case $\bar{x}_2^i \geq \bar{x}_2$ Summing up our results so far, whenever firm i chooses a slightly larger capacity for the late market than everybody else, its expected profits are given by

$$\begin{aligned}\Pi_2^i(\bar{x}_2^i, \bar{x}_2) = &\int_0^{N\bar{x}_2} \frac{\mu^2}{N\alpha} f(\mu)\,\mathrm{d}\mu \\ &+ \int_{N\bar{x}_2}^{\hat{\mu}_a} \left[\frac{1}{\alpha}\left(N\mu - (N-1)N\bar{x}_2\right)\right](\mu - (N-1)\bar{x}_2)\, f(\mu)\,\mathrm{d}\mu \\ &+ \int_{\hat{\mu}_a}^{\hat{\mu}_b} \left[\frac{v}{\alpha} + \frac{2(N-1)\mu - (N-1)N\bar{x}_2 - (2N-3)N\bar{x}_2^i}{2\alpha(N-2)} - c\right] \bar{x}_2^i f(\mu)\,\mathrm{d}(\mu) \\ &+ \int_{\hat{\mu}_b}^{2} \left[\frac{1}{\alpha}\left(v - \frac{N\bar{x}_2^i}{2}\right) - c\right] \bar{x}_2^i f(\mu)\,\mathrm{d}\mu.\end{aligned}$$

and the marginal return to capacity is given by

$$\begin{aligned}\frac{\mathrm{d}\Pi_2^i(\bar{x}_2^i, \bar{x}_2)}{\mathrm{d}\bar{x}_2^i} = &-\frac{1}{N}\left(\frac{v}{\alpha} - \frac{(4N-1)\bar{x}_2^i - (N-1)\bar{x}_2}{2\alpha} - c\right) N\bar{x}_2^i f(\hat{\mu}_a) \\ &+ \int_{\hat{\mu}_a}^{\hat{\mu}_b} \left[\frac{v}{\alpha} + \frac{2(N-1)\mu - (N-1)N\bar{x}_2 - 2(2N-3)N\bar{x}_2^i}{2\alpha(N-2)} - c\right] f(\mu)\,\mathrm{d}\mu \\ &+ \int_{\hat{\mu}_b}^{2} \left[\frac{1}{\alpha}\left(v - N\bar{x}_2^i\right) - c\right] f(\mu)\,\mathrm{d}\mu.\end{aligned}$$

Evaluating equation (7) at $\bar{x}_2^i = \bar{x}_2$ gives us expression (3), so $\Pi_2^i(\bar{x}_2^i, \bar{x}_2)$ is differentiable at $\bar{x}_2^i = \bar{x}_2$. So in any symmetric equilibrium firms must set period 1 prices such that that all customers on the period 1 market buy a ticket, i.e. we get $\bar{x}_2 = 1/N$. This implies that

$$\begin{aligned}\frac{\mathrm{d}^2\Pi_2^i(\bar{x}_2^i,\bar{x}_2)}{\mathrm{d}\left(\bar{x}_2^i\right)^2} = &-\left(\frac{v}{\alpha}-\frac{2(4N-1)\bar{x}_2^i-(N-1)\bar{x}_2}{2\alpha}-c\right)f(\hat{\mu}_a)\\ &-\left(\frac{v}{\alpha}+\frac{(N-1)\bar{x}_2}{2\alpha}-\frac{2N\bar{x}_2^i}{\alpha}-\frac{\bar{x}_2^i}{\alpha(N-2)}-c\right)f(\hat{\mu}_a)\\ &-\frac{(N-1)N^2\bar{x}_2^i}{(N-2)4\alpha}f(\hat{\mu}_b)-\int_{\hat{\mu}_a}^{\hat{\mu}_b}\left[\frac{(2N-3)N}{(N-2)\alpha}\right]f(\mu)\,\mathrm{d}\mu-\int_{\hat{\mu}_b}^{2}\frac{N}{\alpha}f(\mu)\,\mathrm{d}\mu\end{aligned}$$

is negative for all $v > 4 + \alpha c$. So expected period 2 profits are concave in $\bar{x}_2^i$ over the interval $[0, 2]$ and it is sufficient to look at local deviations in the price p_1^i.

Uniqueness Again, let us show that the equilibria are unique for the respective values of μ_2. The proof follows along the same lines as in case $\bar{x}_2^i \leq \bar{x}_2$. In particular, firms that have the same capacity available must still charge the same price. In case aggregate supply exceeds aggregate demand, it follows directly that the equilibrium is unique. Instead, consider the case where $\mu_2 > \hat{\mu}_a$. Assume that firm i charges a higher price. This can only be optimal if firm i is still able to sell its full capacity in equilibrium. That requires firms j to set a strictly higher price, too. But if all firms charge strictly higher prices, some firms are unable to sell their full capacity and have an incentive to reduce prices. Instead, assume that firm i charges a lower price. If firm i is not to exceed its capacity, firms of type j have to set strictly smaller prices, too. But if firms of type j set prices smaller than p_2, firms sell more than their capacity, which can not be part of an equilibrium. Again, we can show that the prices charged by all firms of type j are uniquely determined, too.

B. Routes

Table 3. Routes

Origin	Destination	Origin	Destination	Origin	Destination
Aberdeen	Manchester†	Lisbon	Amsterdam‡	Paris	Dublin‡
Malaga	Madrid‡	London	Bordeaux†	Paris	Hamburg‡
Amsterdam	Barcelona†	London	Frankfurt	Paris	London
Amsterdam	Zurich‡	London	Hannover	Paris	Madrid†
Athens	Budapest‡	London	Prague‡	Paris	Marseille†
Athens	London‡	London	Sofia	Paris	Prague‡
Barcelona	Lyon‡	London	Zurich‡	Paris	Stockholm‡
Belgrade	Vienna‡	Liverpool	Amsterdam	Paris	Turin†
Berlin	Helsinki‡	Lyon	Madrid†	Paris	Valencia†
Berlin	Vienna	Madrid	Barcelona†	Paris	Warsaw‡
Bilbao	Paris‡	Madrid	Copenhagen‡	Palermo	Turin‡
Bologna	Madrid†	Madrid	Lisbon†	Prague	Helsinki‡
Bordeaux	Madrid†	Madrid	Milan‡	Prague	Milan†
Bordeaux	Nantes	Madrid	Stockholm‡	Prague	Rome†
Brussels	Leeds	Madrid	Valencia†	Rome	Nice
Brussels	London	Madrid	Zurich‡	Rome	Vienna‡
Budapest	Munich‡	Milan	Copenhagen‡	Stockholm	Berlin†
Bucharest	Milan	Milan	Duesseldorf‡	Stockholm	Duesseldorf†
Copenhagen	Geneva	Milan	Frankfurt‡	Stockholm	Oslo
Copenhagen	Helsinki‡	Milan	Lyon	Stuttgart	Milan†
Duesseldorf	Athens†	Milan	Paris	Strasbourg	Paris
Edinburgh	Manchester	Moscow	Budapest†	Toulouse	Brussels‡
Frankfurt	Innsbruck‡	Munich	Athens†	Toulouse	Paris‡
Frankfurt	Istanbul†	Munich	Madrid†	Vienna	Amsterdam
Frankfurt	Madrid†	Munich	Paris†	Vienna	Barcelona†
Frankfurt	Moscow‡	Munich	Vienna	Vienna	Frankfurt
Frankfurt	Paris	Naples	Milan‡	Vienna	Lyon†
Frankfurt	Toulouse†	Nice	Brussels‡	Vienna	Paris
Hannover	Amsterdam	Nuernberg	Amsterdam	Zurich	Frankfurt†
Hamburg	Warsaw‡	Oporto	Paris‡	Zurich	Mallorca†
Leipzig	Munich	Paris	Copenhagen‡		

Notes.— †marks routes in the subsample where the destination is one degree Celsius warmer than the origin; ‡marks routes in the subsample where the destination is one degree Celsius colder than the origin.

References

Armstrong, M. and J. Vickers (2001), "Competitive price discrimination," RAND Journal of Economics, 579–605.

Borenstein, S. (1985), "Price discrimination in free-entry markets," The RAND Journal of Economics, 380–397.

Borenstein, S. and N. Rose (1994), "Competition and Price Dispersion in the US Airline Industry," Journal of Political Economy, 102:653–683.

Brueckner, J., N. Dyer, and P. Spiller (1992), "Fare determination in airline hub-and-spoke networks," The Rand Journal of Economics, 309–333.

CAA (2011), "A survey of passengers at Birmingham, East Midlands, Gatwick, Heathrow, Luton, Manchester and Stansted in 2011," CAA Website.

Dana, J. (1998), "Advance-purchase discounts and price discrimination in competitive markets," Journal of Political Economy, 395–422.

——— (1999a), "Equilibrium price dispersion under demand uncertainty: the roles of costly capacity and market structure," The RAND Journal of Economics, 30:632–660.

——— (1999b), "Using yield management to shift demand when the peak time is unknown," The Rand Journal of Economics, 30:456–474.

——— (2001), "Monopoly price dispersion under demand uncertainty," International Economic Review, 42:649–670.

European Commission (2007), "Competition impact of airline code-share agreements: Final report," EC Website.

Fraport (2011), "Frankfurt Airport Air Traffic Statistics 2011," Company Website.

Gale, I. and T. Holmes (1992), "The efficiency of advance-purchase discounts in the presence of aggregate demand uncertainty," International Journal of Industrial Organization, 10:413–437.

——— (1993), "Advance-purchase discounts and monopoly allocation of capacity," The American Economic Review, 83:135–146.

Goolsbee, A. and C. Syverson (2008), "How do incumbents respond to the threat of entry? Evidence from the major airlines," The Quarterly Journal of Economics, 123:1611–1633.

Holmes, T. (1989), "The effects of third-degree price discrimination in oligopoly," The American Economic Review, 79:244–250.

Kreps, D. and J. Scheinkman (1983), "Quantity precommitment and Bertrand competition yield Cournot outcomes," The Bell Journal of Economics, 14:326–337.

Maggi, G. (1996), "Strategic trade policies with endogenous mode of competition,"

The American Economic Review, 237–258.
Morrison, S. and C. Winston (1990), "The dynamics of airline pricing and competition," The American Economic Review, 389–393.
Nocke, V., M. Peitz, and F. Rosar (2010), "Advance-purchase discounts as a price discrimination device," Journal of Economic Theory.
Puller, S., A. Sengupta, and S. Wiggins (2008), "Testing theories of scarcity pricing and price dispersion in the airline industry," Workig paper.
Stavins, J. (2001), "Price discrimination in the airline market: The effect of market concentration," Review of Economics and Statistics, 83:200–202.
Stole, L. (2007), "Price discrimination and competition," Handbook of Industrial Organization, 3:2221–2299.
Thompson, H., O. Renard, and J. Wright (2007), "Mobile termination," Access Pricing: Theory and Practice, Amsterdam: Elsevier.

CURRICULUM VITAE

Personal

Date of birth:	March 28, 1983
Nationality:	German

Academic Positions

2013–	Assistant Professor, Toulouse School of Economics

Education

2013	Ph.D. in Economics, University of Munich
Fall 2012	Visiting Ph.D. Student, Northwestern University
2010–2011	Visiting Ph.D. Student, Northwestern University
2007	Diploma in Economics, University of Munich
2002–2005	Undergraduate studies in Economics and Political science, University of Tübingen

Fields of Interest

Primary:	Macroeconomics, Theory
Secondary:	Contract Theory, Political Economics

Presentations

Seminars

Chicago Fed (2013), NYU Stern (2013), University of Warwick (2013), IIES Stockholm (2013), Sciences Po (2013), Toulouse School of Economics (2013), Università Bocconi (2013), EIEF Rome (2013), Universitat Pompeu Fabra (2013), University of Mannheim (2013), Northwestern University (2010, 2012), Kellogg School of Management (2010), University of Munich (2009–2012).

Conferences

TIGER Forum on Information Processing in Macroeconomics and Finance (Toulouse, 2013), Warwick Conference on Financial Markets and the Real Economy (Venice, 2013), EDGE Jamboree (Munich, 2012), SFB TR 15 Workshops (2009, 2010).

Teaching

- Contract Theory (Ph.D. level), University of Munich (teaching assistant), 2010, 2012
- Advanced Game Theory (Ph.D. level), University of Munich (teaching assistant), 2011
- Managerial economics (undergraduate), University of Munich (teaching assistant), 2012
- Game Theory (undergraduate), University of Munich (teaching assistant), 2010
- Microeconomics (undergraduate), University of Munich (teaching assistant), 2008–2010, 2013

ROBERT ULBRICHT
Munich, Germany
May 2013