UNIVERSITY OF CALIFORNIA,
IRVINE

Firms, Organization, and Capital Market Frictions

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Economics

by

Kurt James Horner

Dissertation Committee:
Guillaume Rocheteau, Chair
Eric Swanson
Michael McBride

2019

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGMENTS

I would like to thank the many faculty and grad students who contributed advice, encouragement and critique during the development of this dissertation. Specific thanks are due to committee members Guillaume Rocheteau, Eric Swanson and Michael McBride, particularly Guillaume and Eric for chapter 3, and Michael for chapters 1 and 2.

For chapter 1, specific thanks are due to Michael McBride, Jiawei Chen, and Linda Cohen on the UCI faculty for advice both technical and conceptual.

For chapter 2, specific thanks are due to Michael McBride, Stergios Skaperdas, Jean-Paul Carvalho on the UCI faculty for advice both technical and conceptual. I would also like to thank the attendees at the 2018 ASREC conference, and in particular Laurence Iannaccone and my discussant Alex Ray. Chapter 2 also benefited from discussions with and commentary from recent graduate students Kara Dimitruk, Aria Golestani, Zachary Schaller, Nishtha Sharma, Patrick Testa, Marshall Urias, and Cole Williams, as well as all of the attendees at Irvine's Theory, History, and Political Economy workshop for invaluable feedback during presentations.

For chapter 3, I would also like to thank Randall Wright and Pablo Ottonello for encouragement and helpful correspondence, as well as Michael Choi from UCI. Chapter 3 also benefited from discussions with and commentary from recent graduate students and postdocs Paul Jackson, Lucie Lebeau, Florian Madison, Mario Silva, and Marshall Urias, as well as all of the attendees at Irvine's weekly Macro workshop, too numerous to list, who have given feedback along the way.

CURRICULUM VITAE

Kurt James Horner

EDUCATION

Doctor of Philosophy in Economics **2019**
University of California *Irvine*

Bachelor of Science in Architectural Engineering **2000**
California Polytechnic State University *San Luis Obispo*

TEACHING EXPERIENCE

Teaching Assistant **2013–2019**
University of California *Irvine*

Learning Assitance Pilot Program **2015–2016**
University of California *Irvine*

Instructor-of-Record, Global Economy **Summer 2018**
University of California *Irvine*

WORK EXPERIENCE

Structural Engineer **2000–2010**
Taylor & Syfan Consulting Engineers *San Luis Obispo, CA*

PAPERS

Smaller is Better? The Implications of Organizational Form **2019**
Working Paper

Norm Enforcement in Faith Communities **2019**
Working Paper

Capital Reallocation, Search-in-Use, and Growth **2019**
Working Paper

CONFERENCE PRESENTATIONS

Capital Reallocation, Search-in-Use, and Growth **Dec 2017**
Southwest Search and Matching Conference

Norm Enforcement in Faith Communities **Mar 2018**
Association for the Study of Religion, Economics, and Culture

ABSTRACT OF THE DISSERTATION

Firms, Organization, and Capital Market Frictions

By

Kurt James Horner

Doctor of Philosophy in Economics

University of California, Irvine, 2019

Guillaume Rocheteau, Chair

Firms and other organizations are typically modeled in economics as a function that combines some set of inputs to produce desired output. Such organizations, in the real world, are often constrained relative to such a simple model by frictions in output markets, frictions in the input markets, and internal frictions due to sub-optimal organization. The first of these frictions (output markets) are perhaps the best explored in the literature. As such, my work seeks to add to the literature exploring the latter two issues. The first chapter of this dissertation examines firm organizational structure as a coordination game with private information, and demonstrates conditions where joint ventures between firms should be favored over mergers. The second chapter examines religious organizations using a club good model where members have norm-based utility functions, and finds a large variety of equilibirium organizational models. The final chapter is a search model of physical capital markets, allowing for search-in-use (an analog to search-on-the-job in macro-labor models), which shows two channels by which capital market frictions can diminish growth. Together, the papers represent the beginning of a research path examining previously neglected constraints on economic growth.

Chapter 1

Smaller is Better?: The Implications of Organizational Form

Coordination decisions between firms with complementary products are examined while allowing for managers to choose among organizational forms. Due to private information about coordination costs, contracts (joint-venture) can cause coordination failure when coordination is first best, while merger can cause coordination when coordination is not first best. Extensions where an organizational form is banned or made the conventional form are explored. Decisions are more likely to be first best when a convention is imposed, and a convention that favors contracts over mergers is most optimal.

1.1 Introduction

It is common to treat the firm as simply a production function that converts inputs to outputs. The internal organization of the firm is often taken as a given, when such matters are discussed at all. Yet, at a most fundamental level, a large firm with multiple divisions or product lines can be thought of as a bunch of smaller firms placed under a common organization. We need an explanation for why this organization might be chosen over a contractual relationship between separate firms. To the extent that models discuss whether firms will merge or remain separate, it is often assumed that one decision or the other is always more optimal (typically merging). While various attempts have been made to explain organizational form, the explanations vary widely and there is no settled consensus. This paper attempts to bring us closer to an eventual consensus by examining how organizational choices are made.

The model framework to be used here is similar to that from Hart and Holmstrom (2010). A key conclusion of their model is that both integration (merger) and non-integration (contract) involve deviations from the first best social optimum. If two firms that are considering coordinating [1] try to remain separate, they will sometimes miss opportunities to successfully coordinate. Yet, if they merge, they will now sometimes coordinate when they should not do so. Two sets of questions naturally arise. First, how do managers come to agree upon an organizational form and what factors influence that outcome? Second, can policy measures be taken that bring overall coordination decisions closer to the first best?

This paper describes a model of two managers considering having their two firms coordinate. The managers can have their firms remain *independent*, can draw up a *contract* that shifts profits between the firms, or they can *merge* their firms into a single firm. The model is simulated over a range of parameter values and for a variety of alternative rules governing

[1] Coordination could mean a joint-venture on a new product or improving complementarity between their existing products. This is not to be confused with collusion between firms producing substitutable products, since that would imply that the firms' gains come at the expense of consumer welfare.

the availability of the organizational forms. These alternative rules are compared to the first best outcome and to each other.

There are several main results of interest. First, that biasing the managers towards either mergers or contracts, improves outcomes relative to not doing so. I will refer to this institutional bias as a *convention* in favor of one organizational form or the other. Second, that having an environment where both mergers and contracts can occur is at least as good as banning one organizational form or the other, and usually better. Finally, that the best outcomes occur under a convention biased toward contracts.

This paper is inspired in part by the experience of craft industries in Emilia-Romagna, Italy. This region, one of Italy's 22 administrative districts, was one of the poorest regions in Italy just after the Second World War. The region now has the highest per capita income in all of Italy, a third higher than the EU average. Emilia-Romagna is unusual for having a manufacturing heavy economy with very small firms, an outcome which would seem to contradict a great deal of the economic literature about firm size. The region's unusual business environment could be explained by their policy of providing public support for small businesses (particularly new firms) funded by local tax revenues. Alternatively, some have suggested that the network of small firms in the region is founded in local culture. Regardless of the cause of this organizational convention, the model in this paper presents provides a potential explanation for how this environment might be responsible for the region's economic success.

1.2 Literature

The unique contribution of this paper is to explore the choice of organizational form. In particular, the firm managers in my model will end up with one of three decision rules to handle a coordination problem, which correspond to independent, contractual and merged organizational forms. While comparisons between organizational forms have been done,

both empirically and theoretically, how managers come to agree upon an organization has not been well explored. This literature review is broken up into three subsections.

First, I review empirical literature on organizational form to clarify the phenomena described by the model. There is a vast literature on merger policy, not explored here, which focuses on the effect of mergers on the consumer. However, this paper is focused on the managers of firms and how they come to choose an organization, which means exploring more obscure aspects of the literature. Due to the economic literature tending to be focused on mergers, much of the empirical literature on networked business environments (resembling the contract form in my model) is drawn from sociology.

Second, I survey theories of the firm and how that relates to organizational choice. Any theory of why firms exist should also be able to hint at why mergers would occur, or why they might not occur when a viable contractual option is available. The three organizational forms in this paper also bear a resemblance to the market / network / hierarchy trichotomy popular in the sociology literature.

Finally, I review recent work related to organizational choice. Readers who want to reference the most direct influences on the model environment should turn to that subsection.

1.2.1 Empirical

A review of the merger literature can be found in Andrade et al. (2001). It is widely known that mergers tend to occur in waves, and that these waves seem to be triggered by shocks to industry structure (major cost changes, either regulatory or resource driven). The primary equilibrium in my model is consistent with this phenomenon, in that firm managers determine their desired organizational form based on an assessment of their own cost structure. While not directly explored in the model, a systematic change in costs could trigger a different organizational outcome.

Davidson and Ferrett (2007) suggests that mergers can occur to exploit complementarities in R&D costs. The model below assumes that coordination generates additional revenue due to complementary products. The exact nature of how those revenues are generated is not explored, and the model would be consistent with an R&D based motivation for coordination. Mergers are not always successful, a situation explored in Weber and Camerer (2003) using an experiment involving distinct corporate cultures. The model below features private costs that cannot be negotiated over, which can lead to mergers creating a net loss for one or both parties.

The relative desirability of mergers versus contractual relationships is central to this paper. Gugler and Siebert (2007) uses evidence from the semiconductor industry to show that joint ventures are not necessarily any less efficient than mergers. Given the potential market power effects of mergers, they conclude that joint ventures might be preferable to mergers. Similarly, Jandik and Kali (2009) provides evidence that as the quality of legal systems increases, firms tend to prefer joint ventures to mergers.

There is a significant sociological literature contrasting contractual versus hierarchical business environments. Networks between firms in the New York City garment industry are explored in Uzzi (1997), showing a large number of firms interacting contractually rather than merging. Similarly, Lazerson (1988) conducted a survey of small manufacturing firms in the Emilia-Romagna region of Italy, again showing a stable contractual environment. The American computer industry was examined in Saxenian (1996), and concluded that loose networks in California's Silicon Valley outperformed the merger friendly environment along Route 128 in Massachusetts.

We should hesitate to advocate contractual arrangements in all cases. Schrank and Whitford (2011) document numerous cases where networked economies seem to underperform, primarily where actors with low degrees of competency transmit poor methods through the network leading to increased costs. The results from the model in this paper do indicate

that simply banning mergers as an organizational option would not be optimal, even if there are broad advantages from promoting contractual arrangements.

1.2.2 Theoretical

In the broadest terms, theories of the firm break down into four categories, shaped by the relative emphasis of two dichotomies. The first dichotomy is the control vs. transaction costs distinction, where the firm is described alternately as either a structure that facilitates decision authority (control) or one that allows the participants to overcome transaction costs from using markets. The second dichotomy is whether or not the firm arises from ex-ante incentive alignment or a need to have ex-post decision governance. The typology above is strongly informed by Gibbons (2005).

In the case of control theories, the seminal work is Knight (1921), where the firm is described as a mechanism for an entrepreneur to bear the risks while reaping the rewards of a successful project. In this view, the employees trade their potential gains for a consistent wage. When applied to mergers, this would suggest that firms merge in order to create a clear decision-maker that can override individual managers (prevent hold-up).

For transaction cost theories, the key paper is Coase (1937), where the non-market structure of the firm overcomes transaction costs that would be too burdensome in a market. The firm is, in this view, an equilibrium between tasks that can be handled by markets and those best handled by command. In the realm of mergers, this implies that the costs of merging are weighed against the cost of coordinating separately and that the lower cost option wins out. In a Coase framework the trade offs that are resolved by firms involve different types of costs, while in a Knight framework, firms resolve differences in time preference between agents.

For ex-ante incentive theories, the firm's organization is chosen in order to ensure that the

intended projects are carried out, even when some people in the organization are reluctant. In Holmstrom and Tirole (1991), ownership of key assets confers the incentive to make the project successful, and in turn induces the owners to pay employees enough to ensure success. In Grossman and Hart (1986), ex-ante bargaining leads to property-rights in the firm accruing to the largest investor. In the context of coordination decisions, the ex-ante incentive framework would suggest that organizational choice will be driven by the party that gains most from coordinating trying to maximize the chance of coordinating.

For ex-post decision theories, the firm is a mechanism to overcome barriers to coordination where its governance structure is simply what is needed for proper decisions after the firm is formed. In Williamson (1971), the manager's fiat removes rent-seeking by subordinates. In Klein (1996), firms are created as an adaptation to overcome uncertainty. When talking about merger decisions, the ex-post framework would imply that mergers occur when needed to ensure coordination (implying high uncertainty as to what the individual managers will do once costs are revealed).

It is also worth noting a sort of anti-theory of the firm, most strongly delineated in Alchian and Demsetz (1972). Here the firm is just a contract by another name and organizational form is considered chimerical. In practice, this view places a padlock on the "black box" of the firm, preventing further examination. The insights on the role of monitoring in Alchian and Demsetz (1972) are quite interesting, but their sweeping dismissal of organizational concerns perhaps goes too far.

Finally, the model environment bears some similarities to the market / network / hierarchy typology from the sociological literature first described in Powell (1990). In particular, the contract organizational form in this paper can be viewed as analogous to a "network" environment. As noted in the empirical review above, sociologists have drawn attention to thriving regions that tend to organize collaboratively rather than by mergers. It is hoped that this paper will encourage greater attention to such environments by economists as well.

My model will follow in the ex-ante incentive tradition, by depicting organizational choice as a means to incentivize coordination when the realization of costs is not yet known. I also allow organizational forms to vary in the types of errors they are prone to, which resembles the transaction costs framework. However, given that my organizational forms also vary on the basis of who retains the final decision to coordinate or not, the model is also consistent with the control framework. I will also examine whether it is helpful to set a conventional organizational form (whether chosen by policy or culture), and whether this brings decisions closer to the first best.

1.2.3 Recent Work on Organizations

As noted above, this paper borrows much from Hart and Holmstrom (2010), which unsurprisingly falls in the ex-ante incentive camp, given the authors. In their paper, the actual decision of whether to merge or use contracts is not considered. Instead the focus is on how likely a particular organizational form will be to cause over or under coordination relative to the first best. An extension of that paper, Legros and Newman (2013), does consider organizational decisions and finds that not only does this cause cusps to occur in the supply curve, but that for a range of consumer demand, a market would even see a mixture of merged and unmerged firms. This opens up the potential of explaining the distribution of firm size in the economy.

Hart and Holmstrom (2010) is also notable for being relatively agnostic along the control vs. contract axis. While their depiction of organizational decisions are reminiscent of the power/certainty trade-off from Knight (1921), the fact that integration faces unique drawbacks (over-coordination) that are the reverse of non-integration (under-coordination) is more reminiscent of the Coasian view, where markets and hierarchies have offsetting bene-

fits and flaws.

Another recent paper that informs this one is Van den Steen (2010), which comes down firmly on the side of the Knightian control perspective. In that paper, the firm is one of two possible equilibria when determining whether two agents should work together on a project. The model here has some similarities to Hart and Holmstrom, in that there are two agents whose actions are distinct but where a gain from coordinating can occur, and the question is whether and how coordination should occur. The main distinction is that Van den Steen is essentially only considering half of the outcome space that Hart and Holmstrom consider (the half where coordination is socially optimal). This leads to a conclusion that the firm solves the coordination problem without delving into possible drawbacks.

The interpretation of a key parameter in my model will draw from Hart and Moore (2008) which describes how much of our organizational arrangements likely stem from a need to define beforehand what sorts of payments can be made between parties. I will refer to the costs that can be described in an agreement as being "contractible." The fact that parties can rarely, if ever, describe all of their costs causes sub-optimal decisions to be made and there is no reason to assume that such errors can be easily re-negotiated once an organizational form has been adopted.

Although the model in this paper takes the revenue gain from coordination as a given, the horizontal merger model in Huck et al. (2004) is relevant. The authors use organizational structure to explain mergers. Their merged firms enforce a Stackelberg style order of production, not unlike how my merged firm commands coordination. Going outside of economics to political science, Jung and Lake (2011) uses a simulation model to investigate the effectiveness of organizational hierarchy. Their main result is hierarchy provides a certainty of cooperation that can override other costs of hierarchy, which resembles the bias toward coordination described in Hart and Holmstrom (2010) that also appears in the model below.

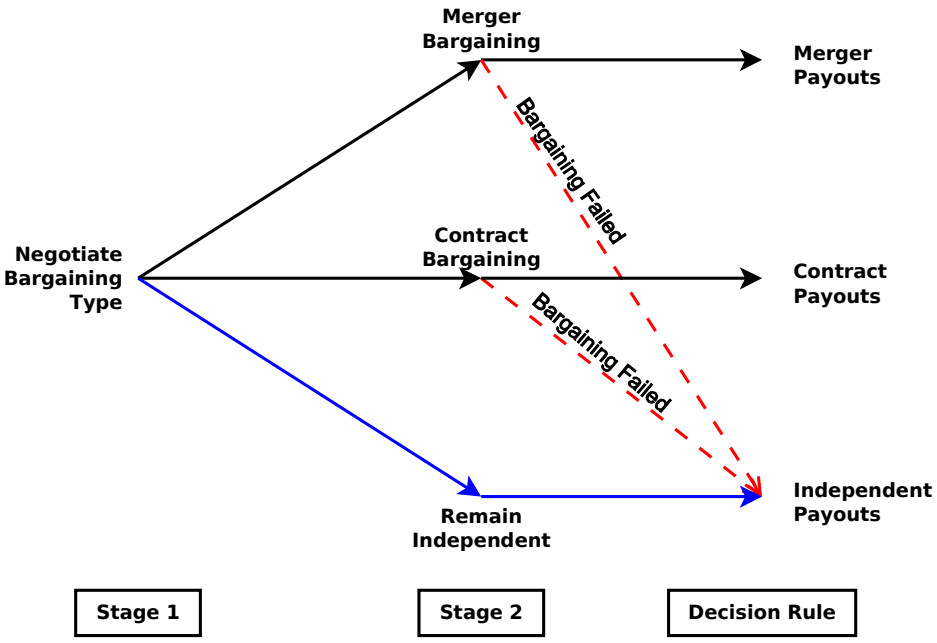

Figure 1.1: Game Diagram

1.3 Model

There are two firms (A and B) that can coordinate their operations or not. By coordinating, the two firms' products are more likely to be purchased together causing both firms to enjoy increased sales. Each firm has a manager who gains profits from the firm but also may incur private costs from coordinating. The managers will attempt to use one of two organizational alternatives to maximize the chance of successful coordination.

To determine organizational form, the two managers will play a two stage game that will result in a particular organizational form whose decision rule will then be implemented and payoffs realized. The structure of the game is depicted in figure 1.1. The managers have two possible agreements (contract or merger) that can improve the chances of successful coordination, relative to remaining independent. Under independence, coordination occurs if both managers expect their individual payouts to be non-negative. The same is true under contract, except that the two managers bargain over a share of the increased profits that will

occur under coordination and make their decision based on the adjusted payouts. Finally, under a merger, a *headquarters* is created to make the coordination decision, which has the ability to impose coordination on the two managers. As with a contract, the managers bargain over profit shares should coordination occur under merger. The bargaining and the decision rules will be described in detail below.

1.3.1 Managers

Each firm has a manager who maximizes utility of the form:

$$U_i = \pi_i + x_i \tag{1.1}$$

- π_i are the profits realized by manager i
- x_i are private costs borne solely by manager i (negative)

Focusing on the coordination decision, managers maximize ΔU_i which is the difference between the manager's utility without coordinating and the manager's utility when the two firms coordinate (however they do so). This has the convenience of setting $\Delta U_i = 0$ when coordination does not occur, and it is not necessary to define the profits and private costs to the managers under normal operations. The payouts of the game are in terms of ΔU_i.

When coordination occurs, there will be an increase in revenues and costs per:

$$z_i = r_i + c_i \tag{1.2}$$

- z_i is manager i's net gain from coordinating with other firm
- r_i is the change in manager i's revenue when coordinating (always non-negative) [2]

[2]We could allow for negative revenue values, but since coordination disincentives already arise from the negative costs, negative revenue is redundant and complicates the analysis.

- c_i is the change in manager i's costs when coordinating (always negative) [3]

The nature of the costs is best thought of as increased costs for that firm that only occur in the coordination case. These could be organizational restructuring costs, training costs, unique materials or even new equipment. These costs can also include psychic or interpersonal costs that are difficult to enumerate. I also refer to z_i as a net *gain* rather than profit, to emphasize that a coordination decision that is profitable in an accounting sense may still be a net loss in this model.

The reader may notice that c_i seems to include x_i. Specifically, define $\Delta x_i = (1-\theta)c_i$ where $\theta \in (0,1)$. This parameter[4] represents the degree to which the costs of coordinating are *contractible*. It is likely that at least some of the costs of coordinating (particularly psychic costs) cannot be sufficiently enumerated to allow the parties to craft an agreement that mitigates them. The degree of contractability is assumed to be symmetric – that there are no disparities in bargaining power or monitoring ability. This assumption will be relaxed later.

The contractability parameter (θ) will assist in interpreting the results. For example, if $\theta = 1$, this would be a situation where all costs associated with coordination can be negotiated and the term x_i would vanish from the managers' utility functions causing them to make their decisions purely on the basis of realized profits. Similarly, if $\theta = 0$, this would be a situation where all of the costs of coordinating are private costs borne by the managers that cannot be negotiated over.

Since a manager's private costs are $(1-\theta)c_i$, the contractible profits would be $r_i + \theta c_i$ which

[3]The descriptions of the variables here differs from Hart and Holmstrom, particularly in that costs are now *all* forms of cost, pecuniary or otherwise. Their model would be equivalent to $r_i = \pi_i$ and $c_i = x_i$.

[4]Conceptually, this is similar to the contractability parameter in Hart and Moore (2008).

is of course, equal to π_i. Thus, a manager's change in utility from coordinating would be:

$$\Delta U_i = \underbrace{r_i + \theta c_i}_{\pi_i} + \underbrace{(1-\theta)c_i}_{x_i} \tag{1.3}$$

Given the above, the first best social optimum will be to coordinate if $z_A + z_B \geq 0$. Even if one of the two managers would take a loss from coordinating, as long as the positive z is of larger magnitude than that for the manager which would take a loss, coordination is socially optimal.

1.3.2 First Stage

At the start of the game, the revenue gains from coordinating (r_i), the *distribution* of costs ($F(c)$) and the contractability parameter (θ) are all common knowledge. The individual realizations of the costs (c_i) are, however, private information known by each manager. Payouts will be based on the realized state of the world, which is defined by a *pair* of cost draws $\{c_A, c_B\}$. Since each manager knows only their own cost draw, the state of the world is not known until after organizational form is chosen.

The first stage of the game has two players (the managers), each with a strategy that maps the cost distribution to their indicated choice for entering contract or merger bargaining ($S_i : F(c_i) \rightarrow \{Contract, Merger\}$) and payouts that are dependent in part on the outcome of the bargaining. The two managers' indicated preferences are compared to determine which second stage node they end up at. The baseline game grid for stage one is: If the two managers agree on which form to bargain over, they enter that node of the overall game. If not, both firms remain independent. The next three sections describe the second stage nodes of the overall game.

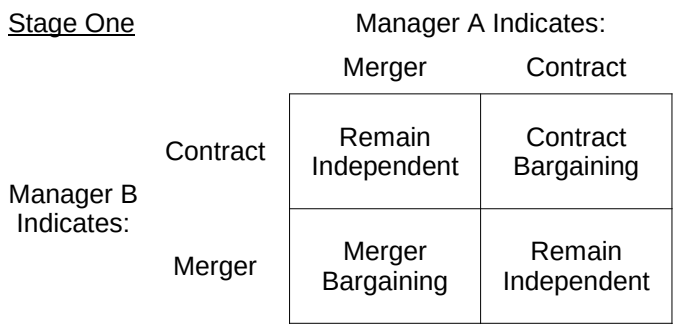

Figure 1.2: Baseline Game Grid

1.3.3 Independence

Under independence, the two managers make no formal agreements about coordinating. If they both decide to do so, coordination occurs, but if either one opts out, coordination does not occur. Thus, the two managers will coordinate if:

$$z_A \geq 0 \quad and \quad z_B \geq 0 \tag{1.4}$$

As noted above, under the first best, either of z_A or z_B can be negative as long as the magnitude of the positive z is larger. In this case, however, if *either* z is negative, the manager in question will choose not to coordinate. This means that for certain values of the parameters, coordination will be first best, but fail to occur.

The reverse is not true, since any case where coordination is *not* first best necessarily involves one or both z values being negative, which would never result in coordination given non-integration. Thus, remaining independent can cause coordination failure to occur even when it would be first best to coordinate, but it never causes coordination when such is not the first best.

1.3.4 Contract

In this case, the firms still do not merge, but the managers negotiate a contract to govern their coordination decision. The managers bargain over shares of the profit gain from coordinating with each other. As under independence, if both agree to coordinate under the terms of the contract, coordination occurs. Otherwise, coordination fails and gains are zero.[5]

The contractability parameter ($\theta \in (0,1)$) now becomes important, since the profit sharing agreement can only take into account the contractible profits, The private costs of each manager will still be felt if coordination occurs.

The contractible profits for each manager would be:

$$z_{\theta i} = r_i + \theta c_i \tag{1.5}$$

The managers then bargain over a share of joint contractible profits (γ):

$$\max_{\gamma \in [0,1]} \left(\sum_{s=1}^{N} p_s (U_{As} - d_{As}) \right) \left(\sum_{s=1}^{N} p_s (U_{Bs} - d_{Bs}) \right) \tag{1.6}$$

s.t. constraint that both sums are positive [6]

- $U_{As} = \begin{cases} \gamma(z_{\theta A} + z_{\theta B}) + (1-\theta)c_A & if\ U_{As} \geq 0\ \&\ U_{Bs} \geq 0 \\ 0 & otherwise \end{cases}$

[5] Another departure from Hart and Holmstrom is the use of profit shares to motivate coordination rather than imposing a shading cost. This is done for two reasons. First, shading costs would cause the independent case to be strictly dominant over the other organizational forms and second, having deadweight losses from coordination failure is not essential to the results that follow.

[6] Any value of γ that created a negative expected surplus would not be accepted by that manager. Without this constraint two negative sums could create a corner solution. Given the fact that private costs $(1-\theta)c_i$ are negative for any $\theta \neq 1$, a γ value of zero or one would always be worse than independence for one of the two managers. Corner solutions are thus not incentive compatible.

- $U_{Bs} = \begin{cases} (1-\gamma)(z_{\theta A} + z_{\theta B}) + (1-\theta)c_B & if\ U_{As} \geq 0\ \&\ U_{Bs} \geq 0 \\ 0 & otherwise \end{cases}$
- d_{As} = Payout to A in state s under Independence
- d_{Bs} = Payout to B in state s under Independence
- γ is the share of joint contractible profits directed to manager A
- N is the number of possible states
- p_s is the probability of state s occurring

If either manager determines that their expected payout is higher under independence, then bargaining will fail and independence will result. The above bargaining problem is a generalization of Nash bargaining, where the solution is *as if* a neutral third party proposed a profit share based only on the common knowledge available to both managers. An alternative interpretation is that the two managers bargain without revealing any information about their costs other than what has already been communicated by their play in stage one. If a manager did reveal such information, it would only strengthen the bargaining position of the other manager by giving them asymmetric information. A more detailed model of the bargaining process is avoided since the complexities of contract bargaining are not the focus of this paper.

The existence of a contract increases the number of cases where coordination occurs. As θ increases, coordination is more likely to occur since the private costs of each manager are less likely to have greater magnitude than their profit share. On the other hand, over-coordination will still be avoided since both managers retain the power to veto coordination if their payout would turn out to be negative. The contract case collapses to the first best at $\theta = 1$. At this value of θ, the surplus of both managers will be positive whenever coordination *is* first best, and for no other cases. Thus, a contract where $\theta = 1$ would perfectly

resolve the coordination problem.

1.3.5 Merger

In this case, the firms combine into a single firm and create a *headquarters* which makes the decision whether to coordinate. In this scenario, headquarters makes the coordination decision based on joint contractible profits (refer to the definition in the previous subsection). If private costs are sufficiently high, this can result in negative payouts to both managers.

Headquarters will compel coordination if:

$$z_{\theta A} + z_{\theta B} \geq 0 \tag{1.7}$$

This results in a payout to manager A of $\phi(z_{\theta A} + z_{\theta B}) + (1-\theta)c_A$ and a payout to manager B of $(1-\phi)(z_{\theta A} + z_{\theta B}) + (1-\theta)c_B$ where ϕ is a bargained profit share. Due to the fact that costs (c) are negative, it is always true that:

$$z_{\theta A} + z_{\theta B} \geq z_A + z_B \tag{1.8}$$

which means that a merged firm will always coordinate when that outcome is first best.

Profit shares (ϕ) are determined via:

$$\max_{\phi \in [0,1]} \left(\sum_{s=1}^{N} p_s (U_{As} - d_{As}) \right) \left(\sum_{s=1}^{N} p_s (U_{Bs} - d_{Bs}) \right) \tag{1.9}$$

s.t. constraint that both sums are positive

- $U_{As} = \begin{cases} \phi(z_{\theta A} + z_{\theta B}) + (1-\theta)c_A & if\ z_{\theta A} + z_{\theta B} \geq 0 \\ 0 & otherwise \end{cases}$

- $U_{Bs} = \begin{cases} (1-\phi)(z_{\theta A} + z_{\theta B}) + (1-\theta)c_B & if\ z_{\theta A} + z_{\theta B} \geq 0 \\ 0 & otherwise \end{cases}$
- d_{As} = Payout to A in state s under Independence
- d_{Bs} = Payout to B in state s under Independence
- ϕ is share of contractible profits paid to manager A
- N is the number of possible states
- p_s is the probability of state s occurring

As with contract bargaining, this is a generalized Nash solution that uses the common knowledge available to both managers. The only significant difference between contracts and mergers in this model is the decision rule used to determine whether coordination occurs. This ensures that any differences in outcomes are driven by differences in the way in which the two organizational types make decisions and *not* by differences in the payout structure.

Note that since headquarters is only concerned about costs that are contractible, it will sometimes choose coordination when it is *not* first best to do so (over-coordination). The fact that mergers remove under-coordination only to cause over-coordination, creates trade offs between the organizational forms.

A key assumption above is that headquarters does not receive any of the profits resulting from coordination. If headquarters receives a fixed salary regardless of their decision, or if headquarters is simply a name for the joint decision making process of the former managers, then the above payouts would be correct. If we consider headquarters to be a separate agent that also receives some share of the profits, both the decision inequality and the payouts would have to be changed to reflect compensation to headquarters when coordination failed. However, since paying headquarters would reduce payouts to the managers, the case where

headquarters does not receive any of the gain from coordinating represents a best case merger scenario, which would be most likely to be chosen by the managers and most likely to result in positive payouts.

1.3.6 Solution Concept

Each manager is seeking to maximize their utility. By backward induction, they determine expected payouts from each decision rule given the expected bargaining outcome, and then determine the choice in stage one that would get them to the desired bargaining game.

Before drawing their costs (c_i), each manager determines when they will submit contract or merger bargaining depending on the realization of their costs. The obvious strategy type is a cutoff strategy where for low cost draws they submit contract or merger and for high cost draws they submit the opposite form. For purposes of clear description, I will for the remainder of this paper define manager B as being willing to coordinate regardless of their cost draw, while manager A is sometimes willing and sometimes not. Put another way, the revenue gain for B (r_B) is larger in magnitude than their highest possible cost. I will sometimes refer to manager B as the *enthusiastic* manager, and manager A as the *reluctant* manager, reflecting their relative attitudes toward coordination.

In order to properly calculate the bargaining games, the managers will need to be aware of how the set of possible states has shrunk given the indicated bargaining preference of the other manager. This would allow them in turn to calculate expectations in the first stage and develop a strategy. There are two ways in which this can be achieved, either of which are sufficient. We can either assume equilibrium knowledge on the part of both managers or have the managers disclose their strategy at the end of stage one. Either of these will result in the two managers being able to agree on which states are still possible to arrive at when they enter bargaining.

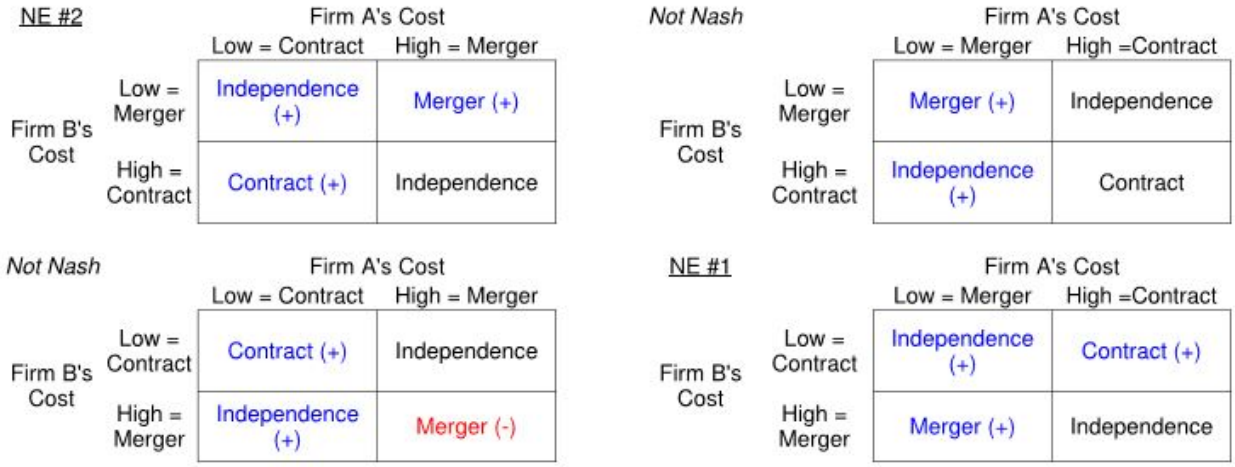

Figure 1.3: Cutoff Strategy Comparisons

For the baseline case, whenever the two managers submit different choices for which bargaining game to enter they simply remain independent. In figure 1.3 the various strategy combinations are depicted. Note that the Nash equilibria involve the two managers playing the opposite cutoff strategies (e.g. Contract corresponds to low costs for one firm and high costs for the other). Negative payouts can occur if a merger occurs when both managers have a high cost draw. The negative payouts are more likely for the reluctant manager (A) than for the other, which makes Nash equilibrium one more plausible to arrive at than equilibrium two (since the payouts for A are always higher). In the simulation results both equilibria are calculated, although equilibrium one weakly Pareto dominates the other. As with any cutoff strategy, the two managers will choose a cutoff such that the expected value on either side of the cutoff is equal (or as close to equal as possible when using a discrete cost space as used in the simulation results). Based on the chosen strategies, the range of possible costs is narrowed to one of the four quadrants as shown in figure 1.3. This means that the bargaining games will only cover a subset of the overall cost space.

One might consider an equilibrium strategy could be obtained with a more complex mapping than a cutoff. However, since expected payouts in all organizational forms are weakly monotonically increasing as c_i decreases, this is highly unlikely. Even if there were multiple

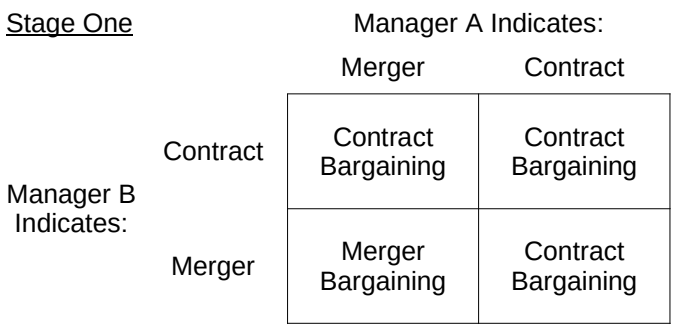

Figure 1.4: Contract Convention

regions where a manager chose merger, the expected cost when choosing merger would still have to be higher or lower than the expected cost when choosing contract, or otherwise their choice in stage one would have no effect on the bargaining. However, the fact that cutoff strategies do result in equilibria is sufficient to proceed.

1.3.7 Alternative Rules for First Stage

Per figure 1.2, when the two managers disagree on which type of bargaining game to enter, they remain independent. But one can easily imagine a legal or cultural environment where either contracts or mergers are the conventional means of solving a coordination problem. As such, in the results that follow, I will also consider both a contract convention (see figure 1.4) and a merger convention (figure 1.5). With a contract convention, the two managers enter merger bargaining if they both indicate as such, and they end up in contract bargaining otherwise. Note that this removes the lower path from figure 1.1. Independence only occurs in the case of failed bargaining. With a merger convention, the two managers enter contract bargaining if they both indicate as such, and they end up in merger bargaining otherwise. As with the other convention, this removes the lower path from figure 1.1.

Once implemented, these conventions alter the outcomes in the Nash equilibria. Note that

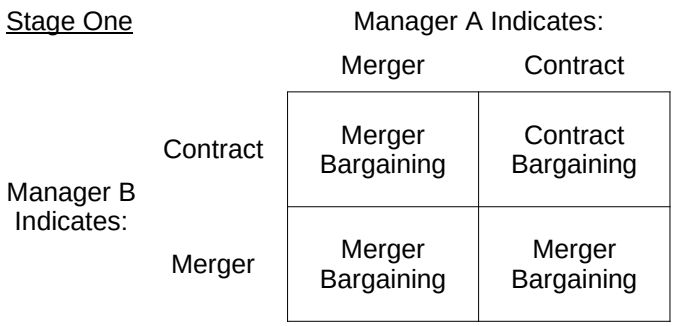

Figure 1.5: Merger Convention

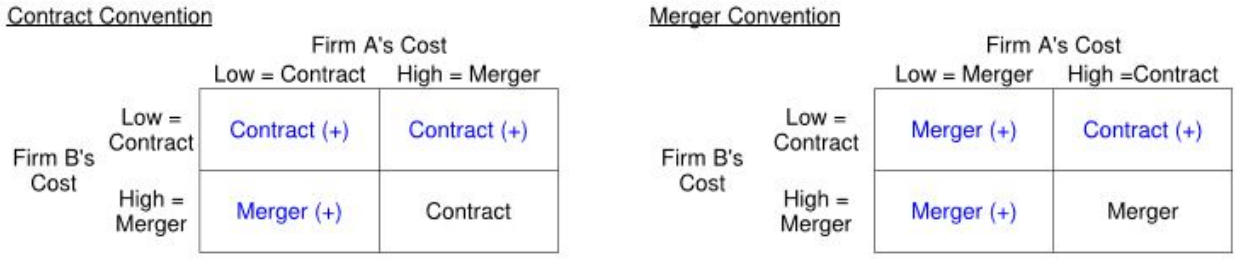

Figure 1.6: NE1 under the Conventions

in the case where both managers have high costs, bargaining is very likely to break down and result in independence (and thus non-coordination).

Another possible option is to ban either contracts or mergers, essentially eliminating the first stage of the game and sending the managers directly into bargaining for whichever organizational form remains. In this case, the bargaining will take place over the entire cost space since they no longer gain information from seeing each others' bargaining choices.

This results in five different versions of the game each reflecting a different possible institutional environment. The contract convention and merger ban environments would resemble a networked economy with lots of small firms collaborating with each other, while the merger convention and contract ban environments would resemble the more typical industrial economy with large firms performing a combination of tasks. The baseline case can be thought to resemble developing economies where the rules for solving coordination problems are not

yet well defined.

1.3.8 Analytic Difficulties

As will be seen in the results that follow, an environment that favors contract will be closer to the first best than one that favors mergers and both will better than having no convention at all. To make the effects more clear, the organizational choice game needs to be solved. Due to the multiple cases where payouts are conditional, the model is not congenial to analytic results and thus investigating the model via simulation yields is more fruitful. The next section will describe simulation output for this game.

To understand why clean analytic results are difficult to obtain, consider a case where the cost distribution consists of only two values (high and low). In this case, if manager A adopts a strategy that chooses merger bargaining under one value and contract bargaining under the other, then their choice of bargaining would give manager B complete information. If, at the same time, manager B chose the same bargaining option under both values, they would now have a information advantage which they could exploit. Manager A would be better off choosing the same bargaining option regardless of their cost draw. Thus, the state space must be sufficiently large such that both managers can adopt a cutoff strategy without giving the other manager knowledge of the exact state of the world.

Using a simulation with a large state space will produce a mixture of organizational forms depending on the parameter values and the specific cost draws. If the state space is small enough, the first stage of the game essentially vanishes. Given that the primary research question concerns how managers choose *between* organizational firms, the state space needs to be large enough for more than one organizational form to be possible. Using simulations, we can describe the model with the share of random draws that achieve a particular outcome, and comparisons can be made in terms of particular organizational forms being more or less

likely.

1.4 Simulation of Organizational Choice

To demonstrate the range of coordination outcomes, we will arrange the simulation to create a wide range of cost draws that result in potential coordination failure. Draw each manager's cost of coordinating c from a discrete uniform $U(0,1)$ with number of increments n. Set the revenue gains from coordination to $r_B = 1$ and $r_A = 2 - \sqrt{3}$, meaning that B will always be enthusiastic about coordinating and A will frequently be reluctant. The latter revenue amount is chosen to make the proportion of draws where coordination is obviously good for both managers equal to the number of draws where coordination is *not* first best. The remaining $\sim 44\%$ of draws are cases where A is reluctant but coordination is first best. See the appendix for a robustness check using alternate revenue amounts.

Below is the output from simulations covering the full range of the contractability parameter (θ) from $0-1$ using m even increments. The results below used $n = 100$ for the cost space and $m = 50$ increments for θ. For each θ value, both managers determine an optimal cutoff for both Nash equilibria, then n draws from the cost distribution are made for each manager. For efficiency, the optimal cutoff is determined by a type of hill-climbing algorithm where cutoffs are alternately adjusted until the jointly optimal cutoffs are located. For each pair of draws from the cost distribution the cutoff strategy for each manager is implemented and they enter the appropriate bargaining game (where applicable). Outcomes are recorded and the output below reflects the proportion of draws where particular outcomes occurred.

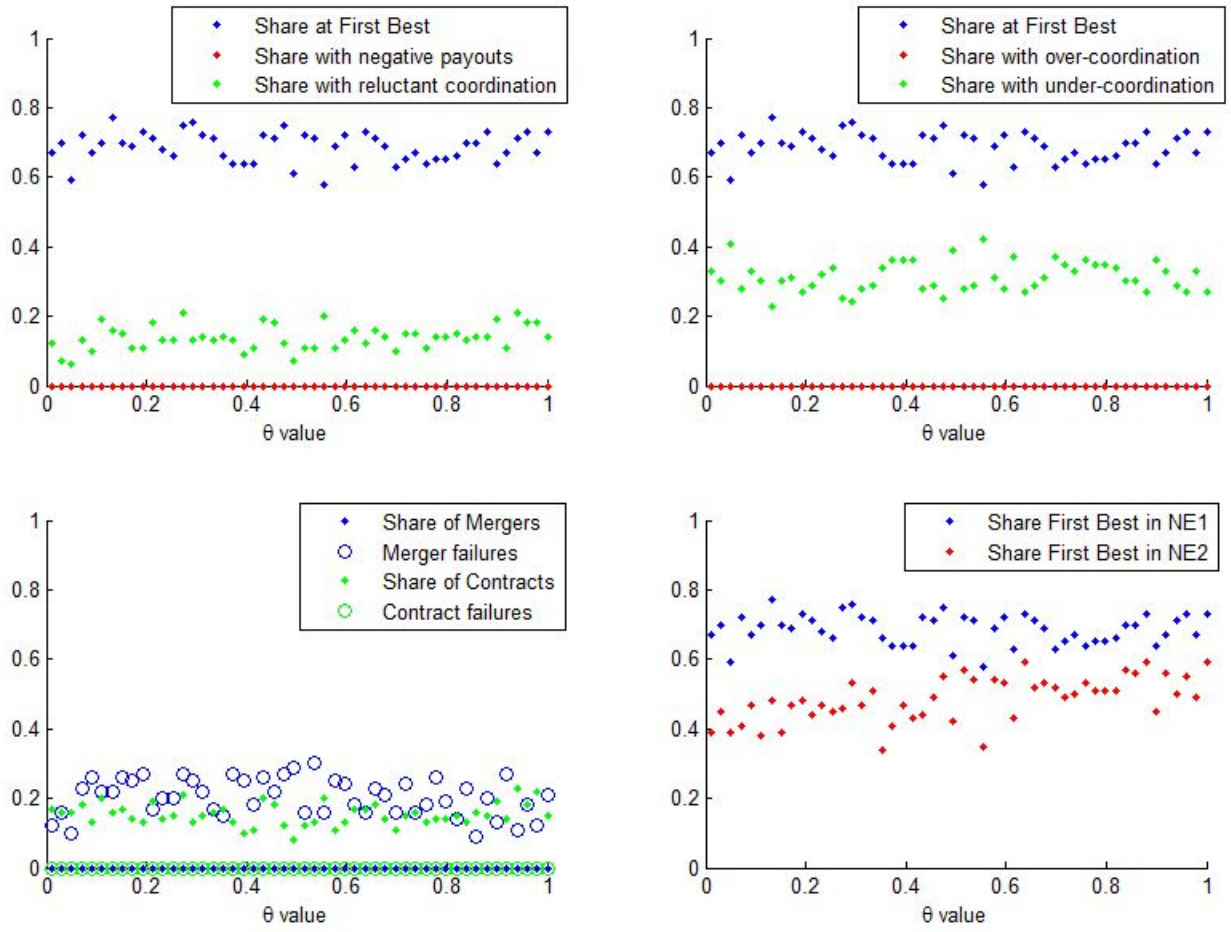

Figure 1.7: Baseline - No Convention

1.4.1 Simulation Results

For all of the results diagrams, the horizontal axis is the value of the contractability parameter (θ) and the vertical axis is the share of draws resulting in the indicated outcome. First, we use a baseline case where no conventional organizational form is imposed. Due to the two quadrants that result in the organizational form being independence, outcomes do not achieve the first best in all cases, even when $\theta = 1$. We can also see that equilibrium two is always further from the first best outcome than equilibrium one (meaning we should expect manager A to favor mergers when their costs are low).

Note that given the form of the payouts, since equilibrium one is never further from the first

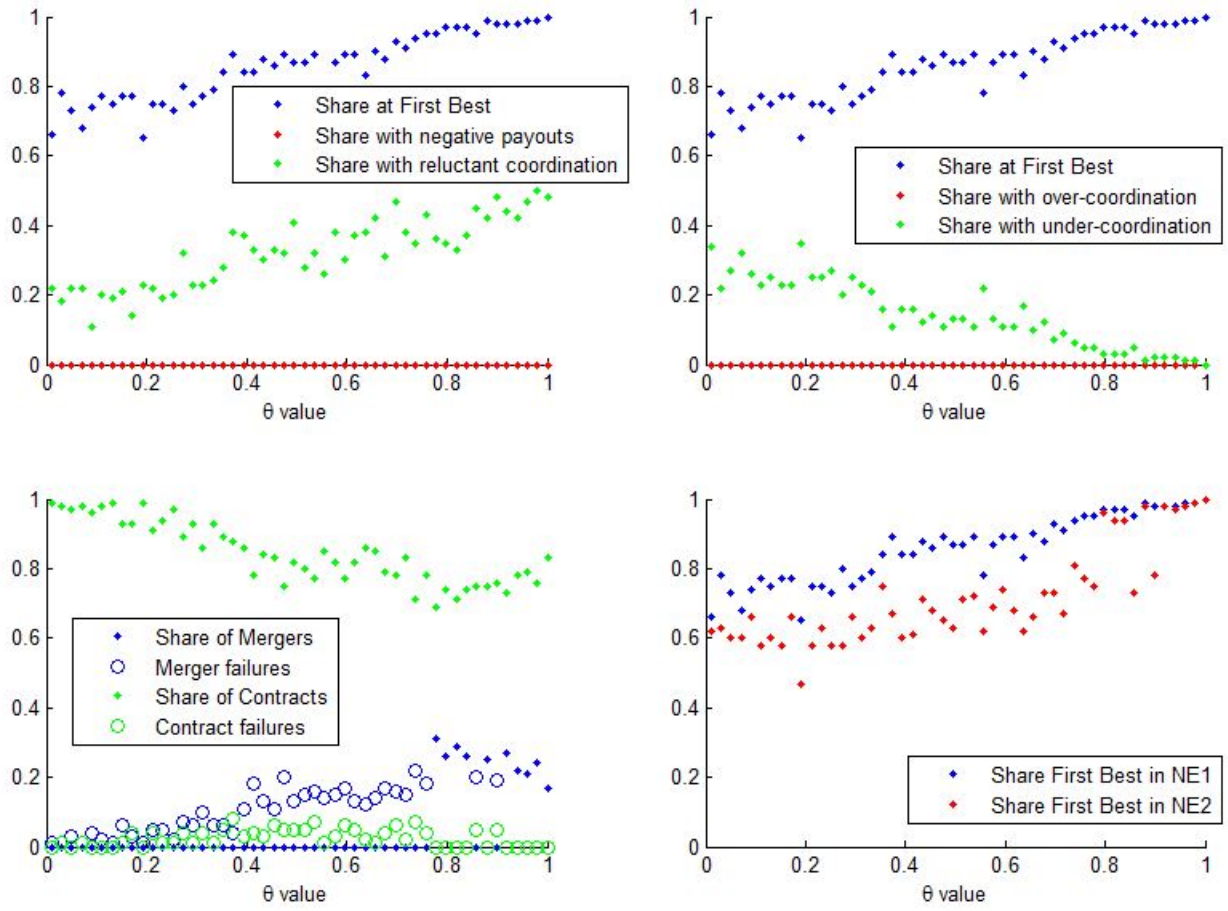

Figure 1.8: Contract Convention

best than equilibrium two, that the sum of the payouts are also always at least as good. Thus, equilibrium one is weakly Pareto dominant and this gives us good reason to focus on this equilibrium for analysis. The equilibrium where manager A chooses merger bargaining when their costs are low, will be superior in all of the cases that follow as well. When we set contracts to be the conventional form of bargaining, the outcomes improve considerably. Most importantly, the share of draws resulting in a first best outcome steadily increases and equals one when $\theta = 1$. Note that until θ is high, merger bargaining breaks down and independence occurs when both managers indicate merger. In this environment, a successful merger requires that both manager's costs be mostly visible and describable to the other (θ is high). Mergers that cause over-coordination never occur. As always, equilibrium one yields

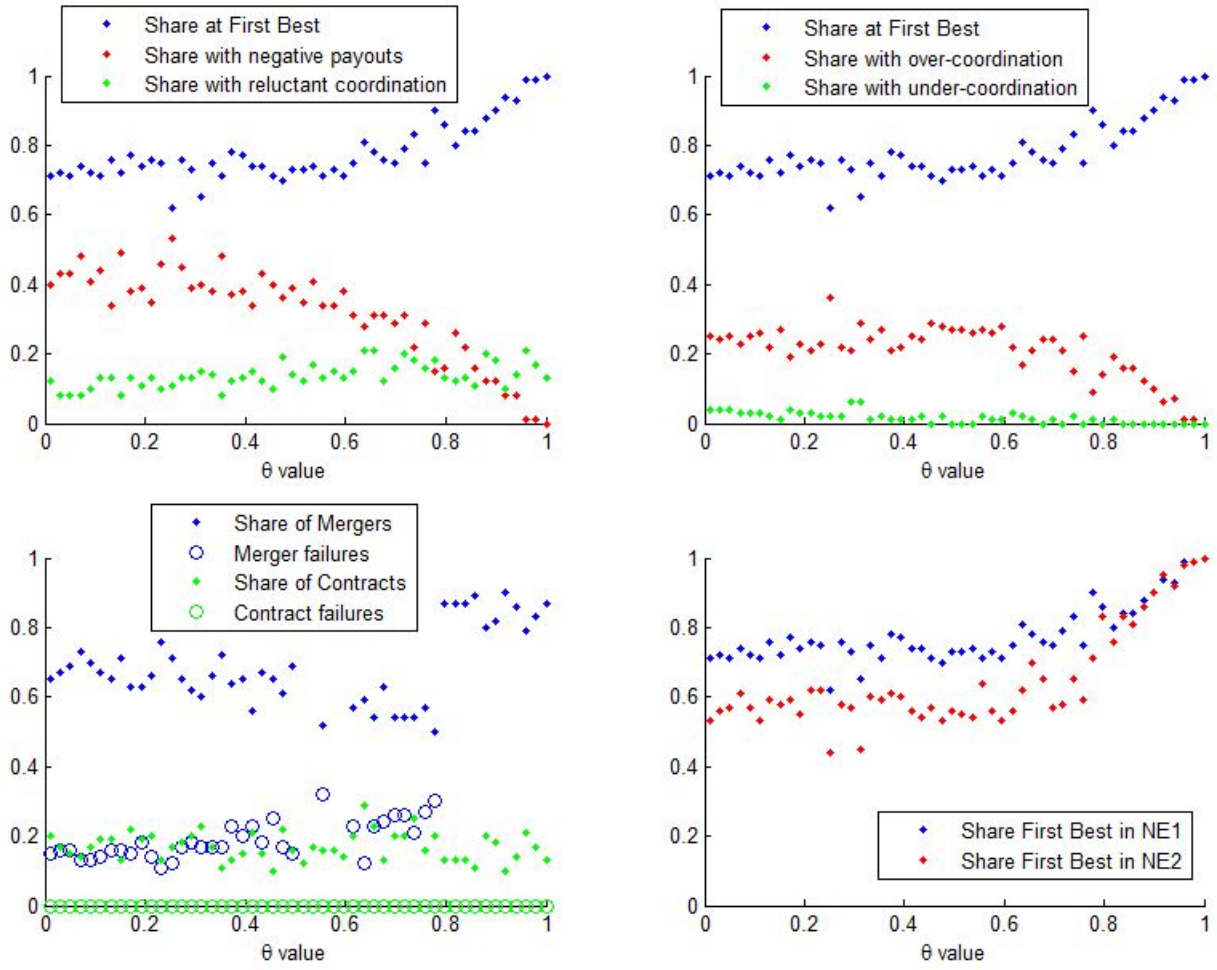

Figure 1.9: Merger Convention

better outcomes than equilibrium two. With a merger convention, outcomes are also better than with no convention. Until θ is sufficiently high, at least some amount of the merger negotiations will break down and result in independence. Contracts, on the other hand, will successfully be created. In this environment, under-coordination is far less common, but instead there is a large amount of over-coordination. In an environment where mergers cannot occur, outcomes are fairly similar to the contract default. However, the discussion of figure 1.12 below will show that this environment is not quite as likely to achieve the first best as the contract convention. If contracts cannot occur, the outcomes are similar to the merger convention, with the difference being that under-coordination never occurs and over-coordination is maximized. If we compare all of the various possible rules environments,

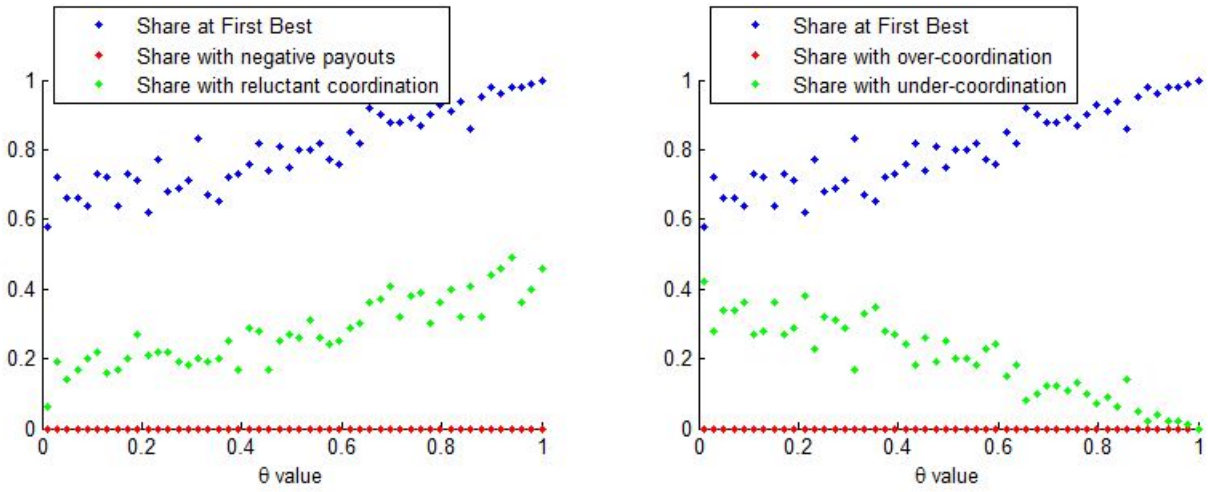

Figure 1.10: Ban on Mergers

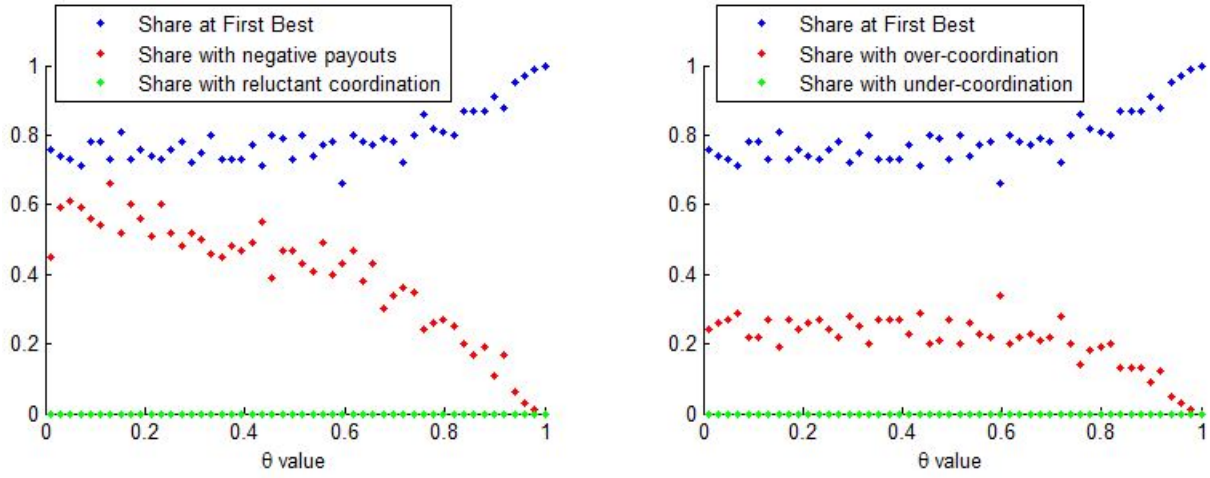

Figure 1.11: Ban on Contracts

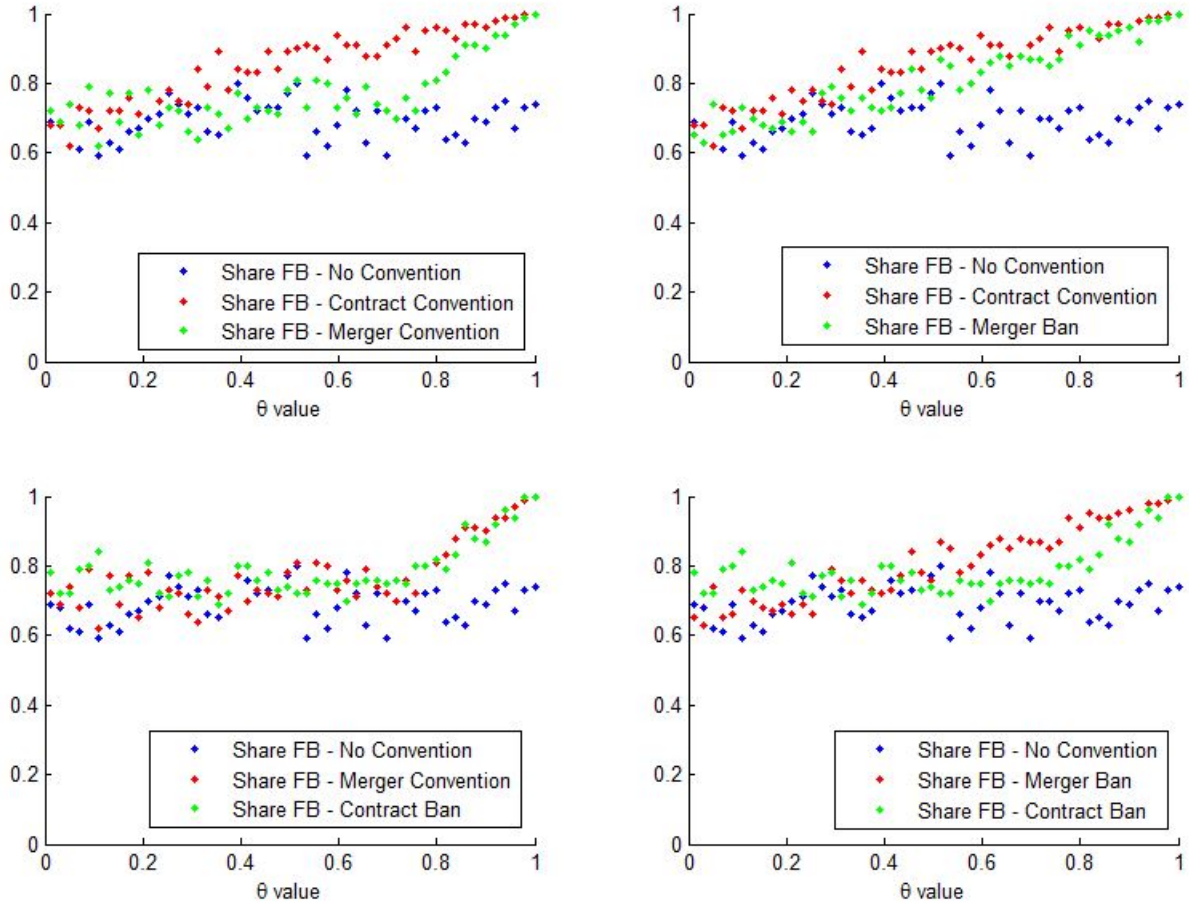

Figure 1.12: Rules Comparisons

we can rank their likelihood of achieving the first best outcome as follows:

1. Contract Convention
2. Merger Ban
3. Merger Convention or Contract Ban (no discernible difference)
4. No Convention (baseline)

Of course the merger and contract ban environments are not particularly reasonable models for any actual institutional framework in the real world. They are included here primarily for illustrative purposes.

1.4.2 Discussion of Results

Perhaps the most important result of the simulations is that the worst outcomes result when there is not a conventional organizational form. Having any kind of conventional mechanism for solving coordination problems between firms is better than having none at all. There are some potential implications here for development economics. A developing country might be inclined to spend time and effort finding a "best" model for future growth. But, with regard to organizational concerns, this hesitation would appear to be quite costly. Industries with very high contractability (i.e. very visible operations and easily quantifiable costs) will achieve notably better outcomes with either convention. When one considers that high θ is more likely to be found in industries where materials and capital costs dominate (rather than labor and administrative costs), a lack of an organizational convention might have a disproportionate impact on manufacturing versus service industries, making such an environment particularly damaging for developing economies.

The other main result is that the optimal institutional environment overall is one that favors

contracts between firms rather than mergers. Most of the developed world currently seems to have a merger convention environment and much of the industrial organization literature takes for granted the optimality of this approach. Local anomalies, like the northern Italian industrial model mentioned in the introduction, challenge this assumption. The model in this paper suggests that the success of this networked business environment may reflect genuine economic advantages rather than some localized special circumstances.

There are however some objections one might make to a contract convention being the optimal organizational environment. First, is that coordination problems governing complementary goods may not be the only problem that organizations solve. Mergers may be a more effective means of managing multiple production centers for the same good. Second, for very high θ values the gap between a merger convention and a contract convention is not very large, and the upheaval from transitioning from one convention to another may not be worth the gain. Finally, the model above assumes that an organizational form imposes no unique costs of its own. One could object that contracts impose additional implementation costs that are greater than those present in a merger. On the other hand, recall that under a merger we have assumed that headquarters does not receive a share of the profits, and that if it did this would make mergers less attractive for the managers. Nonetheless, whether there are systematic differences in the cost of operating in different organizational forms is beyond the scope of this paper.

The last potential objection to the results concerns whether the results are sensitive to different parameter values. The particular revenue values used above were chosen for clarity of presentation, but the reader can refer to the appendix for results using alternative revenue amounts. The main conclusions still hold – that a convention is preferable, and that the contract convention is closest to the first best.

1.4.3 Asymmetric Private Costs

Up to this point I have assumed that the proportion of the managers' costs that is contractible is the same for both managers. This assumption can be relaxed to consider the case where the two managers have different levels of private costs (i.e. different θ values). We might expect this to happen when the products of the two firms have very different production processes. (An example might be smart phones and software applications for use on phones.) In such a scenario, we should expect asymmetric contractability to improve outcomes relative to both managers having the lower θ value, but reduce them relative to both managers having the higher θ value. Giving the lower θ value to the reluctant manager (A) should produce results further from the first best than if manager B has the lower θ value.

First consider the case where $\theta_B = 0.5 * \theta_A$, meaning that manager B will always have greater private costs than manager A. The results here are very similar to what occurs where both managers have the higher θ value. Under a contract convention, the results are closer to the first best than if both managers had the lower θ value, but not as good as if they both had the higher value. Specifically, using results from figures 1.8, 1.13 and 1.16, the percentage of draws that achieve the first best when $\theta_A = 1$ & $\theta_B = 0.5$ is 97% and the percent first best when $\theta_A = 0.5$ & $\theta_B = 1$ is 95%. When $\theta_A = \theta_B = 1$ the percent first best is 100%, and when $\theta_A = \theta_B = 0.5$ the percent first best is 87%. Note that raising the reluctant manager's contractability (θ_A) is more effective at improving outcomes when using a contract convention. Under a merger convention, the results are also closer to the first best than if both managers had the lower θ value, but not as good as if they both had the higher value. Specifically, using results from figures 1.9, 1.14 and 1.17, the percentage of draws that achieve the first best when $\theta_A = 1$ & $\theta_B = 0.5$ is 81% and the percent first best when $\theta_A = 0.5$ & $\theta_B = 1$ is 77%. When $\theta_A = \theta_B = 1$ the percent first best is 100%, and when $\theta_A = \theta_B = 0.5$ the percent first best is 72%. Note that raising the enthusiastic manager's contractability (θ_B) is more effective at improving outcomes when using a merger

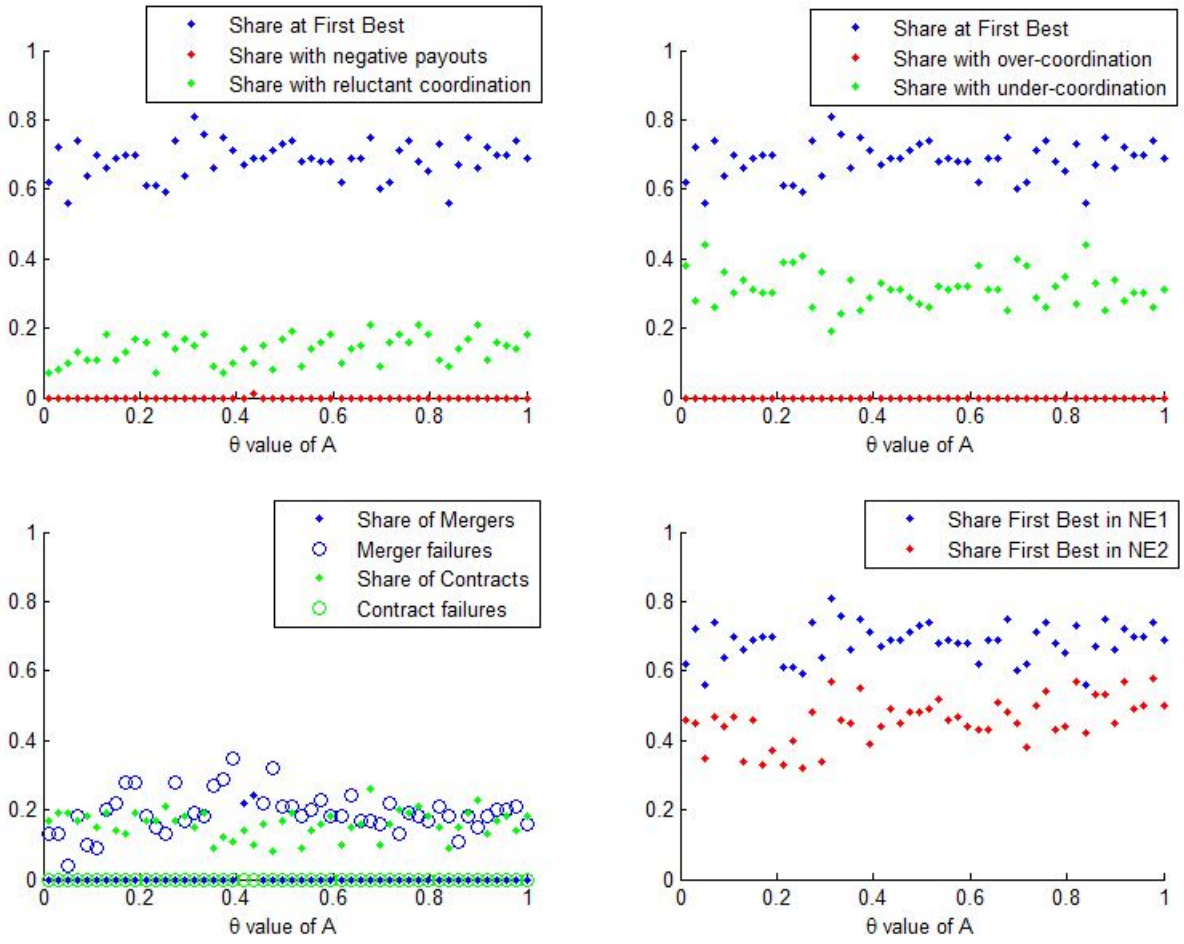

Figure 1.13: Baseline where B has more private costs

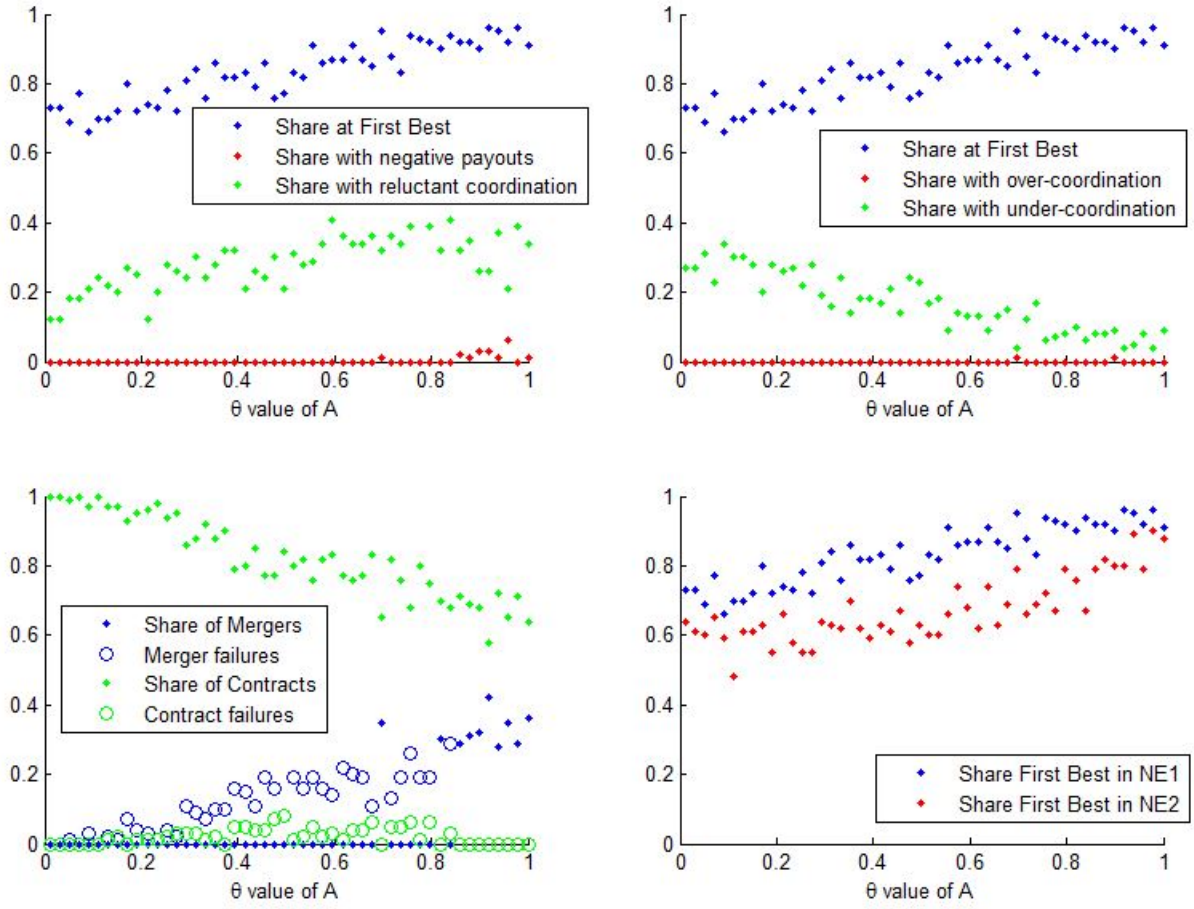

Figure 1.14: Contract Convention, B has more private costs

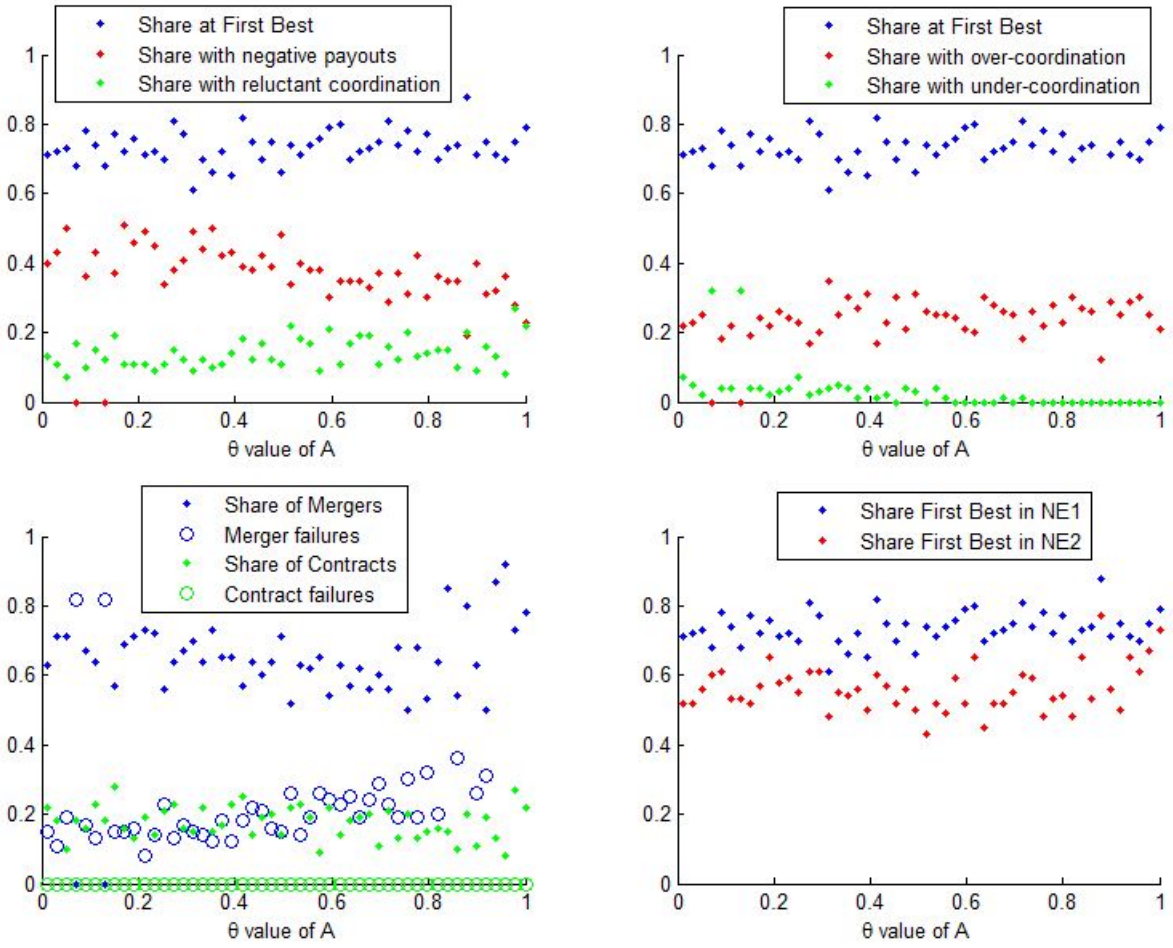

Figure 1.15: Merger Convention, B has more private costs

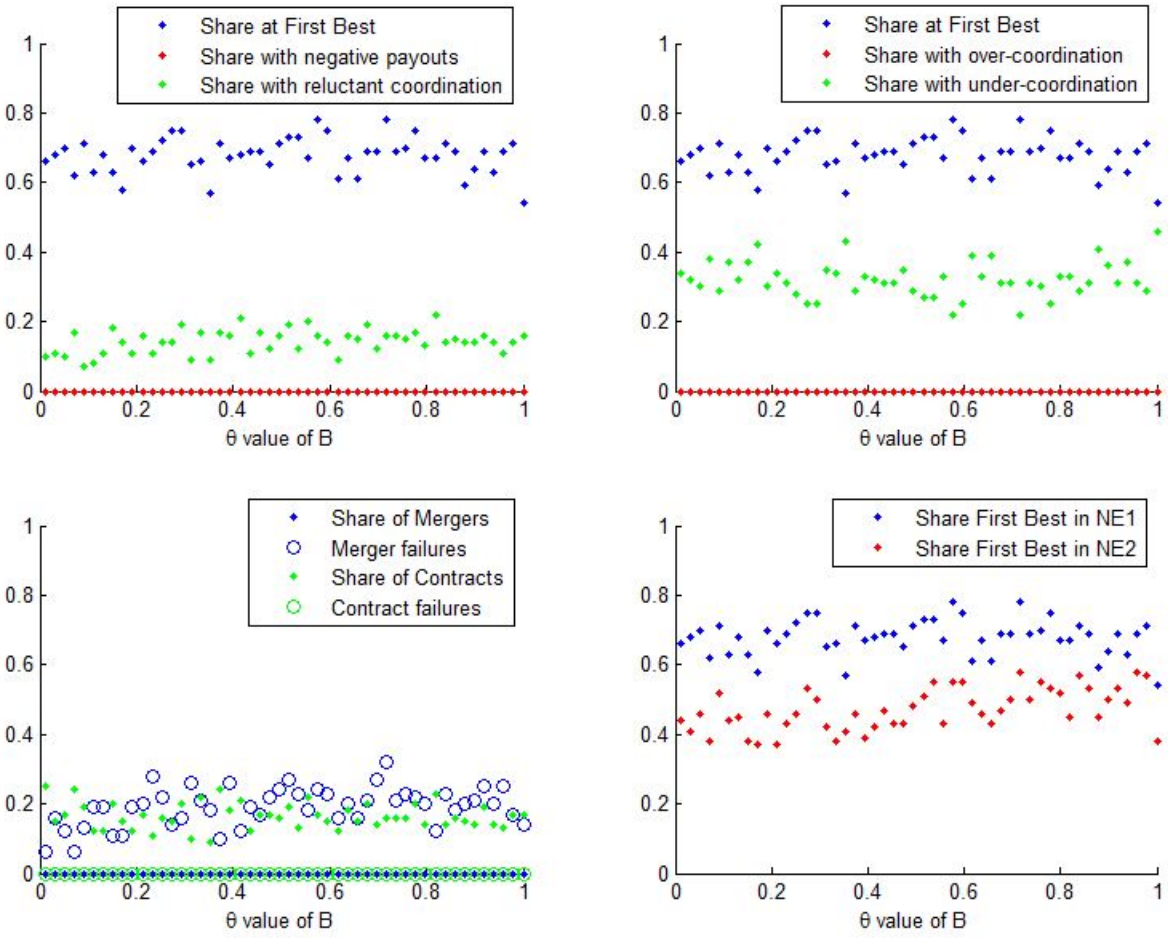

Figure 1.16: Baseline where A has more private costs

convention.

Now consider the case where $\theta_A = 0.5 * \theta_B$, meaning that manager A will always have greater private costs than manager B. As the above figures indicate, outcomes are improved by an increase in θ for only one firm. The effect is minimal under the baseline case. Under a contract convention an increase in the reluctant manager's contractability (θ_A) improves outcomes more than a comparable increase in enthusiastic manager's contractability (θ_B) would. The reverse is true under a merger convention.

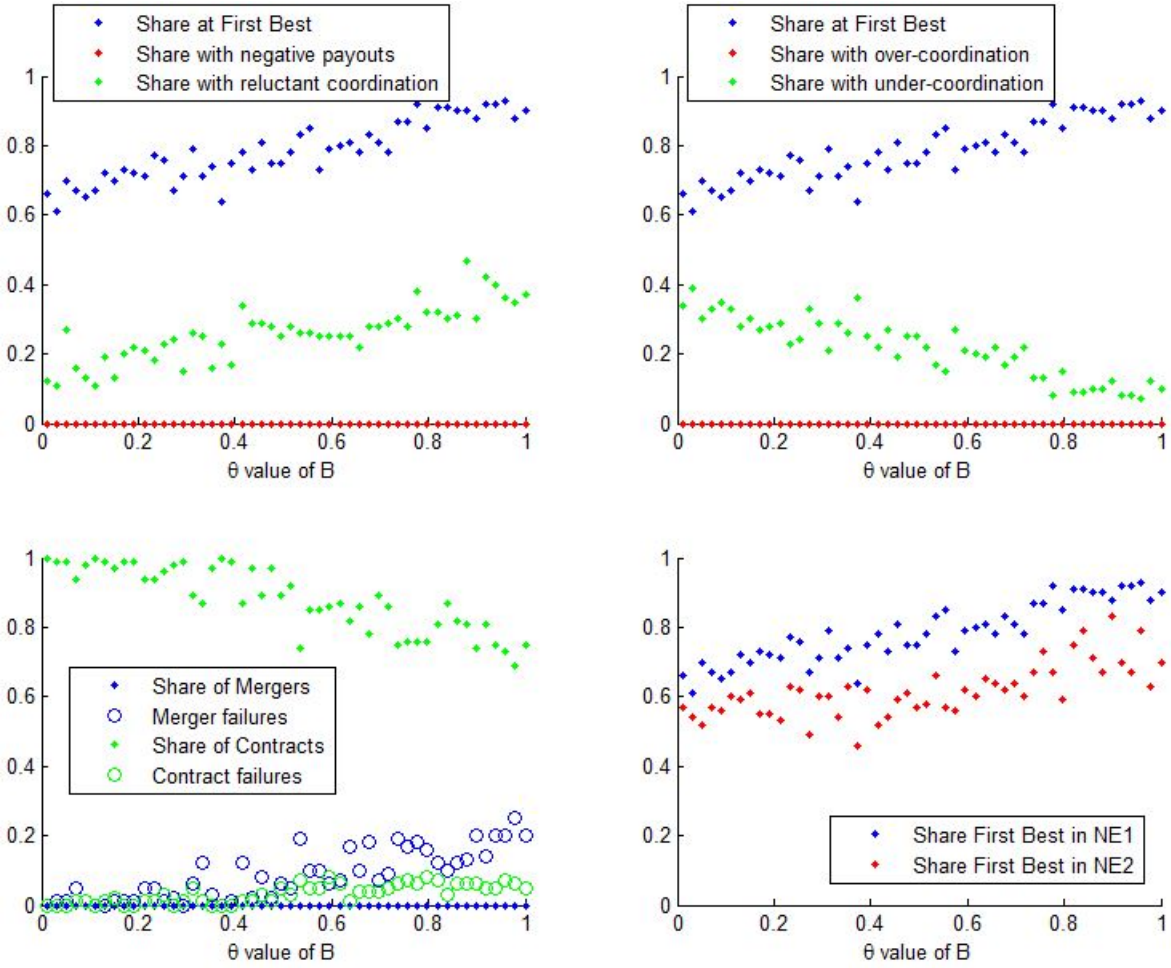

Figure 1.17: Contract Convention, A has more private costs

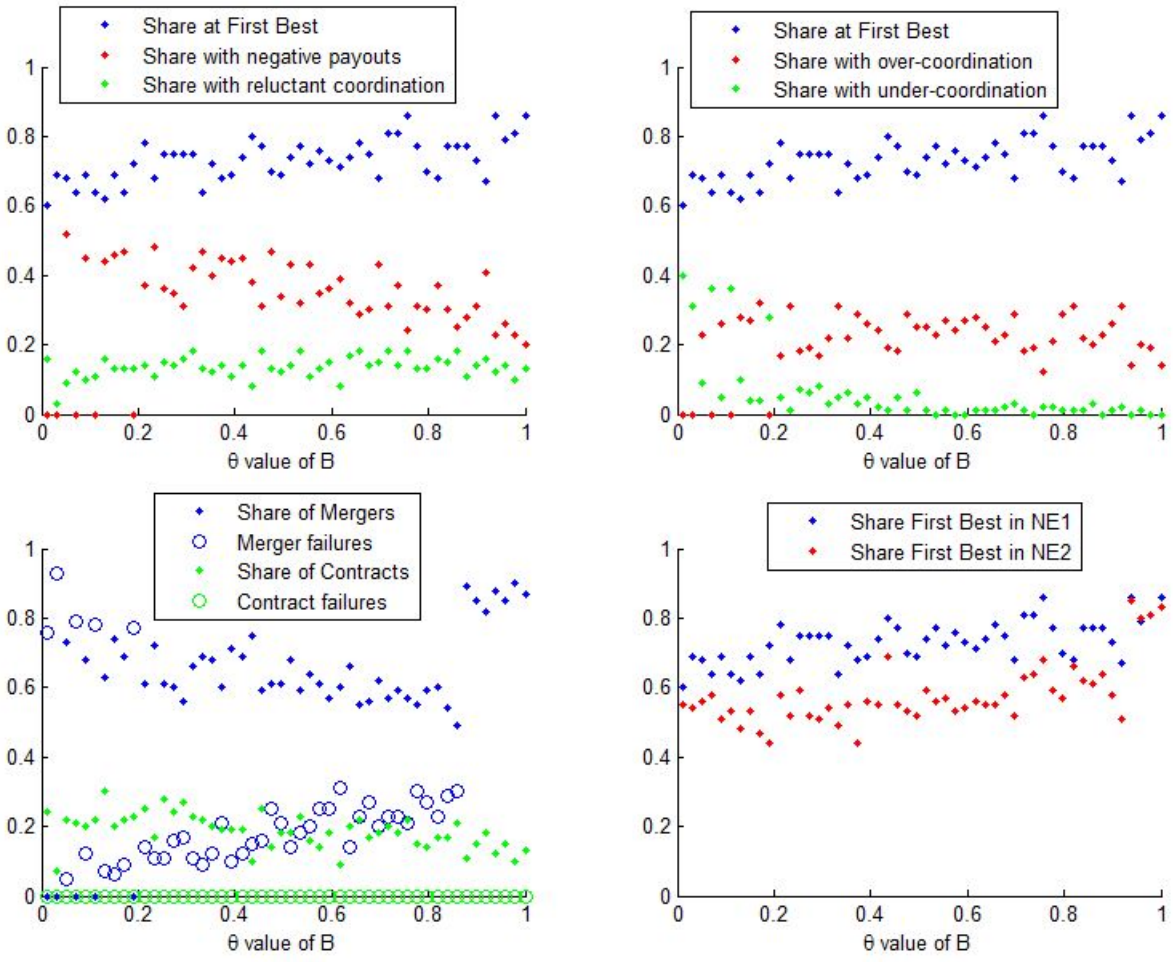

Figure 1.18: Merger Convention, A has more private costs

1.5 Conclusion

I draw three main conclusions from the analysis. First, that setting a convention that biases managers toward contracts or mergers improves outcomes by making independence less likely. Second, that it is not optimal to ban contracts or mergers and mandate a particular organizational outcome (behavior should be nudged, but not forced). Finally, a convention that makes the contracts the standard means of solving coordination problems brings outcomes closer to the first best than one favoring mergers.

In addition, there are some secondary results of interest. The degree of private costs experienced by the managers do not have to be symmetric, and an improvement in the contractability of only one manager's costs improves outcomes. However, it is important *which* manager has the lower private costs. If the manager who is more enthusiastic about coordinating has lower private costs, this improves outcomes for a merger convention rather more than it does for a contract convention. Conversely, if the more reluctant manager has lower private costs, this improves outcomes more for a contract convention. I also observe (see Appendix) that a wider gap in the revenue gains between the two firms increases the benefit of imposing a convention.

A number of extensions and concerns arise. The model here is a one-shot game, and a repeated game with a large population of managers could potentially yield different results. For example, it is unclear whether in a repeated environment that mergers would be more optimal (by providing long term certainty), or be less optimal (due to some managers getting trapped in bad deals). Experimental tests are also worth exploring, especially since the model above uses risk neutral agents and experimental subjects with some degree of risk aversion might produce significantly different results. Lastly, any attempt to demonstrate the theory via an econometric investigation would hinge upon finding a sufficient proxy for the contractability parameter.

The results are broadly consistent with the empirical literature, with a rich variety of outcomes that replicate the typical merger friendly environment of the developed world, the less common and less well-understood contractual environments, as well as the relative lack of effective coordination in developing economies. On the other hand, the model does not explain how these organizational rules get determined and further research would be needed to address policy changes that could induce a shift from one rule set to another. Nonetheless, the model here should be sufficient to demonstrate that organizational form is neither neutral nor irrelevant to economic outcomes.

Chapter 2

Norm Enforcement in Faith Communities

I introduce a model of religious organizations, where members of a population choose whether to enforce material contributions to the group, and also how much to commit to those contributions. Expected effort along each dimension follows a norm framework, where deviations from the norm generate disutility. Depending on the relative strength of the norms and the effectiveness of enforcement, the nature of the equilibrium varies. An expanded model distinguishes doctrinal practices from material commitment to provide further insight. The equilibrium types of the static model are shown to be steady states of a dynamic norm selection process. An extension treats doctrinal norms as fixed, and provides explanations for the formation of sects, as well as observed distinctions between new and mature religious denominations.

2.1 Introduction

There are a vast array of religious groups, exhibiting wide variations in size of congregation, the requirements of membership, and the benefits enjoyed by members. Economists, and other social scientists studying religion, have long observed a correlation between the strictness of a faith group and the amount of benefits conferred. But simply having a code of conduct does not ensure compliance, and members of a faith community exhibit diversity in how faithful they are to the rules, whether strict or not. A substantial existing literature has analyzed strictness as a mechanism for preventing free-riding in a club good environment.

But what is strictness? It could be described as a social distance from secular behavior, but this only seems to be clear and useful in developed world settings where secular behavior is the default. One could also describe strictness as the amount of time or money demanded of members, which is convenient for analysis, but would neglect a variety of phenomena, such as veiling or circumcision, that impose neither temporal nor pecuniary costs. The degree of strictness might take into account the relative visibility of regulated behaviors to outsiders. Strictness could also describe the severity of penalties for non-conformity, or the degree to which members must share the labor of ensuring conformity.

Clearly, the concept of strictness could be divided into a great many components, many of which could provide added insight. For this paper, I propose decomposing strictness into several components. I refer to effort that boosts the likelihood of punishment for failing to contribute as *enforcement*. All other effort provided, whether material or observant, falls into the category of *commitment*. In the full model, I further subdivide material effort from doctrinal observance, with the former still called commitment and the latter designated as *doctrine*.

My analysis uses the framework of a norm utility model, where the amount of effort members are expected to contribute to each category is a particular level called a *norm*. Deviation

from these norms generates disutility in proportion to how much a member deviates from the norm. Initially, I will take the norms as exogenous, but the dynamic treatment will explore a process by which norms might be set. In addition to disutility from failing to follow norms, members take disutility from the effort they individually contribute, but gain utility from a club good that takes as its inputs all of the effort contributed by the group.

When considering how social norms are maintained, one has to first recognize that social norms do not typically arise by consensus. Rather, the set of "proper" behaviors is often disputed, and individuals can differ dramatically in their desire to deviate from acceptable behavior. Some degree of non-conformity appears unavoidable. The key to maintaining a set of norms lies in the response to those individuals who refuse to conform. At least some non-conformists must be punished for doing so, and some members of a group must take up the task of identifying and punishing people who disobey the norms. In other words, a social norm must be *enforced*.

The need for enforcement motivates a unique aspect of the model environment – the use of a punishment function that modifies the disutility that members experience from violating certain norms. In the basic model, punishment affects all commitment (material or not), while in the full model it modifies the norm disutility from failing to observe doctrine. The degree of punishment is determined by the total enforcement effort, causing the norms to interact with each other.

Also, since individuals vary, we must consider individual responsiveness to the demands placed on them by a faith. Members of a congregation will be more or less *sensitive* to these norms, and the distribution of norm sensitivity will in large part determine the amount of effort within the group.

I employ the model environment to examine a number of phenomena commonly addressed in studies of religion These include why sects break off from established religions, and why spreading a faith to new communities can change how the faith is practiced. The model can

also provide insight about why new religions often demand large material commitments or punish infractions harshly, while mature faiths with wide acceptance frequently have many members simply observing doctrine without providing any other form of effort.

2.2 Literature

The modern economics of religion literature starts with the club goods model of Iannaccone (1992). Here religious organizations are described as providing a club good, but where members must also pay the costs of unusual doctrinal practices. The ensures that only serious participants belong to the faith, meaning that religious organizations' distinctive practices screen out free riders. This is not the first economic work on religion. As early as Azzi and Ehrenberg (1975) there were models of consumer choice among religions. Iannaccone's paper is, however, the seminal paper in the club good strand of the literature.

The religious club good model is featured in many papers, and notable recent contributions include Berman (2000) which analyzes the practices of Haredi Judaism as a generous club good that offsets substantial compliance costs. In Chen (2010), the growth of Islam in Indonesia in the wake of the late 1990s Asian financial crises is linked to the services provided to members. The services provided by faith groups are often competitive with state services, and can act as a substitute. See Hungerman (2005) for analysis of church service provision in US context, and Iyer et al. (2014) for a look at social service provision by faiths in India. The benefits of faith groups need not be wholly material, and can nclude a sense of access to supernatural gains, as discussed in Iannaccone and Berman (2006).

However, other papers suggest that passive conformity and varying levels of effort play an important role in religious movements. Examination of the middle eastern veiling movement in Carvalho (2013) demonstrates that the spread of distinctive religious practices do not neatly correspond to high levels of piety. See also Patel (2012) for an analysis of how the

signaling quality of veiling has been diminished as it has become more common, which is consistent with the dynamic version of the model where, when norms become more salient to the population, the intensity of norms tends to fall. The level of service provided by a faith group also need not be uniform as detailed by McBride (2007), an exploration of membership tiers within the Mormon faith, demonstrating the importance of a distribution of religious effort within a denomination.

The model takes as exogenous the degree to which members are more or less responsive to religious norms, which is potentially problematic since some people actively choose their faith affiliation, and might select in or out of a group. However, a study of the Catholic child abuse scandal of the 2000s by Hungerman (2013) measured the degree of substitution between faiths by using the scandal as a natural experiment. The results imply a high degree of substitution which, at least for the larger faiths within a society, reduces concern about self-selection. A possible contrast comes from Barros and Garoupa (2002), a theory paper where churches choose strictness such as to attract desired types from the population, however since the model below allows for types who exhibit no effort one could still interpret the distribution as exogenous, if the model population is viewed not as the congregation, but rather the local public in general. As a counterargument, Levy and Razin (2012) present a model where faiths actively shape individual beliefs, a likely phenomenon that is neglected in the model below.

In the sociological literature there has been substantial debate about how to classify religious groups, particularly newer ones. The most influential view comes from Wallis (1984), where one key aspect of a faith is whether it sees itself as apart from the world (i.e. nonbelievers) or whether it sees itself as existing within the world. Newer religions are much more likely to view their members as a special group set apart from the rest of society, and this can be thought of as reflecting a larger material or doctrinal commitment. The doctrinal rigidity section discusses possible explanations for these characteristics of new religious groups.

The model below is a norm utility model. Such models have frequently been used in the experimental literature to examine rule following behavior. A basic version of these models can be found in Bicchieri (2006). Agents have utility of this form:

$$U_i(x_i) \;=\; g(x_i, x_{-i}) \;-\; \gamma_i \beta_i (\hat{x} - x_i)^2$$

Here $\hat{x}$ is the norm, γ_i is i's norm salience parameter, x_{-i} is all other person's contributions, and β_i reflects i's beliefs about how likely others will comply with the norm. In this paper, I will use a commonly known distribution of γ values, and use the simplification that $\beta_i = 1$ for all i. This simplification is warned against in Bicchieri's paper, however in the setting of this paper, it is less controversial. If we assume that the same people who respond strongly to a religious norm also have confidence that it will be followed, then we could say that $\beta_i = f(\gamma_i)$ where $f(.)$ is an increasing function. In that case a composite parameter combining γ and β would be mathematically indistinguishable from letting $\beta_i = 1$ for all i. I contend that confidence in the viability of a religious norm is reasonably likely to be correlated with with how compelling that norm is to an individual. As such, this simplification is not pernicious.

From the norm literature there are a number of observations that arise in this work as well. In Horne (2003), internalization of norms is shown to be widespread which is consistent with any equilibrium where the norm for enforcement is low or the share of people performing such effort is low, but Horne (2004) points out that some enforcement is nonetheless necessary. Norms are a powerful tool for overcoming collective action problems, and per B (2007), they can sustain cooperation even for conduct that is costly or atypical for the society at large. Lpez-Prez (2008) details an experiment that explains some seemingly non-rational behavior commonly discovered in experimental settings as due to norms of reciprocity that subjects bring into the experiment with them. This attests to the powerful influence of expected conduct on individual effort, and the experience of disutility when failing to live up to those

expectations.

2.3 Basic Model

There is a faith community with a member population of size N. Members of the community choose the amount of effort they wish to contribute amongst two tasks. The first task is their commitment to the group (c_i) which contributes to collective production of a club good. The second task is enforcement (e_i), which increases the chance that members who fail to commit will be punished for doing so, and also contributes to the club good. The model is static.

Members are expected to commit and enforce according to exogenous norms ($\hat{c}$ and $\hat{e}$). Deviation from the norm generates disutility, but members vary in their sensitivity to these norms (γ_i). These non-negative norm salience values are drawn from a distribution whose density function has domain $\{\gamma \in \mathbb{R} | 0 < \gamma < \infty\}$.

The club good produces utility for each member according to a function $g(\tilde{c}, \tilde{e})$ which is increasing and concave in the total amount of commitment effort ($\tilde{c}$) and non-decreasing and non-convex in the total amount of enforcement effort ($\tilde{e}$), and where $g(0, 0) = 0$. The reader might wonder whether enforcement effort should be omitted from the club good function. While in many cases the benefits enjoyed by a faith group are probably separable from effort spent enforcing doctrine, this is not always the case. For example, visible enforcement effort could increase members' sense of belonging, which in turn would enhance their enjoyment of the club good. Regardless, having enforcement effort be an argument of the club good function is the more general case.

The norm disutility when a member fails to meet their commitment to the club good is adjusted by a punishment function $p(\tilde{e})$, that is increasing in total enforcement effort, and

where $p(0) = 1$. The punishment function can be interpreted as the likelihood of being caught shirking commitment, increased guilt from shirking commitment, or the severity of punishment – or a combination of those things. The exact nature of the punishment is abstracted in the model. It is important to note that by making $p(0) = 1$ we are saying that an unenforced commitment norm will reduce to a standard norm utility model with a club good and a single norm.[1] With enforcement, the marginal cost of commitment falls, inducing higher commitment effort.

Thus, members will choose positive commitment and enforcement levels to maximize the following utility function:

$$U_i(c_i, e_i) \;=\; g(c_i, c_{-i}, e_i, e_{-i}) \;- p(e_i, e_{-i})\gamma_i(\hat{c} - c_i)^2 \;- c_i \;- \gamma_i(\hat{e} - e_i)^2 \;- e_i \tag{2.1}$$

Note that c_{-i} is sum of all other members' commitment effort, and similarly for e_{-i} and enforcement effort. This utility function measures all costs and benefits for members relative to the typical utility they would enjoy if not a member of the faith community, and enjoying the typical lifestyle of someone in the society at large. This baseline should not be interpreted as a modeling assumption that the faith group exists within a secular society. The society at large might reflect a religious perspective as well. This does mean that a religious group modeled in one cultural context might have different parameter values when analyzed in another context.

Maximizing an individual's utility results in the following optimal efforts:

$$c_i^* \;=\; \max\left[\hat{c} \;-\; \frac{1 - g_c}{2p(\tilde{e})\gamma_i}\;,\; 0\right] \tag{2.2}$$

$$e_i^* \;=\; \max\left[\hat{e} \;-\; \frac{1 - g_e}{2\gamma_i} \;-\; \frac{p_e}{2}(\hat{c} - c_i)^2\;,\; 0\right] \tag{2.3}$$

[1]Note that if club good requires enforcement input, say if it were a Cobb-Douglas function, then $p(0)$ would result in zero club good production and thus zero utility for all members.

Note that as the population increases, the individual contributions to the club good and punishment function diminish. Thus, the limit case as $N \to \infty$ is:

$$c_i^* \;=\; \max\left[\hat{c} \;-\; \frac{1}{2p(\tilde{e})\gamma_i}\;,\; 0\right]$$

$$e_i^* \;=\; \max\left[\hat{e} \;-\; \frac{1}{2\gamma_i}\;,\; 0\right]$$

From the above we can see that a member's norm salience (γ_i) is critical to determining whether they contribute a non-zero effort to either task. Members for whom the norms are highly salient will produce efforts very close to the normative amounts. This also means that there is a cutoff for both commitment and enforcement, and individuals with norm salience below those values will contribute zero effort, but depending on the norm values and the punishment function, one cutoff or the other could take a higher value.

Rearranging the optimal effort equations gives the cutoff conditions, where γ_c is the norm salience where an individual is indifferent about contributing commitment effort, and γ_e is the norm salience where an individual is indifferent about contributing enforcement effort:

$$\gamma_c \;=\; \frac{1-g_c}{2p(\tilde{e})\hat{c}} \tag{2.4}$$

$$\gamma_e \;=\; \frac{1-g_e}{2\hat{e} \;+\; p_e\,\hat{c}^2} \qquad if\;\; \gamma_e \leq \gamma_c \tag{2.5}$$

$$\gamma_e \;=\; \frac{1-g_e}{4\hat{e}} \;+\; \frac{\sqrt{p(\tilde{e})^2(1-g_e)^2 + 2\hat{e}p_e(1-g_c)^2}}{4p(\tilde{e})\hat{e}} \qquad if\;\; \gamma_e > \gamma_c \tag{2.6}$$

As population increases the enforcement cutoff is increasingly determined primarily by the level of the enforcement norm, while the commitment cutoff is determined by the commitment norm and the shape of the punishment function. This implies that a smaller congregation will have a higher proportion of members participating in both commitment and enforcement.

The total effort expended on commitment and enforcement would thus be:

$$\tilde{c} \;=\; \sum_{i=1}^{N} \max\left[\hat{c} \;-\; \frac{1-g_c}{2p(\tilde{e})\gamma_i}\;,\; 0\right] \tag{2.7}$$

$$\tilde{e} \;=\; \sum_{i=1}^{N} \max\left[\hat{e} \;-\; \frac{1-g_e}{2\gamma_i} \;-\; \frac{p_e}{2}(\hat{c}-c_i)^2\;,\; 0\right] \tag{2.8}$$

Neither of these can be evaluated without explicitly defining the norm-salience distribution, but it will always hold that $\tilde{e}$ is increasing in $\hat{e}$ and that $\tilde{c}$ is increasing in $\hat{c}$. It also clear that a rise in the enforcement norm will always increase total commitment. However, a rise in the commitment norm would cause enforcement to fall slightly, although this effect diminishes as the congregation gets large.

However, as noted above, a larger congregation implies smaller p_e and thus the commitment norm will have decreasing effect on total enforcement as the congregation grows. For a given norm-salience distribution and a sufficiently large congregation, total enforcement effort is determined by the enforcement norm while total commitment effort is determined by the commitment norm and the shape of the punishment function.

2.3.1 Equilibrium Types

We can now classify the possible equilibria into distinct types depending on the relative position of the enforcement and commitment cutoffs. The simplest equilibrium is one where $\gamma_e > \bar{\gamma}$, which would mean that no one contributes enforcement effort and thus, per equation (2), the model reduces to a simple norm utility model. I designate these equilibria Type Zero. Since this is an outcome indistinguishable from a secular social club, this is not a particularly interesting type of equilibrium for the subject under consideration. The key result is that at least some members of the faith group must have high norm-salience in

order for an outcome resembling a religion to occur in equilibrium.

If $\gamma_e \leq \bar{\gamma}$ and $\gamma_e < \gamma_c$, the result is a Type EC equilibrium. Members with $\gamma_i \leq \gamma_e$ will provide neither enforcement nor commitment effort. Members with $\gamma_e < \gamma_i \leq \gamma_c$ will provide enforcement effort but no commitment effort. Finally, members with $\gamma_i > \gamma_c$ will provide both. The Type EC equilibrium is distinguished by this middle category of "cheating" enforcers, who do not contribute commitment effort. There are also two subtypes of these equilibria. First, subtype $EC(ce)$ where $\hat{c} \leq \hat{e}$ and $c_i \leq e_i$ for all members, and second, subtype $EC(ec)$ where $\hat{c} > \hat{e}$ and $c_i > e_i$ for members with the highest norm-sensitivity.

If $\gamma_e \leq \bar{\gamma}$ and $\gamma_e > \gamma_c$, the result is a Type CE equilibrium. Members with $\gamma_i \leq \gamma_c$ will neither enforce nor commit. Members with $\gamma_c < \gamma_i \leq \gamma_e$ will provide commitment effort but not enforcement. Finally, members with $\gamma_i > \gamma_e$ will provide both. The Type CE equilibrium is also distinguished by the members with mid-range norm-sensitivity, in this case "passive" commitment. There are also two subtypes of Type CE equilibria. First, subtype $CE(ce)$ where $\hat{c} < \hat{e}$ and $c_i < e_i$ for members with high norm-sensitivity, and second, subtype $CE(ec)$ where $\hat{c} \geq \hat{e}$ and $c_i \geq e_i$ for all members.

Figure 2.1 depicts the four subtypes (neglecting the uninteresting Type Zero). Note that the capital letters of a type designation indicate lowest to highest effort cutoffs, and the parenthetical lower case letters indicate lowest to highest norm levels. Alternatively, the capital letters can be read as highest to lowest share of the congregation contributing that type of effort.

2.3.2 Discussion

As noted above, and typical for club good models, a larger congregation will have a smaller proportion of members contributing either type of effort. Given that new religious groups must, of necessity, start out small – this result is consistent with the known phenomena of

new religions taking up more of their members time and resources, as well as being harsher to those who go against doctrine.

The four equilibrium types provide broad categories of faith groups. To distinguish them, I will describe each type using an example of a Christian denomination. Type $EC(ce)$ has a portion of members who enforce but do not commit, and the enforcement norm is higher. This is the region of nondenominational evangelicals, where material and doctrinal commitments are modest but sin and proper living are of high concern. Type $EC(ec)$ also has members who enforce but do not commit, but here the commitment norm is higher. This resembles the Mormons, where material and doctrinal commitments are high for those who opt into them (high salience members). Type $CE(ce)$ includes members who commit but do not engage in enforcement, and the enforcement norm is higher. This is the region of Pentecostals, particularly Assemblies of God, which exhibit high rates of giving, distinctive worship practices but enforcement effort is more concentrated in clergy and key members of a congregation. Finally, type $CE(ec)$ also has members who commit but do not enforce, but the commitment norm is higher. This resembles Catholic churches, where there are many unique doctrinal practices and substantial charities, but enforcement is of relatively low priority.

2.4 Full Model

A reasonable criticism of the basic model is that it conflates members' lack of material commitment (time and money) with failure to observe the unique practices (dress, diet, conduct, etc.) of a particular faith. The basic model treats these two types of effort as having equal impact on the benefits, and also being equally responsive to punishment or admonition when effort is low. As a refinement, let us call these unique practices the *doctrine* of a faith group, with a distinct doctrine norm ($\hat{d}$) and individual doctrinal effort d_i. In addition let

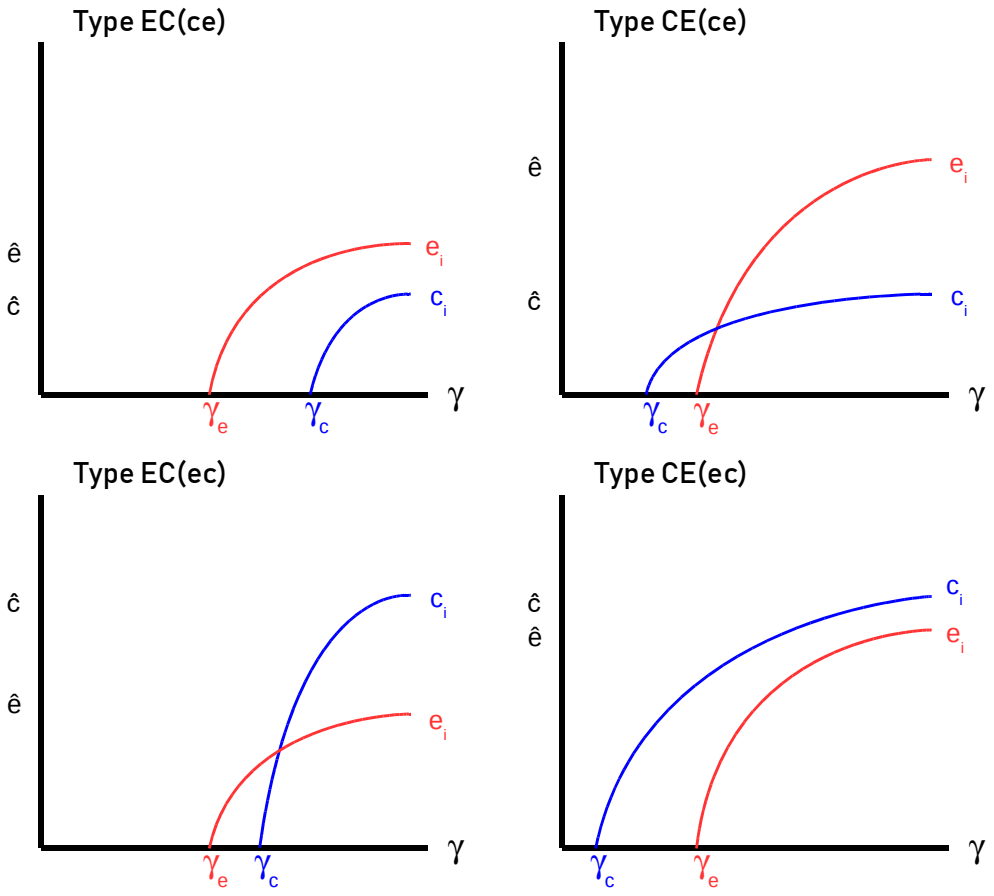

Figure 2.1: Basic Model Equilibrium Types

the punishment function only modify the norm disutility of doctrine. The individual utility function becomes:

$$U_i(c_i, d_i, e_i) \;=\; g(c_i, c_{-i}, d_i, d_{-i}, e_i, e_{-i}) \;-\gamma_i(\hat{c} - c_i)^2 \;- c_i \;- p(e_i, e_{-i})\gamma_i(\hat{d} - d_i)^2 \;- d_i \;- \gamma_i(\hat{e} - e_i)^2 \;- e_i \quad (2.9)$$

Maximizing an individual's utility results in the following optimal efforts:

$$c_i^* \;=\; \max\left[\hat{c} \;-\; \frac{1 - g_c}{2\gamma_i}\;,\; 0\right] \quad (2.10)$$

$$d_i^* \;=\; \max\left[\hat{d} \;-\; \frac{1 - g_d}{2p(\tilde{e})\gamma_i}\;,\; 0\right] \quad (2.11)$$

$$e_i^* \;=\; \max\left[\hat{e} \;-\; \frac{1 - g_e}{2\gamma_i} \;-\; \frac{p_e(\hat{d} - d_i)^2}{2}\;,\; 0\right] \quad (2.12)$$

As with the basic model, the optimal efforts imply cutoff conditions defining the minimum norm sensitivity needed for an individual to contribute that type of effort. These cutoffs are:

$$\gamma_c \;=\; \frac{1 - g_c}{2\hat{c}} \quad (2.13)$$

$$\gamma_d \;=\; \frac{1 - g_d}{2p(\tilde{e})\hat{d}} \quad (2.14)$$

$$\gamma_e \;=\; \frac{1 - g_e}{2\hat{e} \;+\; p_e\,\hat{d}^2} \qquad if\;\; \gamma_e \leq \gamma_d \quad (2.15)$$

$$\gamma_e \;=\; \frac{1 - g_e}{4\hat{e}} \;+\; \frac{\sqrt{p(\tilde{e})^2(1 - g_e)^2 - 2\hat{e}p_e(1 - g_d)^2}}{4p(\tilde{e})\hat{e}} \qquad if\;\; \gamma_e > \gamma_d \quad (2.16)$$

For purposes of characterizing equilibrium types it is useful to consider a limit case where

$N \to \infty$, meaning that the congregation is sufficiently large that individual efforts do not notably change the output of the club good or the value of the punishment function. This simplifies the cutoff points to:

$$\gamma_c = \frac{1}{2\hat{c}} \qquad \gamma_d = \frac{1}{2p(\tilde{e})\hat{d}} \qquad \gamma_e = \frac{1}{2\hat{e}}$$

It is also important to note that if one were to assume a club good function where doctrinal effort or enforcement, or both, have no effect, then simpler cutoff values for doctrine and/or enforcement would result even without the limit case assumption.

As in the basic model, total effort in any category is increasing in the corresponding norm (meaning $\tilde{c}$ is increasing in $\hat{c}$, etc.), and an increase in the enforcement norm will also drive up doctrinal effort via increased punishment. However, an increase in the doctrinal norm will reduce enforcement effort, with this effect diminishing as congregation size grows.

2.4.1 Equilibrium Types

I will again use a notation for categorizing equilibrium types of the form $CDE(cde)$, where the capital letters are lowest to highest effort cutoffs and the lower case letters are lowest to highest norm value. Recall that the capital letters can also be read as highest to lowest share of the membership contributing a particular type of effort. This would seem to result in 36 different equilibrium types where at least some members of the congregation contribute all three types of effort, and at least some members contribute no effort. However, many of these equilibrium types can be ruled out.

There are also numerous cases where due to a very high minimum norm salience or very low maximum norm salience where the above types would be truncated in some way. For several reasons, I will focus on the equilibrium types with a wide range of norm salience. First, since I am using the limit case to simplify the cutoff criteria, this would also mean

that the population would encompass the entire domain of the norm salience distribution (all positive finite numbers), which would rule out truncated versions of the equilibrium types. Second, equilibria where one of the three types of effort doesn't exist do not resemble the organizations being analyzed. A faith group with no doctrinal practices, or no enforcement of them, or no material commitments (not even worship time) is no faith at all. Finally, as pointed out in McBride (2015), even very strict faiths have at least some free riders, thus we can neglect cases that lack them.

By comparing the cutoff expressions and performing proofs by contradiction, many of the combinations of cutoffs and norm rankings cannot actually occur. After removing the impossible types, we are left with only 12 relevant equilibrium types. Three of these[2] I will call a *marginal commitment* equilibrium, meaning that the member contributing marginal effort is adding to material commitments. Six of these[3] I designate as *marginal doctrine* equilibrium types. The last three[4] are *marginal enforcement* types. In the extensions in later sections, we will be able to narrow our focus even further.

Figure 2.2 depicts the marginal enforcement and marginal commitment equilibrium types, and Figure 2.3 depicts the marginal doctrine equilibrium types:

[2]These are $CDE(dec)$, $CDE(edc)$, and $CED(dec)$, found in the right column of Fig. 2
[3]These are $DCE(ecd)$, $DCE(edc)$, $DCE(dec)$, $DEC(ced)$, $DEC(cde)$, and $DEC(dce)$, found in Fig. 3
[4]These are $ECD(dec)$, $EDC(cde)$, and $EDC(dce)$, found in the left column of Fig. 2

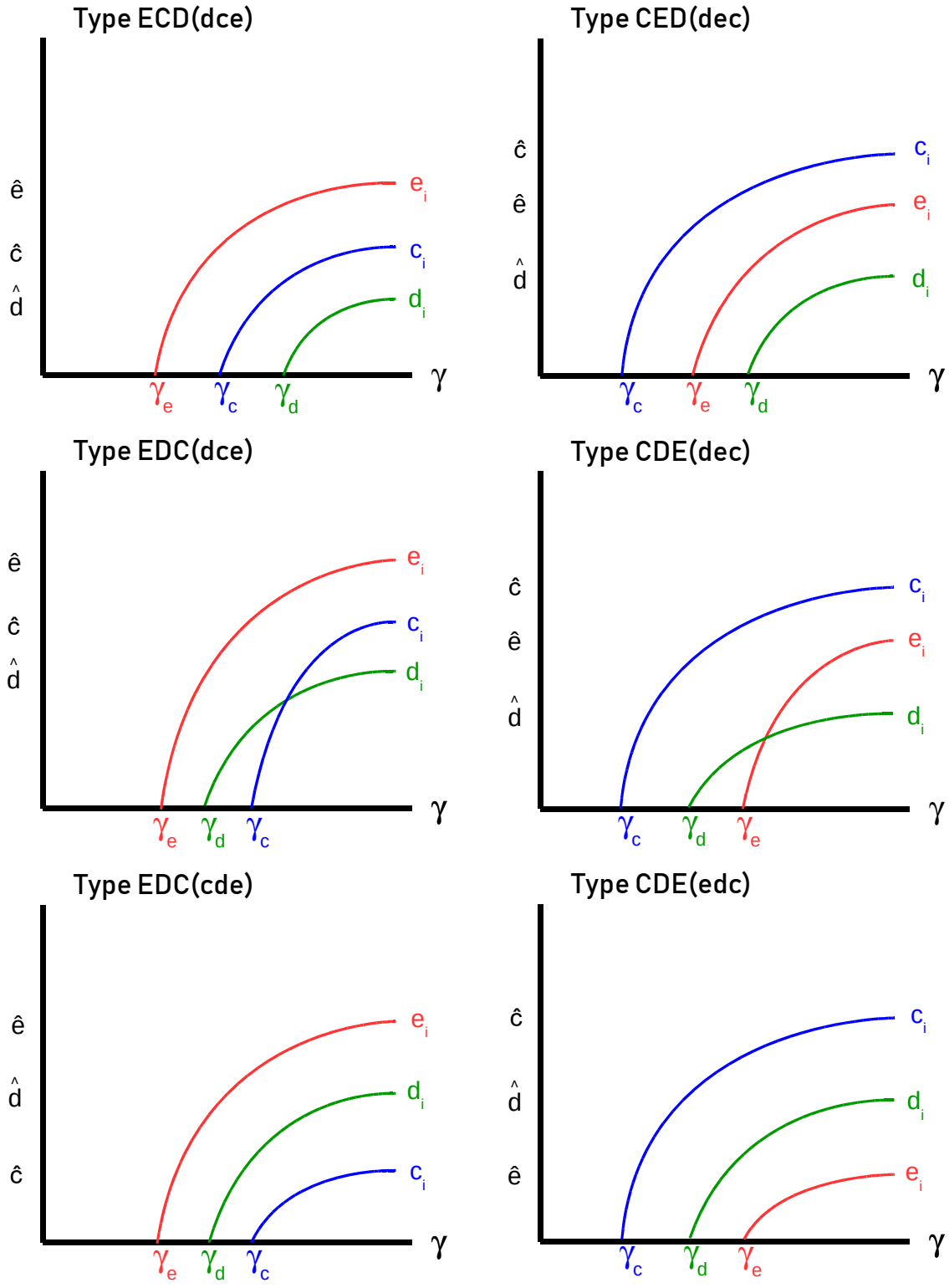

Figure 2.2: Full Model: E- and C- Equilibrium Types

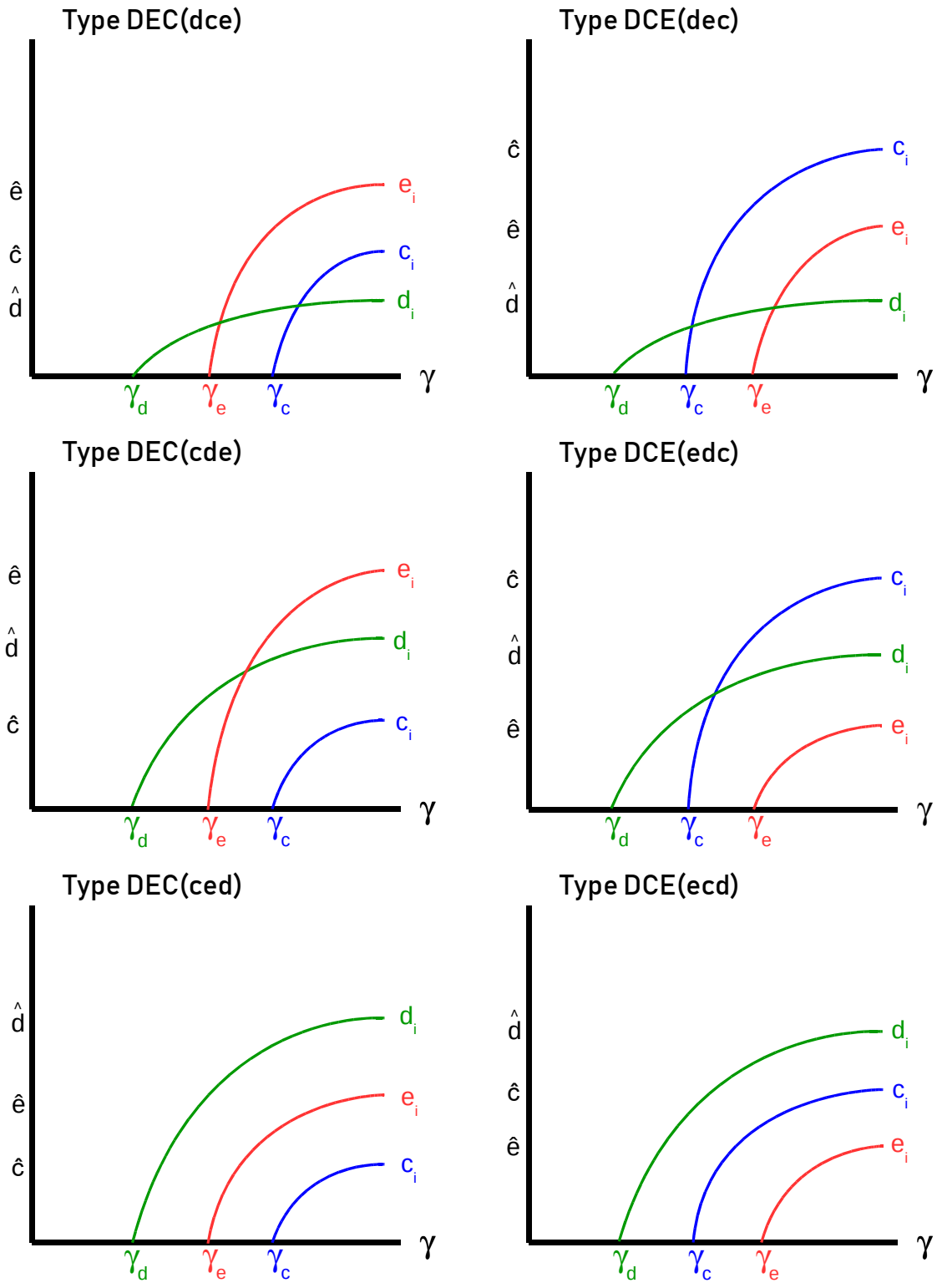

Figure 2.3: Full Model: D- Equilibrium Types

If the number of types seems extravagant, it may help the reader to notice that half of the equilibrium types involve enforcement being consistently higher than commitment (left columns), and the other half feature the reverse (right columns). The remaining variation has to do with whether and how doctrinal effort intersects the other two response functions.

2.5 Dynamic Endogenous Norms

Per above, the full model with exogenous norms yields a set of static equilibrium types. But religious norms are not truly exogenous, and we would expect the norms within a group to adjust over time to their collective desires, meaning that the distribution of norm salience types would ultimately shape the norms. To model this in the most straightforward fashion, assume an infinite horizon, discrete time setting with no discounting, and let the norms adjust via laws of motion of the following type:

$$\hat{c}_{t+1} = \frac{\sum_{i=1}^{N}(\hat{c}_i - \hat{c}_t)c^*_{i,t}}{\sum_{i=1}^{N} c^*_{i,t}} + \hat{c}_t \tag{2.17}$$

In the above equation, $\hat{c}_t$ and $\hat{c}_{t+1}$ are the current period and next period norm values, respectively. Also, $\hat{c}_i$ is the preferred norm value of individual i and $c^*_{i,t}$ is the current effort of individual i toward that type of activity. A similar law of motion would exist for $\hat{d}_{t+1}$ and $\hat{e}_{t+1}$. An individual's preferred norm values are those which would maximize their utility.

The implied mechanism is that the people who put out effort of a particular type have influence on that norm, to the degree of their current contribution. If the sum in the numerator of the first term is positive, this means that contributors prefer a higher norm and the norm will rise. The reverse occurs if the sum is negative. Obviously, this abstracts from any organizational constraints that might slow down the process of norm adjustment

or make certain individuals have more or less weight than their contributions would imply.

To clarify the relationship between norm preference and norm salience, I will again use the simplification of a "large" congregation, and have agents be fully aware of each others' reaction functions. Thus, to obtain preferred norms an individual would maximize their utility taking into account all members responses to the norms. A member whose maximizing norms would cause them to contribute to all three inputs would solve:

$$U_i(\hat{c}_i, \hat{d}_i, \hat{e}_i) \;=\; g(\tilde{c}, \tilde{d}, \tilde{e}) \;+ \frac{1}{2\gamma_i} \;+ \frac{1}{4p(\tilde{e})\gamma_i} \;- \hat{c} \;- \hat{d} \;- \hat{e} \tag{2.18}$$

Recall that tildes represent total effort of that type, and note that the norm disutility terms collapse into constants that are a function of the individuals norm salience. Solving for the optimal norms yields:

$$g_{\hat{c}} = \; 1 \qquad\qquad g_{\hat{d}} = \; 1 \qquad\qquad g_{\hat{e}} = \; 1$$

This means that for individuals who do not intend to free ride on any input, their preferred norms are such that the marginal benefit of increasing the norm will equal their marginal cost of effort. However, not all individuals will prefer norms such that they contribute to all club good inputs. To illustrate, a member whose maximizing norms would cause them to free ride on all inputs would solve:

$$U_i(\hat{c}_i, \hat{d}_i, \hat{e}_i) \;=\; g(\tilde{c}, \tilde{d}, \tilde{e}) \;- \gamma_i \hat{c}^2 \;- p(\tilde{e})\gamma_i \hat{d}^2 \;- \gamma_i \hat{e}^2 \tag{2.19}$$

$$g_{\hat{c}} = \; 2\gamma_i \hat{c}_i \qquad\qquad g_{\hat{d}} = \; 2p(\tilde{e})\gamma_i \hat{d}_i \qquad\qquad g_{\hat{e}} = \; 2\gamma_i \hat{e}_i$$

In the limit case, the terms do not interact with each other and thus intermediate cases where an individual free rides on one input but not others would use the free rider condition for that input, and the contributor conditions for the others. As such, it is clear that preferred norms go to infinity as γ approaches zero, and fall as γ increases, but once the marginal

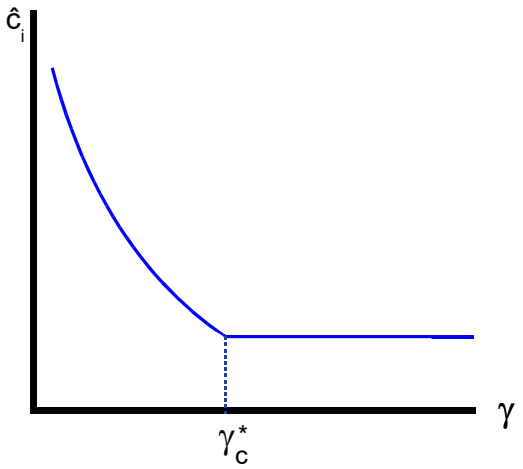

Figure 2.4: Preferred Commitment Norms

benefit equals the cost of effort, the preferred norm will be flat for all γ above that point. For example:

Thus, there will exist an equilibrium path leading to a unique steady state which will correspond to one of the static equilibrium types, and where the norm values will be such that the marginal benefit from increasing the norm equals the marginal cost of effort for all those contributing. For any norm that is below the steady state value, all contributions will come from individuals who desire the norm to be higher[5], causing it to move upwards. For any norm that is above the steady state value, all contributions will come from individuals who desire the norm to be lower, causing it to move downwards.

[5]Note that at the steady state value, the marginal contributor is the member whose desired norm is the same whether assuming they contribute or not. Thus a norm level lower than that will necessarily involve the marginal contributor having a higher norm salience, and all contributors desiring the steady state norm value (the flat portion of the curve).

2.6 Doctrinal Rigidity

It may seem unrealistic to allow the doctrinal norm to adjust as freely as the commitment and enforcement norms. While doctrinal practices do change, expectations about time and money contributed, or the degree to which doctrinal offenses are condemned should reasonably adjust much faster. To simulate this in a simple fashion, assume the dynamic model is at a steady state and that the doctrinal norm is fixed in the short run. Applying a permanent shock to the distribution of norm salience would place the model on a path to a new steady state where $\hat{d}$ remained the same, but $\hat{c}$ and $\hat{e}$ would change.

If the distributional shock causes average norm salience to fall, this means that contributions would fall for all inputs. Since the club good function is concave in its inputs, this means the marginal benefit of higher norms will rise, causing the norm values to rise to reduce marginal benefit back to marginal cost. But since $\hat{d}$ is fixed, only the commitment and enforcement norms will rise. If the shock is large enough, the new steady state may have a different equilibrium type. The reverse process would occur for a rise in average norm salience.

If we leave the doctrinal norm fixed and reduce average norm salience, eventually only six of the equilibrium types will be viable steady states. Specifically, these are the six types where the doctrinal norm is the lowest norm (top two rows of Fig. 2.2, and top row of Fig. 2.3). Because the commitment and enforcement norms both move up (and to the same degree), the ranking of those norms will remain consistent. Thus, if a faith group starts in an equilibrium in one of the left columns of Fig. 2.2 or Fig. 2,3, the new steady state will also be in the left column of those two figures, and vice versa. Which of these types result is going to be determined by the strength of the punishment function. For a very strong punishment function, expect the top row of Fig. 2.3. For a weak punishment function, expect the top row of Fig. 2.2.

If we instead increase norm salience, eventually only two of the types will be viable, specifi-

cally $DEC(ced)$ and $DCE(ecd)$ (the bottom row of Fig. 2.3). These two equilibrium types are characterized by doctrinal observance being far more likely, and more substantial than either material commitment or enforcement. Since this results from very high norm salience, we can identify these equilibrium types as the pattern of mature faiths with high cultural acceptance.

To clarify, movement *across* the columns is only possible if changes are made to the club good function, altering the relative productivity of commitment versus enforcement. Movement *within* the columns of figures 2.2 and 2.3 can occur due to changes in average norm salience, or from changing the effectiveness of the punishment function.

2.6.1 Discussion

Using the above observations about how doctrinal rigidity interacts with changes in norm sensitivity, we can examine a number of examples that illustrate the theory. Before discussing specific examples, it is worth examining the secularization thesis, and why the advance of modernity has not uniformly diminished the role of religion in public life. If we describe secularization as a general decline in norm salience for all faiths, and assume relatively rigid doctrines, the response will be a rise in commitment and enforcement effort. In other words, the religious groups that survive secularization pressure will expect more material contributions from their members and expend more effort policing members' behavior. For example, in the United States, the rise of evangelical Christianity in the wake of the Sexual Revolution is thus not paradoxical at all – it is predicted by the theory.

The fact that high commitment and enforcement norms are associated with low norm salience has implications for new religious movements (NRMs). First, it explains why many NRMs treat those who violate their doctrines very harshly, and also the common phenomenon of communal living arrangements amongst these groups. Second, when such groups mature

and become more accepted, the earlier strict practices begin to decline. The history of the Methodists provides a useful example. Originally a reform movement within the Anglican faith, Methodists broke away in the late 18th century[6] Early Methodists promoted plain dress, periodic fasting, and abstinence from alcohol and gambling. By the mid-19th century, Methodism was the largest Christian denomination in the United States and after the Civil War it was the denomination most associated with the Temperance movement. Over the course of the 20th century, however, the strict rules were relaxed in most congregations. For example, the current United Methodist Church stance on alcohol is merely to discourage it – restrained consumption is acceptable.

We can also use the interaction between norm salience and doctrinal rigidity to explain the formation of sects. Contrast the history of the Unification Church of Sun Myung Moon with the history of Soka Gakkai Buddhism. As described below, these two faiths emerged contemporaneously in northeastern Asian countries and attempted to expand into North America at a similar time as well. However, the Unification Church, ironically given its name, has splintered into competing sects, while Soka Gakkai has remained unified. This cannot be explained as being due to the Unification Church being larger, since Soka Gakkai has more members (12 million) than all variations of the Unification Church (3 million) combined[7]. The difference in outcomes lies in a doctrinal reform that Soka Gakkai undertook during its overseas expansion.

The Unification Church, or more formally, the "Holy Spiritual Association for Unification of World Christianity", started in mid-20th century Korea. The faith had a charismatic founder, Sun Myung Moon (1920-2012), who was a staunch anti-Communist, perhaps in part due to his imprisonment by the North Korean regime in the late 1940s. Before and during the Korean War, Moon had been a Presbyterian, but post-war he founded the Unifi-

[6]This is partly due to Methodism taking hold in the American colonies. In England, Methodism largely existed as an upgrade or add-on to Anglican worship. But in the American context, the faith became distinct since many experienced no other way to be Anglican other than the strict manner preferred by Methodists.

[7]Both of these numbers are from these faiths' own figures, and are likely inflated. Nonetheless, the relative size is clear.

cation Church. By 1970, the faith began to expand outside of South Korea. In the United States, the Unification Church was plagued by scandals involving right wing political connections, accused of brainwashing members and was often labeled a cult. Several sects have formed in the wake of Moon's death, with the two most prominent being associated with Moon's widow Hak Ja Han, focused in Korea, and another with one of Moon's sons, Hyung Jin Moon, focused in the United States. In the context of the model, sect formation came about as a result of the faith experiencing lower norm salience in the Western world, and these congregations evolved into a different equilibrium type.

Soka Gakkai, Japanese for "Value Creation Society", is a variation of Nichiren Buddhism, started in 1930s Japan. Soka Gakkai had a pair of charismatic founders, Tsunesaburo Makiguchi and Josei Toda, both of whom were imprisoned by Imperial Japan for criticizing State Shinto. Makiguchi died in prison, and Toda continued the movement and oversaw its rapid growth post-war. In the 1970s, the faith began expanding overseas, but implemented a series of doctrinal reforms paring the faith down to key essentials, and designating many practices optional. See Dawson (2001) for more detail. Soka Gakkai has spread to more than a hundred countries, and has yet to experience a break-away sect. Because the doctrinal reforms did not end any practices, just rendered them optional, it meant that observing the faith would be less burdensome outside of a Japanese context. The drop in norm salience that overseas congregations would face was thus reduced, allowing them to remain in the same equilibrium type in the new cultural context. Thus, the newer overseas temples would still have the same relative emphases on commitment, enforcement and doctrine as the older temples in Japan – inhibiting sect formation.

As noted above, the equilibrium types where the commitment norm exceeds the enforcement norm can only shift into the reverse scenario due to a change in the club good function that alters the relative importance of these two inputs. While outside the scope of the model,

faith groups may have some control over this weighting[8]. If such changes are difficult or uncommon, then once we identify a faith group as having high relative commitment or high relative enforcement, its equilibrium type is then a matter of the norm salience distribution it faces, and relative effectiveness of punishment.

Specifically, for a high relative enforcement club good, high norm salience will push a faith toward type $DEC(ced)$, while low norm salience with weak punishment yields type $ECD(dce)$ and low norm salience with strong punishment yields type $DEC(dce)$. For a high relative *commitment* club good, high norm salience will push a faith toward type $DCE(ecd)$, while low norm salience with weak punishment yields type $CED(dec)$ and low norm salience with strong punishment yields type $DCE(dec)$.

For an example of the last type, the Amish offset substantial doctrinal restrictions on dress and technology by use of intentional communities with high material commitment requirements. Enforcement is strong, but young people are allowed a lot of leeway (the rite of passage called Rumspringa) and anyone can return to the faith after a period of lax behavior. Technological restrictions make these material contributions essential to the faith remaining viable, especially since children born into the faith who encounter the wider world would have decreasing norm salience as time goes on. The $DCE(dec)$ type is a basin among the equilibrium types that will occur when low norm salience persists, and the punishment function is very effective.

A notable contrast can be found among Jehovah's Witnesses, who also encourage a high degree of material commitment and enforcement. There are an estimated 20 million Witnesses worldwide, but only about 8 million are considered "publishers" – i.e. people engaged in preaching to non-members (a time-based mode of commitment). More than this, there are substantial internal mechanisms for punishment, including several tiers of shunning up to,

[8]Mathematically, the club good function could be Cobb-Douglas and the exponent parameters would be adjustable. However, this would seem to entail a major change in community organization, likely to be even less common than doctrinal changes.

and including, outright expulsion from the faith. Doctrinal requirements for modest dress, patriarchy, and restraint of vices are widely observed, placing the Witnesses at odds with larger society. This would be an instance of the $DEC(dce)$ equilibrium, similar to the equilibrium for the Amish described above, but with the relative importance of enforcement and commitment reversed.

With high norm salience, the doctrine norm will take precedence. An example of this scenario is the cultural dominance of Catholicism in the Boston Archdiocese, where about half the local population identifies as Catholic. Many members exhibit little commitment outside of attending the occasional mass, and do nothing to enforce doctrine, yet culturally the church continues to have massive influence – even in the wake of recent scandals. This would be an instance of the $DCE(ecd)$ equilibrium. Of course, if norm salience were to fall enough, this equilibrium would shift. Although, there have been some recent reorganizations in this Archdiocese, there is not yet signs of escalating enforcement or commitment norms.

As a final example, consider the case of evangelical megachurches. Such churches are frequently found in the suburbs of rapidly growing urban areas, where they can draw members from a wide area. Few members are asked to do any more than attend a weekly sermon, which is usually plain spoken and light on theological terminology. Doctrinal requirements are mostly in the realm of private conduct, and offering plates are rarely used. On the other hand, the private conduct that is regulated (mostly related to sexuality) is frequently the topic of sermons and a major focus of members. This would be the $ECD(dce)$ equilibrium, where enforcement dominates the other norms, and is a product of low norm salience (drawing members from a wide area of a secularizing society) and where punishment is weak.

2.7 Conclusion

This paper builds upon the familiar economics of religion framework, where an organization providing a club good uses strict practices to screen out free riders, by decomposing the concept of strictness. Key to the expanded environment is the addition of enforcement, where some of members' effort increases the penalty for failing to contribute. In addition, since different portions of the population will contribute different types and amounts of effort, the environment results in a multi-dimensional classification system for religious groups. The dynamic extensions pare down the classifications to a handful of stable outcomes that would be most commonly observed.

The model indicates that a mature faith whose norms are highly salient to the population will likely exhibit marginal contributors following the doctrines of the faith but whose material and enforcement efforts are negligible. By contrast a faith whose norms have low salience among the population can exhibit marginal enforcement or marginal material commitment, if the effect of punishment is weak. In other words, we would expect most faiths to exhibit marginal doctrinal observance, except in the case where norm salience is low *and* punishment is weak.

Finally, the use of a norm-utility framework helps neatly explain variations in effort within a faith community. The norms reflect an ideal for members to aspire to, but actual effort falls short to varying degrees based on individual norm salience. Changes in the distribution of norm salience, either due to secular trends or due to expansion of a faith into new cultural contexts, can result in a new equilibrium that transforms the faith group or causes a schism.

Chapter 3

Capital Reallocation, Search-in-Use, and Growth

I develop a growth model with a decentralized market for physical capital that allows for reallocation of both idle and currently active units. Capital is reallocated among firms with a common technology, yet generates a dispersion of rental rates, consistent with empirical observations of commercial real estate markets. In equilibrium some measure of capital is always idle due to search frictions, and this reduces the endogenous savings rate. Steady-state output is found to be increasing in the degree of "search-in-use" – the extent to which reallocation occurs for capital that is currently active. Extensions allow for endogenous savings and transfers to increase saving. Calibration demonstrates that the model can easily account for observed savings rates below that implied by a standard neoclassical growth model.

3.1 Introduction

Effective allocation of capital to productive use is vital to long-term economic growth, as well as providing the incentive for capital accumulation. Many physical capital markets exhibit a chronic share of capital that is idle, as well as dispersion in rents that are not easily explainable by heterogeneity in the capital itself. My proposed theory is that frictions associated with capital reallocation diminish the incentive to save, create a dispersion in rates of return on capital, and that policies that diminish these frictions will lead ultimately to higher growth.

Recent literature has established that markets for physical capital exhibit substantial frictions, and have large secondary markets. The paper most responsible for increased interest in these markets is Eisfeldt and Rampini (2006), which estimated that reallocation of existing capital amounts to 30% of annual expenditures on physical capital and found substantial dispersion in Tobin's q, which implies dispersion in the value of capital. Slightly earlier, Ramey and Shapiro (2001) demonstrated that capital goods are often on the market while in operation for an extended period before sale and transfer occurs, a situation correlated with frictions due to specificity of capital.

In Kurmann and Petrosky-Nadeau (2007), commercial real estate vacancy rates are among many examples given of capital market frictions, and studies such as Tse and Fischer (2003) show that high commercial vacancies are an international phenomenon. This example of idle capital is particularly intriguing, since office and retail space is among the most homogeneous of capital goods. Location and amenities do vary, of course, but variations in rents are found even with the same building (see the discussion of concessions in the literature section).

The application of search models to capital markets is relatively new, and most attempts so far have placed capital goods in one of two states: not-searching but active, or searching but idle. However, in the labor search literature we find that allowing for search-on-the-job can

provide a richer explanation of wage dispersion and capture the phenomenon of job-to-job transitions with no intervening period of unemployment. Comparably, in capital markets one frequently finds *search-in-use*, where currently active capital is also being offered for sale or rent in secondary markets, as well as dispersion in rental rates. Thus it seems natural to adapt search-on-the-job models into a model of search-in-use.

The basic structure of the model is a Solow-type growth model with a frictional capital market similar to the labor market model found in Burdett and Mortensen (1998). The frictional capital market is coupled with a frictionless labor market both for tractability and to allow the growth features of the model to mimic the widely known Solow model. Most macro models with both labor and capital make the reverse simplification – a frictional labor market and frictionless capital markets. This modeling choice is further justified below by providing evidence that frictions are at least as significant in capital markets as in labor markets.

Idle capital in the model is directly analogous to unemployment in labor models, in that capital owners are searching for firms to use their capital but are currently unmatched. This is not to be confused with capacity utilization where a firm has access to capital that they are currently declining to use. Capacity utilization would reflect capital goods that are "outside the capital force" to continue the labor market analogy. An example of a measure that would properly capture "unemployed" capital is vacancy rates for non-residential real estate, which is the primary setting in mind for the model.

The model will include several known features of physical capital markets. First, there will be equilibrium unemployment, caused by search frictions. Changes in use will be described partly by exogenous deactivation (similar to separation in labor models) but also by search-in-use, where units of capital can find increasingly better rental rates. Finally, the structure of rents will exhibit dispersion.

3.2 Literature and Motivating Evidence

This section will first address the existing evidence for substantial frictions in markets for physical capital. After establishing these features, I will discuss recent theoretical explorations of capital markets, most of which are presently working papers.

3.2.1 Empirical Evidence

General evidence of capital market frictions can be found in Eisfeldt and Rampini (2006). As noted above, reallocation is a sizable share of capital expenditures. Somewhat surprisingly, reallocation is procyclical, i.e. used capital markets "seize up" during recessions. This phenomenon strongly implies search frictions in capital markets, since otherwise one would expect used capital to be quickly reallocated as firms restructure in response to economic conditions. While the authors do not provide direct evidence of rental rate or price dispersion, they do show persistent dispersion in Tobin's q (stock value over asset value). In the model below, profits over capital stock would exhibit dispersion as a side effect of the underlying rent dispersion, which is consistent with this data.

The phenomena of idle capital is related to capital misallocation studied in Hsieh and Klenow (2009), although the proposed mechanism is different. Hsieh & Klenow's model does not have idle capital, but instead has capital that is located in less productive firms. By comparing the dispersion of marginal products across countries they find that much of the cross-country differences attributed to TFP may in fact be allocative inefficiencies. Similarly, the model below would imply that countries with identical technology could still have substantially different levels of output. The institutional problems (corruption and legal monopoly) that H&K draw attention to, could also create differences in capital market frictions that would create similar performance gaps.

A well-documented aspect of capital goods that can generate search frictions is specificity. The study of aerospace plant closings in Ramey and Shapiro (2001) found that although the highly specialized equipment used in that industry was difficult to sell cross-industry, such sales did occur – typically with a haircut. Also, equipment up for sale nearly always continued in use while the owner searched for a buyer.

The above papers establish the basic contours of capital markets. The following stylized facts will be elaborated further below:

- Substantial frictions with persistent idle capital
- Robust rental and secondary markets (establishing search-in-use)
- Dispersion of rental rates and prices

To establish the degree of frictions present, I obtained national level data on non-residential real estate from Reis, Inc. As can be seen below, vacancy rates in these markets are large and exceed the labor market unemployment rate over the entire series. Vacancy rates in the $5 - 20\%$ range are not unique to the U.S., and have been found in the UK and Australia (Tse and Fischer (2003))an even in rapidly developing economies like China (Ke and White (2013)). Another relevant aspect of non-residential real estate is the widespread practice of concessions. This is where property owners reduce the asking rent for a property in order to lure and retain tenants. Concessions data is anonymous, but aggregated such that an effective rent can be calculated relative to the asking rent. This practice is so common that commercial real estate over the period for which data was available, the average effective rent is only 85% of the asking rent. The practice of concessions is masking a high degree of rent dispersion, even across tenants in the same building (which is why survey respondents insist on the data being anonymous).

Vacancy rates appear to act as a signal for new construction, per Eppli and Shilling (1995).

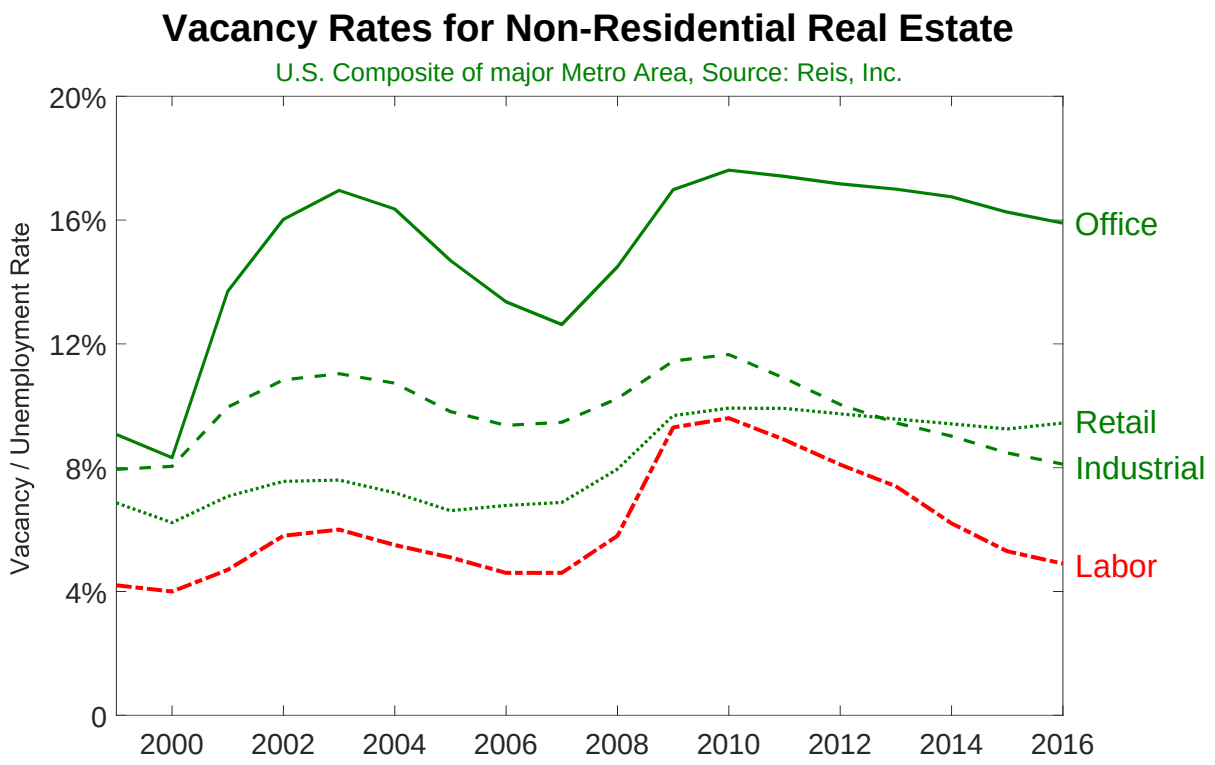

Figure 3.1: Vacancy Rate Comparison

However, the lag between initiation of a project and finished construction often leads to vacancies even for brand new properties. Rent levels for real estate are inversely related to vacancies (Wheaton et al. (1997)), but there is also variation due to time-on-the-market (Piazzesi et al. (2013)). This demonstrates the core nature of rent dispersion in the model, that rents are higher when there is a brief vacancy (search-in-use) versus a long vacancy (separation).

As noted above, Kurmann and Petrosky-Nadeau (2007) point to commercial real estate vacancy rates as evidence of capital market frictions. In addition, the authors use inventory data from manufacturers of physical capital to determine a hazard rate for delivery of new capital. That these rates are measurably below unity (as low as 0.7 for heavy machinery) indicates that even *new* capital is subject to search frictions in initial allocation. Other general evidence for capital price dispersion can be found across sectors as in Arnade and

Gopinath (1998), and contrasting new versus used prices as in Lanteri (2017). Price data in Lantieri also demonstrates higher volatility in the price dispersion of used capital.

Additional evidence comes from secondary markets for commercial aircraft. Similarly to the real estate data, Pulvino (1998) finds that aircraft sales prices vary inversely with search time. The secondary markets for aircraft are quite robust due to leasing by the first user of the aircraft, per Gilligan (2004). According to Gavazza (2011), used aircraft sales are more than three times the volume of new purchases, even though the search times are usually months long. In a follow-up, Gavazza (2016) compared aircraft secondary markets to a hypothetical Walrasian benchmark and estimated that 18% of commercial aircraft are misallocated at any given time.

The choice of a single depreciation rate for the model, regardless of whether capital is active or idle is a convenient simplification, but also has some empirical support. Per Boucekkine et al. (2009), obsolescence is a major component of depreciation, and this will eventually render capital worthless regardless of use.

3.2.2 Theoretical Literature

As noted above, the model below draws from Burdett and Mortensen (1998), a labor search model where a distribution of wages forms among homogeneous workers entirely due to search frictions and differences in firm size. However the quantity of workers is a continuum of fixed size, and the main difference in the model below is that units of capital (the equivalent of a BM worker) are a stock that grows with investment and shrinks with depreciation. In Robin and Roux (2002), the authors resolved a concern that the Burdett & Mortensen dispersion result was dependent on the matching technology, showing that only an economy where job finding rates were perfectly proportional to firm size would result in a degenerate distribution.

There are a wide variety of other labor search models, many of which employ the framework from Mortensen and Pissarides (1994). The model below avoids this framework since the baseline model has single worker firms (analogously, a single unit of capital) and a common wage rate (i.e. rental rate), and deviations from this require the addition of ex ante heterogeneity. However, since capital good heterogeneity can manifest through everything from scale to irreversibility to productivity differences, the specific form of any ex-ante heterogeneity will drive the results of the model. Using a framework where all of these differences are subsumed into the matching technology results in a rent distribution even when capital goods are homogeneous.

In Bai et al. (2011) a matching market for final goods causes variable demand and this is combined with each firm having a location and a matching market for capital goods. Their model yields plausible capacity utilization dynamics, and is one the first papers using search frictions to examine capital markets. Another notable model is Kurmann (2014), where a capital goods matching market incentivizes firms to overinvest in capital. While technically all capital in Kurmann's model is active (there is just an inefficient quantity), the model is suggestive of how search frictions could be a useful paradigm for describing capital markets. This overinvestment result also appears in a subsequent paper, Kurmann and Rabinovich (2018) – a variation on Lagos and Rocheteau (2009). In both of these models, firms are using overinvestment to improve their bargaining position.

Ottonello (2015) directly labels capital unused due to allocative frictions as unemployed, an intentional analogy to labor markets. Their basic model has a series of submarkets, each with its own market tightness, and firms select which to enter in order to obtain capital. There is an extension which yields dispersion, but this is reliant on adding ex-ante heterogeneity. The allocations are efficient due to directed search, where both capital owners and firms have complete control over which submarket they enter.[1]

[1]The haircuts for cross-industry sales of capital mentioned in Ramey and Shapiro (2001) suggests free entry into submarkets may be too strong an assumption. On the other hand, my model places bargaining

Another recent capital search model is Shi and Cao (2015). They investigate the procyclical behavior of capital reallocation (which is the opposite of what many models would predict), and show that a frictional capital market can result in less clearing in recessionary conditions. Unlike other attempts to model these markets, new capital enters its life in an active state and is later able to be sold in secondary markets[2]. Dispersion in Tobin's q is addressed in Dong et al. (2016), a capital search model where firms are able to vary in size. Like all of the above models, there is no search-in-use, and all dispersion is created by ex-ante heterogeneity (in this case, in productivity).

3.3 Basic Model

In all versions of the model, agents are infinitely lived and represented in continuous time. There exists a unit continuum of firms that are non-discounting profit maximizers with a Cobb-Douglas production technology accepting labor and capital inputs. There also exists a measure L of households that are non-discounting consumption maximizers. Finally, there is a unit measure of non-discounting, profit maximizing intermediaries that convert household savings into capital and pay a dividend to the households. The intermediaries rent their capital to the firms, and households provide labor to the firms for wages. In the basic model, households have an exogenous savings rate s and consume the remainder of their income. Many of these assumptions will be relaxed in the extensions.

power in the hands of capital users which would seem to err in the opposite manner.

[2]Frictionless matching of new capital is not empirically accurate either, but new capital allocation certainly faces *fewer* frictions than used capital.

3.3.1 Environment

The firms' production technology is:

$$Y = zK^{\alpha}L^{1-\alpha}$$

where z and α are common to all firms, K is the amount of capital rented by the firm and L is the amount of labor the firm employs. Labor is obtained through a frictionless market at wage w. To obtain capital, firms post an offer of a rental rate r. The offer distribution will be denoted $P(r)$, and the distribution of capital actively used in production will be called $A(r)$.

Intermediaries attempt to find firms to rent their capital to, and inactive units of capital will encounter firms at Poisson rate λ. A share of active capital $\eta \in (0, 1)$ will also search[3], thus active capital encounters new firms at the rate $\eta\lambda$. Active capital becomes inactive at an exogenous deactivation rate d. The share of capital that is vacant (idle), v, will be determined endogenously.

All of the above is directly analogous to the wage posting mechanism and matching rates featured in Burdett and Mortensen (1998). A key difference is that the supply of the input good is endogenous. Specifically, the intermediaries accumulate a stock of capital that depreciates at rate δ and this is funded by accepting investment of households' savings in return for paying them a dividend i per unit of capital created. The investment market is frictionless, and savings are converted to capital $1:1$. Finally, households save an exogenous share of total output s. The figure below depicts the model environment.

[3]If $\eta \geq 1$, this could result in the reservation rent being zero or even negative. In a labor market model, this situation (akin to an unpaid internship) is potentially reasonable, but capital markets do not appear to exhibit non-positive rents and thus I exclude these parameter values as unrepresentative of the phenomena of interest.

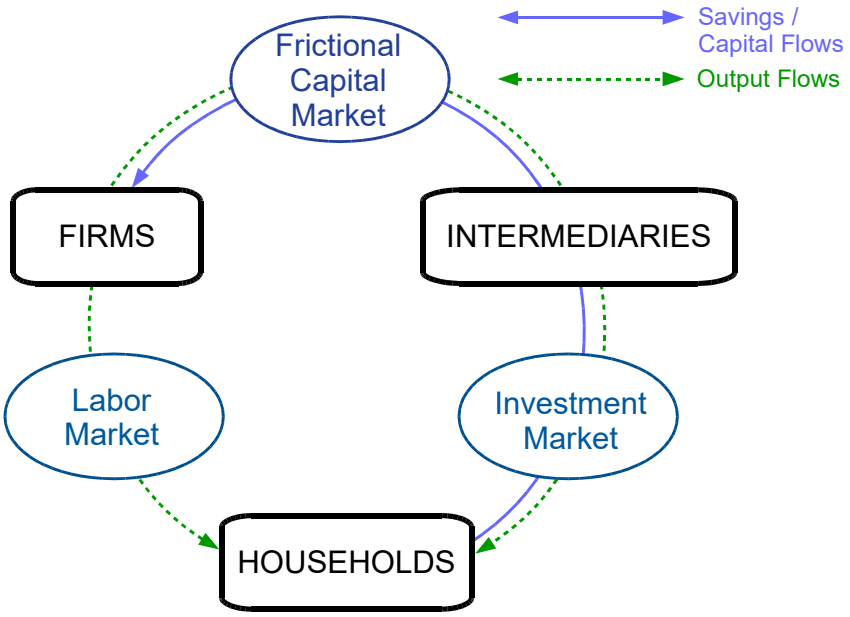

Figure 3.2: Model Diagram

3.3.2 Equilibrium

Given the growth model elements of the environment, I focus on a steady-state equilibrium where active capital per worker is constant. A steady state implies a stationary rent distribution, wage level, and a specific vacancy rate for capital, as described below. Using the production function and aggregate quantities, we can express a law of motion for the overall capital stock:

$$\dot{K} \;=\; sz\left((1-v)K\right)^{\alpha} L^{1-\alpha} - \delta K \tag{3.1}$$

Thus steady state active capital per worker (k_a) will be:

$$k_a \;=\; \left(\frac{sz(1-v)}{\delta}\right)^{\frac{1}{1-\alpha}} \tag{3.2}$$

This capital per worker expression is just the familiar Solow result reduced by the effect of vacancies. To pin down the vacancy rate, we can observe the flow of capital from the active (K_A) to inactive (K_I) state.

$$\dot{K}_I \;=\; dK_A - \lambda K_I - \delta K_I + szK_A^{\alpha} L^{1-\alpha} \tag{3.3}$$

Since $\dot{K}_I = 0$ in a steady state, and $v \equiv K_I/K$ by definition:

$$v^* \;=\; \frac{d+\delta}{\lambda + d + \delta} \tag{3.4}$$

The vacancy rate expression is, unsurprisingly, very similar to that found in labor search models where the equilibrium unemployment rate is the factors causing separation, divided by separation plus the finding rate. Equations (2) and (4) will be useful for substitution later. Turning to the firm's problem, their profits are

$$\pi(r) \;\;=\;\; z\, n(r)^{\alpha} \ell(r)^{1-\alpha} \;-\; r\, n(r) \;-\; w\ell(r) \tag{3.5}$$

where firms offer a rent r to obtain capital quantity $n(r)$ and hire labor quantity $\ell(r)$. Since the labor market is frictionless, there will be a single wage equal to the marginal product of labor and this in turn means that active capital per worker will equalize across all firms in equilibrium. This is true along any point on an equilibrium path, not just at a steady state. This means that a firm's choice of rent will, on its own, fully determine their level of profit. At the steady state, $\frac{n(r)}{\ell(r)}$ would equal k_a for all r.

When a unit of capital encounters a firm, the intermediary owner will have the option to accept or reject the firm's rent offer. This implies a reservation rent for inactive capital ($\underline{r}$) and that any offers made to active capital will need to exceed the current rent in order to be accepted. As a result of the latter (intermediaries "trading up" for higher rates of return), a firm can potentially obtain more capital by offering a higher rental rate. The steady state

flow of capital in and out of active status up to a particular offered return r is:

$$\underbrace{v\ \lambda\ P(r)}_{inflows} \ = \ \underbrace{(1-v)A(r)\ \{d+\delta+\eta\lambda[1-P(r)]\}}_{outflows} \tag{3.6}$$

Rearranging this yields an expression for the active capital distribution in terms of the posting distribution:

$$A(r) = \frac{v\lambda P(r)}{(1-v)(d+\delta+\eta\lambda[1-P(r)])} \tag{3.7}$$

The law of motion for capital flows through firms offering rent r would be:

$$\dot{n}(r) \ = \ -\{d+\delta+\eta\lambda[1-P(r)]\}\ n(r) \ + \ v\lambda K \ + \ (1-v)\eta\lambda A(r)K \tag{3.8}$$

At a steady state, where $\dot{n}(r)=0$, and substituting for $A(r)$ yields capital per firm in terms of the posting distribution:

$$n(r) \ = \ \frac{v\lambda(d+\delta+\eta\lambda)K}{\{d+\delta+\eta\lambda[1-P(r)]\}^2} \tag{3.9}$$

Since $P(r)$ is a cdf, we observe that capital per firm will be increasing in r. If there were a discontinuity in the posting distribution, $n(r)$ would be weakly decreasing in r. However, a discontinuity in the rent distribution is not incentive compatible. Consider a discontinuity from r_1 to r_2 with $r_2 > r_1$. A firm at r_2 would be able to poach capital from any firm posting a rent between $\underline{r}$ and r_1. But this would still be the case if that firm lowered its rent offer to some value r' where $r_1 < r' < r_2$ and the firm's profits would increase since their cost of capital would fall with no change in the amount of capital they would attract. Thus $P(r)$ must be continuous in equilibrium, meaning that capital per firm is strictly increasing in the rent offer.

Since firms can freely change their rent at any time, and since any $r \geq \underline{r}$ will yield a positive

quantity of capital, paying a rent above the reservation rent must be at least as good as the outcome when offering $\underline{r}$. But if a firm is earning high profits at a higher rent, this will induce other firms to increase their rent above them, which in turn would reduce the amount of capital they are able to attract and lower their profits. Thus, in equilibrium, firms will end up with rent offers such that they are indifferent between continuing to post r or switching to posting $\underline{r}$ forever. With no discounting, this is an equal profit condition. Profit equivalence and the $n(r)$ expression yields the posting distribution:

$$P(r) \;=\; \frac{(d+\delta+\eta\lambda)}{\eta\lambda}\left(1 \;-\; \left(\frac{\alpha z k_a^{\alpha-1} \;-\; r}{\alpha z k_a^{\alpha-1} \;-\; \underline{r}}\right)^{\frac{1}{2}}\right) \tag{3.10}$$

Note that $P(\bar{r}) = 1$ which implies:

$$\bar{r} \;=\; \alpha z k_a^{\alpha-1} \;-\; (\alpha z k_a^{\alpha-1} \;-\; \underline{r})\left(\frac{d+\delta}{d+\delta+\eta\lambda}\right)^2 \tag{3.11}$$

Turning to the intermediaries, we need to determine the reservation rent that they will accept from firms. Given that being matched reduces the arrival of new matches (due to the search-in-use parameter, η), a very low rent offer will be rejected since it will be better to wait for a higher offer. Value functions for a unit of capital require a distinction between inactive capital V_I and active capital at a particular rental rate $V_A(r)$. Without discounting, the intermediary is assessing the expected useful life of their capital, determined by the depreciation rate.

$$\delta V_I \;=\; \lambda\left[\int_{\underline{r}}^{\bar{r}} V_A(x)\, dP(x) \;-\; V_I\right] \tag{3.12}$$

$$\delta V_A(r) \;=\; r \;+\; \eta\lambda\left[\int_{r}^{\bar{r}} V_A(x) \;-\; V_A(r)\; dP(x)\right] \;+\; d\left[V_I - V_A(r)\right] \tag{3.13}$$

Note that $\bar{r}$ is the rental rate offered by the largest firm. Combined with the condition that $V_A(\underline{r}) = V_I$, the above equations can be solved to obtain:

$$\underline{r} \;=\; (1-\eta)\lambda \int_{\underline{r}}^{\bar{r}} \frac{1-P(x)}{d+\delta+\eta\lambda[1-P(x)]}dx \tag{3.14}$$

We can then evaluate the reservation rent condition above to produce an equation relating $\underline{r}$ to $\bar{r}$, that combined with (11) obtains a closed form expression for the reservation rent.

$$\underline{r} \;=\; \frac{\alpha\delta}{s}\left(\frac{\lambda+d+\delta}{\lambda}\right)\left[\frac{(1-\eta)\eta\lambda^2}{(d+\delta+\eta\lambda)^2 \;+\; (1-\eta)\eta\lambda^2}\right] \tag{3.15}$$

And thus, the rent dispersion is:

$$\bar{r} \;-\; \underline{r} \;=\; \frac{\alpha\delta}{s}\left(\frac{\lambda+d+\delta}{\lambda}\right)\left[\frac{(d+\delta+\eta\lambda)^2 \;-\; (d+\delta)^2}{(d+\delta+\eta\lambda)^2 \;+\; (1-\eta)\eta\lambda^2}\right] \tag{3.16}$$

Note that setting the search-in-use parameter (η) to zero will remove all dispersion, and the posting and active capital distributions will both be degenerate. The ex-post heterogeneity in rental rates is entirely due to intermediaries' ability to "trade up" when they receive a higher rent offer. Due to perfect competition in the investment market, the intermediaries will pay a common dividend to all households based on the expected rent paid on active capital.

$$\mathbb{E}r(1-v)K \;=\; iK \tag{3.17}$$

And evaluating $\mathbb{E}r$ by integrating over the active capital distribution yields:

$$\mathbb{E}r \;=\; \alpha z k_a^{\alpha-1}\left[\frac{(d+\delta+\eta\lambda)^2 \;-\; (d+\delta)^2 \;-\; \frac{v}{1-v}\eta\lambda^2}{(d+\delta+\eta\lambda)^2 \;+\; (1-\eta)\eta\lambda^2}\right] \tag{3.18}$$

The households in the basic model have an exogenous savings rate and are price takers in both the labor and investment markets, so their behavior is automatic. However, since s is

Factor	Increasing in:	Decreasing in:
Output	α, s, z, λ	d, δ
Vacancy Rate	d, δ	λ
Rent Dispersion	α, η, d, δ	s, λ

Figure 3.3: Comparative Statics

defined as the share of total output, the above equilibrium solution has assumed that the output share going to the firms is no larger than $1 - s$. This is guaranteed to hold for any $s \leq 1 - \alpha$, since the labor share of income will always flow to the household, regardless of the amount of frictions in the capital market. Given the typical observed values of the labor share and savings rate, this is a trivial restriction.

Altogether, an equilibrium is a tuple (v, K, i, w, $\underline{r}$, P, π), such that both the labor and investment markets clear, and where the capital market follows:

$$\pi(r|\underline{r}, P) = \pi \quad \forall \ r \ on \ P$$

$$\pi(r|\underline{r}, P) \leq \pi \quad otherwise$$

At a steady state, equations (2), (4), and (16) will hold as well.

3.3.3 Analysis

As expected, the basic model results in equilibrium idle capital (the vacancy rate), for any positive, finite search rate λ. Also, as the search speed falls (capital market frictions increase), steady state output is reduced. The existence of search-in-use creates an endogenous rent dispersion, however in this environment the search-in-use parameter (η) does not affect overall output. In the next section, endogenous savings will create a link between search-in-use and steady state output. The basic model yields intuitive comparative statics.

Note that the direct effect of increased savings is to make it easier for firms to obtain

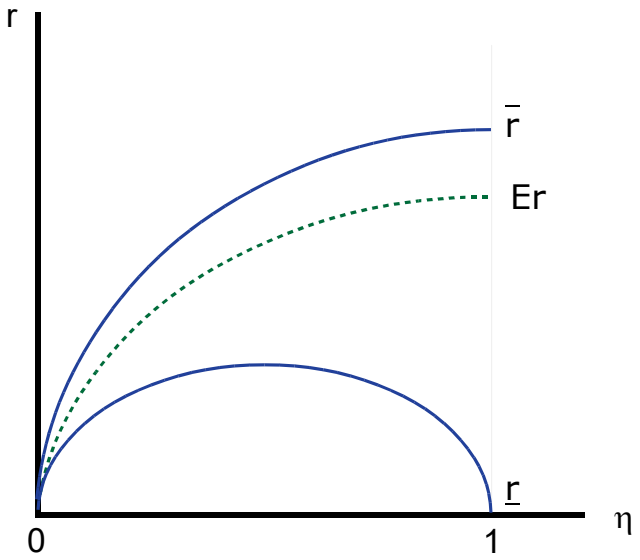

Figure 3.4: Dispersion and Search-in-Use

capital, thus reducing their need to bid up to capture larger quantities and reducing rent dispersion. Specifically, both the reservation and maximum rents fall, but the maximum rent falls faster. Search-in-use (η) is crucial to generating dispersion, but greater matching efficiency will diminish dispersion. The figure below clarifies the relationship between the search-in-use parameter and the dispersion. While the reservation rent has a non-monotonic relationship, the range increases monotonically in η. Also, the distribution is increasingly skewed toward the ceiling rent as η increases.

The basic model would yield transition dynamics familiar from any Solow-type model. A parameter change causes most endogenous variables to adjust over time to the new steady-state. One major exception is, of course, a change in the savings rate, which would cause an immediate change in household consumption followed by slow adjustment to the steady state. Less obvious is that changes to either search parameter (η, λ) will also cause immediate changes, since those parameters affect the value of capital for intermediaries causing the reservation rent to jump.

3.4 Endogenous Savings

Two endogenous savings models are provided below. The first imposes future discounting on the households only, while the second has all agents discounting. They are presented separately since the latter case notably alters the equilibrium in the capital market from that described in the basic model, by causing firm profits to be increasing in firm size. Presenting the two cases separately will clearly distinguish between the results changed by endogenous household saving, and the results changed due to discounting by firms and intermediaries.

In all cases, agents with discounting shall have CRRA utility in consumption. However, since capital can only be created by the intermediaries, firms will have no incentive to save. In a later extension, I experiment with transferring firm profits to the households in order to boost capital growth. The present version of the model can be interpreted as a complete separation of production decisions from capital accumulation decisions. This is admittedly an extreme assumption, but the claim that these two decisions are perfectly harmonized in actual economies is similarly suspicious. The transfer extension will explore the full range of possibilities and clarify in what ways capital market frictions will continue to reduce output even if households are given all of firm profits.

3.4.1 Household Discounting

If households alone discount the future at rate ρ, they face the problem:

$$\max \int_0^\infty e^{-\rho t} \frac{c(t)^{1-\theta}}{1-\theta} dt \tag{3.19}$$

$$s.t. \;\; \dot{k}(t) = (i(t) - \delta)k(t) \; + \; w(t) \; - \; c(t) \tag{3.20}$$

Per the basic model, the household is actually accumulating a stock of assets linked to the capital held by the intermediaries which pays a dividend for the life of the capital. For simplicity, the household budget constraint is written in terms of capital since the two are always equivalent. Using the household's problem, we can form a Hamiltonian in terms of c and k:

$$H(c,k,\chi) \;=\; \frac{c^{1-\theta}}{1-\theta} \;+\; \chi[(i-\delta)k \;+\; w \;-\; c] \tag{3.21}$$

Solving this yields the consumption path:

$$\frac{\dot{c}}{c} \;=\; \frac{1}{\theta}\,[\rho \;+\; \delta \;-\; i] \tag{3.22}$$

Thus at the steady state $i \;=\; \rho + \delta$, and per the intermediary equilibrium (17) and expected rent (18), this can be combined with the vacancy rate (4) and active capital per worker (2) to obtain the savings rate.

$$s \;=\; \frac{\alpha\delta}{\delta+\rho}\underbrace{\left[\frac{\eta\lambda(d+\delta+\eta\lambda)}{(d+\delta+\eta\lambda)^2 \;+\; (1-\eta)\eta\lambda^2}\right]}_{\psi} \tag{3.23}$$

The reader may notice that the portion to the left of ψ is the standard result for a neoclassical growth model. The effect of capital market frictions is encapsulated within ψ. Observe that $\psi \leq 1$ and that:

$$\lim_{\lambda\to\infty} \psi \;=\; \eta$$

And since $\eta = 1$ would be where search-in-use is just as fast as matching idle capital, this means that $\psi = 1$ when frictions are removed, yielding the standard result. Setting $\eta = 1$ where λ is still finite reduces ψ to $1 - v$, or the rate at which capital is active. Even with perfect search-in-use, savings would still be below the neo-classical result.

The shares of total output going to each type of agent are $\alpha(1-\psi)$ to firms, none to interme-

Factor	Increasing in:	Decreasing in:
Savings Rate	α, η, d, δ	λ^4, ρ
Output	α, z, η, λ	d, δ, ρ
Vacancy Rate	d, δ	λ
Rent Dispersion	η, d, δ, ρ	λ

Figure 3.5: Statics - Endogenous Savings

diaries, and $\alpha\psi + (1-\alpha)$ to households. Endogenous savings produces some adjustments to the comparative statics: Increasing the search-in-use parameter (faster reallocation of capital) now increases steady state output. Greater search-in-use means higher expected rent, and capital thus becomes more valuable. As for transition dynamics, with the savings rate endogenized, any parameter change will cause an immediate change in consumption.

3.4.2 General Discounting

If all agents have preferences of the form given in (19), this alters the structure of the capital allocation market. As before, we are solving for the steady state equilibrium. The value function of a firm posting rent r at some candidate steady state would be:

$$\rho V(n(r)) \;=\; z\, n(r)^{\alpha} \ell(r)^{1-\alpha} \;-\; r\, n(r) \;-\; w\ell(r) \tag{3.24}$$

Firms still choose their rent level such that they are indifferent between posting r or switching to posting the reservation rent. Let $\hat{V}(n(\underline{r}))$ be the value of switching to the reservation rent, and let $\pi(\underline{r})$ be the profit flow to a firm posting the reservation rent. Standard dynamic programming arguments imply:

$$\rho\hat{V}(n) \;-\; \hat{V}'(n)\dot{n} \;=\; \pi(\underline{r}) \tag{3.25}$$

Since firms choose r such that $V(n) = \hat{V}(n)$ at the steady state, the equilibrium value of a firm is:

$$V(n) = \frac{\pi(\underline{r})}{\rho + d + \delta + \eta\lambda}\left[n(r) + \frac{\lambda v K}{\rho}\right] \tag{3.26}$$

The expression for firm size, $n(r)$, is unchanged from the basic model. This means that unlike the basic model where profits are equal across all firms, profits are increasing in the rent offer. This result is intuitive since, with discounting, the windfall from a firm lowering their rent offer needs to be balanced against the lower profits they would earn in the future. As ρ increases, the profit premium for larger firms increases. Using this profit function and the firm size expression one can derive the posting distribution and active capital distribution.

$$P(r) = \frac{(d + \delta + \eta\lambda)}{\eta\lambda}\left[1 - \left(\frac{\alpha z k_a^{\alpha-1} - r}{\alpha z k_a^{\alpha-1} - \underline{r}} - \frac{\rho(r - \underline{r})}{(d + \delta + \eta\lambda)(\alpha z k_a^{\alpha-1} - \underline{r})}\right)^{\frac{1}{2}}\right] \tag{3.27}$$

$$A(r) = \frac{(d + \delta + \eta\lambda)}{\eta\lambda}\left[1 - \left(\frac{n(\underline{r})}{n(r)}\right)^{\frac{1}{2}}\right] \tag{3.28}$$

The model with discounting for all agents reduces the rent dispersion relative to the basic model. This can be seen by comparing (27) to (10). Since the rent dispersion is reduced, this implies that expected rent is decreasing in the discount factor (ρ). This in turn means smaller dividends paid to households (i), and reduced incentive to save. The expressions are considerably more cumbersome, and only suitable for numeric analysis. For purposes of this paper it is only necessary to note that a model where only households discount the future is likely to somewhat overestimate the rent dispersion.

3.5 Transfer to Households

As noted above, a reasonable objection can be raised about firms consuming their profits. In practice, firm managers do not simply binge on their profits, but often either reinvest those profits or transfer them to shareholders. Thus, the model above, when calibrated, would likely result in far too low of a savings rate. This is in fact what we will see in the next section. In this section, I introduce a simple transfer of profits from the firm to the household.

I will start with the endogenous saving model where only households discount (3.1) and include a tax rate $\tau \in [0,1]$ on Firm profits. All of the taxed profits are transfered to the households as $T = \frac{\tau\Pi}{L}$ where Π is total firm profits. At $\tau = 0$, this is identical to the model above. At $\tau = 1$, this is equivalent to endowing the households with a balanced portfolio of shares in the firms. The calibration section includes a discussion of the possible causes of intermediate values of τ.

The transfer changes the household budget constraint to:

$$\dot{k}(t) = (i(t) - \delta)k(t) + w(t) + T(t) - c(t) \tag{3.29}$$

$$T(t) = \frac{\tau}{L}\left[z((1 - v(t))K(t))^{\alpha}L^{1-\alpha} - \mathbb{E}r(t)(1 - v(t))K(t) - w(t)L\right] \tag{3.30}$$

This, in turn, adjusts the household Hamiltonian:

$$H(c,k,\chi) = \frac{c^{1-\theta}}{1-\theta} + \chi\left[(i-\delta)k + (1-\tau)w + \tau z(1-v)^{\alpha}k^{\alpha} - \tau\mathbb{E}r(1-v)k - c\right] \tag{3.31}$$

Solving this yields a more complicated consumption path:

$$\frac{\dot{c}}{c} \;=\; \frac{1}{\theta}\left[\rho \;+\; \delta \;-\; i \;-\; \alpha\tau z(1-v)^{\alpha}k^{\alpha-1} \;+\; \tau\mathbb{E}r(1-v)\right] \tag{3.32}$$

The steady state savings rate is recovered using the same methods as previously described, and is written in terms of the composite parameter ψ for clarity.

$$s \;=\; \frac{\alpha\delta}{\delta+\rho}\left[\psi \;-\; \tau(\psi-1)\right] \tag{3.33}$$

If we consider the case of $\tau = 1$, the savings rate will be equivalent to the neo-classical growth model solution. Output would still be reduced due to some capital being idle. Thus, the undersaving that occurs for $\tau < 1$ results from an imperfect connection between the profit maximization decisions of the firm and the investment decisions of households interacting with frictions in the capital market.

3.6 Calibration

To assess validity, I perform a calibration exercise using the transfer extension of the model. The years 1999 to 2016 are used since that is the range for which U.S. commercial real estate vacancy rates were available. Lacking a good proxy for the contact rate (λ), the observed vacancy rates[5] are used to back out this value. Using the median calculated λ and the typical commercial lease duration of four years, a value for the search-in-use parameter (η) was inferred. Essentially, this assumes that typical lease length is determined by the likelihood that active capital becomes matched.

Observed savings rates are determined using gross private saving and nominal GDP from FRED, which averages 19.9%, and is countercyclical. The targets are the rent dispersion

[5]Data obtained from Reis, Inc. covering a composite of all U.S. major metro areas.

Parameter	Value	Source
α	$0.357 - 0.435$	BLS labor share data
d	$0.069 - 0.096$	BLS establishment data
δ	0.1	Eisfeldt & Rampini, 2006
η	0.23	Described above
λ	$0.908 - 2.238$	Described above
ρ	$0.012 - 0.04$	Otrok, 2001 and E&R

Figure 3.6: Calibration Parameters

ratio (highest over lowest rent) and the value of τ needed to achieve the observed savings rates. The assumed discount rate has a significant effect on the optimal saving rate, and this affects the τ estimate. The high discount rate value is taken from Eisfeldt and Rampini (2006), while the low value is from Otrok (2001). These imply a mean optimal frictionless saving rate of 28.9% and 36.2% respectively.

Observed vacancy rates and dispersion ratios use a composite of office and retail real estate scaled based on square feet per employed worker. The dispersion ratio used is based on twice the gap between asking and effective rents. This is admittedly a coarse measure of dispersion, but if we assume rough symmetry of the distribution, and that the asking rent represents the maximum this should be a fair approximation. In addition, since this measure is a ratio, any heterogeneity due to differences in locations or amenities are averaged out.

The calibrated dispersion ratios that result are all somewhat larger than the observed values. The mean calibrated dispersion ratio is 1.712, while the mean observed ratio is 1.441. As noted above, a version of the model that neglects discounting by the firms yields larger dispersion than one that does. Since the calibration uses the simpler model, this is expected.

With regard to savings rates, the calibrated model with no transfers to households ($\tau = 0$) would yield savings rates that are far too low, with a mean savings rate of 9.8% for the low discount rate, and 7.9% for the high discount rate. To produce the observed savings rates would require the mean transfer share (τ) to be 0.380 for the low discount rate, and 0.569

for the high discount rate. The transfer shares also exhibit a countercyclical pattern, which may hint at the underlying cause of these mid-range values for τ.

The reader might have expected that τ should be very close to one, since real world firms either reinvest their profits or pass dividends to shareholders who engage in saving behavior. Under this perspective, the calibrated transfer shares are surprising, suggesting that a substantial amount of firm profits are simply consumed. Several explanations suggest themselves. Principal-Agent problems could be causing firm managers to act out of accordance with shareholder interests. Alternatively, there could be search costs in the capital market unaccounted for in the model[6]. Finally, there could be organizational issues where reinvestment by firms is less optimal than broad-based investment by households, and the mid-range value of τ is reflecting this. Of these possibilities, the last is most likely to have a countercyclical pattern since dividends drop in economic downturns. Nonetheless, the current study can only highlight this interesting puzzle, and note that frictions in the process of capital reallocation will exert a depressing effect on investment.

3.7 Conclusion

This paper provides a model of idle physical capital, which most closely resembles markets for non-residential real estate. While the model is based on a seminal labor unemployment model, the addition of depreciation and investment is necessary to capture the unique nature of capital markets. It is, of course, impossible for households to generate more workers except in the sense of population growth, which is too slow of a process to matter in unemployment dynamics. Such is not the case with capital goods, where the creation of new goods, and their eventual breakdown or obsolescence, is pivotal to the path of the economy.

[6]A version of the model with a free entry condition for the firms combined with a cost of posting would yield broadly similar results, except that the entire portion of output currently going to firm profits would be spent on the posting costs, as firms would enter until profits were zero.

The model generates two channels by which steady state output is reduced from what would be implied by a neo-classical benchmark. First, the fact that a share of capital is always idle decreases production at all times. The second channel is that capital market frictions can reduce the incentive to invest (undersaving). While this second channel can be shut down by assuming a perfect transfer of firm profits to households, calibration suggests that such a model would be inaccurate. The exact mechanisms that give rise to this are left to further research. For this paper, it is sufficient to point out that frictional capital markets in a growth model environment can result in undersaving.

Bibliography

Alchian, Armen and Harold Demsetz (1972), "Production, Information Costs and Economic Organization." *American Economic Review*, 777–795.

Andrade, Gregor, Mark Mitchell, and Erik Stafford (2001), "New evidence and perspectives on mergers." *The Journal of Economic Perspectives*, 15, pp. 103–120, URL http://www.jstor.org/stable/2696594.

Arnade, Carlos and Munisamy Gopinath (1998), "Capital adjustment in u.s. agriculture and food processing: A cross-sectoral model." *Journal of Agricultural and Resource Economics*, 23, URL https://EconPapers.repec.org/RePEc:ags:jlaare:31167.

Azzi, Corry and Ronald Ehrenberg (1975), "Household allocation of time and church attendance." *Journal of Political Economy*, 83, 27–56, URL https://EconPapers.repec.org/RePEc:ucp:jpolec:v:83:y:1975:i:1:p:27-56.

B, Pedro Dal (2007), "Social norms, cooperation and inequality." *Economic Theory*, 30, 89–105, URL http://www.jstor.org/stable/27822462.

Bai, Yan, Jose-Victor Rios-Rull, and Kjetil Storesletten (2011), "Demand Shocks that Look Like Productivity Shocks." 2011 Meeting Papers 99, Society for Economic Dynamics, URL https://ideas.repec.org/p/red/sed011/99.html.

Barros, Pedro and Nuno Garoupa (2002), "An economic theory of church strictness." *Economic Journal*, 112, 559–576.

Berman, Eli (2000), "Sect, subsidy, and sacrifice: An economist's view of ultra-orthodox jews*." *The Quarterly Journal of Economics*, 115, 905–953, URL http://dx.doi.org/10.1162/003355300554944.

Bicchieri, Cristina (2006), "The grammar of society: The nature and dynamics of social norms."

Boucekkine, Raouf, Fernando del Ro, and Blanca Martnez (2009), "Technological progress, obsolescence, and depreciation." *Oxford Economic Papers*, 61, 440–466, URL +http://dx.doi.org/10.1093/oep/gpn016.

Burdett, Kenneth and Dale T. Mortensen (1998), "Wage differentials, employer size, and unemployment." *International Economic Review*, 39, 257–273, URL http://www.jstor.org/stable/2527292.

Carvalho, Jean-Paul (2013), "Veiling." *The Quarterly Journal of Economics*, 128, 337–370, URL http://EconPapers.repec.org/RePEc:oup:qjecon:v:128:y:2013:i:1:p:337-370.

Chen, Daniel (2010), "Club goods and group identity: Evidence from islamic resurgence during the indonesian financial crisis." *Journal of Political Economy*, 118, 300–354, URL http://www.jstor.org/stable/10.1086/652462.

Coase, R H (1937), "The nature of the firm." *Economica*, 4, 386–405.

Davidson, Carl and Ben Ferrett (2007), "Mergers in multidimensional competition." *Economica*, 74, pp. 695–712, URL http://www.jstor.org/stable/4541567.

Dawson, Lorne (2001), "The cultural significance of new religious movements: The case of soka gakkai." 62.

Dong, Feng, Pengfei Wang, and Yi Wen (2016), "A search-based framework of capital reallocation." *China International Conference in Finance*.

Eisfeldt, Andrea and Adriano Rampini (2006), "Capital reallocation and liquidity." *Journal of Monetary Economics*, 53, 369–399, URL https://EconPapers.repec.org/RePEc:eee:moneco:v:53:y:2006:i:3:p:369-399.

Eppli, Mark J. and James D. Shilling (1995), "Speed of adjustment in commercial real estate markets." *Southern Economic Journal*, 61, 1127–1145, URL http://www.jstor.org/stable/1060745.

Gavazza, Alessandro (2011), "The role of trading frictions in real asset markets." *American Economic Review*, 101, 1106–43, URL https://EconPapers.repec.org/RePEc:aea:aecrev:v:101:y:2011:i:4:p:1106-43.

Gavazza, Alessandro (2016), "An empirical equilibrium model of a decentralized asset market." *Econometrica*, 84, 1755–1798, URL http://dx.doi.org/10.3982/ECTA10847.

Gibbons, Robert (2005), "Four formal(izable) theories of the firm?" *Journal of Economic Behavior and Organization*, 58, 200–245.

Gilligan, Thomas W. (2004), "Lemons and leases in the used business aircraft market." *Journal of Political Economy*, 112, 1157–1186, URL https://EconPapers.repec.org/RePEc:ucp:jpolec:v:112:y:2004:i:5:p:1157-1186.

Grossman, Sanford J. and Oliver D. Hart (1986), "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration." *Journal of Political Economy*, 94, 691.

Gugler, Klaus and Ralph Siebert (2007), "Market power versus efficiency effects of mergers and research joint ventures: Evidence from the semiconductor industry." *The Review of Economics and Statistics*, 89, pp. 645–659, URL http://www.jstor.org/stable/40043091.

Hart, Oliver and Bengt Holmstrom (2010), "A Theory of Firm Scope." *Quarterly Journal of Economics*, CXXV.

Hart, Oliver and John Moore (2008), "Contracts as Reference Points." *Quarterly Journal of Economics*, CXXIII, 1–48.

Holmstrom, Bengt and Jean Tirole (1991), "Transfer Pricing and Organizational Form." *Journal of Law, Economics, {&} Organization*, 7, 201–228.

Horne, Christine (2003), "The internal enforcement of norms." *European Sociological Review*, 19, 335–343, URL http://www.jstor.org/stable/3559552.

Horne, Christine (2004), "Collective benefits, exchange interests, and norm enforcement*." *Social Forces*, 82, 1037–1062, URL http://dx.doi.org/10.1353/sof.2004.0041.

Hsieh, Chang-Tai and Peter J. Klenow (2009), "Misallocation and manufacturing tfp in china and india." *The Quarterly Journal of Economics*, 124, 1403–1448, URL http://dx.doi.org/10.1162/qjec.2009.124.4.1403.

Huck, Steffen, Kai A. Konrad, and Wieland Mller (2004), "Profitable horizontal mergers without cost advantages: The role of internal organization, information and market structure." *Economica*, 71, pp. 575–587, URL http://www.jstor.org/stable/3548981.

Hungerman, Daniel (2005), "Are church and state substitutes? evidence from the 1996 welfare reform." *Journal of Public Economics*, 89, 2245–2267, URL https://EconPapers.repec.org/RePEc:eee:pubeco:v:89:y:2005:i:11-12:p:2245-2267.

Hungerman, Daniel M. (2013), "Substitution and stigma: Evidence on religious markets from the catholic sex abuse scandal." *American Economic Journal: Economic Policy*, 5, 227–53, URL http://www.aeaweb.org/articles?id=10.1257/pol.5.3.227.

Iannaccone, Laurence (1992), "Sacrifice and stigma: Reducing free-riding in cults, communes, and other collectives." *Journal of Political Economy*, 100, 271–91, URL http://EconPapers.repec.org/RePEc:ucp:jpolec:v:100:y:1992:i:2:p:271-91.

Iannaccone, Laurence R. and Eli Berman (2006), "Religious extremism: The good, the bad, and the deadly." *Public Choice*, 128, 109–129, URL https://doi.org/10.1007/s11127-006-9047-7.

Iyer, Sriya, Chander Velu, and Melvyn Weeks (2014), "Divine competition: Religious organisations and service provision in india." Cambridge working papers in economics, Faculty of Economics, University of Cambridge, URL https://EconPapers.repec.org/RePEc:cam:camdae:1409.

Jandik, Tomas and Raja Kali (2009), "Legal systems, information asymmetry, and firm boundaries: Cross-border choices to diversify through mergers, joint ventures, or strategic alliances." *Journal of International Business Studies*, 40, pp. 578–599, URL http://www.jstor.org/stable/25483386.

Jung, Danielle F. and David A. Lake (2011), "Markets, hierarchies, and networks: An agent-based organizational ecology." *American Journal of Political Science*, 55, pp. 972–990, URL http://www.jstor.org/stable/23025131.

Ke, Qiulin and Michael White (2013), "A tale of two chinese cities: The dynamics of beijing and shanghai office markets." *The Journal of Real Estate Portfolio Management*, 19, 31–48, URL http://www.jstor.org/stable/24885523.

Klein, Benjamin (1996), "Why Hold Ups Occur: The Self-Enforcing Range of Contractual Relationships." *Economic Inquiry*, XXXIV, 444–463.

Knight, Frank (1921), *Risk, Uncertainty, and Profit*. Harper and Row, New York.

Kurmann, Andr (2014), "Holdups and overinvestment in capital markets." *Journal of Economic Theory*, 151, 88 – 113, URL http://www.sciencedirect.com/science/article/pii/S0022053114000222.

Kurmann, Andr and Nicolas Petrosky-Nadeau (2007), "Search frictions in physical capital markets as a propagation mechanism." *SSRN Electronic Journal*.

Kurmann, Andr and Stanislav Rabinovich (2018), "Dynamic inefficiency in decentralized capital markets." *Journal of Economic Theory*, 173, 231 – 256, URL http://www.sciencedirect.com/science/article/pii/S0022053117301266.

Lagos, Ricardo and Guillaume Rocheteau (2009), "Liquidity in asset markets with search frictions." *Econometrica*, 77, 403–426, URL http://dx.doi.org/10.3982/ECTA7250.

Lanteri, Andrea (2017), "The market for used capital: Endogenous irreversibility and reallocation over the business cycle (january 17, 2016)." Working Paper 207, Economic Research Initiatives at Duke (ERID), URL https://ssrn.com/abstract=2733849.

Lazerson, Mark H. (1988), "Organizational growth of small firms: An outcome of markets and hierarchies?" *American Sociological Review*, 53, pp. 330–342, URL http://www.jstor.org/stable/2095642.

Legros, Patrick and Andrew F. Newman (2013), "A price theory of vertical and lateral integration." *Quarterly Journal of Economics*, 128, 725–770.

Levy, Gilat and Ronny Razin (2012), "Religious beliefs, religious participation, and cooperation." *American Economic Journal: Microeconomics*, 4, 121–51, URL http://www.aeaweb.org/articles?id=10.1257/mic.4.3.121.

Lpez-Prez, Ral (2008), "Aversion to norm-breaking: A model." *Games and Economic Behavior*, 64, 237 – 267, URL http://www.sciencedirect.com/science/article/pii/S0899825608000031.

McBride, Michael (2007), "Club mormon: Free-riders, monitoring, and exclusion in the lds church." *Rationality and Society*, 19, 395–424, URL https://doi.org/10.1177/1043463107083736.

McBride, Michael (2015), "Why churches need free-riders: Religious capital formation and religious group survival." *Journal of Behavioral and Experimental Economics*, 58, 77 – 87.

Mortensen, Dale T. and Christopher A. Pissarides (1994), "Job creation and job destruction in the theory of unemployment." *The Review of Economic Studies*, 61, 397–415, URL http://www.jstor.org/stable/2297896.

Otrok, Christopher (2001), "On measuring the welfare cost of business cycles." *Journal of Monetary Economics*, 47, 61–92, URL https://EconPapers.repec.org/RePEc:eee:moneco:v:47:y:2001:i:1:p:61-92.

Ottonello, Pablo (2015), "Capital unemployment, financial shocks, and investment slumps." In *2015 Meeting Papers*, 1153, Society for Economic Dynamics.

Patel, David S. (2012), "Concealing to reveal: The informational role of islamic dress." *Rationality and Society*, 24, 295–323, URL https://doi.org/10.1177/1043463112440683.

Piazzesi, Monika, Martin Schneider, and Johannes Stroebel (2013), "Segmented housing search."

Powell, W (1990), "Neither market nor hierarchy." *The sociology of organizations: classic, contemporary, and critical readings*, 315, 104–117.

Pulvino, Todd C. (1998), "Do asset fire sales exist? an empirical investigation of commercial aircraft transactions." *Journal of Finance*, 53, 939–978, URL https://EconPapers.repec.org/RePEc:bla:jfinan:v:53:y:1998:i:3:p:939-978.

Ramey, Valerie and Matthew Shapiro (2001), "Displaced capital: A study of aerospace plant closings." *Journal of Political Economy*, 109, 958–992, URL https://EconPapers.repec.org/RePEc:ucp:jpolec:v:109:y:2001:i:5:p:958-992.

Robin, Jean-Marc and Sbastien Roux (2002), "An equilibrium model of the labour market with endogenous capital and two-sided search." *Annales d'conomie et de Statistique*, 257–307, URL http://www.jstor.org/stable/20076350.

Saxenian, AnnaLee (1996), *Regional Advantage: Culture and Competition in Silicon Valley and Route 128*. Harvard University Press.

Schrank, Andrew and Josh Whitford (2011), "The anatomy of network failure." *Sociological Theory*, 29, pp. 151–177, URL http://www.jstor.org/stable/23076400.

Shi, Shouyong and Melanie Cao (2015), "Endogenously procyclical liquidity, capital reallocation, and q." 2015 Meeting Papers 100, Society for Economic Dynamics, URL https://EconPapers.repec.org/RePEc:red:sed015:100.

Tse, Raymond Y. C. and Dominique Fischer (2003), "Estimating natural vacancy rates in office markets using a time-varying model." *Journal of Real Estate Literature*, 11, 37–45, URL http://www.jstor.org/stable/44103687.

Uzzi, Brian (1997), "Social structure and competition in interfirm networks: The paradox of embeddedness." *Administrative Science Quarterly*, 42, pp. 35–67, URL http://www.jstor.org/stable/2393808.

Van den Steen, Eric (2010), "Interpersonal Authority in a Theory of the Firm." *American Economic Review*, 100, 466–490.

Wallis, R. (1984), *The Elementary Forms of the New Religious Life*. International library of sociology, Routledge & Kegan Paul, URL https://books.google.com/books?id=V3wOAAAAQAAJ.

Weber, Roberto A. and Colin F. Camerer (2003), "Cultural conflict and merger failure: An experimental approach." *Management Science*, 49, pp. 400–415, URL http://www.jstor.org/stable/4133947.

Wheaton, William C., Raymond G. Torto, and Peter Evans (1997), "The cyclic behavior of the greater london office market." *The Journal of Real Estate Finance and Economics*, 15, 77–92, URL https://doi.org/10.1023/A:1007701422238.

Williamson, Oliver E. (1971), "The vertical integration of production : Market failure considerations *." *The American Economic Review*, 61, 112–123.

Appendix to Chapter One

To check whether the results are unique to the particular parameter values chosen, we can try two alternative sets of revenue values. First, we can widen the gap between the enthusiastic and reluctant managers by using $r_A = 0$ and $r_B = 3 - \sqrt{3}$. Note that the sum of the revenue values remain the same in order to ensure a clean comparison. Second, we can remove the gap between the two managers by setting both $r = (3 - \sqrt{3})/2$.

For the first check, with the gap in revenues maximized, we can see from the output below that outcomes are still closer to the first best with a convention than without, and that the relative performance of the two conventions and the ban environments is unchanged.

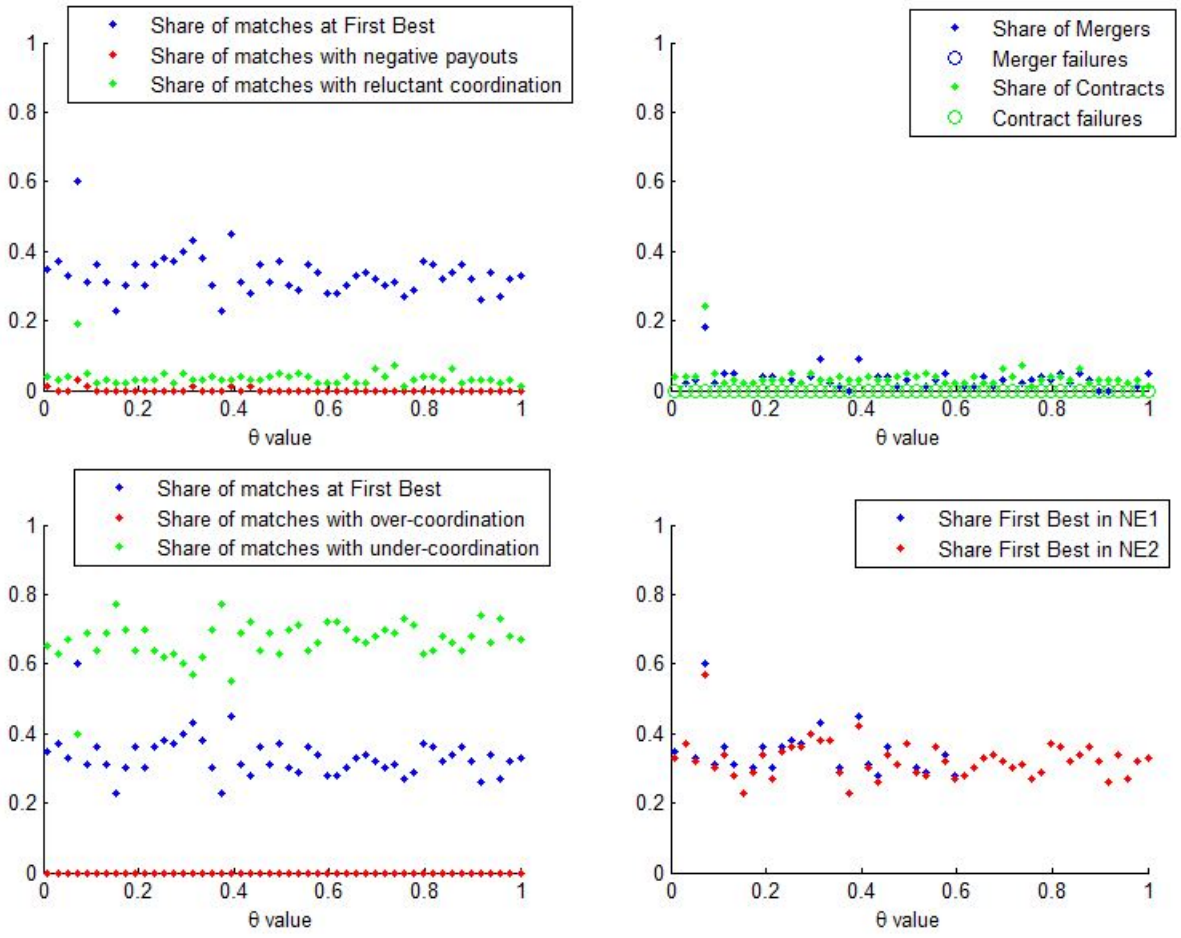

Figure 3.7: Baseline, maximum revenue gap

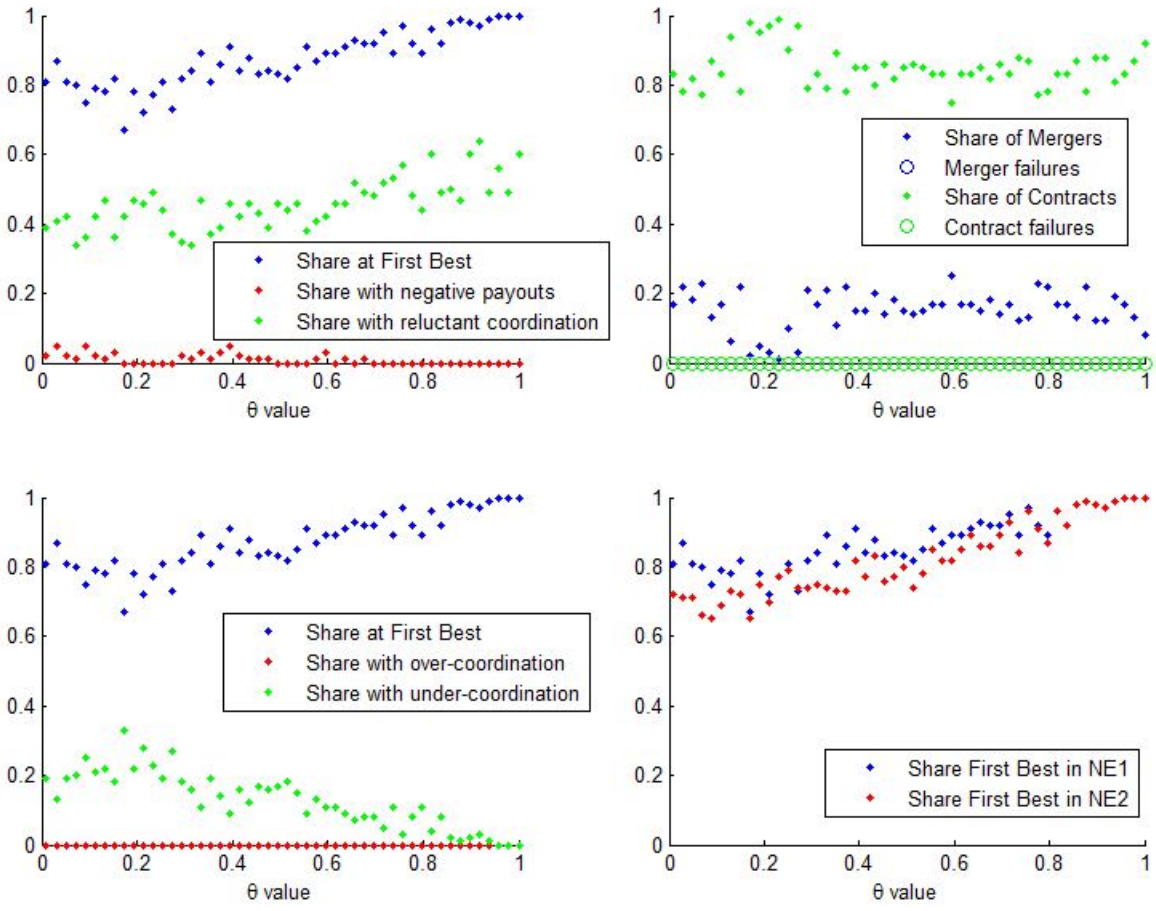

Figure 3.8: Contract Convention, maximum revenue gap

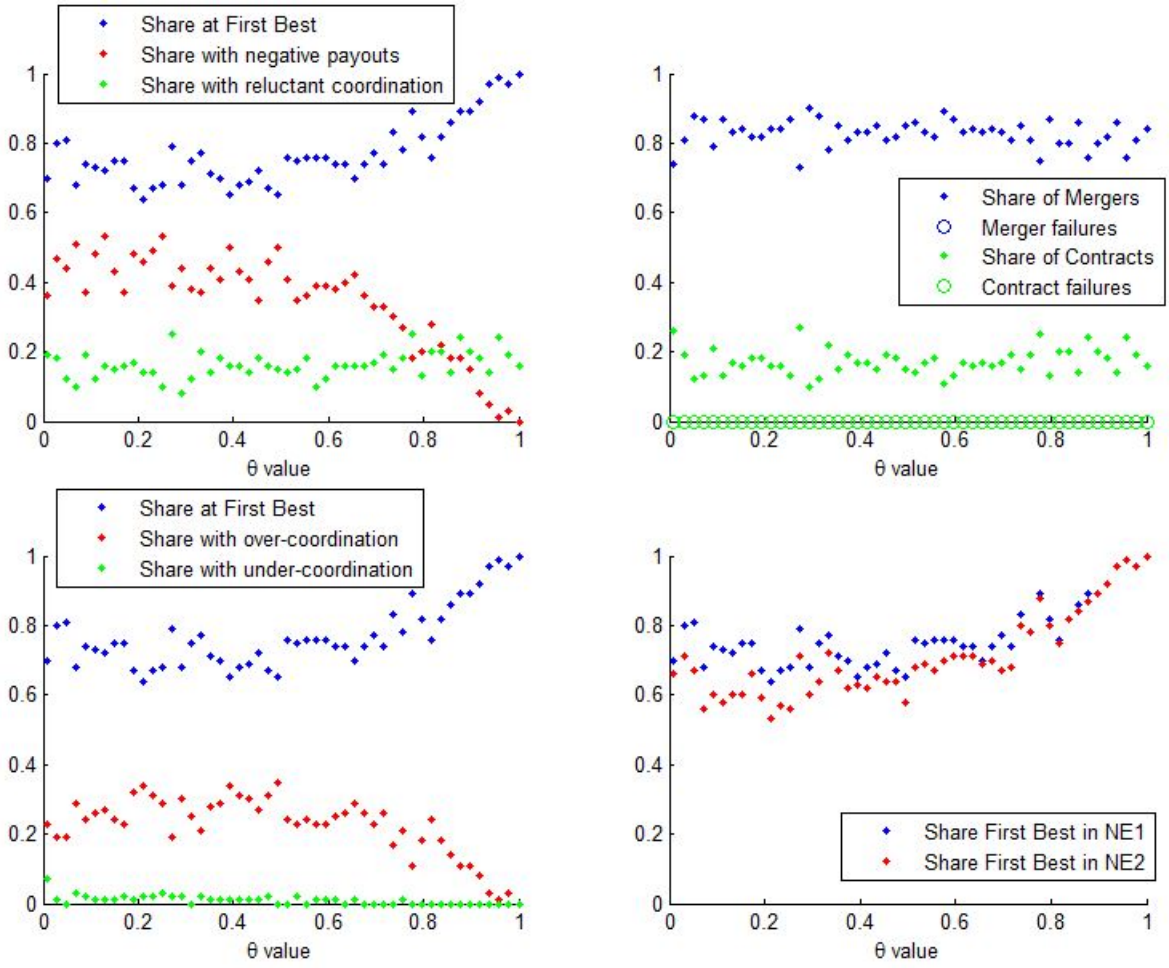

Figure 3.9: Merger Convention, maximum revenue gap

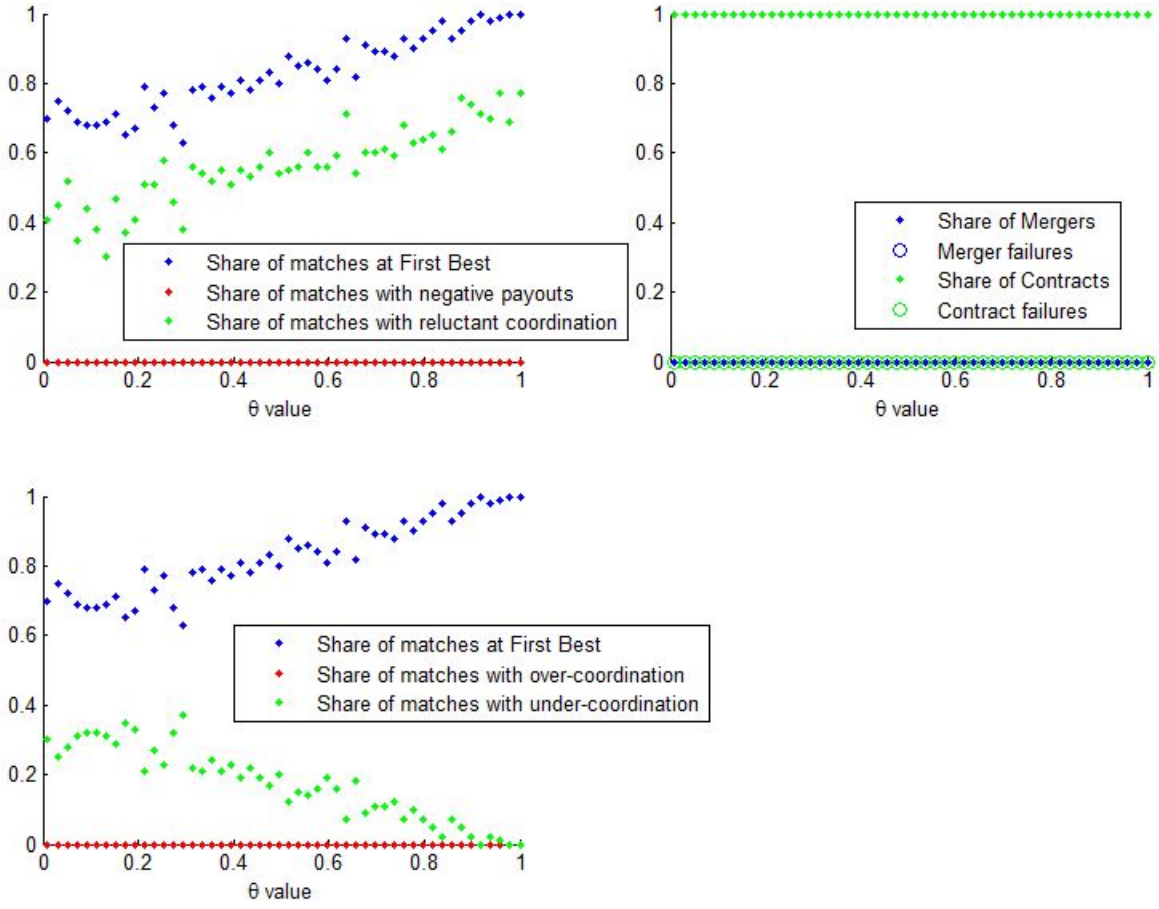

Figure 3.10: Merger Ban, maximum revenue gap

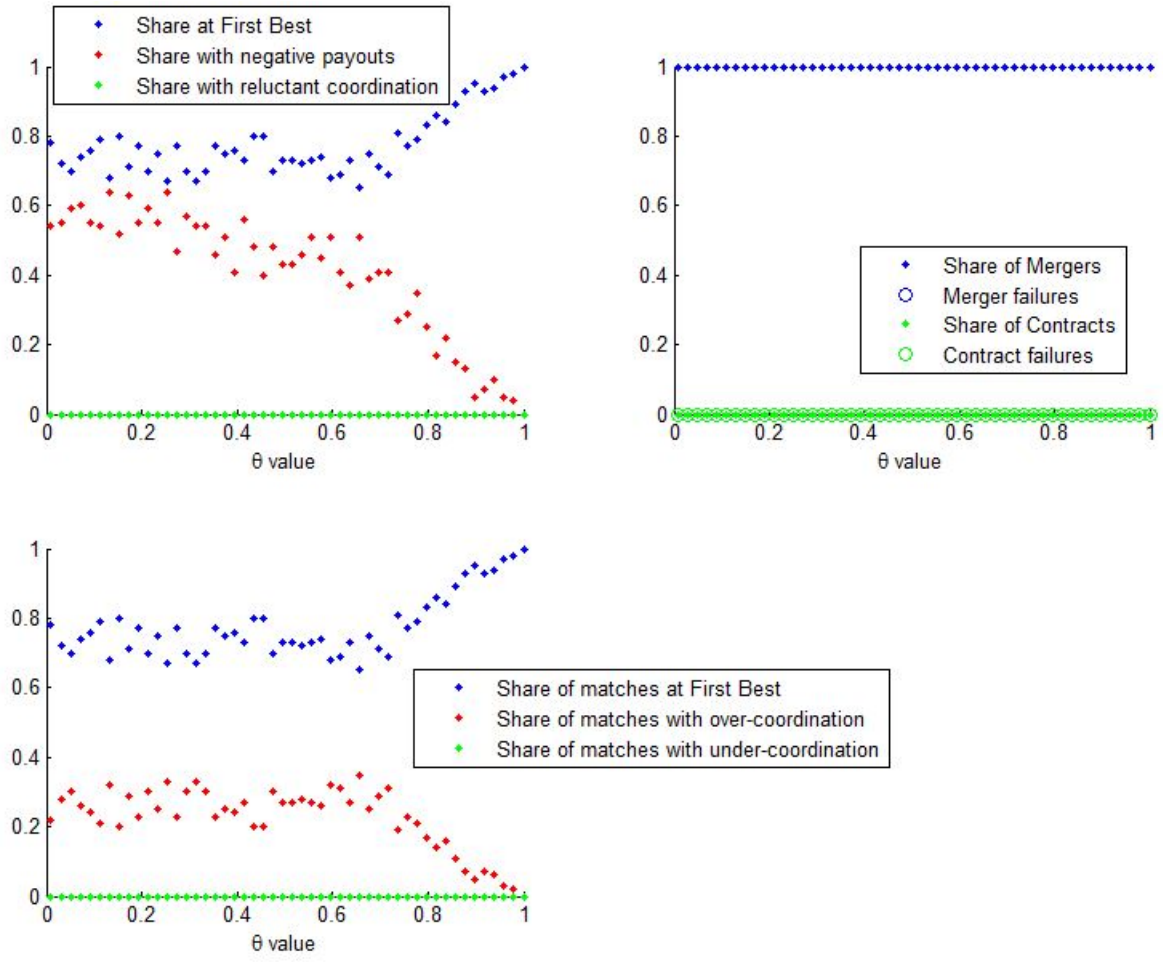

Figure 3.11: Contract Ban, maximum revenue gap

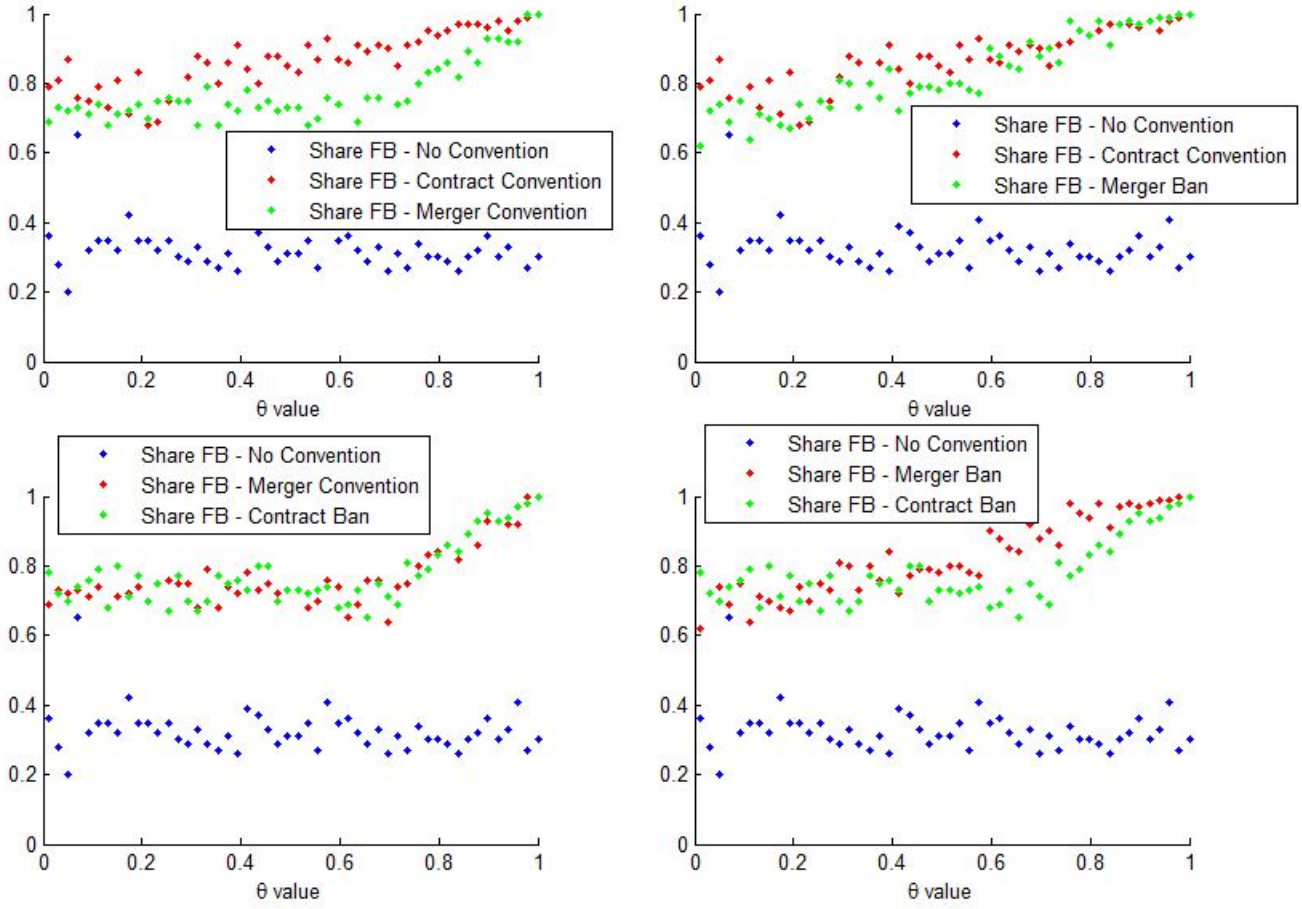

Figure 3.12: First Best Comparison, maximum revenue gap

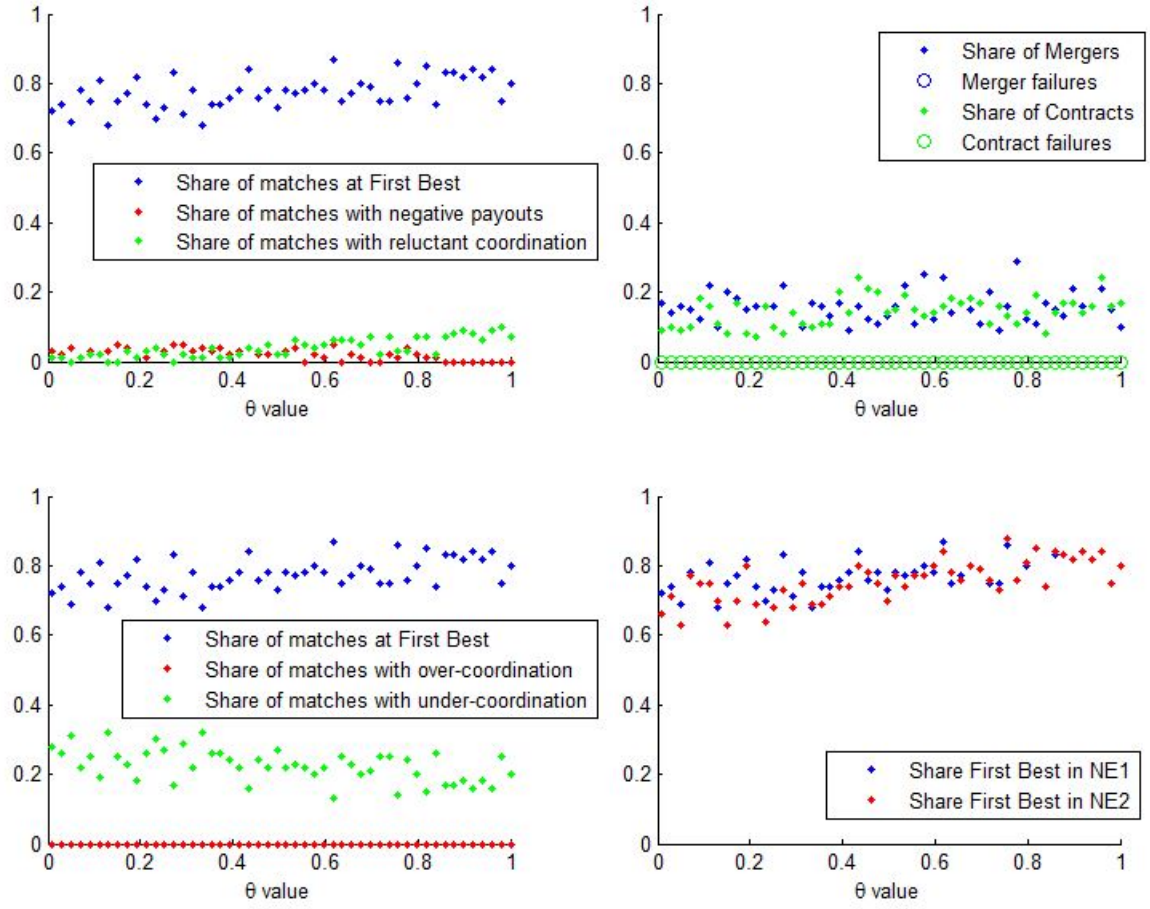

Figure 3.13: Baseline, maximum revenue gap

For the second check, with the gap in revenues removed, the differences in outcomes between the various alternative rules are much narrower, but the relative ranking remains. Thus, the benefits of adopting a contract or merger convention increase as the degree of asymmetry between the firms increases.

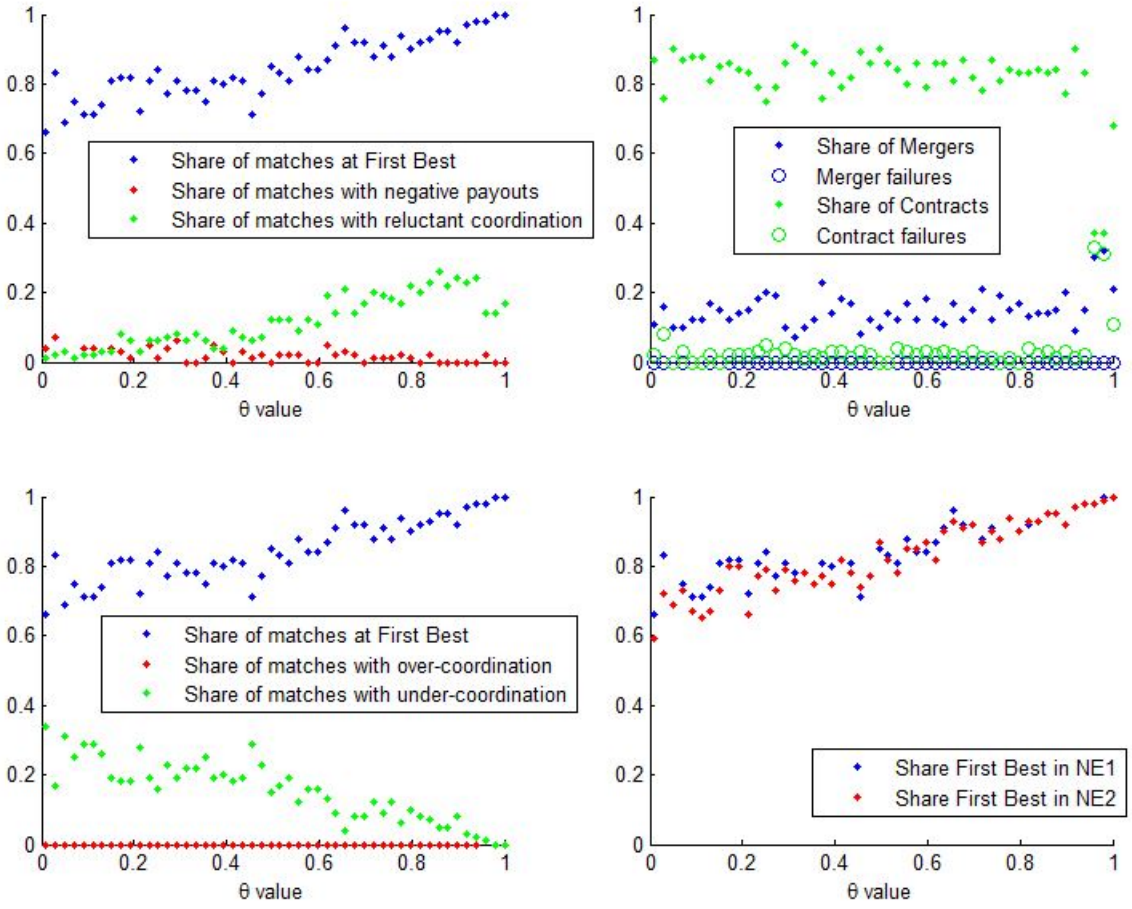

Figure 3.14: Contract Convention, maximum revenue gap

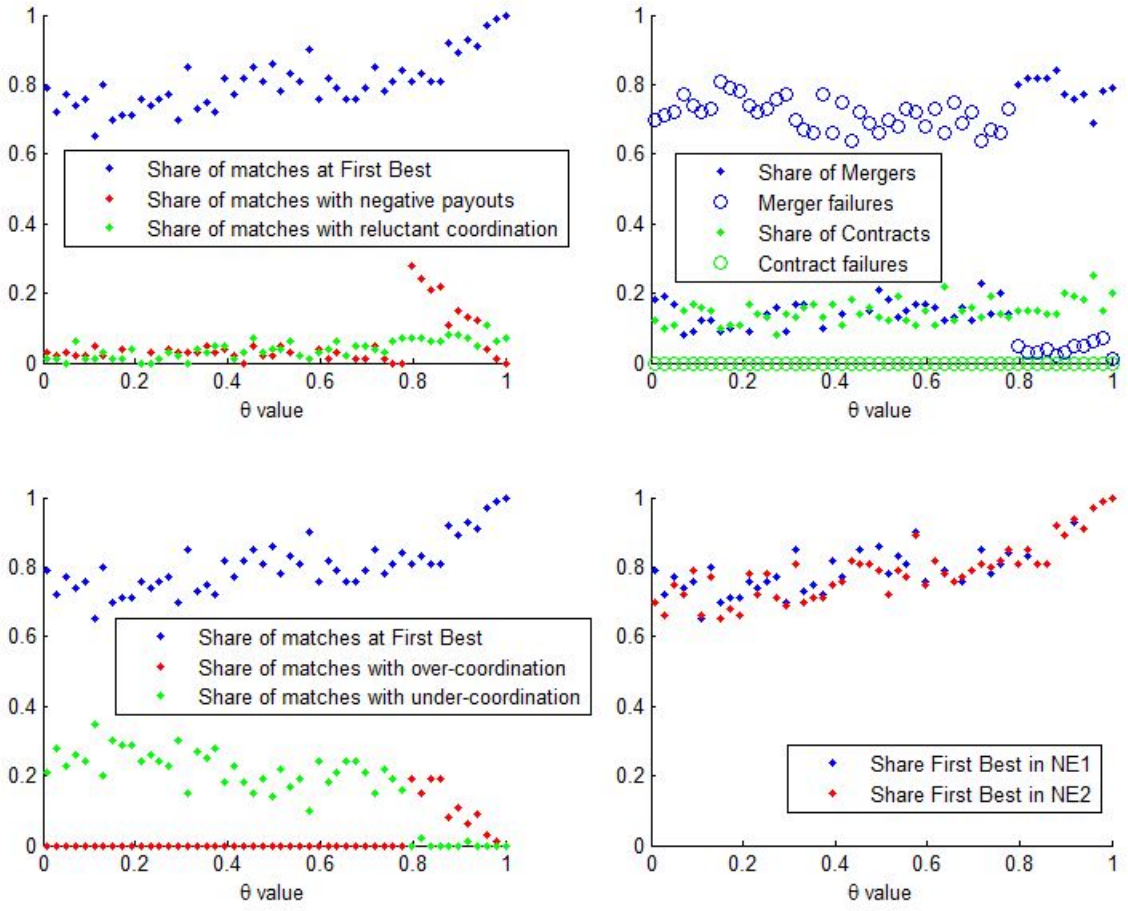

Figure 3.15: Merger Convention, maximum revenue gap

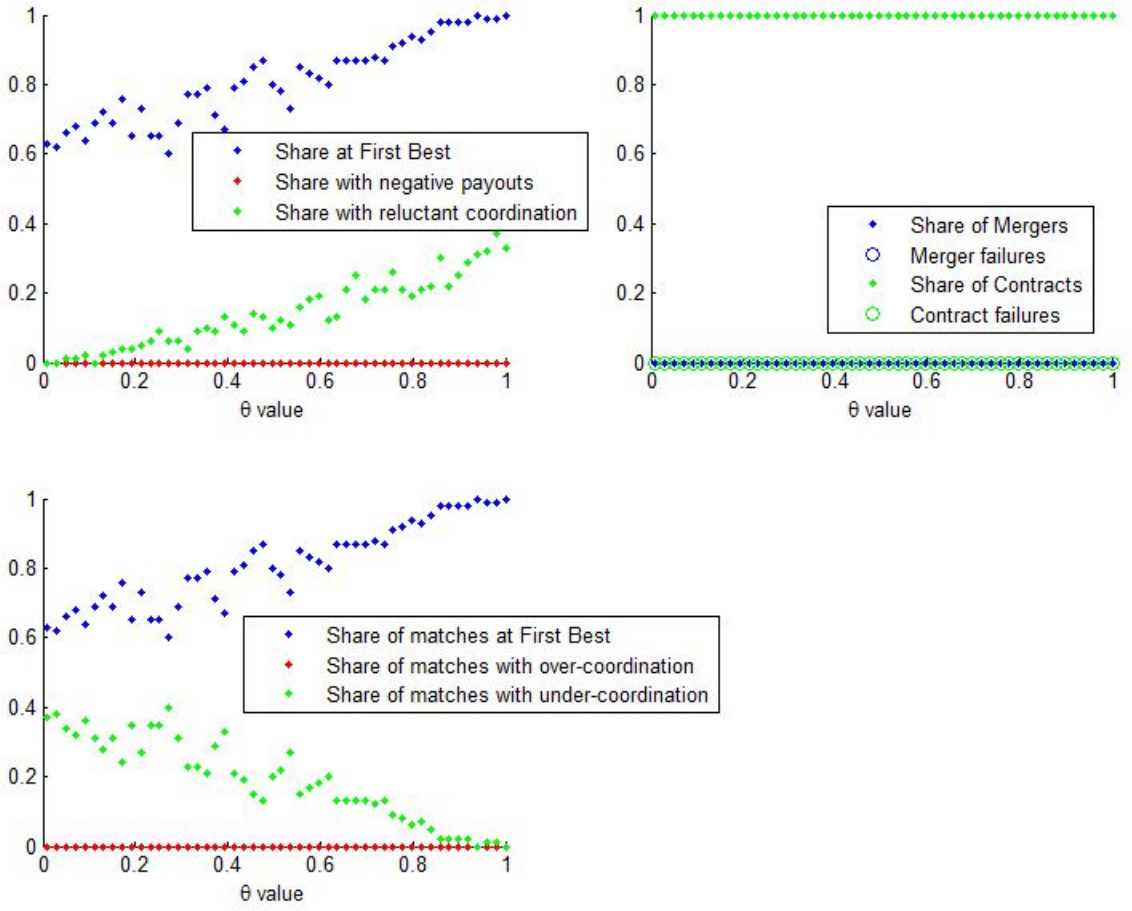

Figure 3.16: Merger Ban, maximum revenue gap

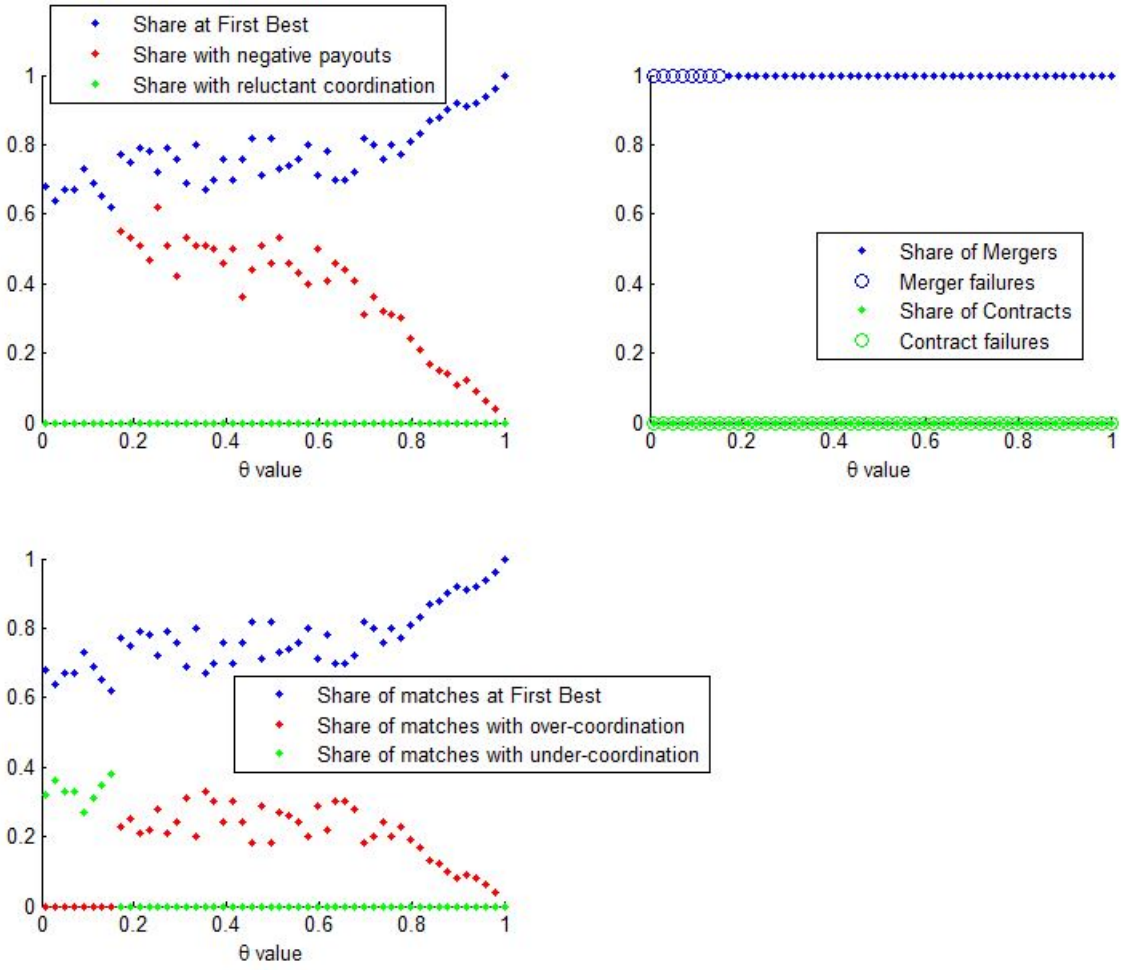

Figure 3.17: Contract Ban, maximum revenue gap

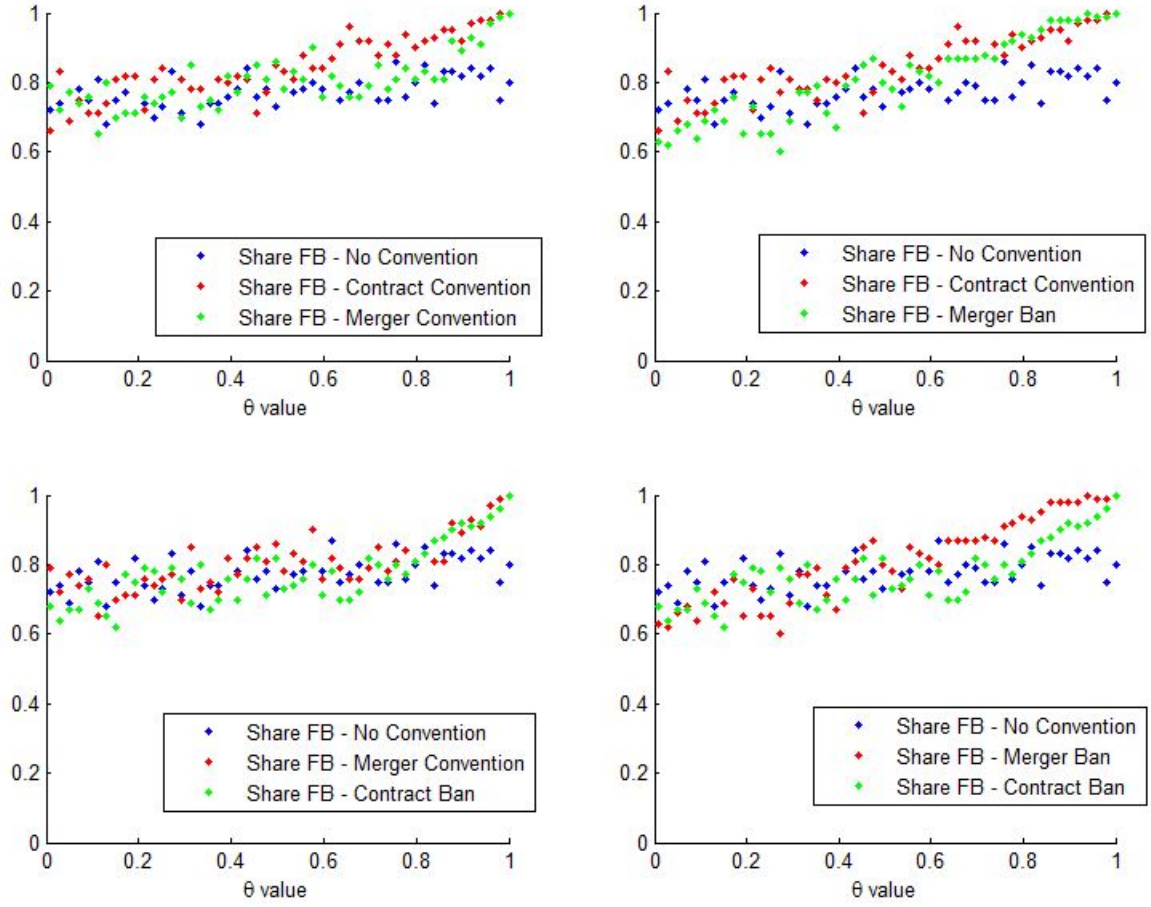

Figure 3.18: First Best Comparison, maximum revenue gap

Appendix to Chapter Three

Endowing firms with equal productivity is a useful simplification, and although the calibration focuses on aggregate data over an entire country, a more detailed empirical analysis would need to take into account variations in firm productivity. As such, I provide here changes to the model that would be needed to allow for heterogeneity in productivity.

The simplest case is to have two productivity values. The environment would be adjusted such that firms have productivity z_i with $i \in \{L, H\}$, and where $z_H > z_L$. The share of firms of the high type would be σ, and the posting distributions for the high and low types would be $P_H(.)$ and $P_L(.)$, respectively. This implies that:

$$P(r) = \sigma P_H(r) + (1 - \sigma) P_L(r)$$

Proposition: For $\eta > 0$, $\bar{r}_H > \underline{r}_H = \bar{r}_L > \underline{r}_L$. In other words, the overall posting distribution is continuous, but all high type postings are greater than all low type postings in equilibrium. As a proof, assume that there is some $r_2 \geq r_1$, where r_2 is in $P_L(.)$ and r_1 is in $P_H(.)$. Given profit maximization, this implies:

$$(z_H - z_L) n(r_1)^{\alpha} \ell(r_1)^{1-\alpha} \geq (z_H - z_L) n(r_2)^{\alpha} \ell(r_2)^{1-\alpha}$$

The above requires $r_1 \geq r_2$, which can only hold at $r_1 = r_2$, thus demonstrating that the equilibrium must be continuous and non-overlapping in productivity.

Non-overlap of the postings and a process similar to the derivation of the basic model yields an expression for the posting distribution:

$$P_i(r) \;=\; \frac{(d+\delta+\eta\lambda)}{\eta\lambda}\left(1 \;-\; \left(\frac{\alpha z_i k_i^{\alpha-1} \;-\; r}{\alpha z_i k_i^{\alpha-1} \;-\; \underline{r}_i}\right)^{\frac{1}{2}}\right)$$

Where k_i is capital per worker used by type i firms. Due to the single wage, capital per worker must vary by firm type in order to ensure a common marginal product of labor across all workers, which means that $k_L > k_H$. This new posting distribution has a larger dispersion than the basic model. If we were to add additional types, the dispersion would continue to widen. An empirical analysis using firm level data would thus need to take into account dispersion caused by productivity heterogeneity in order to isolate dispersion caused by search frictions.

Note that by definition, $P(\bar{r}_L) \;=\; P(\underline{r}_H) \;=\; 1-\sigma$. If we select z_H and z_L such that aggregate productivity is the same as the single z in the basic model, the average rents will be the same, meaning that the dividend paid to households will be the same. This in turn, will cause no change to the saving decision. Thus, heterogeneity in firm productivity expands the rent distribution but has no other effects.