Optimal Contracting and Spatial Competition among Financial Service Providers

Gustavo Joaquim [*] Robert Townsend [†] Victor Zhorin [‡§]

December 31, 2018

Abstract

We present a contract-based model of industrial organization for markets characterized by information and other frictions (Moral Hazard, Adverse Selection, Limited Commitment etc.) and different market structures (Monopoly, Oligopoly, Competition), the latter driven by spatial costs, logit errors, and number of financial service providers. We show this method can be applied to understand and quantify the impact of spatial and technological changes in the banking sector in emerging market countries. We derive a likelihood estimator for the structural parameters that determine contracting frictions and market structure, but also establish methods, depending on counterfactuals of interest, that do not need to specify both. We illustrate our framework using simulated data, illustrating competition of local, relationship based banks versus less-informed national banks with a spatial cost advantage. Using real data from banks and entrepreneurs in the Townsend Thai Data, our results indicate that reducing spatial costs by 50% is equivalent to increasing consumption by 4.85%, which we compare to other policies. Our larger goal is to develop an operational, broadly applicable toolkit for empirical work.

[*] MIT, e-mail: gpgj@mit.edu

[†] MIT and NBER, Corresponding author: rtownsen@mit.edu

[‡] University of Chicago, e-mail: vzhoring@chicago.edu

[§] We thank Nikhil Agarwal, Varadarajan Chari, David Donaldson, Glenn Ellison, Amy Finkelstein, Thomas Holmes, Ariel Pakes, Christopher Phelan, Michael Whinston, and seminar participants at Minneapolis Fed, MIT, Harvard, University of California/Berkeley-Haas for very useful comments, ASSA and SAET conferences. This work used the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant number OCI-1053575. We acknowledge the University of Chicago Research Computing Center for support of this work. We gratefully acknowledge research support from the National Institute of Child Health and Human Development (NICHD), the research initiative "Private Enterprise Development in Low-Income Countries" (funded jointly by the CEPR and the DFID), the John Templeton Foundation, the CFSP at the University of Chicago (funded by Bill & Melinda Gates Foundation), and the BFI through the Macro Financial Modelling Fellowship.

1 Introduction

We have data on villages that were randomly selected for various forms of financial intermediation. As a substantial number of baseline villages had no formal financial service provider (FSP), a subset was randomly selected to receive a bank that has been offering credit and insurance targeting entrepreneurs. Contrary to our expectations, we observe that the villages that received the FSP services had lower average consumption than non-treated and a higher standard deviation of consumption in the cross section of the treated relative to the non-treated villages. What can explain this apparent puzzle?

Additionally a random subsample of treated intermediated villages were selected for an additional program. A new big data based screening system was introduced, essentially eliminating potential adverse selection on observables. To our surprise, welfare increased in the villages where the new bank and the screening system were introduced, while it decreased in villages where only the screening system was implemented (but not the new bank). What can explain this disparity?

We show that, in a model with risky production where FSPs can offer both credit and insurance, it is challenging to infer welfare changes from changes in average/standard deviation of observables - production and consumption - when contracts offered by FSP change with the interventions. Moreover, the model shows that the introduction of a screening system that moves an economy from adverse selection to full information can have different effects on welfare depending on the level of competition in the intermediation market. With adverse selection, the FSPs does not know how to differentiate the entrepreneurs, and thus it cannot extract the rents it would otherwise in a full information case.

Our context is illustrative but the problem is general. Typically, key sectors in the economy are characterized by contracting frictions and varied degrees of competition among contract providers. Salient examples are health care, finance and insurance. However, although economists have long been interested in problems of competition and, *separately*, contracting, there are few frameworks that consider of these questions together. We provide a more extensive literature review below. Settings where the products offered have inherent risk, unobserved types or are dynamic, can be challenging to understand without a specific framework. Yet, at the same time, parts of the system can be ideitified without imposing much structure, depending on the counterfactuals of interest. The conceptual framework makes that clear, too.

This paper develops a framework to solve, estimate and apply models of simultaneous competition and contracting. We *solve* a contract-based model of industrial organization that allows us to consider in a unified way both different information frictions (moral hazard, adverse selection, both) and a variety of market structures (monopoly, imperfect competition etc.). The model has implications for profits and market share of contract providers and for the distribution of consumption, income and capital of agents in the economy. This allows us to use a likelihood method to *estimate* the deep parameters of the economy. We show how our framework can be *applied* to analyze the impact of the spread of the banks and increased financial access in emerging market countries. Our larger objective in this paper is to develop a tool kit, an operational empirical framework. That is, our ultimate goal is to do for industrial organization and contract theory what Doraszelski and Pakes (2007) did for industrial organization and steady state dynamics. So we do not shy away from reporting what we know about computation.

As an overview of our method, we construct a theoretical framework focused on utilities generated by contracts rather than the contracts themselves. This draws on a theory literature using promised utilities as a key state variable (Green (1987), Spear and Srivastava (1987)). Once we define the model in terms of utilities, most of the usual toolbox of competition is at our disposal. The framework is then divided in two building blocks: the utilities and profits frontier and the market structure. The frontier, as in Karaivanov and Townsend (2014), represents the profits of an contract provider for a given level of utility of the client. Then, in the next step, the market structure pins down the market share of a contract provider given the utilities a given contract provider and its competitors are offering. The division of the model into these two building blocks is not only pedagogical, but makes economic sense: changes in contracting frictions only alter the initial frontier block and all strategic interactions are contained in the market structure block.

For more detail about the model, we analyze contracting between a group of entrepreneurs and a set of financial intermediaries. Entrepreneurs in the model are risk-averse households running small and medium enterprises (SME) in need of external credit and insurance. Financial Intermediaries are risk-neutral banks that provide contracts for households and compete with each other. The financial regimes we consider: full information (complete insurance/perfect credit); unobserved effort (moral hazard) but with complete information on types; limited commitment to repay loans (limited commitment) and unobserved types (adverse selection). The market structure we consider is based on a demand system where SMEs have, in addition to the basic spatial structure, idiosyncratic preferences for intermediaries that generate logit market shares. In the framework proposed, it is possible to guarantee existence and uniqueness of the Nash equilibrium among intermediaries. This Nash equilibrium can be easily computed numerically through an iterative algorithm.

Typically banks in developing countries are geographically sparsely located, with relatively few branches and few banks operating in a given area. Travel to branches is non trivial in terms of time, repeated customer visits, and visits of credit officers to the field. We thus focus on the absence of centralized markets and concentrate on bank lending and competition among relatively few banks. The actual structure of observed bank contracts (credit and insurance arrangements for SMEs) is not simple, i.e. does not fit the stylized contracts of theory, of borrowing at interest with collateral and fixed term payments, with presumed repayment but allowing for default. Instead, typical contracts offered by banks represent a blend of credit and insurance, e.g, loans are rolled over, some interest is forgiven, and indeed there are well known and explicit contingencies under which an effective indemnity is paid and some or all of principal is written off (as if paid with the indemnity) . In sum, our formulation is not only more flexible, it is more realistic, and it matches real observable outcomes (income, savings, consumption etc.) to the model.

We conduct several counterfactual exercises in our model to understand the terms of loans offered and how each contracting friction and the market structure affects the equilibrium contracts and real outcomes. We focus first on the results for full information and moral hazard/limited commitment.[1] Using a spatial model as Hotelling (1929), D'Aspremont et al. (1979), Prescott and Visscher (1977), we provide several numerical results. First, we illustrate how changes in the equilibrium utilities changes real outcomes - such as production and consumption - in a heterogeneous way across SMEs that are spatially separated. Second, local competition among providers can significantly increase utilities, yet interestingly, more so under moral hazard and limited commitment than under full information. Third, reduction in spatial costs can increase or decrease welfare of SMEs, as it creates local monopolies, which are able to charge more for financial services. Forth, we show that the way market shares changes when contracts (and utilities) change is a key determinate of welfare. If SMEs are not likely to change FSPs based on which contracts they offer (either through regulation, lack of financial literacy etc.), more competition or reduction in spatial costs are not effective to increase welfare and reduce the financial costs. These exercises are not only useful to understand the inner workings of the model, but they illustrate important mechanisms in reality. One can think of advancements in the banking sector as new branches, reduction in spatial costs (through technology), or changes in the elasticity of demand (through contract platforms, for instance). All of these changes increase competition. Our model contributes to an understanding of who benefits the most and how to quantify which policy change (e.g. spatial costs vs network branch extension) is more effective in increasing welfare - and how this depends on imperfect competition and the underlying financial frictions.

We focus on the case of adverse selection separately, given its empirical relevance and theoretical complications. We tackle this problem through two different ways. First, we propose a more analytical based form where we impose enough structure in the model such that incentive compatibility constraints only bind for adjacent types. Second, we propose a brute force numerical method to compute equilibrium utilities in more general settings. We analyze the implications of the model in two relationship lending scenarios: two intermediaries competing, but the SMEs types are known only for one of them, and local vs national banks competing. Local banks are informed, but national banks can have spatial advantages (e.g., through a well developed app). Our results indicate that relationship

[1]In the context of our model, adverse selection is different from other contracting frictions due to the fact that the utility offered to one type of agent potentially changes the frontier for other types of agents.

lending can significantly increase within-region inequality between those that have access to credit coming from better information about them, and those who don't. This effect is larger for larger spatial costs, where markets are more isolated. We also show that if market shares significantly change with utilities, most bad types use financial services from the uniformed provider, bringing with it systemic risk consequences. This creates a rationale for not inducing too much competition in the banking sector. Moreover, we show that even when subject to different information constraints/spatial costs, local and national banks can exists if competition is not fierce and spatial costs are relatively low.

Our framework implies an equilibrium among financial service providers. This equilibrium makes endogenous the distribution of promised utilities faced by households/firms as an outcome, along with branch locations, profits, and market shares. The structure can thus be estimated with full information maximum likelihood techniques, comparing both information regimes and market structures. We show the construction of the likelihood functions in two different datasets: (i) data on FSPs market share, as is typically available, and (ii) data on the locations of FSPs and consumption, income and capital for SMEs, from the Townsend Thai project, which is a unique dataset.

One can use market share data as with the usual logit model (Berry (1994)), but instead of assuming a linear utility based on observables, we assume that the utility is generated by the equilibrium in the model. More interestingly, we show how to recover the contracting frontier from market share data only. For that, we show that variation in spatial configuration and competition across markets allows us to identify the frontier non-parametrically. This allows researchers to conduct market structure counterfactuals without having to take a stand on which contracting friction is relevant. We showcase our methodology in two counterfactual exercises using simulated data: changes in spatial costs and the introduction of one additional FSP in a location.

Using the Thai data set on the locations of FSPs and consumption, income and capital for SMEs, we develop a numerical method based on Bresnahan and Reiss (1991) and Karaivanov and Townsend (2014). We use a model of entry Bresnahan and Reiss (1991) with information of number of banks in each location. We extend the methodology of Karaivanov and Townsend (2014) which maps unobserved equilibrium utilities to equilibrium contracts in the model, which implies a joint distribution for consumption, capital and income. We discuss how to estimate the model in the presence of unobserved heterogeneity (such as initial asset positions), or the equilibrium utility themselves (which bypasses the need to define the market structure block of the model).

We apply our general method to the the Townsend Thai Data on bank locations, travel time between banks and villages micro level data on SMEs in each village - consumption, income and capital - to estimate the structural parameters of our model. Our counterfactual exercises indicate that reducing spatial costs and the variance of idiosyncratic preference shocks by 50% (one at a time) can increase welfare by, respectively, 4.85% and 15.36%.

Bank entry has a limited effect on welfare in our sample. Overall, our results suggest that to increase welfare, policy makers need to guarantee that markets shares change when utility offerings change. This means policy makers should pursue policies that make SMEs more likely to choose better financial products, such as financial literacy, platforms where financial services can be easily compared and bank correspondents, rather than simply increasing the number of FSPs.

Our broad motivation for this research is both positive and normative. On the positive side we seek to understand better the industrial organization of financial service providers in terms of both the geography of branches and expansion over time as well as in terms of the actual loan/insurance contracts which are offered. On the normative side, we seek to answers policy questions such as the coexistence of local and national banks and the role of information and competition (Petersen and Rajan (1995)); the impact of deregulation which alleviates artificial geographic or policy/segmentation boundaries (Brook et al. (1998), Demyanyuk et al. (2007)); and the welfare and distributional consequences of different market structures, different obstacles to trade (information, trade costs) (Koijen and Yogo (2012), Martin and Taddei (2012)), and the interaction of these obstacles with market structure. Our broad goal in this paper is to develop a toolkit to answer those questions.

Related Literature. Our work is close in spirit to the work of Einav et al. (2010), Einav et al. (2013) on health care, Lester et al. (2018) for insurance and Einav et al. (2012) on auto loans except that we try to make few restrictions on contracts to see how far we can get and we add in a market structure model for competition side in financial services. For that, we bring two literatures together.

First, as in Karaivanov and Townsend (2014), we move beyond fixed contracts and compute solutions for arbitrary information regimes. [2] Karaivanov and Townsend (2014) shows how to estimate financial/information regimes for SME's, distinguishing moral hazard constrained lending and insurance, versus more limited contracts, using Townsend Thai project data on consumption, income, investment, and capital stock, at a point in time and over time as in the panel.

Second, we use a simplified version of the supply side for financial service providers as in Assuncao et al. (2012). Assuncao et al. (2012) uses data on the timing and location of the opening of new branches for both the commercial banking sector and government banks (in the same setting, Thailand). When there are only a few branches around, households would need to travel relatively long, time consuming distances to get to a branch or choose to not participate in the (formal) financial system. As new banks/branches enter, the market

[2] Lustig (2010), following Berry et al. (1995) and Berry (1994), assumes products are characterized by a low, finite dimensional vector of product characteristics. In this literature, the characteristics of products are the observed contracts (deductible, benefit programs), and the price (with a separable, negative impact on households utility). The demand from households for these characteristics is driven by extreme value additive errors and by random coefficients. The latter is characterized by a distribution of diversity of types in the population. For an alternative presentation of those methods see the excellent review paper of Einav et al. (2010), which goes beyond medical markets.

catchment areas effectively evolve. The key point is that a "market" is not a fixed object with heterogeneous characteristics and the environment is not modeled as being in a steady state. Here we report on work to bring these these two strands together with both the location of bank branches and the contracts they offer as endogenous (though our framework allows for regulatory restrictions if we choose to further restrict the environment exogenously), to match the contracts we see in reality and allow for those we do not see out of equilibrium. [3]

Moreover, our model is at the same time flexible in terms of contracts and can be taken to the data in several different forms. Therefore, we move beyond calibrated examples of models of Azevedo and Gottlieb (2018) and others. We illustrate this in the Townsend Thai Data, using a numerically efficient method. We additionally show that depending on the counterfactual of interest, one does not have define all the blocks of the model: either the contracting block in some case or competition blocks in other cases, useful of course in empirical applications.

An alternative approach to modeling imperfect competition would be the one in Lester et al. (2018), which allows for analytic solutions of contracts and equilibrium for the Adverse Selection case. Here, we opt for a model where market power comes from spatially separated SMEs and FSPs, and idiosyncratic preferences (which deliver a logit demand system) to generate market power. We opt for this version due to evidence that distance is relevant in financial contracts (Nguyen (2018)) and to directly speak to most of the IO literature on discrete choice.

Though this paper is specifically aimed at developing methods that could be applied to the financial institutions of Thailand and other emerging market countries, we believe these same methods could be applied to other markets and more developed countries. For example, we have a lot in common with (and interest in) the parallel work on medical contracts and health markets in the U.S. To be more specific, as a lead example, there is a literature on selection into private health care providers, for example, under the Medicare Advantage program. Outcomes vary across local geographic markets which vary in the evident degree of competition (number of providers) and in the contracts offered (degree of product variety); see Dunn (2009) on selection and competition in Medicare Advantage, Einav et al. (2013) on behavior responses under exogenously introduced variation in medical contracts.

Our welfare metrics also differ from some, if not all, of the applied literature. Even in oligopoly, financial firms operate on obstacle-constrained Pareto frontiers, without price distortions or loss of social surplus. The equilibrium outcomes on these frontiers, the contracts, and the division of gains between households and firms, are determined by obstacles and the degree and nature of market competition. In particular, we do not employ the welfare metric of an unconstrained utilitarian social planner but rather respect the information incentive and truth telling constraints and also the timing in most competitive market struc-

[3] In Assuncao et al. (2012), the authors propose and solve a dynamic game between banks. Here, we simply take the branch locations as given, and focus on the implications on contracts.

tures.

The paper is organized as follows. In Section 2 we discuss the interpretation of reduced form evidence mentioned at the outset, in settings with contracting under financial frictions and market power. In Section 3 we discuss how to write the model in terms of utilities and present the two building blocks: the contracting of the model and the market structure. We show that the equilibrium is well defined in this model and conduct a few exercises to illustrate the implications of the model for contracts and for consumption and income dynamics. We then move to the specific case of adverse selection in Section 4 and discuss the model's implications for relationship lending. After presenting the theoretical framework, we show how to construct likelihood functions, in Section 5 for two different types of datasets. One is based on market share data, more common in the IO literature. The other is based on the Townsend Thai data, that uses household level data. We use the likelihood functions for the Townsend Thai Data in Section 6 and provide parameter estimates and counterfactual experiments. Finally, Section 7 concludes the paper and points on directions to future research.

2 Motivational Evidence

Before presenting our framework, we discuss how to interpret reduced form data in settings with market power in intermediation and contracting. We focus on two research questions. First, we explore the effects on consumption and production of introducing a new FSP into a given region (varying the degree of competition). Second, we study the effects of the introduction of a screening system (like a credit registry) that essentially eliminates selection on observables village wide (i.e., accessible to all FSPs).

2.1 Introduction of FSP in Villages

We have data on 500 villages, half of which were randomly selected to receive a new FSP that provides credit and insurance to entrepreneurs. Each village has on average 70 households. The villages that did not receive the additional FSP are the control group, while the villages that receive it are the treatment group. We are interested in the following question: what the effects of this introduction on welfare of villages?

We consider two different subsamples. First, we focus on an area where, before the intervention, there was no FSP in a village - both in the treatment and controls groups. This set is comprised of 100 villages in each of the control and treatment groups. To answer our proposed question, we start by computing the average and standard deviation of consumption and production. The results are in Table 1.

When we compare the treatment group, that is, the villages that now have a FSP, with the control group of no FSPs, we find that average consumption *decreases*, an unexpected result.

However, if insurance is relevant in these villages, then the drop in consumption might be a premium, so we also examine the standard deviation of consumption. Yet, the standard deviation in consumption *increases* in the treatment group. As can be seen in column 3 of Table 1, these differences are statistically significant. Simultaneously with all of this, production increases. A potential conclusion could be that the introduction of the FSP was not welfare enhancing for these villages. Although production increases, FSPs charged too much for financial services and entrepreneurs are worse off than they originally were. Apart from behavioral explanations, this is of course a puzzle. Entrepreneurs had the *option* of using intermediation services, or not. Thus their welfare should not decrease. We refer to this puzzle as the *consumption-production puzzle*.

Table 1: Outcomes from Randomly Introducing a FSP: No FSPs in Baseline

	Control	Treatment	Difference
Avg. Consumption	2.2089	2.1344	-0.0745***
	(.0223)	(0.0257)	(0.0216)
Std. Dev. of Consumption	1.8780	2.1590	0.2810***
	(0.0193)	(0.0309)	(0.0290)
Avg. Production	2.2089	6.0863	3.8775***
	(.0223)	(0.0756)	(0.0652)
Std. Dev. of Production	1.8780	6.3856	4.4624***
	(.0193)	(0.0820)	(0.0839)

Standard errors in parenthesis, computed through 1,000 bootstrap resamples from collected sample of 200 villages in this subsample where originally there were no FSP. Each village has on average 70 households, which we are aggregating over to generate averages and standard deviations. *** denotes 1 % significance.

To begin to address the puzzle, we compare changes in consumption *conditional* with changes in production, by running a regression, as in Eq.(1) at the household level. For a household h, in village v, we compute changes in consumption as a function of production, a dummy for treated villages and the interaction of treatment and differences in production.

$$\Delta c_{h,v} = \beta_0 + \beta_1 \mathbb{1}_{v \in T} + \beta_2 \Delta p_{h,v} + \beta_3 \mathbb{1}_{v \in T} \Delta p_{h,v} + \eta_h \tag{1}$$

where c is consumption, p is production, Δ is the difference post and pre intervention, superscripts a village is in the treatment group if $\mathbb{1}_{v \in T} = 1$. We recover $\hat{\beta}_1 \approx 1.04$, $\hat{\beta}_3 \approx -1.25$, both significant at 1%. At the average production level, the effect of a new FSP on consumption is negative, and consumption decreases by more for those that produce relatively more, which adds to the puzzle.

In our second subsample, both the control and treatment groups had at least one FSP before the intervention. This subsample has 150 villages in each of the treatment and control groups. We present the outcome statistics in Table 2. There is now no puzzle in this setting. Consumption increases on average, and the changes in its standard deviation are simply because consumption is larger on average, [4] as would be expected in the first place when thinking about a setting with production where risk and insurance is second order. What can explain the differences between Table 1 and Table 2?

Table 2: Outcomes from Randomly Increasing FSP Competition in Control and Treatment Villages.

	Control	Treatment	Difference
Avg. Consumption	2.1344	2.7442	0.6098***
	(.0257)	(0.0331)	(0.0074)
Std. Dev. of Consumption	2.1590	2.7759	0.6169***
	(.0309)	(0.0397)	(0.0088)
Avg. Production	6.0863	6.1039	0.0176
	(.0756)	(0.0755)	(0.0250)
Std. Dev. of Production	6.3877	6.3865	-0.0055
	(.0803)	(0.0802)	(0.0323)

Standard errors in parenthesis, computed through 1,000 bootstrap re-samples from original collected sample of 300 villages in this subsample where originally there were no FSP. Each village has on average 70 households, which we are aggregating over to generate averages and standard deviations. *** denotes 1 % significance.

2.2 Introduction of Screening System

We are interested in answering the following question now: what is the welfare effect of introducing a village wide screening system (that is, that all FSPs have access to) that virtually eliminates adverse selection on observables?

In our subsample of 300 villages that originally had a FSP operating, we first selected a group of 150 to receive a new FSP. We sequentially select a random set of village to receive the screening system. We have now 4 types of villages: those that randomly were assigned to receive a new FSP (or not) and those that were assigned to receive the new screening system (or not). In each of these subgroups, we end up with 75 villages. Given our results from the previous section in terms of consumption and production dynamics, we put some structure onto the problem and calculate welfare though inputs on consumption and hours.

Table 3 reports differences in welfare for treated and untreated villages. In Column 1, we report the difference for the case where we compare villages where there was a new FSP introduced. In Column 2, we report the difference for the case where we compare villages where there was no new FSP introduced - just the screening system. More specifically, we define a village that received a new financial service provider as $v \in T$ and $v \in S$ for the villages that receive the screening system. What we show in Table 3 (Column 1) is given in

[4] In particular, we can compute the same statistics for the coefficient of variation (average over standard deviation) of consumption. The difference in this case is not statistically significant.

Eq. (2), while in Column 2 we simply replace $v \in T$ for $v \notin T$.

$$\Delta W \equiv \sum_{v \in T, v \in S} W_v - \sum_{v \in T, v \notin S} W_v \quad \text{and} \quad \Delta W \equiv \sum_{v \notin T, v \in S} W_v - \sum_{v \notin T, v \notin S} W_v \tag{2}$$

where W_v is the average welfare of households in village v (that we infer welfare from a structural model of consumption and production).

When there is the introduction of the new FSP, we see that welfare *increases* by eliminating information problems in intermediated markets, as one would expect. When there is no introduction of new FSPs, however, we see that the result is exactly the opposite. Household welfare falls significantly as a result of the introduction of the screening system. We denote this as the *information structure puzzle*. What can explain the difference between the results in the different subsamples?

Table 3: Introduction of a Village Wide Screening System: Welfare Changes

	New FSP	No new FSP
ΔW (Eq. 2)	0.0940***	- 0.2662***
	(0.0106)	(0.0113)

We have four subsets of villages depending on if there was or not the introduction of a new FSP and the screening system, each with 75 villages in it. In this table, we compare the welfare of villages in the same FSP setting, but with different screening technologies as in Eq. (2). Standard errors computed through 1,000 bootstrap repetitions. *** denotes 1 % significance.

2.3 Taking Stock

Consumption-Production puzzle. Although the movements in consumption, production, an welfare are presented as a puzzle, the data used to compute the moments in Table 1 and the regression results are generated through an experiment that is run in a model, with **model generated data**. The model features entrepreneurs that have a risky production process and are risk averse. Entrepreneurs are heterogeneous in their productivity, which is unobserved by the econometrician, but observed by FSPs. FSPs compete to provide credit *and* insurance.

The average consumption is reduced in Table 1 as entrepreneurs prefer to insure consumption - and pay for it. Variation in consumption, however, increases because most of the variability in consumption comes from the changed cross sectional heterogeneity in productivity of entrepreneurs, as contracts change before and after the intervention, from autarky to ones offered by FSPs, and not from risk in production. Variation in consumption does drop dramatically for each type. The details of the model are in Appendix A. In particular, the consumption equivalent gains in welfare in treatment villages with respect to control villages due to the intervention are of 91.45% - a large effect yet still not successfully estimated from consumption. This explains the *consumption-production* puzzle.

Eq.(1) delivers a negative estimate of β_3 not only because of the theoretical results of the model, the moments of the Table 2, but due to endogeneity. Even though we have a perfect experiment (since we simulate in the model), changes in production and the error, η_i depend on the unobserved productivity of entrepreneurs and, thus, we would need some instrument, or pre-intervention data estimating TFP, in order to estimate $\hat{\beta}_3$ correctly. In settings where outcomes depend on an unmodelled heterogeneity of individuals, it is not enough to randomize across villages to get rid of endogeneity in a regression, as the error is potentially also a function of the entrepreneurs types - and thus correlated with the regressor.

In the data displayed in Table 2 we do not observe the puzzle. In the model that generates the data, contracts do not change in more competitive markets, with more FSPs [5] only the price of intermediation changes as market power changes. Intermediation gains are divided among agents and FSPs in proportion to market power for every type. This proportionality factor is a function of type, but the overall weighted average increases. There is no variance for each type as FSPs can pool and eliminate idiosyncratic risk. Thus, in this case, changes in average consumption perfectly track changes in welfare.

The key message here is that if competition does not change contracts, then experimental evidence is enough to identify the welfare the effects of the intervention. If contracts do change, however, reduced form evidence is not sufficient. In our more general model, competition does change contracts and, thus, we need the model to interpret the data.

Information Structure puzzle. The data on the *information structure puzzle* are also model generated. There are two types of entrepreneurs in each village, $\theta_L < \theta_H$, now unobserved by the FSP and the econometrician. The distribution of types of entrepreneurs is the same in treatment and control villages. We assume that both regions have a market power in intermediation indexed by $\omega \in (0,1)$, where $\omega = 0$ is perfect competition and $\omega = 1$ is a monopolist. We leave the details and equations of the data generating process of Table 3 to Appendix B.

The difference between the results in Table 3 comes from differences in market power in the underlying economies. With new FSPs, all villages (treatment and control) have a relatively more competitive intermediation sector (average of $\omega = .3$), while with no new FSPs all villages are in an economy with market power in intermediation ($\omega = .7$). If the FSPs have enough market power (high ω), *agents are better off in an environment with AdS.* As the FSPs cannot distinguish between the agents, it cannot extract the rents of the full information case.

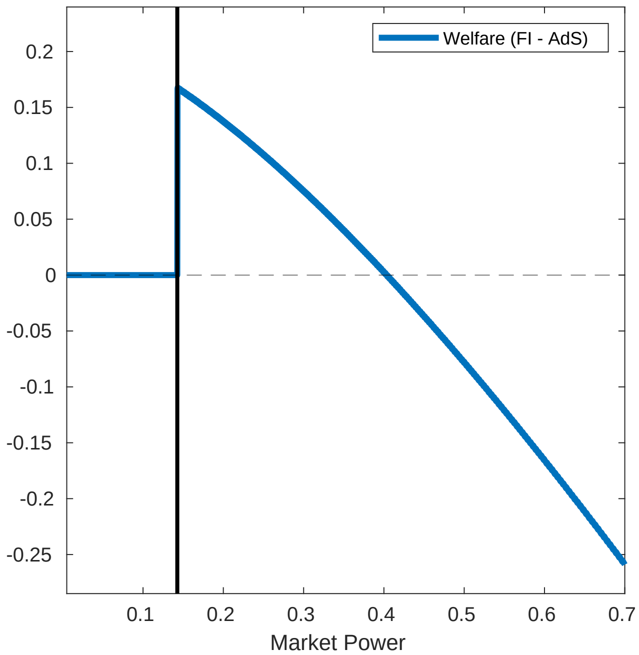
We show the difference between welfare [6] in the AdS selection case versus the Full information in Figure 1 for various levels of market power. The vertical line is the minimum

[5] This is not a general statement, but rather the outcome of a very specific model. See Appendix A for more details.

[6] We compute total welfare as the simple average of individual welfare.

level of market share such that the adverse selection selection constraint is binding. Adverse selection is only not binding if the market is competitive, that is, if ω is low. With little competition, the transfers for each type are sufficiently different - since intermediaries keep most of the surplus of the trade - that no type wants to take the quality-transfer pair of the other. From Figure 1 it becomes clear that we cannot extrapolate the effects of changing the information structure without taking into account the market structure.

Figure 1: Welfare Effect: The Introduction of the Screening System (From AdS to FI)



Note: Market power in this case comes from the elasticity of demand. $\nu = 0$ is a perfectly competitive economy, while $\nu = 1$ is the perfectly monopolist case. Welfare differences between adverse selection and full information for economies with contracting and competition. FSPs provide entrepreneurs contracts with leverage and insurance, but charge to do so according to their market power. Entrepreneurs can have high or low productivity, which is unobserved to the FSP in the AdS case. Welfare shown here is the average welfare of entrepreneurs in the economy. See Appendix B for details.

One example where this conclusion is relevant for policy is the introduction of credit score systems. Brazil, for instance, is in the process of introducing a credit score system (*Cadastro Positivo* in Portuguese), but the banking sector is extremely concentrated, with one of the highest spreads in the world (39.37 p.p. annually). [7] Our analysis suggests that the credit system could make entrepreneurs worse off if ω is high. Studies based on other countries and settings cannot be extrapolated without taking into account the market structure of the banking sector.

Takeaway. In our first example, when contracts are changing, it is challenging to correctly estimate the welfare effects of the intervention. In our second example, we change gears and consider that welfare is observed, but that the researcher is trying to interpret changes

[7] 5-Bank Asset Share around 85 % in 2016, Source: WDI).

in the information regime (from adverse selection to full information). There is an external validity problem that comes from market power. If market power is high, adverse selection is welfare increasing, since it reduces the ability of FSPs to extract surplus. This means we need a market structure model to interpret the evidence, which is exactly what we develop in this paper. We develop a model of welfare and competition that is at the same time flexible in terms of contracting (Section 3) and can be mapped into micro-level data (Section 5), allowing for the estimation of structural parameters and counterfactuals (Section 6).

3 Theoretical Framework

The theoretical framework is composed of two building blocks, which we denote as the frontier and the market structure. The frontier is defined as the profits of a FSP given that a contract must provide a given level of utility for an agent. The market structure defines the market share of a specific financial provider given a utility that it is offering. Profit for an intermediary is a multiplication of the two building blocks: profits it would have contracting with an agent (frontier) and market share (market structure). The frontier of the model is presented and compared for two contracting regimes: Full Information and Moral Hazard. The market structure is presented assuming a logit demand system, and we establish existence and uniqueness of a Nash Equilibrium in utilities.

Note that both the frontier and market structure are defined in terms of utilities, and not contracts. We change the contracting space from contracts (base on interest rate, collateral etc.) to utilities in the model for two reasons. First, contracts can have multiple and intricate dimensions: maturity, fixed and floating interest rate, covenants etc., while utility is a unidimensional object. As a unidimensional object, a representation in utility space allows for most IO tools designed for price (also an unidimensional object) to be applied in our setting. Second, our methodology in utilities allows us to easily encompass classic models of lending and borrowing with models of insurance and risk sharing. This expands the real of applications to insurance markets, healthcare or any sector characterized by incomplete markets.

The separation of the model in building blocks is not only pedagogical but carries an economic meaning and relates on the techniques used to solve the model. For different contracting frictions (e.g., Moral Hazard vs Limited Commitment), only the frontier block changes. For a different demand system from agents or a different number of banks, only the market structure changes. Importantly, given general conditions in the utility of agents, a logit demand system guarantees uniqueness and existence of a Nash Equilibrium through a contraction argument.

The key difference of the framework with the usual models of competition is the frontier. The market structure block is standard in the literature of IO. The reason why the frontier is different is due to the fact that it encompasses the contracting frictions we want to analyze.

In a standard Cournot model of competition, for instance, the frontier would be defined by price minus marginal cost. In our framework, the frontier will be defined by the solution of the contracting problem.

3.1 The Frontier

In this subsection we construct the profit of a FSP when the contract that it is offering provides a certain level of utility for agents. We start this section by arguing that we can move from the space of contracts to the space of utilities. In the textbook model of industrial organization, this step is not needed: the profit is simply price minus average cost (times quantities). In a model of contracting, however, the price minus cost of the profit function is more complex, since we must take into account the agent type and reaction to a contract. In the utility space, however, the profits of a FSP are represented by a Pareto frontier: the profit is the maximum profit that can be generated conditional on offering a level utility. The actual contract can then be recovered from argmax of the optimization problem. From this point forward, we thus refer to the the profit function as the *Frontier*.

To assume that we can move from contracts to utilities, we must assume that: (i) no contract that simultaneously generates higher profits for FSPs and higher utilities for agents exists and (ii) there are no two contracts that offer the same profit of a FSP and same utility for an agent. The first assumption is natural: it does not make sense for a contract to exist if there is a different contract that is both better for FSPs and agents simultaneously; that is, they need to be on the frontier. The second condition means that two different contracts must be different in a key variable for either FSPs or agents in our model. Condition (ii) is trivially satisfied, for instance, in a world where consumers are risk averse and FSPs are risk neutral. See Appendix D for a mathematical formulation of these ideas. From now on, we focus on utilities.

We consider an economy populated by output-producing households running small and medium enterprises (SME) in need of external credit and insurance. Households come into the economy with a capital $k \in K$ and a type $\theta \in \Theta$, and a utility $\mathbb{U}(c, z|\theta)$ for consumption $c \in C$ and an effort $z \in Z$. There is a production technology $P(q|k, \theta, z)$ available to all agents that determines the probability of output q being observed conditional on capital k and effort z [8]. Type θ is potentially a vector, and both preferences and the production function can dependent on it. We assume output and capital are observable and, thus, the contract can be made conditional on it. Define the profit of an intermediary that offers to type θ and capital k an expected utility $u \in W$ by $S(u|k, \theta)$. For now, we exclude the Adverse Selection (AdS) problem and assume θ is observed by FSPs (we specifically tackle AdS models in Section 4). The problem of FSPs that defines the frontier is given by Eq.(3). In this static

[8] We assume that $\forall\ k, \theta, z\ P(.)$ has full support. This avoids perfect information extraction from observed outcomes

contracting problem, the FSP prescribes the level of effort z and capital $k^{'}$ to be used in production and, once output q is realized, the level of consumption $c(q)$. The interest rate is r. (fixed, as a small open economy). The FSP can acquire the depreciated capital, $(1-\delta)k$, over and above $k^{'}$ (or the reverse, provide capital, if $k^{'}$ is larger). Later in this section we discuss dynamic extensions.

$$S(u|k,\theta) \equiv \max_{c(q),z,k^{'}} \sum_{q} P(q|k^{'},\theta,z)\Big\{q - c(q) + (1+r)\Big[(1-\delta)k - k^{'}\Big]\Big\} \tag{3}$$

s.t.:

$$\sum_{q} P(q|k^{'},\theta,z)\mathbb{U}\left(c(q),z|\theta\right) = u \tag{4}$$

$$\Gamma\Big(c(q),z,k^{'}|k,\theta\Big) \leq 0 \tag{5}$$

where Γ is a general representation of the contracting frictions, i.e., a set of frictions the contract must satisfy. Eq. (4) is the *Promise Keeping Constraint* and by varying u, we can construct the frontier of $S(.|k,\theta)$ points subject to this constraint.

To guarantee that the set of constraints is convex and to guarantee a solution, we write the above problem in the lottery space over discrete grids (as in Prescott and Townsend (1984) and, more recently, Karaivanov and Townsend (2014)). The discrete grids can be seen as a technological constraint or an approximation. The idea of the methodology is that instead of choosing allocations, the FSP chooses a probability distribution over allocations for each SME or equivalently a mixture for a certain group of clientele. More specifically, assume C,Z,K are discrete grids. In mathematical terms, the problem of FSPs is as in Eq. (6). [9]

$$S(u|k,\theta) \equiv \max_{\pi(c,z,q,k^{'})} \sum_{c,z,q,k^{'}} \pi(c,z,q,k^{'})\Big\{q - c + (1+r)\Big[(1-\delta)k - k^{'}\Big]\Big\} \tag{6}$$

s.t. Eq. (7)-(10).

$$\sum_{c,z,q,k^{'}} \pi(c,z,q,k^{'}) = 1,\ \ \pi(c,z,q,k^{'}) \geq 0 \tag{7}$$

$$\sum_{c,q,z,k^{'}} \pi(c,q,z,k^{'})\mathbb{U}(c,z|\theta) = \overline{u} \tag{8}$$

$$\sum_{c} \pi(c,\overline{q},\overline{z},\overline{k}^{'}) = P(\overline{q}|\overline{k}^{'},\theta,\overline{z}) \sum_{c,q} \pi(c,q,\overline{z},\overline{k}^{'}|k,\overline{u}),\ \ \forall(\overline{q},\overline{z},\overline{k}^{'}) \in Q \times Z \times K \tag{9}$$

[9]Note that there is an abuse of notation by using S,Γ in both problems. It is expected that S and Γ are different same across Eq. (3) and Eq.(6) due to the economics behind it - ability to offer lotteries - and the numerical approximation of the discrete grid if the true model has continuous supports for variables.

and the contracting frictions [10]:

$$\mathbf{\Gamma}(k,\theta)\,\pi \leq 0 \tag{10}$$

where $\mathbf{\Gamma}$ is a matrix. Eq. (7) is the condition that the probability elements are non-negative and sum to one. The constraint in Eq. (8) is the lottery version of the Promise Keeping Constraint. The constraint in Eq. (9) is the *Mother Nature* constraint. It limits the probability elements such that they respect the distribution given by the production function, P. [11]

We mainly use two contracting frictions in this paper: Limited commitment (LC) and Moral Hazard (MH). Both could be binding, or only one, or neither - as in the case of full information. In this case, the constraints in Γ are:

$$\sum_{c,q,k'} \pi(c,q,\overline{z},k')\mathbb{U}(c,\overline{z}|\theta) \geq \sum_{c,q,k'} \pi(c,q,\overline{z},k')\mathbb{U}(c,\hat{z}|\theta)\frac{P(q|k',\theta,\hat{z})}{P(q|k',\theta,\overline{z})}\ \forall\, \overline{z},\hat{z}\in Z,\ \forall\theta \tag{12}$$

$$\mathbb{U}(\rho\overline{q},\overline{z}|\theta) \leq \sum_{c} \pi(c,\overline{q},\overline{z},\overline{k}')\mathbb{U}(c,\overline{z}|\theta),\ \forall\, \overline{q},\overline{z},\overline{k}' \in Q\times Z\times K,\ \forall\theta \tag{13}$$

Eq. (12) is the *Incentive Compatibility Constraint,* it guarantees that, when effort is not observed, it is optimal for the agent to execute the effort recommended by the FSP. Eq .(13) simply states that if the FSP can recover $(1-\rho)$ of the output, the utility offered are such that the household has incentives to repay if it can keep the remaining ρ share of the output. The idea is that the household cannot default on k' (imagine that the bank lends for a household to buy a tractor and uses the tractor as collateral), but can run away in the end with a share of the income q generated in production (q). Note that this is one of the many possible ways of writing a LC constraint. [12]. As a benchmark, we also use the Full information (FI) problem, for which again the only constraints are given by Eq. (7)-(10).

The value function represents the profit of the FSP. Graphically, we expect it to look as in Figure (3). The concavity of S in u comes from the risk-neutrality of the FSP and risk-aversion of households. The *argmax* of the problems are the probabilities, π, which are a function of θ,k and u themselves, that is $\pi(c,q,z,k'|u,k,\theta)$.

The advantage of the methodology is that once we have $S(u|k,\theta)$, we can use all the IO techniques to solve and estimate models. In this paper, we provide a specific application to financial services where geography plays an important role, but one could apply the model

[10] Note that our restriction before as of the form: $\Gamma(c(q),z,k'|k,\theta)\leq 0$. Now, however, we are writing this as a linear constraint, i.e.: $\mathbf{\Gamma}(k,\theta)\pi\leq 0$. All constraints can be written this way. This allows the problem in Eq. (6) to be a Linear Programming problem in π, which can be easily solved numerically.

[11] To understand why it is written this way, note that it is equivalent to:

$$P(\overline{q}|\overline{k}',\theta,\overline{z}) = \frac{\sum_c \pi(c,\overline{q},\overline{z},\overline{k}')}{\sum_{c,q}\pi(c,q,\overline{z},\overline{k}')} \tag{11}$$

which is simply saying that the marginal distribution of q is consistent with the production function, P.

[12] For instance, an alternative would be to assume that capital can be partially recovered or that it introduces some type of leverage constraint.

to several other contracting problems or frictions, i.e., other S's. Finally, the transformation to utilities and the frontier concept also provide an exciting avenue for estimation of the frontier, that is: if all we need to know about the friction block is related to the frontier S, is there a way of estimating the frontier S *without* specifying the specific friction? We provide initial results on this in Section 5.1. As previously mentioned, the AdS case is more complex, and we tackle it specifically in a different section.

Feasible Utility Levels Given a grid for consumption, C, and for effort, Z, the contracting formulation we use implies endogenous levels of minimum and maximum utility: the minimum utility for a non-MH regime that can be assigned to a household is the value of consuming the lowest possible value of consumption and exerting the maximum value of effort. On the other hand, the minimum utility for a MH regime is assigning a minimum level of consumption, which is them followed by a household decision of exerting the minimum level of effort. With LC, the minimum value of consumption is ρq_{min}, that is, the non-recoverable share of the minimum level of production q. The maximum feasible utility in FI, MH and LC the utility with maximum consumption and minimum effort. Mathematically, the min and maximum utilities are as in Eq. (14)-(14).

$$w_{min} = \begin{cases} U(c_{min}, z_{max}), \text{if } FI \\ U(c_{min}, z_{min}), \text{ if } MH \\ U(\rho q_{min}, z_{max}), \text{if } LC \\ U(\rho q_{min}, z_{min}), \text{if } LC + MH \end{cases} \tag{14}$$

$$w_{max} = U(c_{max}, z_{min}) \tag{15}$$

3.1.1 Numerical Example and Optimal Contracts

To illustrate the frontier pictorially and the contracts, we present some numerical examples. We parametrize the utility function as:

$$\mathbb{U}(c, z \mid \theta) = \frac{c^{1-\sigma}}{1-\sigma} - \theta z^{\varphi} \tag{16}$$

where the type of a household, θ, represents a multiplier in the cost of exerting effort. For now, we focus on a unique type θ and normalize it to $\theta = 1$. We come back to multiple types θ in Section 4. We use the parameter values and grid for the contracting variables as in Tables 4-5.

We solve four different versions of the contracting problem in this section. First, a Full Information version without any contracting friction. Second, a version with MH only. Third, a version with LC only. Finally, a problem that combines MH and LC. Table 6 summarizes the problem and constraints. We leave the detailed discussion on computation later, when

Table 4: Parameter Values

Parameter	Constraint	Role
σ	1.5	Risk Aversion
φ	2	Disutility of Effort
θ	1	Effort Multiplier
ρ	.25	Share of Non-Recoverable Assets

Parameters for a utility function given by: $\mathbb{U}(c,z\,|\,\theta) = \frac{c^{1-\sigma}}{1-\sigma} - \theta z^{\psi}$.

Table 5: Grids

Variable	Grid	# Points	Points
Q	[0.04, 1.75]	5	$10th, 30th, ..., 90th$ p-tile in data
K	[0, 1]	5	$10th, 30th, ..., 90th$ p-tile in data
Z	[0,1]	3	uniform
C	[0.001, 1.75]	64	uniform
W	$[w_{min}, w_{max}]$	150	uniform

Grids and grid sizes based on the Townsend Thai Data (Section 6) and as in Karaivanov and Townsend (2014). The grid for consumption has enough points to guarantee that the frontier is smooth.

we discuss the numerical method.

As can be seen in Figure 2, the frontier for both levels of capital is decreasing and concave in utility. The higher the level of utility that must be offered for an agent, the lower the level of profit for a bank. Moreover, as the agent has a concave utility function and the FSPs is risk-neutral, higher levels of utility require marginally higher losses in profits. Moreover, the profit of FSPs is larger at any given utility for larger values of capital.

For a given level of capital, Figure 3 shows how the frontier is different for different frictions. As we input more frictions, the profit of FSPs decreases due to extra constraints in the contracting problem. Note, moreover, that under LC regimes, there is a significant loss in terms of feasible utilities that can be offered. This is due to the fact that to achieve this low values of utility, one would need to decrease consumption too much and agents would

Table 6: The Contracting Problem: FI, MH, LC and MH + LC

Solve: Eq. (6) s.t. Eqs. (7), (8), (9) , and:

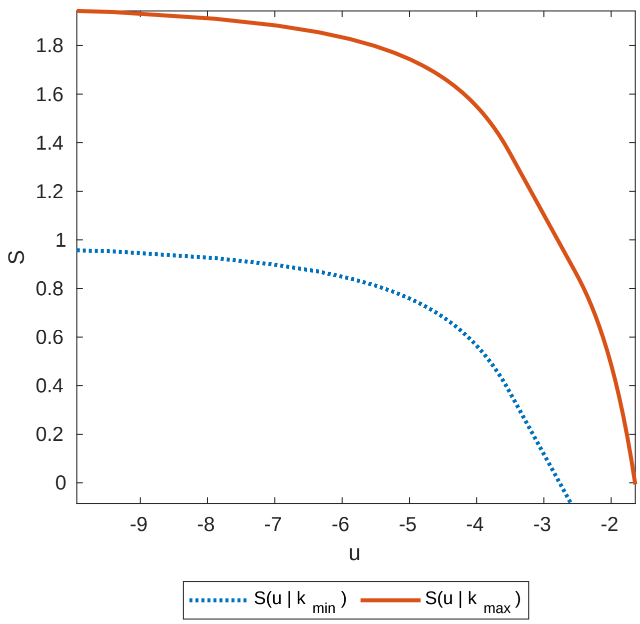
Friction	Constraint(s)
Full Information (FI)	-
Moral Hazard (MH)	Eq. (12)
Limited Commitment (FI)	Eq. (13)
MH + LC	Eqs. (12) and (13)

Linear programming problems solved with the Gurobi Linear Solver for Matlab for 150 utility levels equally spaced between $[u_{min}, u_{max}]$.

simply avoid paying back.

Figure 4 shows how the behavior of capital allows us to differentiate the behavior of capital between MH and non-MH models. If there is moral hazard, leveraging the project is a good way to increase risk in outcomes and, thus, increase effort for high values of utility. Note in Figure 5 how MH also induces the FSP to increase the standard deviation of consumption. For MH, the FSP needs to create risk to incentivize effort. [13]

Figure 2: The Profit Function for different levels of capital k



Note: the Profit function of FSPs under FI for two levels of capital, the 90th percentile and 10th percentile of what is observed in the Townsend Thai Data, which are k_{max} and k_{min} respectively. Linear programming problems solved with the Gurobi Linear Solver for Matlab for 150 utility levels equally spaced between $[u_{min}, u_{max}]$. See the Townsend Thai Data (Section 6) for details on the data and production function P.

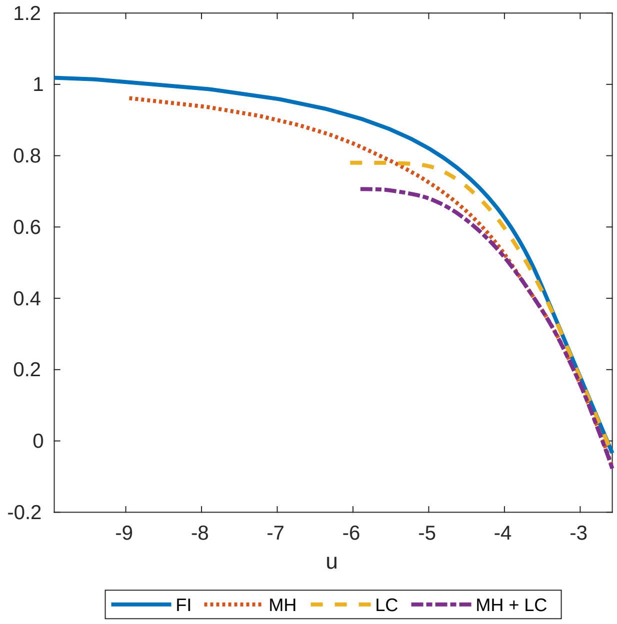
3.1.2 Extensions and Limitations

There are several ways in which the frontier can be extended to included new features. There are also several assumptions we must make such that the frontier is a valid representation of the contracting problem. In this subsection, we discuss possible extensions and limitations of the methodology.

In terms of extensions, one can consider dynamic contracting or a more parametric form of contracting. Dynamic contracts can be included if there is full or no commitment by both sides (households and banks). We provide here the full commitment version, but the problem can be re-adapted for no-commitment contracting. In a full-commitment case,

[13] Without the grids, we should expect the standard deviation of consumption to be zero under full information. However to achieve some levels of utility the FSP must use a non-degenerate lottery.

Figure 3: The Profit Function of FSPs as a Pareto Frontier: FI, MH, LC and MH + LC



Note: Standard deviation of consuption, effor and capital for implied contracts for four different contracting regimes: FI, MH, LC and MH + LC. Linear programming problems solved with the Gurobi Linear Solver for Matlab for 150 utility levels equally spaced between $[u_{min}, u_{max}]$. For this picture: $\theta = 1$ and k is the median k in the Townsend Thai Data (Section 6) for details on the data and production function P.

we follow Spear and Srivastava (1987) and use the promised utility representation. The idea is that we include future utility, $w^{'}$, as a choice variable and satisfy a promise keeping constraint to this variable in the next period. Including $w^{'}$ as a choice variable is consistent with choosing lotteries over it to represent future promises. We also assume here that θ now explicitly follows a Markov Process as in most applications. This allows us to write the problem recursively, as in Eq. (17). Let β^{FSP} be the discount rate of the FSP and β^{HH} the discount rate of households.

$$S(u|k,\theta) \equiv \max_{\pi(c,z,q,k^{'},w^{'})} \sum_{c,z,q,k^{'},w^{'}} \pi(c,z,q,k^{'},w^{'}) \left\{ S_0 + \beta^{FSP} \sum_{\theta^{'}} \mathbb{P}\left[\theta^{'} \mid \theta\right] S(w^{'}|k^{'},\theta^{'}) \right\} \tag{17}$$

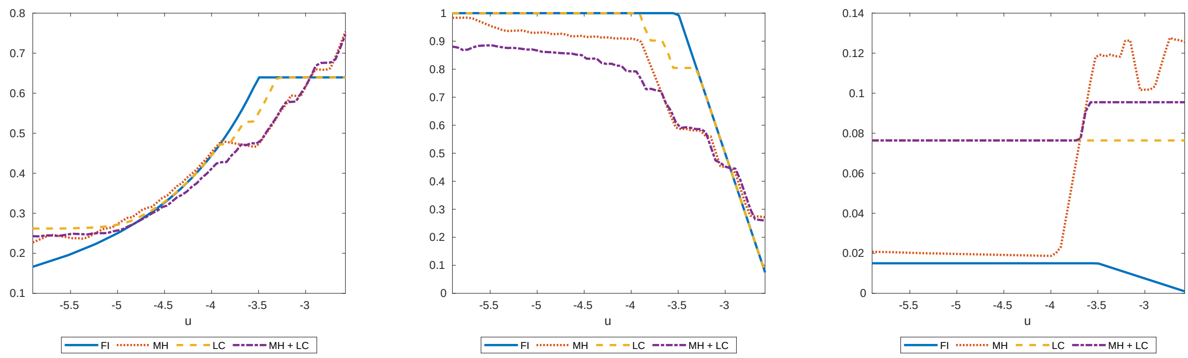
for $S_0 \equiv q - c + (1+r)\left[(1-\delta)k - k^{'}\right]$, s.t. the constraints on probabilities, mother nature, contracting frictions (which are similar to the static case) and:

$$\sum_{c,q,z,k^{'},w^{'}} \pi(c,q,z,k^{'},w^{'}) \left[\mathbb{U}(c,z|\theta) + \beta^{HH} w^{'}\right] = u \tag{18}$$

Eq. (18), together with the recursive term in the objective function in Eq. (17) guarantee that a promised utility for tomorrow, $w^{'}$, will be delivered tomorrow as a constraint such as

Figure 4: Contracts: expected consumption (c), effort (z) and capital (k) for varying levels of utility

(a) Consumption (b) Effort (c) Capital



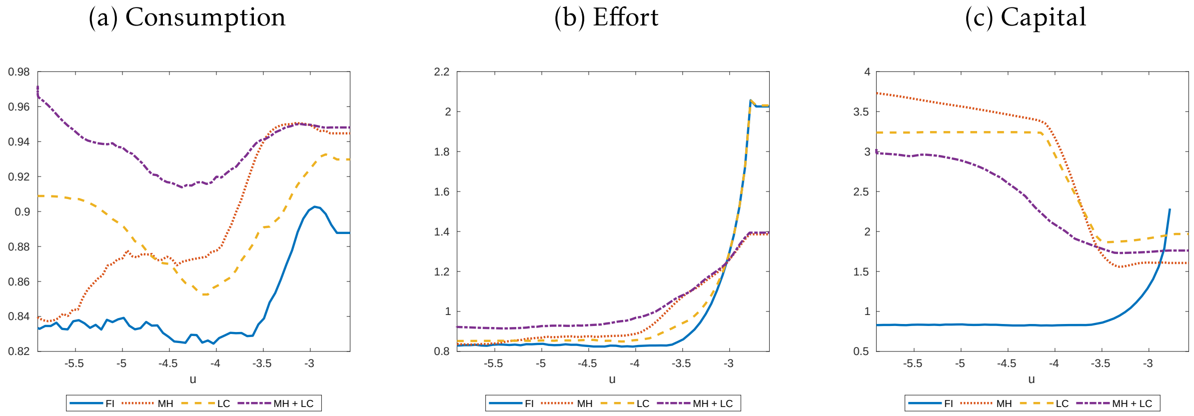
Note: The expected consumption, effort and capital for four different contracting regimes: FI, MH, LC and MH + LC. Linear programming problems solved with the Gurobi Linear Solver for Matlab for 150 utility levels equally spaced between $[u_{min}, u_{max}]$. For this picture: $\theta = 1$ and k is the median k in the Townsend Thai Data (Section 6) for details on the data and production function P.

Eq.(18). This contract specifies a path for utilities and capital and, conditional on those, at any moment choices of consumption and effort as in the static problem. Note that it all can be reduced to a single number, u, at the present due to full-commitment by households and FSPs to stay in the contract.

One can also do a more parametric version of the above problem that is similar to the one used in in Moll et al. (2017). In a more general form, any problem that can be written as a choice of lotteries over a discrete set with linear restrictions (which is a very general requirement) can be solved in the methodology and can be easily implemented in practice by changing the contracting block of our framework.

There are, however, limitations. The methodology does *not* encompass cases where the frontier itself depends on the strategy of the competitors. The two main examples are common agency and renegotiation models. In a common agency setting, the effort required by each FSP that relates with a specific household interacts with the contracts other FSPs are offering. In a renegotiation setting where the household has no commitment, the frontier would be not only a function of the current offered utility, but also the utility competitors are offering at any possible moment in the future going forward. The methodology is still useful since we could view the competitors strategy in the 'type' θ we used. However, this would be only feasible to be solved numerically only in a few specific settings. Overall, the issue of common agency and renegotiation is still very model dependent (e.g., Handely, Hendelz, and Whinston (Handely et al.)) and one would have to proceed case by case.

Figure 5: Contracts: Standard Deviation in consumption (c), effort (z) and capital (k) for varying levels of utility



Note: The standard deviation consumption, effort and capital for four different contracting regimes: FI, MH, LC and MH + LC. Linear programming problems solved with the Gurobi Linear Solver for Matlab for 150 utility levels equally spaced between $[u_{min}, u_{max}]$. For this picture: $\theta = 1$ and k is the median k in the Townsend Thai Data (Section 6) for details on the data and production function P.

3.2 Market Structure

In this subsection we focus on the market structure where FSPs compete with each other. Our model features a logit demand system and spatial differentiation between FSPs. We focus on competition in utilities *given* location. Within this market structure, we show that there is a unique Nash Equilibrium in utilities, which can be computed through an iterative algorithm.

More specifically, there are P independent markets in the economy. For each market $p = 1, ..., P$, there are B_p FSPs, located at a position $x_b \in \mathbb{R}^2$, $b = 1, ..., B_p$. Households reside in villages, denoted by $v = 1, ..., V_p$ in each market. Each village v has a population of N_v. We denote individual markets as *Maps*. A *Map* consists of the location of banks and households and the travel time between any two points in a province.

We assume in this paper that given a map configuration, competition among financial service providers generates the same output - regardless of entry order, identity of the financial service providers etc.. This is not an innocuous assumption. For instance, it does not hold in a dynamic competition model (e.g., Stackelberg), where there is a leader-follower dynamic. A paper that takes the sequence of entry into account would be very close to Assuncao et al. (2012) on the entry of private financial institutions vs BAAC in the Thai economy. However, we make this assumption to simplify the competition part of the model and focus on the interactions of the competition with the contracting frictions of the previous section. We discuss later in this section how more complex models of competition can still be solved within our framework.

As discussed earlier, instead of focusing directly on competition over contracts, we focus on competition in terms of **offered utilities**. This transformation in the choice space for FSPs reduces significantly the complexity of the competition game: instead of choosing a multidimensional vector of product characteristics, the FSP chooses the utility that the agent derives from the contract - and then figures out the optimal contract. As we reduce the choice space of each FSP to a unidimensional element (utility), most of the toolbox of industrial organization applies.

As we assume capital k and type θ are observed, the competition is in each level of (k,θ) separately. We come back to hidden type models in Section 4. As markets are independent, and in the interest of simplifying the notation, we drop market (p) and capital-type(k,θ) from the notation. We assume that there is a linear spatial cost in our economy ψ. In particular, the value a households located at location x_v attributes to a contract that offers u_b is given by Eq. 19

$$\mathbb{V}(u_b,x_b,x_v) = u_b - \psi t(x_b,x_v) \tag{19}$$

where $t(x,y)$ is the travel time between points x and y in the map. In our case, we use the GIS system and road maps to compute actual travel time. Note, more generally, that $t(x_b,x_v)$ does not need to be a distance based or spatial measure. It can represent any type of heterogeneity among households (preferences, information etc.). For instance, it could how much advertising FSP b runs at village v. Let u_0 denote the outside option of households (u_0 can be a function of (k,θ), but assumed to be the same in all markets). Finally, define u_{-b} as the vector of utilities offered by all FPSs (except b). Define:

$$\varphi_b \equiv \{u_b, u_{-b}, u_0, x_b, x_{-b}, \{x_v\}_v\} \tag{20}$$

as the vector of relevant variables in the profit of a financial service provider b, where the subscript $-b$ denotes the a variable for all other banks in a given province. We assume that the profit of a bank b is given by the surplus for a utility offer times the number of clients served Eq. (21)

$$\Pi(\varphi_b) \equiv S(u_b)\mu(\varphi_b) \tag{21}$$

In our empirical application, we will assume that there is a fixed cost of operating a FSP and a idiosyncratic shock to profits of FSPs. However, as we are considering so far the competition in contracts *given* their locations, we abstract from these.

Demand. The total demand of a financial intermediary b is given by the sum of the local market shares times the size of each market (population wise) in each location $v = 1,...,V$ where households reside [14]

$$\mu(\varphi_b) \equiv \sum_{v=1}^{V} N_v \mu_v(\varphi_b) \tag{22}$$

[14]As we are multiplying be the actual population, this corresponds to the total demand, and not a share. This does not change the problem of FSPs now, since it is simply a constant in the profit function.

The functional form we use for μ_v in the benchmark specification is given by Eq.(23)

$$\mu_v(\varphi_b) = \frac{e^{\sigma_L^{-1}[\mathbb{V}(u_b, x_b, x_v) - u_0]}}{1 + \sum_{\hat{b}=1}^{B} e^{\sigma_L^{-1}\left[\mathbb{V}(u_{\hat{b}}, x_{\hat{b}}, x_v) - u_0\right]}} \tag{23}$$

This is the textbook logit model and. It can be micro-founded with a extreme type 1 idiosyncratic preference shocks by agents to contract with each FSP with mean 0 and variance σ_L at the household level. The key difference from the usual logit is that we are now *structurally* modeling the utility offerings from the contracting problem and an equilibrium among FSPs. We use markets shares as in Eq. (23) for three main reasons. First, it speaks directly to the data. Without idiosyncratic preference shocks, a single FSP would always dominate the market of a given village, which we typically do not see. Second, it allows us to summarize in one parameter, σ_L, dimensions of the model that we are not considering that eventually affect the elasticity of demand. Third, it smooths the demand functions and guarantee existence and uniqueness of an equilibrium utilities. Given this market structure, we do now move on to proving that an equilibrium in utilities exists, is unique and how to compute it.

Equilibrium in Contracts. Given locations for financial intermediaries, $\{x_b\}_{b=1}^{B}$, the equilibrium concept for the solution in utilities we use is a Nash Equilibrium, i.e.:

$$u_b^* = \arg\max_{u \in W} \ \Pi\left(u_b, u_{-b}^*, u_0, x_b, \{x_v\}_v, x_{-b}\right), \ \forall b \tag{24}$$

note that the equilibrium is at the province, capital k, type θ and province p level, i.e.:

$$\{u_b^*(k, \theta | p)\}_{k \in K, \theta \in \Theta, p=1,\ldots,P}$$

but we chose to keep the notation concise.

Lemma 3.1 characterizes the equilibrium properties. It shows that under the assumption that μ is log-concave in u_b and Eq.(26) below holds (which we show to be true in the case of the logits, as in Eq. (23)), the equilibrium exists, is unique and can be computed by an iterative algorithm. We provide an intuitive explanation below. Before proceeding to the result, define $\varphi_b(a)$ as the variables relevant to the FSPs - utility that itself is playing, the vector of utilities that competitors are playing, outside option and locations - as in Eq. (25)

$$\varphi_b(a) \equiv \{u_b - a, u_{-b} - a, u_0 - a, x_b, x_{-b}, \{x_v\}_v\} \tag{25}$$

where in $\varphi_b(a)$ all utilities subtracted by a.

Lemma 3.1. *Let the demand μ, as in Eq.(22), be log-concave in u_b, log-supermodular in (u_b, u_{-b}),* *bouded away from zero and satisfy Eq. (26) $\forall\, a \in \mathbb{R}$:*

$$\mu(\varphi_b(a)) = \mu(\varphi_b(0)) \tag{26}$$

then $\exists!\{u_b^*\}_b$ that satisfies Eq. (24). Moreover, $\{u_b^*\}_b$ can be computed by an iteration of best responses starting at any strategy.

Proof. See Appendix E. ∎

The idea behind Lemma 3.1 can be represented pictorially. Imagine that both μ, S are continuously differentiable and abstract away from corner solutions. For notation purposes, let $\partial_x f(x) \equiv \frac{\partial f(x)}{\partial x}$. Given that μ is log-concave and S is concave, the optimum of $\Pi = S \times \mu$ can be computed by a FOC of the form in Eq. (27)

$$-\frac{\partial_{u_b} S(u_b^*)}{S(u_b^*)} = \frac{\partial_{u_b} \mu\Big(u_b^*, u_{-b}, u_0, x_b, x_{-b}, \{x_v\}_v\Big)}{\mu\Big(u_b^*, u_{-b}, u_0, x_b, x_{-b}, \{x_v\}_v\Big)} \tag{27}$$

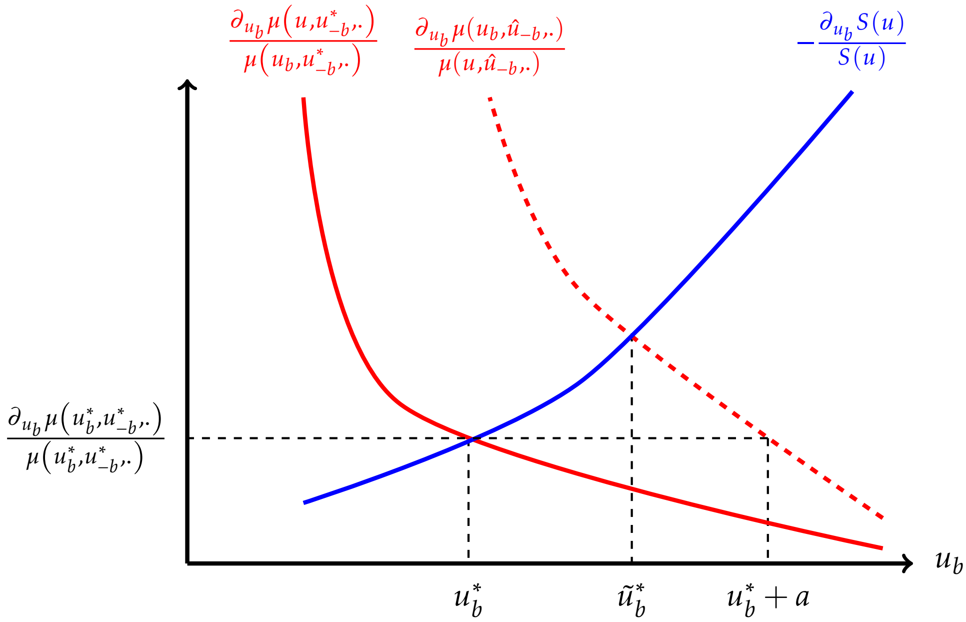
In Eq. (27), the marginal cost of increasing the level of utility by offering a better contract (RHS) is equal to the marginal benefit of a higher market share (LHS). The log-concavity of μ in u_b assumed in Lemma 3.1 guarantees that the RHS of Eq.(27) is strictly decreasing, while the concavity of S guarantees that the LHS is increasing. Pictorially, one can see this trade-off in Figure 6.

Consider that we are in an equilibrium $\{u_b^*\}_b$. Lets focus on a case where all other FSPs play the following deviation $\tilde{u}_{-b} = u_{-b}^* + a$, a a positive constant. As all other FSPs are playing a higher utility and we assume in Lemma 3.1 that μ is log-supermodular in (u_b, u_{-b}), we have that the RHS of Eq. (27) moves upwards. This is the monotonicity property of our equilibrium. Moreover, given that all other FSPs are offering $\tilde{u}_{-b} = u_{-b}^* + a$ and we assume that Eq.(26) applies, we have that by moving the utility a units up, we are back at the same level of market share as in the equilibrium u_b^*. However, as $-S''/S$ is increasing, the new optimum must be at $\tilde{u}_b \in (u_b^*, u_{-b}^* + a)$. This is the monotonicity property of our equilibrium.

Jointly, the monotonicity and discounting guarantee an unique equilibrium that can be computed through an iteration of best response functions.

In Lemma 3.2, we show sufficient conditions that guarantee Lemma 3.1 hold in our framework. The logit itself is log-concave, but the sum of logits in Eq. (22) may not be if there is enough variation in market shares across villages in a given market. This means we have to either bound the role of spatial costs with respect to the logit variance or the relative population between villages. One case in particular where all of this concern of log-concavity is irrelevant and is useful for other researchers is when $\psi = 0$ (i.e., there is no spatial cost). More specifically, if $\psi \leq \overline{\psi}$ as defined in Eq. (30), Lemma's 3.2 condition is satisfied. Moreover, if market shares are always smaller than .5 (i.e., a very segmented market), Lemma's 3.2 condition is satisfied.

Figure 6: Nash Equilibrium: Monotonicity and Discounting



Note: Pictorial representation of maximization of one FSP abstracting from technical details (non-differentiability, corner solutions etc.) as in Eq. (27). The equilibrium, u^*, is where the marginal benefit is equal to the marginal cost (for all FSPs, although the picture denotes only one). u^* is the baseline equilibrium and $\tilde{u}_b = u^*_b + a$ is the best responses to the deviation if all other FSPs increase their utilities played by a.

Lemma 3.2. *If for all banks $b, \hat{b}$ and for any two villages, $v, \hat{v}$, the spatial cost ψ and logit variance, σ_L imply that the market share at the village level, $\{\mu^b_v, \mu^b_{\hat{v}}, \mu^{\hat{b}}_v, \mu^{\hat{b}}_{\hat{v}}\}$ satisfies*

$$\sum_{i\in\{v,\hat{v}\}, j\in\{v,\hat{v}\}} \left[N_i N_j \mu^b_i (1-\mu^b_i)\mu^b_j \left(2\mu^b_i - \mu^b_j\right)\right] > 0 \tag{28}$$

and

$$\sum_{i\in\{v,\hat{v}\}, j\in\{v,\hat{v}\}} \left[N_i N_j \mu^b_i \mu^{\hat{b}}_i \mu^b_j \left(2\mu^b_i - \mu^b_j\right)\right] > 0 \tag{29}$$

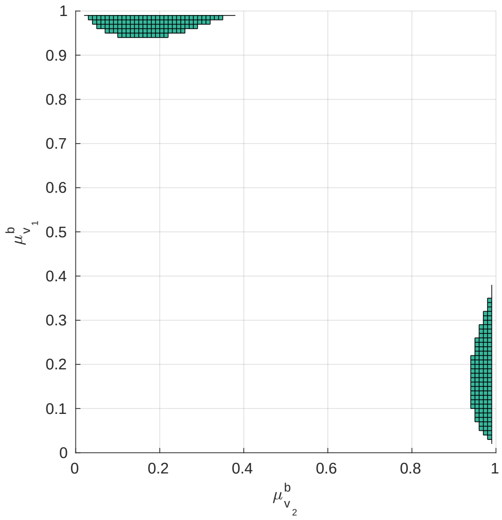
the market share defined by Eqs. (22)-(23) satisfies the conditions of Lemma 3.1. A sufficient condition for Eqs. (28)-(29) is that maximum ($\overline{\mu}$) and minimum ($\underline{\mu}$) market share between villages in the same market at any given level of utilities satisfies $\overline{\mu} \leq \underline{\mu}\left[4+\sqrt{11}\right]$. In terms of (ψ, σ_L), this means

$$\psi \leq \overline{\psi} \equiv \frac{\ln(4+\sqrt{11})\sigma_L}{\left[\max_{b,v}\|x_b - x_v\| - \min_{b,v}\|x_b - x_v\|\right]} \tag{30}$$

Proof. See Appendix F. ■

In practice, all parametric values we tested satisfy this condition. To provide an example, if $N_v = N_{\hat{v}}$, for instance, the regions shadowed in Figure 7 represent the combination of market shares in different villages for a given bank that would fail to satisfy the condition of Eq. (28) in Lemma 3.2.

Figure 7: Conditions on village level Market Shares to violate Log-Concavity



Note: Combination of market shares in villages v_1, v_2 for bank b, $\mu^b_{v_1}, \mu_{v^b_2}$, that guarantee Lemma 3.2 is satisfied assuming $N_{v_1} = N_{v_2}$. Shaded regions represent points where the condition is violated.

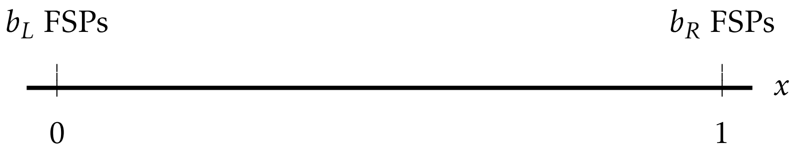
3.3 Comparative Statics

We explore how the spatial configuration, number of FSPs and contracting regime change the equilibrium in our model. Intuitively, reducing spatial costs and introducing FSPs can both increase welfare of villages in equilibrium. The contribution of our theoretical framework is to understand who benefits, how to quantify which policy change is more effective in increasing welfare - and the extent to which this depends on the underlying financial friction. First, we explore how consumption and effort vary spatially in a given equilibrium (Section 3.3.1). Second, we consider the effects of changes in the spatial cost (ψ) and the logit variance (σ_L) (Section 3.3.2). Third, we focus on an increase the number of FSPs in a given location (Section 3.3.3).

Throughout this section, we use the median level of capital observed in the Townsend Thai Data (more details on Section 6). More specifically, the frontier we use for each contracting regime is the one in Figure 3. The spatial configuration is a Hotelling line from $x = 0$ to $x = 1$ where FSPs are located in the extremes, with the set of V villages uniformly distributed in $[0, 1]$ (Figure 8). We denote b_L as the number of FSPs at $x = 0$ and b_R as the number of FSPs at $x = 1$. For this section, we assume that each village has a continuum of entrepreneurs, such that theoretical market shares correspond to actual market shares in the simulated data.

The parameters for the frontier are as in Table 4, while the baseline parameters in market

Figure 8: Spatial Configuration in Comparative Statics Exercises



Note: Representation of the spatial configuration in the numerical exercise. We assume that there are V villages equally spaced between 0 and 1.

structure are as in Table 7 below. To allow for an easier comparison between experiments, we re-scale utilities to be such that a zero utility represents the autarky level and a one utility represents perfect competition with full information level. Spatial costs here are given by $t(x,y) \equiv |x-y|$ for locations $x,y \in [0,1]$.

Table 7: Baseline Parameters used for Comparative Statics Exercises

Parameter	Value	Meaning
ψ	1	Spatial Cost
σ_L	.33	Logit Variance
V	50	Number of Villages
b_L	1	Number of FSPs in $x=0$
b_R	2	Number of FSPs in $x=1$

3.3.1 Heterogeneity Across Villages

Before conducting comparative statics exercises, we first show the equilibrium implications for each village in $[0,1]$. Using the parameters in Table 7, we solve for the equilibrium in utilities (with Lemma 3.1) and recover the implied equilibrium contracts. Note that there is one FSP at the left point, $x=0$, and two the right point, $x=1$. We compute market shares of FSPs and welfare for each village for two different contracting regimes: full information (FI) and moral hazard with limited commitment (MH + LC). The welfare in a given village is the market-share weighted welfare of households in that village, as in Eq.(31). [15] The results are in Figure 9. [16]

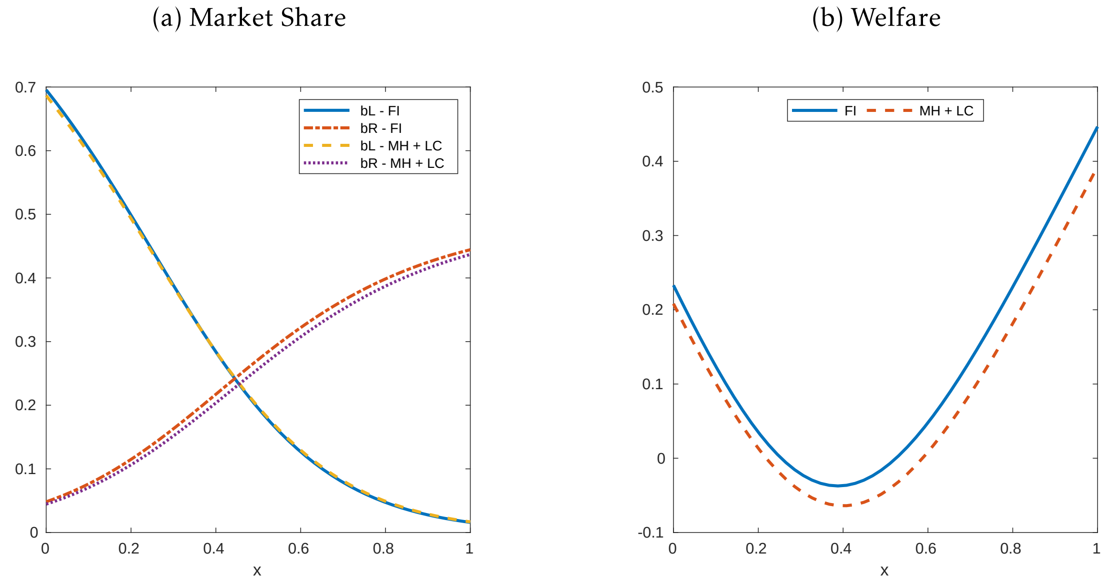
$$W_v(\psi,\sigma_L) \equiv b_L\mu_{v,x=0}\left[u_{v,x=0} - \psi t(x_v,0)\right] + b_R\mu_{v,x=1}\left[u_{v,x=1} - \psi t(x_v,1)\right] \tag{31}$$

where $\mu_{v,x=0}$ is the market share of the b_L FSPs located at $x=0$ for village v, located at x_v, and $\mu_{v,x=1}$ is the market share of one of the b_R FSPs located at $x=1$ for village v.

[15] Recall that we re-scale the utility levels to guarantee that the utility of the outside option is zero. This means we do not have to include the market share of the outside option times its utility in Eq.(31).

[16] Note that Eq.(31) is not the utilities played (which are $u_{v,x=0}$ and $u_{v,x=1}$) and they do not take into account the love of variety from the logit demand system (so is not a *ex-ante* measure of welfare). All measures would give similar results qualitatively.

Figure 9: Market Share and Welfare by villages located in $x \in [0,1]$ for FI and MH + LC



Note: Market shares by village in position x (as in Eq. (23)) and welfare (as in Eq. (31)) in the equilibrium with spatial configuration of Figure 8 and parameters of Table 7.

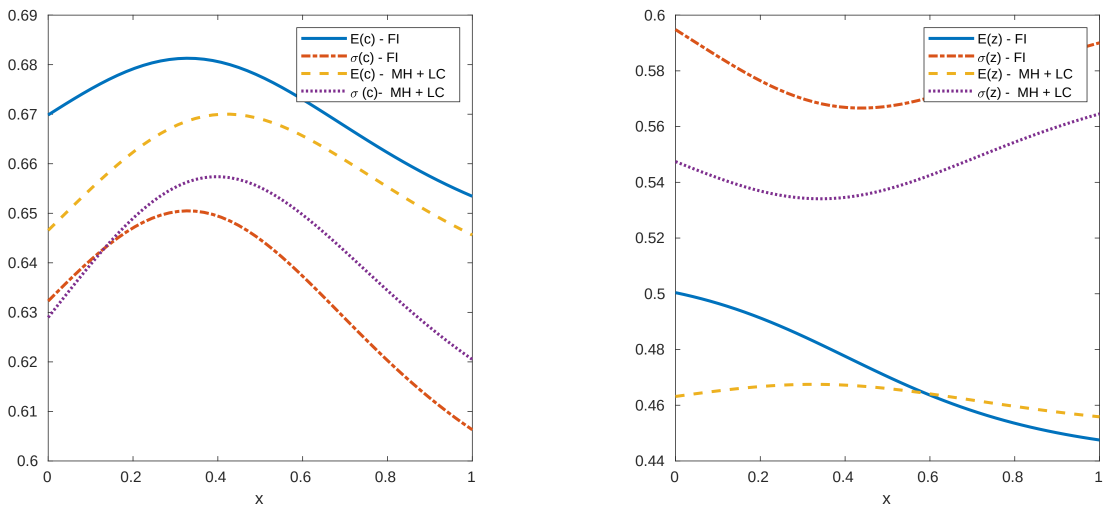
First, see in panel (a) the effects of spatial costs in which FSP provides more of their services for each village. As expected, villages closer to $x = 0$ mostly contract with the FSP in $x = 0$. The key model implication is how this curve decays as distance grows. In our baseline calibration, market share of the FSP at $x = 0$ decays from .7 to .02 in the closest to the furthest village. Second, in panel (b) one can see the effects of *local* competition. Higher utilities are played by the FSPs in $x = 1$, since we use $b_R = 2$ and $b_L = 1$ as our baseline. Third, the difference of utilities across regimes (FI vs MH + LC) is larger when there is more competition. At $x = 0$, where almost 70 % of households contract with the unique FSP at $x = 0$, the utilities are closer in the two regimes than at $x = 1$, where households contract with two FSPs.

Given this difference in utility levels between regimes and its differential spatial effect, we can see that the average and standard deviation of consumption and effort in villages will also be different. Note in panel (a) of Figure 10 that average consumption is always larger under full information contracting, but the difference is reduced closer to where theres is more competition (at $x = 1$, where there are two FSPs). The opposite is true for standard deviation in consumption. These results are a combination of different utility levels implying different contracts (Figures 4 and 5) weighted by different market shares at each region (Figure 9). Note that average consumption behaves in the exact opposite way of utilities. This is a consequence of insurance in the model. Note, for instance, how the standard deviation in consumption is also decreasing closer to the FSPs. The levels and heterogeneous behavior of consumption and effort across villages is what will allows us to identify the structural parameters later on. A model that ignores this spatial variation, as was the case of the model behind the experiment of Section 2, will mistakenly use this cross sectional variation in location as consumption variation, which was the source behind the consumption-production puzzle.

Figure 10: Consumption and Effort by village located in $x \in (0,1)$: average and std. deviation

(a) Consumption

(b) Effort



Note: Consumption/effort average and standard deviation by villages computed using the implied equilibrium utilities - to compute contracts - and market shares - to compute weights. Equilibrium with spatial configuration of Figure 8 and parameters of Table 7.

3.3.2 Spatial Costs and Logit Variance

Section 3.3.1 illustrates the heterogeneity across villages for a given equilibrium. We change gears now to how the equilibrium changes with changes in spatial costs, denoted by ψ, and the the logit variance, denoted by σ_L. For simplicity, we focus in an economy where there is MH + LC in contracting. [17]

Increasing spatial cost can increase or decrease the overall level of profits of FSPs depending on local competition (Figure 11, panel (a)). At $x = 0$, where there is only one FSP, increasing spatial costs has a non-monotone effect on profits. For low values of ψ, the FSP loses market share for a given level of utility and must offer higher levels of utility. For high values of ψ, however, profits increase as the market becomes more segmented, that is, close to a local monopoly. At $x = 1$, where there are two FSPs, increasing the spatial costs always decrease profits, since the local monopoly effect is reduced due to local competition.

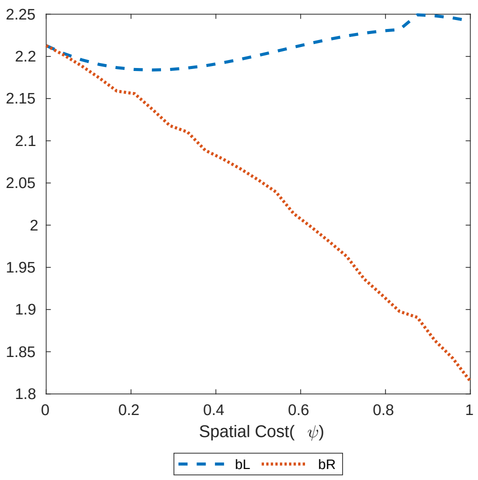
Increasing spatial costs also has non-monotone and heterogeneous effects across vilalges in terms of welfare. To illustrate this result, we compute the welfare of villages situated at

[17] The levels for profits and utilities are different depending on the contracting regime, but the qualitatively insights carry over for all contracting regimes we consider.

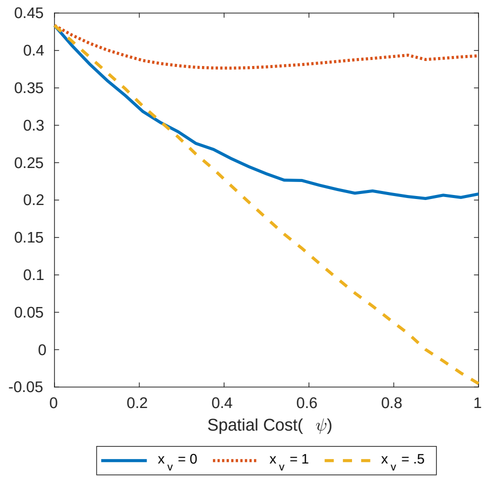
$x \in \{0, .5, 1\}$ as defined in Eq. (31). Without spatial costs, all villages have the same access to the three FSPs, and thus all have the same welfare (panel (b) of Figure 11). As spatial costs increase, a resident of village $x = 0$ not only has to potentially pay larger costs if it wants to visit FSPs at $x = 1$, but the utility being offered by FSPs at $x = 0$ is reducing (due to the creation of the local monopoly). At $x = 1$, where there are two FSPs, local competition eventually increases offered utilities to compensate for the rising spatial costs, which benefits those at $x = 1$ the most. For the households at $x = .5$, however, welfare is strongly decreasing when spatial costs are high, since all FSPs are significantly further away (recall that the welfare here includes the travel costs).

Figure 11: Profits of FSPs and Welfare of villages located in $x \in \{0, .5, 1\}$ as a function of spatial costs

(a) Profits



(b) Welfare

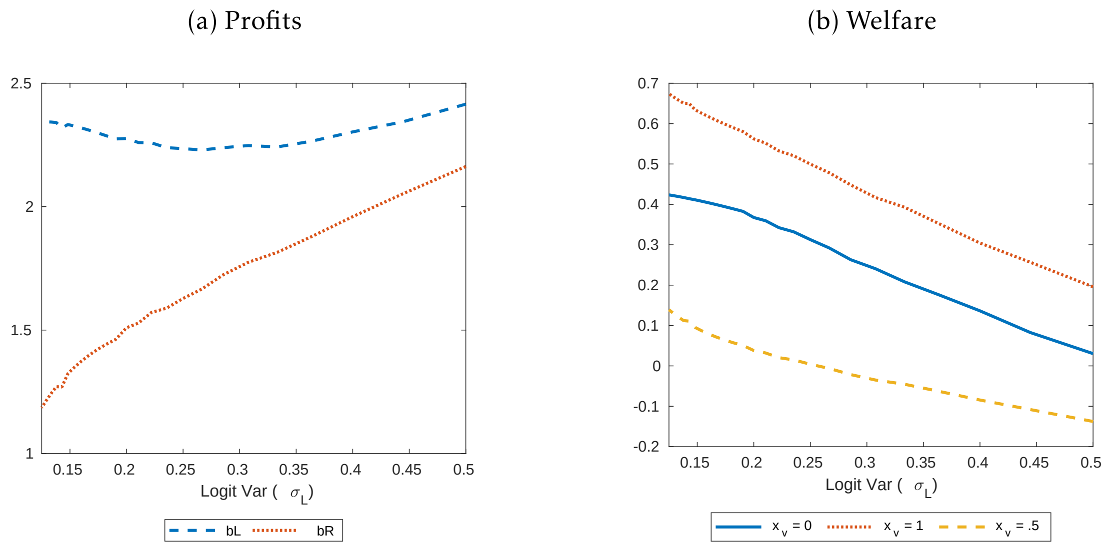


Note: Profits of FSPs and Welfare (as in Eq.(31) for three villages - the ones located in $x \in \{0, .5, 1\}$. Equilibrium with spatial configuration of Figure 8 and parameters of Table 7, changing the spatial cost parameter, denoted by ψ. Contracting frictions are MH + LC.

Figure 12 has the equivalent results for changes in the logit variance, denoted by σ_L. The scale of utility - which in here is exactly pinned down by the inverse of σ_L - is important for the equilibrium since it determines how market share changes with an equivalent change in utility. This can be seen in Figure 6. Changes in σ_L affect the downward sloping curve and, thus, the equilibrium determination.

For larger values of σ_L, market share changes more with larger utility offerings, which means that the marginal incentives of a given FSP to increase utilities in equilibrium is higher. Contrary to what we see with spatial costs, this effect is homogeneous across all villages. The differences in the panel (b) of Figures 11 and 12 is what allows us to identify these two parameters separately in the data. For larger values of σ_L, we should observe a smaller utility across all villages (which we know how to match to consumption data, for instance, as in Figures 4 and 5), while high spatial costs should lead to dispersion in utilities across villages in a given market.

Figure 12: Profits of FSPs and welfare of villages $x \in \{0, .5, 1\}$ with changes the logit variance



Note: Profits of FSPs and Welfare (as in Eq.(31) for three villages - the ones located in $x \in \{0, .5, 1\}$. Equilibrium with spatial configuration of Figure 8 and parameters of Table 7, changing the logit variance, denoted by σ_L. Contracting frictions are MH + LC.

We combine the previous exercises to understand if changes in σ_L and ψ are substitutes or complements, and how this changes with the level of competition in the economy. For simplicity, we change our baseline economy to be symmetric in locations and have one FSP at each location, that is $b_L = b_R = 1$. We compute the overall welfare in the economy as in Eq.(32), and plot the results in Figure 13, panel (a).

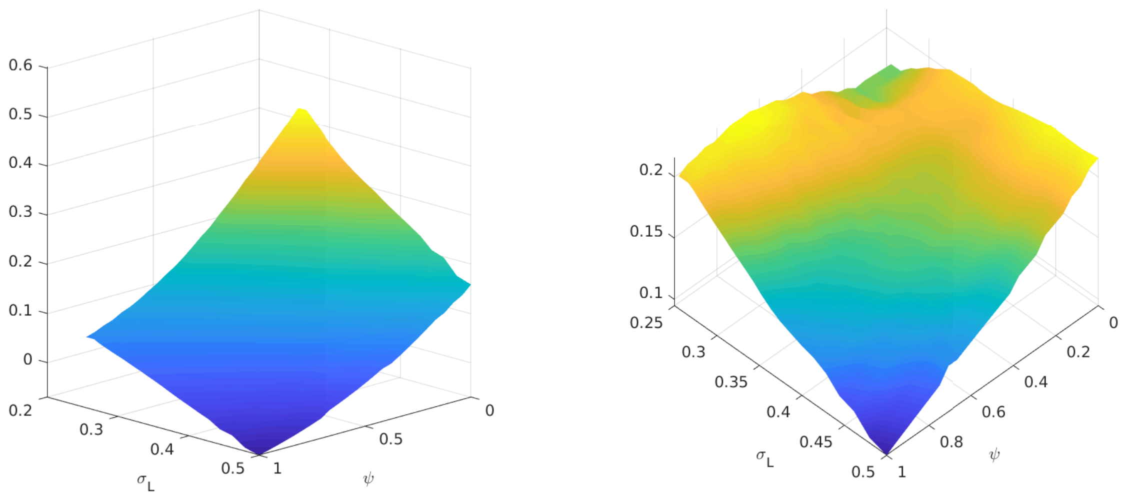
$$\mathcal{W}(\psi, \sigma_L) \equiv V^{-1} \sum_v W_v(\psi, \sigma_L) \tag{32}$$

where $W_v(\psi, \sigma_L)$ is the one defined in Eq.(31).

The effects of reducing spatial costs are more pronounced with lower values of σ_L, which indicates that if utility offerings do not sufficiently change market shares, reduction is spatial costs is also less effective to induce welfare changes. This can be seen in panel (a) of Figure 13. For low values of σ_L, welfare is larger *and* increases by more when spatial costs reduce than for high values of σ_L. In panel (b) we plot the difference in welfare between a more competitive economy with $b_L = 4 = b_R = 4$ with the welfare plotted in panel (a) with $b_L = b_R = 1$. Note that more FSPs increase levels of utilities, since the surface - the welfare differential - is all in positive numbers. This change is not constant across the parametric space of spatial costs ψ and logit variance σ_L. Reduction in spatial costs are passed through more to consumers when competition is higher. The difference is at its highest when either ψ or σ_L are sufficiently small, which is when FSPs have enough profitability to accommodate these changes.

Figure 13: Average Welfare varying spatial costs and logit variance

(a) Welfare with $b_L = 1, b_R = 1$ (b) Welfare Gains with $b_L = 4$, $b_R = 4$



Note: Panel(A): Welfare (as in Eq.(32). Equilibrium with spatial configuration of Figure 8 and parameters of Table 7, changing the logit variance, denoted by σ_L, and spatial costs, denoted by ψ. One FSP at $x = 0$ and one at $x = 1$, that is $b_L = b_R = 1$. Panel (B): Welfare difference between economy with $b_L = 4 = b_R = 4$ and the economy with $b_L = 1 = b_R = 1$. Contracting frictions are MH + LC.

3.3.3 Local Competition

Our last comparative static exercise involves changing the number of banks in a given location. We fix the number of FSPs at $x = 0$ at $b_L = 1$ and consider that FSPs at $x = 1$ can be in $b_R = 1, ..., 8$. As in the previous section, we consider the effects of the introduction of additional FSPs in profits and welfare of households in villages at $x \in \{0, .5, 1\}$. Finally, we show how profits and utilities in equilibrium differ between full information and moral hazard and limited commitment in this case.

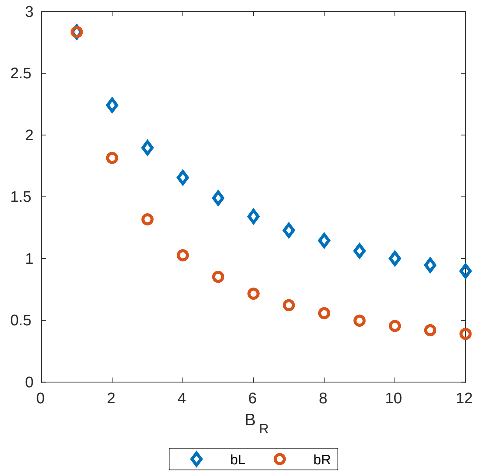
The introduction of FSPs at $x = 1$ increases welfare of the village at $x = 1$ by a significant amount - from 20 % to almost 50 % of the perfect competition full information utility. The effect on the village at $x = .5$ is qualitatively similar, but quantitatively smaller given the distance of this village to this new more competitive locale. In our framework, as we have logit market shares that come from idiosyncratic preferences of households within a village, competition in $x = 1$ can decrease welfare of households at $x = 0$, since some of them prefer to pay spatial costs to visit the FSPs in $x = 1$. This comes from our measure of welfare used. We do not take into account in panel(b) of Figure 14 the idiosyncratic preferences effects (that generate the logit market share), and thus it may seem that welfare is decreasing when in fact it is not. [18]

Finally, our last exercise is to focus on the interaction of contracting and changes in the

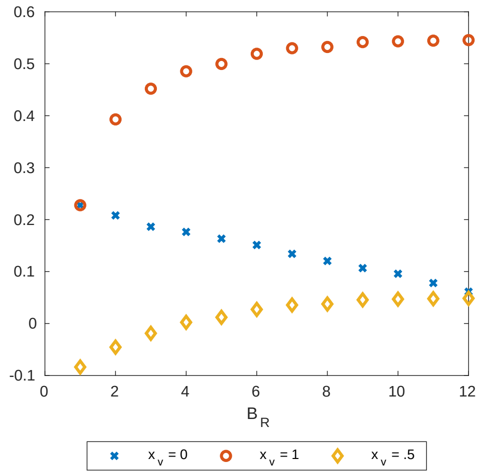
[18]The idiosyncratic preference shocks imply a few households will travel now larger distances. We capture the large distances in our measure of welfare, but not the effect of increased varieties.

Figure 14: Profits of FSPs and welfare of villages $x \in \{0,.5,1\}$ with changes in b_R

(a) Profits



(b) Welfare



Note: Profits of FSPs and Welfare (as in Eq.(31) for three villages - the ones located in $x \in \{0,.5,1\}$. Equilibrium with spatial configuration of Figure 8 and parameters of Table 7, changing the number of FSPs in $x = 1$, denoted by b_R. Contracting frictions are MH + LC.

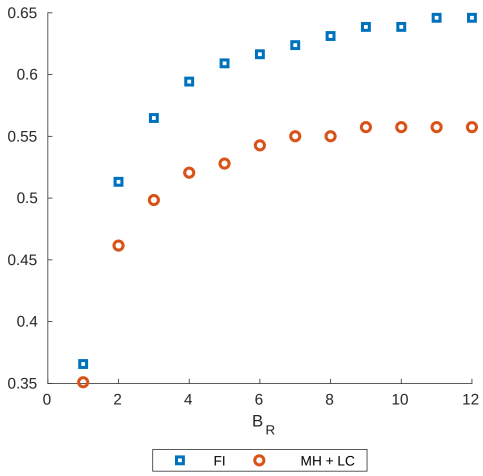
level of competition. We focus on the same comparative statics exercise: we fix the number of FSPs at $x = 0$ at $b_L = 1$ and consider that FSPs at $x = 1$ can be in $b_R = 1,...,8$. We compare the equilibrium utilities played by FSPs at BR under FI and under MH + LC. Not only utility is higher under full information, but the gains from competition are also larger. Our result suggest that local competition can be more or less effective depending on the contracting regime.

4 Adverse Selection

We explore now the case of Adverse Selection, where FSPs do not observe the type θ of households. Adverse selection is more complex than when types are observed because the frontier is now a function of the *contract menu* offered for all types, and not simply the utility offered for one given type θ. We show that under some conditions we can still apply the results of Lemma 3.1 and discuss a robust numerical method to solve models when we cannot.

As an application, consider a specific case of adverse selection with two types that differ only in their cost of exerting effort. We present the results of numerical experiments that allow for different contracting frictions and distance costs for FSPs. These numerical experiments can shed light on the competition of local banks versus national banks and the growth of fintechs. Relationship lending can significantly increase within region inequality between those that have access to credit through previous relationships and those who

Figure 15: Utilities in Equilibrium with changes in the number of banks in b_R for Full Information vs MH + LC



Note: Equilibrium utilities (as in Eq.(24) for FSPs located in $x = 1$. Equilibrium with spatial configuration of Figure 8 and parameters of Table 7, changing the number of FSPs in $x = 1$, denoted by b_R. Contracting frictions are FI (blue curve) and MH + LC (red curve).

don't. This effect is larger for larger spatial costs, where markets are more isolated. If market shares change significantly with utilities, most bad types use financial services from the national uniformed provider, bringing with it systemic risk, a rationale for not inducing too much competition in the banking sector. Even when subject to different information constraints/spatial costs, local and national banks can exists if competition is not fierce and spatial costs are relatively low.

4.1 Theory

Consider now the case where $\theta \in \Theta$ is not observed by the FSP. The FSP knows, however, that in the population the distribution of θ has a cdf $F(\theta)$ (with a p.d.f. $f(\theta)$). For simplicity, we focus on the case of Θ discrete. Given a promised utility level for all types $\{u^\theta\}_{\theta\in\Theta}$, the problem of a FSP for a given capital level k is given by Eq.(33)

$$S^{AdS}\left(\{u^\theta\}_{\theta\in\Theta} \mid k\right) \equiv \max_{\{\pi^\theta(c,z,q,k')\}_{\theta\in\Theta}} \sum_{\theta\in\Theta} \left\{ \sum_{c,z,q,k'} \pi(c,z,q,k') \left\{q - c + (1+r)\left[(1-\delta)k - k'\right]\right\} \right\} f(\theta) \tag{33}$$

s.t. Eq. (7)-(9) (the probabilities and mother nature constraints) and the *Truth Telling* constraint:

$$\sum_{c,q,z,k'} \pi^\theta(c,q,z,k')\mathbb{U}(c,z|\theta) \geq \sum_{c,q,z,k'} \pi^{\hat{\theta}}(c,q,z,k')\frac{P(q|k',\theta,z)}{P(q|k',\hat{\theta},z)}\mathbb{U}(c,z|\theta), \ \forall\hat{\theta},\theta \in \Theta \tag{34}$$

and, potentially, the other contracting frictions (MH, LC etc.). The difference is that now we added the truth telling constraint (Eq. 34). The menu of contracts must be constructed to guarantee that the agent reveals its true type θ when choosing from the menu. From the perspective of a FSP, contract choices cannot be done independently, that is, *the contract offered for a type impacts the frontier of the other type under truth telling*. Note that the constraint in Eq. (34) complicates the problem significantly, since we cannot separate the contracting problem for different types. Without the constraint in Eq. (34), we could separate the sum in θ in independent problems (the case of Section 3.1).

Simplified Case: Ordered Types and Binding Constraints. We first consider a simpler case of AdS where the structure of the problem allows us to use a result similar to Lemma 3.1.

Our first simplifying assumption is that utility is separable between consumption and effort. Our second assumption is that SMEs only differ in one characteristic cost of exerting effort etc.. In particular, we focus on the utility function in Eq. (35) [19]

$$\mathbb{U}(c,z|\theta) \equiv u(c) - \theta v(z), \;\; \theta \in \Theta \tag{35}$$

Eq.(35) provides an ordering of types according to their cost of exerting effort. We denote θ_L as the good (lowest type) and θ_H as the bad (highest type). We additionally assume that the only truth telling constraints that are binding are those of a lower θ taking the contract of a higher one, that is [20]

$$\sum_{c,q,z,k'} \pi^{\theta}(c,q,z,k')\mathbb{U}(c,z|\theta) \geq \sum_{c,q,z,k'} \pi^{\hat{\theta}}(c,q,z,k')\mathbb{U}(c,z|\theta), \;\; \forall\theta, \;\; \forall\hat{\theta} > \theta \tag{36}$$

This is *not* an innocuous assumption. In models of AdS and competition we do not know which constraints are binding. Given that the FSP cannot extract all rents, the parameters of the model (as, for instance, the share of each type in the population) determine the incentives of FSPs to distort the allocation across types. For more details on that see Appendix B, where we make this point mathematically for our simple model that generated the data in Section 2.

With these assumptions, we can prove Lemma 4.1, which is a version of Lemma 3.1 for the case of AdS. The intuition behind Lemma 4.1 is the same as in Lemma 3.1. For instance, if all competitors raise their offerings of utilities, an FSP would like to raise it for both types. As the truth telling constraint requires both to increase at the same time, this is what the FSP ends up doing. Thus, the equilibrium still satisfies monotonicity. An analogous reasoning

[19] We can generalize the assumption to be that SMEs differ only in one characteristic: cost of exerting effort (as here), or risk aversion, productivity etc.. We can also generalize this utility function to be $\mathbb{U}(c,z|\theta) \equiv u(c) - v(z|\theta)$, with $v(z\,|\theta)$ increasing in θ and $\partial_{z,\theta} v > 0$. We focus on the simplest case here for exposition purposes.

[20] The ratio $\frac{P(q|k',\theta,z)}{P(q|k',\hat{\theta},z)}$ does not appear because we assume the only difference between agents is over preferences, as in Eq. (35).

shows discounting.

Lemma 4.1. *Assume that all conditions for Lemma 3.1 are true. Additionally, assume that the utility function is as in Eq.(35). Finally, assume that only the truth telling constraints that potentially bind are those in Eq.(36). Then, a Nash Equilibrium in utilities exists, is unique and can be computed iteratively.*

Proof. See Appendix G. ■

Lemma 4.1 is useful for two reasons. It highlights that under a few conditions, it is possible to solve AdS simply models within our framework. At the same time, however, it also highlights the limitations of analytical frameworks to deal with AdS problems. As a result, we move on to a more general, numerical method.

General Case: To solve for the equilibrium in utilities with the Frontier as in Eq.(33), we use a *distance-to-Nash* algorithm (See Appendix H) for details. The idea of the algorithm is to write the Nash Equilibrium as an optimization problem (instead of a fixed point one). We do not have proofs of existence or uniqueness for the equilibrium in this case, but our numerical method always finds an approximate Nash equilibrium (up to) specified computer precision. This is the method we apply in our numerical examples.

4.2 Application: Relationship Lending and Local vs National Banks

To understand the effects of adverse selection, we focus on a simple case where there are only two types, θ_L and θ_H. The difference between agents of different types is their cost of exerting effort, i.e., in the utility function of Eq.(35)

$$\mathbb{U}(c,z\,|\,\theta) = \frac{c^{1-\sigma}}{1-\sigma} - \theta z^{\varphi}, \ \ \theta \in \{\theta_L, \theta_H\}$$

where $\theta_H > \theta_L$ and refer θ_H is the 'bad' type (bad from the point of view of the FSP). As in Section 3.3, we use a Hotelling line as our spatial configuration with villages uniformly distributed over it. However, we place only one FSP at $x = 0$ and one at $x = 1$ (instead of two at $x = 1$). The asymmetry now comes from the information that each FSP has. We consider two cases. First, we focus on the case where a local FSP does not have any informational problems, while the other is subject to both MH and AdS. We denote this case as *relationship lending*, as we interpret the FSP with full information as having a relationship with the SME. Second, we consider a case where one FSP has the advantage information (as in the relationship lending case), but SMEs do not have to pay travel costs to visit the national FSP. We interpret this case as a local vs national bank situation: the local FSP can be better informed about SMEs - and thus operate in full information - while the national bank has an app or bank correspondents that facilitate access, which we represent in our model by zero spatial costs.

For both cases, we show how equilibrium utilities change with spatial costs and the logit variance. We use the same parameters as in Section 3.1 to generate the frontier,[21] with the addition now of the high type, $\theta_H = 2 > 1 = \theta_L$. The market structure parameters are as in Table 8. To facilitate interpretation, we standardize utilities such that zero represents the autarky utility and unity is the full information, perfect competition level - both for the bad type, θ_H. The difference from the baseline σ_L in Table 7 and Table 8 comes from the fact that now utility scales are naturally different for different types.

Table 8: Baseline Parameters used for AdS Comparative Statics Exercises

Parameter	Value	Meaning
θ_L	1	Low Type
θ_H	2	High Type
f_L	.5	Share of Low Type in each Village
f_H	.5	Share of High Type in each Village
ψ	1	Spatial Cost
σ_L	.1	Logit Variance
V	50	Number of Villages
b_L	1	Number of FSPs in $x = 0$
b_R	1	Number of FSPs in $x = 1$

Case 1: Relationship Lending. We focus first on the relationship lending case. Figure 16 shows equilibrium utilities for each type and Figure 17 the respective market shares and profits of FSP.

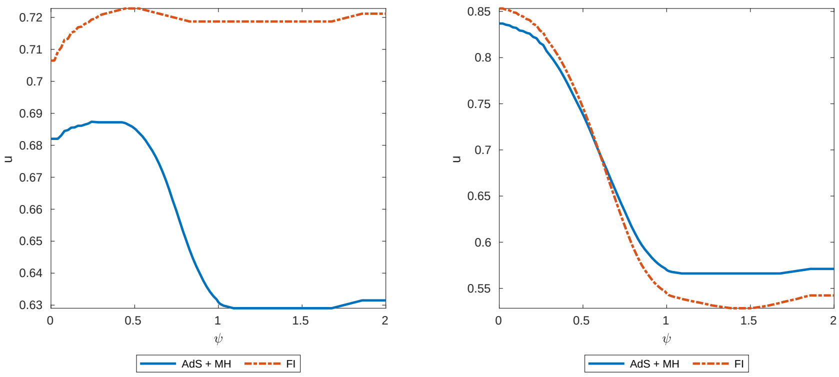
When the spatial cost ψ increases, the utility for the good type increases to partially off set the higher costs, while it decreases for the bad type (through the local monopoly channel). However, when the spatial cost is sufficiently high and the local monopoly channel dominates for both types, we have that the FSP subject to MH + ADS must keep utilities somewhat consistent between the two types, as anticipated, while FSP that contracts under FI does not. This generates an asymmetry in the response of utilities for each type when spatial costs are altered. As we assume types are uniformly distributed across villages, this generates regional inequality: villages closer to the FI FSP are better off on average with rising spatial costs. For these villages, however, inequality across types increases within village.

We repeat the same experiment but for changes in the logit variance σ_L. The results for utilities and market share and profits are, respectively, in Figures 18 and 19. Utilities for the good type, θ_L, are decreasing, while they are hump-shaped for the bad type, θ_H. Note that when SMEs are sensitive to utilities choosing FSPs (σ_L low), the FSP subject to FI has market share advantages in the good type, θ_L, since it can offer a higher utility without having to also increase utilities for the bad type. As the logit variance σ_L increases and both FSPs have

[21] That is, Tables 4-5.

Figure 16: Relationship Lending and Spatial Costs: Equilibrium Utilities

(a) θ_L - Good Type (b) θ_H - Bad Type



Note: Equilibrium utilities played by two FSPs in a Hotelling line. One FSP is located at $x = 0$, while the other is at $x = 1$. The FSP at $x = 0$ contracts under FI, while the one at $x = 1$ contracts under AdS + MH. Parameters for estimation are in Table 8. We solve the equilibrium using the distance to Nash algorithm (Appendix H). The x-axis, ψ, is spatial costs. Utilities are normalzed such that zero is the autarky and one is the FI, perfect competition level for the bad type.

more market power, we observe the opposite.

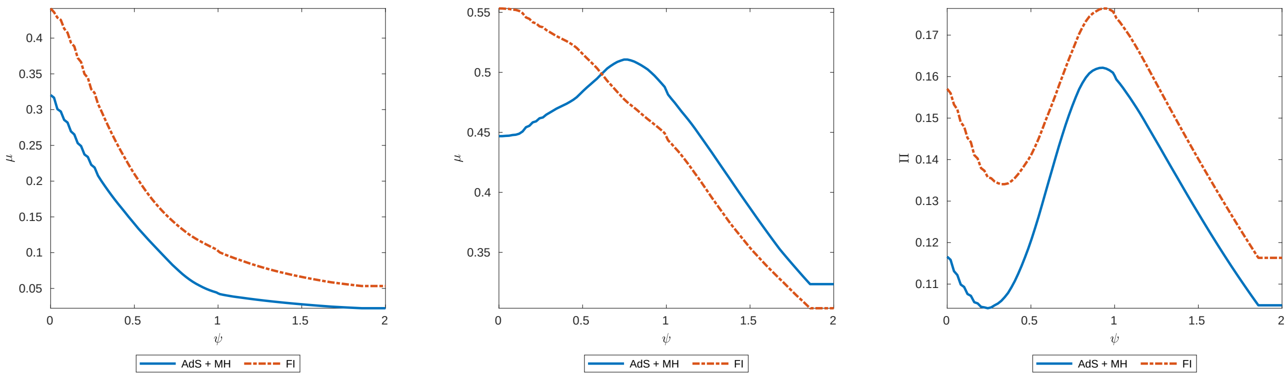
For low values of σ_L, note that we have that the uninformed bank as a very high share of bad clients, which could indicate a worse portfolio (riskier, for instance). In our model, we do not explicitly take this into account, since FSPs are risk neutral and there are no aggregate shocks. The systemic risk this generates (all bad clients with the same FSP) could be relevant to explain macro fluctuations. This is a direction for future research.

Case 2: Local vs National Banks. Our second experiment focuses on local vs national banks. We assume that local banks are contracting with SMEs under FI, but SMEs must pay the cost to visit them. On the other hand, national banks contract under AdS + MH, but provide convenient access to their services.

The results of equilibrium utilities and market shares and profits varying spatial costs are in, respectively, Figures 20-21. Note that local banks always offer higher utilities, but end up with lower market shares and profits if ψ, the spatial costs, are high. The two can co-exist in our model, since each of them will have some advantage (informational vs spatial), as long as spatial costs are not excessively high. When spatial costs ψ increase (and are small to begin with), we observe that the local FSP increases their utility offerings to partially offset this effect. At the same time, national banks can reduce their utility offerings, since SMEs do not pay utility costs to visit the national bank, it is as if competition for the national bank has decreased as a consequence of this increase. If spatial costs ψ are large in the baseline, we have that both banks reduce their utility offerings: the national still due to reduced competition, which allows the local bank to decrease utility offerings as well (although to a lesser degree). In Figure 21 we see the consequences for market shares. The total participation of good types is reduced from 75 % to 37 % (with the rest producing through autarky), while

Figure 17: Relationship Lending and Spatial Costs: Market Shares and Profits

(a) Market Share, θ_L (b) Market Share, θ_H (c) Profits



Note: Market shares and profits implied by the equilibrium utilities of the game between two FSPs in a Hotelling line. One FSP is located at $x = 0$, while the other is at $x = 1$. The FSP at $x = 0$ contracts under FI, while the one at $x = 1$ contracts under AdS + MH. Parameters for estimation are in Table 8. We solve the equilibrium using the distance to Nash algorithm (Appendix H). The x-axis, ψ, is spatial costs.

the total market share of bad types simply transfers from the local to the national bank. This is a non-extreme version of the lemons problem. Competition of an informed (local bank) with an uniformed FSP (national bank) can lead to a reduction in relative participation of good types (θ_L)

5 Taking to the Data: Likelihood and Numerical Method

In Section 3 we developed a theoretical framework to analyze contracting and competition in intermediation. Our framework is at the same time flexible in terms of contracts and easy to solve numerically, due to the Linear Programming formulation of contracts and Lemma 3.1. In this section, we explore how to take the model to the data. For simplicity, we assume here that types are observed (no AdS). [22]. Our ultimate goal is to develop an empirical toolkit for models of competition and contraction that can be used by other researchers.

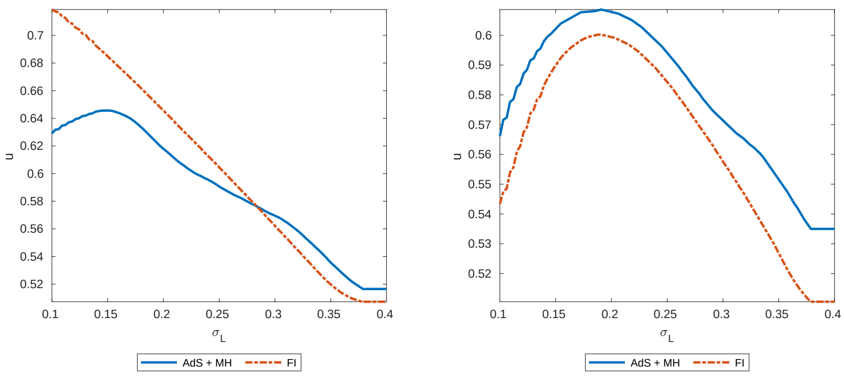
We discuss model implications for the data under two different assumptions for what data is available. First, we discuss how to use the theoretical framework when there is market share data, allowing to run a more structural version of the basic logit regression. We show how to identify the frontier from market share data, which in turn provides a way of computing some counterfactuals *without* a model of contracts.

Second, we discuss how to use data on the number of intermediaries in a given location (as in Bresnahan and Reiss, 1991) and on households to construct a likelihood function that maps model to the distribution of consumption, income and capital of house-

[22] As in Karaivanov and Townsend (2014), it is feasible to extended our methodology for non-observed types.

Figure 18: Relationship Lending and Logit Variance: Equilibrium Utilities

(a) θ_L (b) θ_H



Note: Equilibrium utilities of the game between two FSPs in a Hotelling line. One FSP is located at $x = 0$, while the other is at $x = 1$. The FSP at $x = 0$ contracts under FI, while the one at $x = 1$ contracts under AdS + MH. Parameters for estimation are in Table 8. We solve the equilibrium using the distance to Nash algorithm (Appendix H). The x-axis, σ_L, is the logit variance, which changes market share sensitivity to utilities. Utilities are normalized such that zero are the autarky and one is the FI, perfect competition level for the bad type.

holds, as in Karaivanov and Townsend (2014). Within this framework, we discuss how to deal with unobserved heterogeneity in the data. Even if we do not observe all the lending/borrowing/insurance terms of a household with a bank, what we are interested are the implications for household level outcomes. Therefore, a method that maps model in actual outcomes speaks directly to our main goal in the paper. The structure of likelihood derived allows us to provide a relatively quick numerical method, which we also discuss in this section. Using this numerical method, we provide Monte-Carlo evidence that we can identify the parameters of interest.

5.1 Market Share data

We focus now on a case where we our datasets consists of P provinces (which are our independent markets), indexed by p. [23] In each province, we assume that our datasets contains

1. Locations of villages and banks, and travel time between locations. In our notation, $\{x_v^p\}_v$ for villages, $\{x_b^p\}$ for banks and $t(x_v, x_b)$ for travel time.

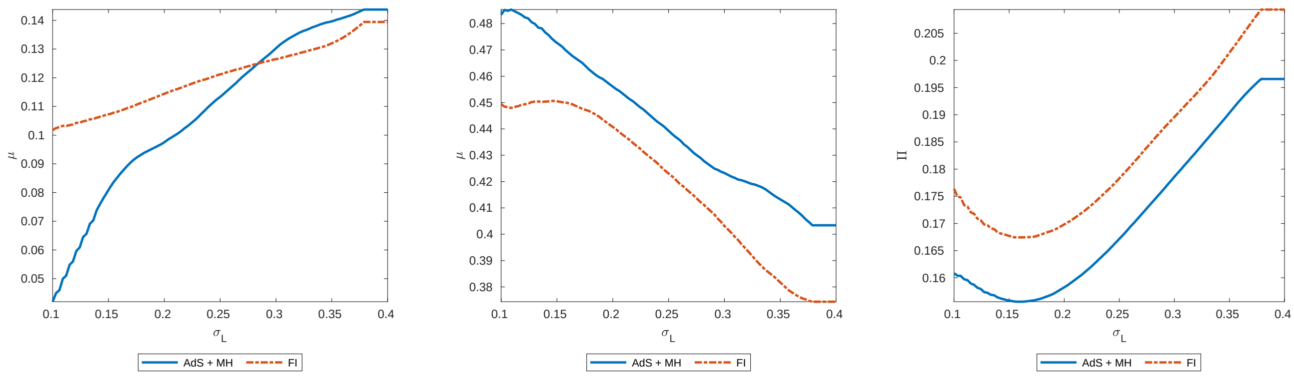
2. Market shares of each bank b in each village v, $\hat{\mu}_{v,b}^p$ (and $\hat{\mu}_{v,0}^p$ for the outside option). We use $\hat{\mu}_{v,b}^p$ for the *observed* market share, while $\mu_{v,b}^p$ is the model implied.

One could observe additional village level characteristics and additionally control for this in our estimation method, but for simplicity in the exposition here we assume that villages

[23] In Section 3 we used a simplified notation without p indexing market shares and equilibrium quantities, but know it is necessary to re-include it, since in estimation we use data from several provinces.

Figure 19: Relationship Lending and Logit Variance: Market Shares and Profits

(a) Market Share, θ_L - Good Type (b) Market Share, θ_H - Bad Type (c) Profits



Note: Market shares and profits implied by the equilibrium utilities of the game between two FSPs in a Hotelling line. One FSP is located at $x = 0$, while the other is at $x = 1$. The FSP at $x = 0$ contracts under FI, while the one at $x = 1$ contracts under AdS + MH. Parameters for estimation are in Table 8. We solve the equilibrium using the distance to Nash algorithm (Appendix H). The x-axis, σ_L, is the logit variance, which changes market share sensitivity to utilities. Utilities are normalized such that zero are the autarky and one is the FI, perfect competition level for the bad type.

are homogeneous apart from their market structure in intermediation and distribution of capital. We assume that locations and market share are observed. Define ζ_S as the set of structural parameters that determine the frontier. In our case, these are the parameters in the utility function, the share of capital that can be recovered in a Limited Commitment regime, etc.. Moreover, define ζ_M as the set of structural parameters on the market structure side. These parameters correspond to the spatial cost, denoted by ψ, and the variance of the logit error, σ_L. Finally, define ζ as the set of structural parameters, $\zeta \equiv \{\zeta_S, \zeta_M\}$.

As the model does not fit perfectly the data, we could add an error to bank-village level market shares and write a empirical version of Eq.(23) as Eq.(37) [24]

$$\ln\left(\mu^p_{v,b}\right) - \ln\left(\mu^p_{0,v}\right) = \sigma_L^{-1}\left[u^p_b(\zeta_S) - \psi t(x^p_b, x^p_v) - u_0(\zeta_S)\right] + \vartheta^p_{b,v} \tag{37}$$

where μ_0 is the market share of autarky, $u^p_b(\zeta_S)$ are the utilities played in equilibrium and $\vartheta^p_{b,v}$ is an exogenous error (does not affect $u^p_b(\zeta_S)$). The utilities in equilibrium and outside option are a function of (i) the parameters that change the frontier, denoted by ζ_S, and (ii) the market structure of the model. [25] Given Eq. (37), we could estimate the structural parameters by the using the IO toolbox of models with discrete choice (e.g., Berry (1994)) [26] We choose to focus on a general and deeper question: given the structure of the model,

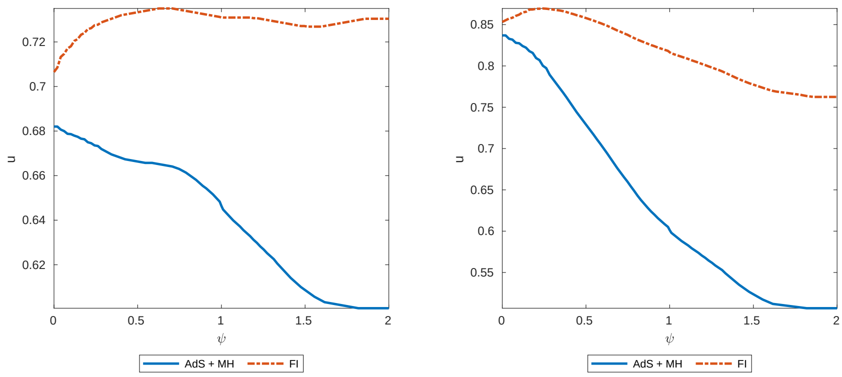
[24]By taking the log at Eq.(23) and the fact that market shares must sum to one (inclusing the outside option).

[25]The outside option is a function of these parameters as they include the utility parameters and we define the outside option as producing under autarky here.

[26]With the difference that instead of assuming a parametric form for the utility (generally linear), we let the model imply what is the equilibrium level of utility given the deep parameters of the economy. Given each

Figure 20: Coexistence of Local vs National Banks and Spatial Costs: Equilibrium Utilities

(a) θ_L - Good Type (b) θ_H - Bad Type



Note: Equilibrium utilities of the game between two FSPs in a Hotelling line. One FSP is located at $x = 0$, while the other is at $x = 1$. The FSP at $x = 0$ contracts under FI, while the one at $x = 1$ contracts under AdS + MH, but SMEs pay not spatial cost to visit $x = 1$(i.e., $t(x_v, 1) = 0$ for any village at x_v). Parameters for estimation are in Table 8. We solve the equilibrium using the distance to Nash algorithm (Appendix H). The x-axis, ψ, denotes spatial costs. Utilities are normalized such that zero are the autarky and one is the FI, perfect competition level for the bad type.

how to use market share data to allow for the identification of the frontier without having to define which contracting frictions is relevant in the data. This is the topic of the next subsection.

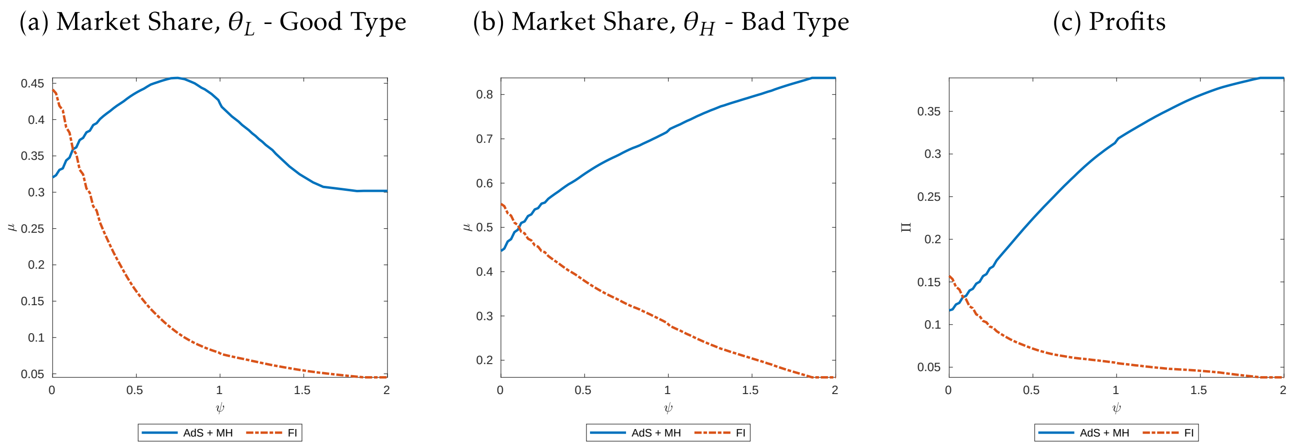
5.1.1 Identification of the Frontier

Instead of assuming that the utility is parametrized by ζ_S, we assume that the frontier is directly depends on it. We let the data pin down the shape of the frontier, i.e., the effects of the contracting friction in our economy without having to define what the contracting friction is ex-ante. This allows for counterfactuals on the market structure side, where the contracting friction is still the same.

Before going through the specifics, we want to discuss the general idea of the method. As in Eq.(27) and Figure 6, we can represent the solution to the problem of the FSP as the intersection of the marginal loss of offering a higher utility with the market shares gains. For FSPs in different provinces, however, the marginal gains in market share, the LHS of Eq.(27), are different. They depend on the spatial configuration of FSPs and the overall competitiveness of the market. On the other hand, the frontier S comes from a fundamental contracting problem and *is not* a function of competition. The separation between these two building blocks is key. For notation purposes, define $\Sigma(u)$ and $\Upsilon_b^p(u)$ as in Eq. (38).

set of parameters, we can solve for the frontier and equilibrium and recover the implied market shares and construct an extreme value estimator based on the observed market shares. We do not explore this idea further in this paper.

Figure 21: Coexistence of Local vs National Banks and Spatial Costs: Market Shares and Profits



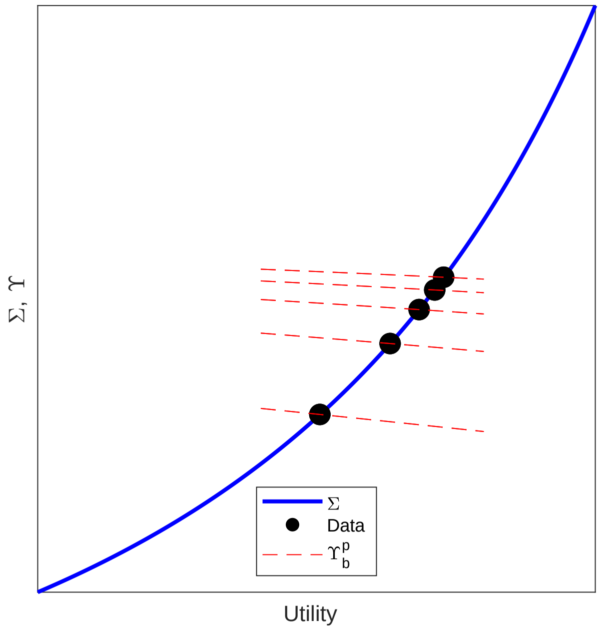
Note: Market shares and profits implied by the equilibrium utilities of the game between two FSPs in a Hotelling line. One FSP is located at $x = 0$, while the other is at $x = 1$. The FSP at $x = 0$ contracts under FI, while the one at $x = 1$ contracts under AdS + MH, but SMEs pay not spatial cost to visit $x = 1$(i.e., $t(x_v, 1) = 0$ for any village at x_v). Parameters for estimation are in Table 8. We solve the equilibrium using the distance to Nash algorithm (Appendix H). The x-axis, ψ, denotes spatial costs.

$$\Sigma(u) \equiv -\frac{\partial_{u_b} S(u_b)}{S(u_b)} \text{ and } \Upsilon_b^p(u) \equiv \frac{\partial_{u_b} \mu\left(u_b, u_{-bp}^{*,p}, .\right)}{\mu\left(u, u_{-bp}^{*,p}, .\right)} \tag{38}$$

With the structure of the model, variation in competition allows us to estimate the curvature of the frontier, as seen in Figure 22. This picture is equivalent to Figure 6 (from our definition of Υ and Σ, but instead of focusing on different utilities for competitors to study equilibrium properties (as in Figure 6), we focus now on different markets with different levels of competition. Both of these represent shifts in the Υ_b^p - but for different reasons. This intuition here is the same as in instrumental variables estimation. We have variation in the market structure that does not directly affect the frontier, only affects the utilities played through the game. With enough variation in competition, we can identify the curvature of the frontier. Without any errors in our model, what we would observe in Figure 22 are the circular dots. They are the intersection of the various curves of $\Upsilon_b^p(u)$ with $\Sigma(u)$, which happens at $u_b^{*,p}$.

Given the idea in Figure 22, we focus first on what we *cannot* identify. We cannot identify the scale of utilities, as it is the case in the usual logit with the methodology of this section. In other words, we cannot identify σ_L and the scale of the frontier jointly. We thus assume that $\sigma_L = 1$ in what follows without loss of generality. We show that even if σ_L (or, more generally, the scale of the utilities) is not identified, there are two key counterfactuals that we can conduct in this case: the introduction of an additional bank and changes in spatial costs (or, more broadly, the spatial development of the banking system.). Although we cannot

Figure 22: First Order Condition of an FSP in Different Provinces



Note: visual Representation of FOC of a FSP b in market p, i.e., Eq. (27) if there are no stochastic terms in the model. For different provinces, p, we observe different points in the curve of the frontier Σ from market shares. With enough variation in competition across provinces, we can identify Σ from the data.

interpret the welfare gains in absolute terms (since we do not have a model of utilities here), we can interpret the welfare gains in terms of the range of utilities *observed* in the data (more details on Section 5.1.2.)

In this subsection, we assume that the model cannot replicate the data perfectly because FSPs do not understand fully the structure of the model (either the frontier or market shares) and/or there is a measurement error in market shares, as in Eq.(37). In particular, we assume that the profit of an FSP is given by Eq. (39) and that the errors FSPs make are province specific.

$$\Pi\left(\varphi_b^p\right) \equiv S(u_b^p)\mu(\varphi_b)\,\chi_b^p(u_b^p), \text{ where } \chi_b^p(u_b^p) \equiv e^{\varsigma_b\left[u_b^p - u_0\right]} \tag{39}$$

Without the error $\psi_{b,p}$, we are back at the profit function defined in Eq.(21). The form of the error in Eq. (39) guarantees that the FOC of a FSP, given what others FSPs are doing, is given by Eq.(40). The difference from Eq.(27) (the version without the error) is now that FSPs do not follow that FOC exactly due to the error

$$-\frac{\partial_{u_b} S(u_b^*)}{S(u_b^*)} + \varsigma_b = \frac{\partial_{u_b}\mu\left(u_b^*, u_{-b}\right)}{\mu\left(u_b^*, u_{-b}\right)} \tag{40}$$

For notation purposes, we define the log difference in market shares as $\omega^p_{v,b-0}$, that is

$$\omega^p_{v,b-0} \equiv \ln\left(\mu^p_{v,b}\right) - \ln\left(\mu^p_{0,v}\right) \tag{41}$$

Moreover, let $\bar{x}^p_b$ represent the mean of a given variable $x^p_{v,b}$ over villages. This will be useful to apply the insight of a fixed effects panel model, where we subtract the within FSP-province mean of a variable and use the variation to estimate the spatial cost.

First, we focus on the identification and estimation of ψ. We can re-write the difference of variables with respect to their means across villages in Eq.(37) as

$$\omega^p_{v,b-0} - \bar{\omega}^p_{v,b-0} = \psi\left[t(x^p_b, x^p_v) - \bar{t}(x^p_b, x^p_v)\right] + \vartheta^p_{b,v} - \bar{\vartheta}^p_{b,v} \tag{42}$$

Thus, intra-province variation in market shares allows us to estimate the spatial cost, which we can do in Eq.(42) through OLS. This is not surprising. Within a market, we expect villages to have different market shares in each FSP due to the travel time between them. Let $\hat{\psi}$ the OLS estimator of ψ in Eq.(42).

Second, we focus on the identification of $\Sigma(u)$. If we identify $\Sigma(u)$, S is identified up to a constant.[27] Given a value of ψ, we can define an estimator for the observed utilities, $\hat{u}^p_b$ as a function of $\hat{\psi}$, as in Eq.(44).

$$\hat{u}^p_b \equiv \bar{\omega}^p_{v,b-0} + \hat{\psi} t(x_{v^p}, x_b) \tag{44}$$

From the FOC of the FSPs in Eq.(40), we can estimate a value for $\hat{\Sigma}_{b,p}$, the value of Σ *in equilibrium* for FSP b in province p given by [28]

$$\hat{\Sigma}^p_b = \hat{\Upsilon}^p_b = 1 - \frac{\sum_{v^p} N^p_v \left(\hat{\mu}^p_{v,b}\right)^2}{\sum_{v^p} N^p_v \hat{\mu}^p_{v,b}} \tag{45}$$

where the second equality comes simply from taking the derivative of market share in Eq. (38).

Although there may be more efficient possibilities to estimate $\hat{\Sigma}$ and $\hat{u}_{b,p}$, we focus on the simple approach of Eqs.(44)-(45) of simply computing sample averages. Once we obtain the vectors $\{\hat{u}^p_b, \hat{\Sigma}^p_b\}$, we can estimate a non-parametric function for $\Sigma(u)$. We explore the exact way to do this in Section 5.1.2.

[27] We can integrate $\Sigma(u)$ to obtain S as in Eq.(43). The constant does not change bank choices at the margin, so can it can be ignored here.

$$S(u) = \text{ct} \cdot \exp\left[\int_{u_{min}}^{u} \Sigma(\nu) d\nu\right] \tag{43}$$

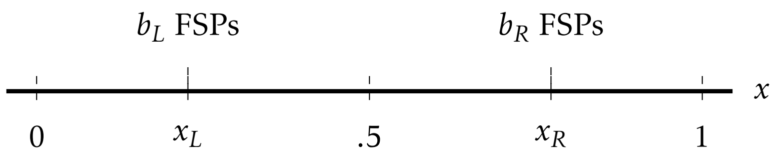
[28] To derive Eq.(45), simply compute the derivative of the logit market shares over itself. The square term comes from the fact that the derivative of the market share at μ is given by $\sigma_L^{-1}\mu(1-\mu)$.

5.1.2 Numerical Example

To showcase the power of the methodology, We present a numerical example. We conduct two counterfactuals: the introduction of an additional bank and changes in the spatial cost, and compare the true and estimated welfare effects.

We simulate P markets with a different spatial configuration and number of FSPs. As in Section 3.3, each market has a Hotelling as in Figure 23. We assume that V villages are uniformly distributed in the $(0, 1)$ and that FSPs are in two different locations, one to the left of the middle point $x = .5$, denoted by x_L, and one to the right, denoted by X_R For simplicity, we assume that $x_R = 1 - x_L$, i.e., the positions of FSPs are always symmetric. We randomly select a position for FSPs and the number of FSPs in x_L and x_R - between 1 and 5 for each location. Each village has a continuum of SMEs.

Figure 23: A Province p in the simulation



Note: visual representation of a sample province in our simulation exercise. We assume that there are V villages uniformly distributed between 0 and 1. We choose x_L random between $[0, .5]$ (and use $x_R = 1 - x_L$) and $b_L, b_R \in \{1, ..., 5\}$ for each province p.

For this numerical exercise, we compute the equilibrium in utilities assuming that there is only one type of entrepreneur and the frontier is given by Eq.(46). This particular functional form is one of many we concave forms we could choose from. We opt to define the frontier in terms of parameters instead of microfounding it (as in Section 3.1) to highlight that it does not matter where the frontier comes from.

$$S(u \mid \zeta_S) \equiv 1 - e^{\zeta_S (u-1)} \tag{46}$$

As we cannot identify the scale of utilities, we assume each FSP can choose a utility $u \in [0, 1]$. We use a iterative best response to find the equilibrium in utilities (See Lemma 3.1).

To be consistent with the previous section, we assume that FSPs do not fully understand the model and each FSP has the error $\chi_b^p(u_b^p)$ that distorts the FOC and that market shares are observed with an error $\vartheta_{b,v}^p$. The baseline parameters we use for the estimation are in Table 9. We compute the standard errors by a non-parametric bootstrap by re-sampling provinces.

We start by using Eq. (42) to estimate the spatial cost by OLS. The true spatial cost is $\psi = 1.5$, while our estimate is $\hat{\psi} = 1.49$ (with standard error .006). The heterogeneity in market shares by the same FSP in different villages is what identifies this parameter. We use Eqs. (44)-(45) to recover a dataset of $\{\hat{u}_b^p, \hat{\Sigma}_u^p\}$. Although we could non-parametrically estimate

Table 9: Parameters used for Numerical Simulation

Parameter	Value	Meaning
ζ_S	2	Concavity of Frontier
ψ	1.5	Spatial Cost
V	50	Number of Villages by Province
P	250	Number of Provinces
u	[0,1]	Possible utility values
σ_ϑ	.25	Std. Dev. in $\vartheta^p_{b,v}$, measurement errors in $\mu^p_{v,b}$
σ_ς	.05	Std. Dev. in ς^p_b, error of FOC of FSPs
b_L, b_R	{1,...,5}	Possible Number of FSPs in Each Location
x_L	[0,.5]	Possible Position of Left Location of FSPs

Note: parameters used to estimate $\Sigma(u)$, the curvature of the frontier, and ψ, the spatial costs.

the curvature of the frontier, we opt for the simplicity of fitting a polynomial regression as in Eq. (47) [29]

$$\hat{u}^p_b = \beta_0 + \beta_1 \hat{\Sigma}^p_u + \beta_2 \left(\hat{\Sigma}^p_u\right)^2 + \beta_3 \left(\hat{\Sigma}^p_u\right)^{1/2} + \nu_{b,p} \tag{47}$$

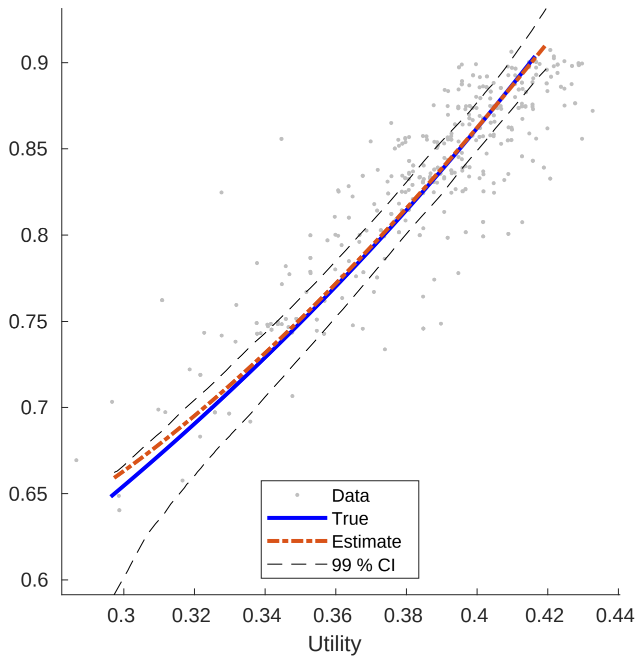
The results are in Figure 24. We can approximate the frontier well overall, and the confidence interval is only large at low levels of utility. This is a consequence of errors $+\varsigma_b$ in the FOC of FSPs impacting more the utility choices in monopolies. In competitive markets, the effect of the errors on utilities are reduced through competition.

We focus now on the counterfactuals. First, we consider the effects of adding FSPs at three different markets. In all markets, FSPs are located at $x_L = 0$ and $x_R = 1$, that is, in the extremes of the Hotelling line. The markets differ however in their baseline number of FSPs, which can be 2, 3 or 4 (in each location). The results are in Table 10. As utility is a cardinal concept and here we are assuming that there is no model as to translate utility gains to consumption gains, we showcase changes in utilities from the policy over the range observed in the data. This means that we can interpret changes in welfare as a percentage of the variation we observe between the minimum and maximum utilities we recover from market shares. This is informative because it can tell how much a policy can add in terms of making some villages without competition in intermediation closer to the villages with competition in intermediation in the sample. As can be seen in Table 10, there are gains from competition form introducing FSPs in markets, but this gain is decreasing with the baseline number of FSPs, as expected. Note that our method does a good job at estimating the effect and providing reasonable confidence intervals to it.

Table 11 has the equivalent results for changes in spatial costs. Not only our model works well, but we can recover the insight of Figure 13 on the Comparative Statics Exercises (Sec-

[29] Is is worth noting that given the structure of the model, we opt for running Eq. (47) with $\{\hat{u}^p_b\}$ as the dependent variable to avoid having the error the estimation of $\{\hat{u}^p_b\}$ to be correlated with $\nu_{b,p}$, the true error in the model.

Figure 24: Identification of Frontier Through Market Share Data



Note: Estimation of $\hat{\Sigma}$ using market share data. Blue solid line represents the true $\Sigma(u) \equiv \frac{\partial S(u)/\partial u}{S}$. Red dashed-dotted line is the estimated. The dashed black lines represent lower and upper bound of the confidence intervals at 1 %, computed with 1000 bootstrap repetitions re-sampling provinces. Grey dots are observed $\{\hat{u}_b^p, \hat{\Upsilon}_b^p\}$.

tion 3.3) where we discussed that reductions in spatial costs can increase welfare by significantly more in more competitive intermediation environments.

5.2 Location of Banks and Household Level data

We develop in this section a likelihood estimator based on household level data (consumption, income and capital) and the location of banks. We develop this estimator for two reasons. First, this is what we do observed in the Townsend thai data and use in our application in Section 6. Second, we want to show how to take contracting models in general to the data. With a model of contracting in utilities and the market structure, the utilities in equilibrium imply contracts, which themselves imply how the joint distribution of consumption, income and capital should be in the data for households.

There are three steps in constructing the likelihood we use here. First, we generalize the problem of banks to include an additive random term, uncorrelated across banks and locations. As in Bresnahan and Reiss (1991), our model will imply the number of banks that *should* be present in each location and, thus, a likelihood function. Second, we extend the analysis in Karaivanov and Townsend (2014) to map contracts played in equilibrium to household income, consumption and capital data. As seen in Section 3, contracts played

Table 10: The Welfare Effect of Additional Banks

Change in FSPs in $\{x_L, x_R\}$	True	Estimated
2 to 3	.1781	.1638
		[.1405, .1868]
3 to 4	.0911	.0935
		[.0870, .1022]
4 to 5	.0434	.0409
		[.0373, .0468]

Note: Welfare effects of including additional banks in each location for three different levels of baseline competition. Each province we analyze has either 2,3 or 4 FSPs in both $x_L = 0$ and $x_R = 1$ in the baseline and we add one bank in both locations. Parameters used to estimate $\Sigma(u)$, the curvature of the frontier, and ψ, the spatial costs, are given by Table 9. Confidence intervals computed with 1000 bootstrap repetitions re-sampling provinces.

Table 11: The Welfare Effect of Reducing Spatial Costs: from $\psi = 1.5$ to $\psi = .75$

Baseline FSPs in $\{x_L, x_R\}$	True	Estimated
2	.6166	.6168
		[.6158, .6183]
3	.6724	.6728
		[.6719, .6737]
4	.7757	.7746
		[.7728, .7768]

Note: Welfare effects of reducing the spatial costs for three different levels of baseline competition. Each province we analyze has either 2,3 or 4 FSPs in both $x_L = 0$ and $x_R = 1$ in the baseline. Parameters used to estimate $\Sigma(u)$, the curvature of the frontier, and ψ, the spatial costs, are given by Table 9. Confidence intervals computed with 1000 bootstrap repetitions re-sampling provinces.

in equilibrium are lotteries. We add a measurement error to household level data (either due to data collection or to the finite grids we define contracting over) and combine the measurement error with the lotteries probabilities to recover what should be observed in household level data in terms of the joint distribution of consumption, income and capital if the model was true (i.e., a likelihood). Third, we must combine the first two steps in one unique likelihood to be optimized. We show that we can maximize the log-likelihood of household level data *given* the number of banks we actually observe in the data plus the log-likelihood of observing that number of banks in a given location.

After constructing the likelihood function, we discuss our numerical method. We show that some parameters can be computed in indirect ways, which speeds up computation significantly. We show that parameters on the market structure (σ_L, ψ) are identified through Monte Carlo experiments. We explore identification of other parameters and provide pseudocodes in Appendix J.

5.2.1 Likelihood of FSP Location Data

We develop the likelihood of FSP location data as in a model of entry (as Bresnahan and Reiss, 1991). The information in the data that is informative about the parameters of the model is the number of banks in each potential entry location in each province. In particular, the next bank to entry in any potential location in province p would have negative profits - and that is why it does not enter. As the model does not perfectly predict the number of banks in each location, we add a random term in profits as Bresnahan and Reiss (1991) and try to maximize the likelihood of the number of banks we observe in each location given the equilibrium of the model.

Before diving into the likelihood, we need to introduce additional notation. Let ζ be the set of structural parameters in out model. Let m^p be a potential location for a new entrant FSP in province p. [30]. As in our model profits are symmetric within location, all banks in a given location would have negative profits, which means we could not have an equilibrium in the first place. Define $\Pi^E\left(B_{m^p}|m^p,\{x_b^p\}_{b\notin m^p}\right)$ as the profits in **equilibrium**, in location m^p of market p given the position of all other banks in other locations $\hat{m}^p \neq m^p$ and that there are B_{m^p} intermediaries at m^p, as in Eq. (48).

$$\Pi^E\left(B_{m^p}|m^p,\{x_b^p\}_{b\notin m^p}\right) \equiv S(u_b^*)\mu\left(u_b^*,u_{-b}^*,x_b^p,x_{-b}^p,u_0\right) \tag{48}$$

If the model perfectly replicates reality, inequalities (49) should hold for all locations m^p in all provinces p. The equilibrium number of banks, B_{m^p}, is such that banks make a positive profit and the marginal bank, that would imply $B_{m^p}+1$ banks, should imply negative profits for all banks in a given location (otherwise it would have entered). [31]

$$\Pi^E\left(B_{m^p}|m^p,\{x_b^p\}_{b\notin m^p}\right) \geq 0 \;\; \cap \;\; \Pi^E\left(B_{m^p}+1|m^p,\{x_b^p\}_{b\notin m^p}\right) < 0, \forall \;\; m^p,p \tag{49}$$

For simplicity, we define the indicator variable $\mathcal{E}(m^p) = 1$ if Eq. (49) is true for location m^p in province p and $\mathcal{E}(m^p) = 0$ otherwise.

As in Bresnahan and Reiss (1991), we add an idiosyncratic location shock to profits given that the model is not flexible enough to match the number of banks in each location. We define the final profits, Π^F, as the profit that includes this idiosyncratic term.

$$\Pi^F\left(B_{m^p}|m^p,\{x_b^p\}_{b\notin m^p}\right) \equiv \Pi^E\left(B_{m^p}|m^p,\{x_b^p\}_{b\notin m^p}\right) + \iota_{m^p}, \;\; \iota_{m^p} \sim \mathcal{N}(c_E,s) \tag{50}$$

where ι_{m^p} is normally distributed with a mean c_E (cost of entry) and variance s, i.i.d. across locations m^p and provinces p. We define the number of banks in each location and the set of

[30] As will be clear later when we discuss the Townsend Thai Data, we use 1997 data to estimate our parameters. The potential locations for FSPs is any location that has a FSP between 1997-2011

[31] Note here that we are assuming that potential entrants can only enter a given location at each period, i.e., there is no joint entry or coordination in the entry game. If we allow for coordination in entry game, the number of deviations - and of equilibrium calculations to compute the likelihood - grows exponentially.

potential locations as the supply side data, denoted by $\mathfrak{S}$ (not to be confused with S for the frontier). As in the previous section, we denote the set of structural parameters as ζ (which now also includes c_E, the cost of entry, and s).

With the stochastic terms in the profit as in Eq. (50), we can assign a probability for the number of FSPs in the data given the structure of model. This will be the likelihood of the supply side. Note, however, that for each $p \in P$, the above system of inequalities is not independent: for a given number of FSPs in a given potential location, the utilities in equilibrium are different in all locations across that province [32]. We have to compute a new equilibrium in the *whole* market for a deviation at *each* location. What we can compute for each p given data on the position of banks and assuming that the marginal entrant is such that they will compete afterwards is given in Eq. (51). The reason we can write the intersection as a multiplication is our assumption that the errors are independent across locations.

$$\mathbb{P}\left(\mathfrak{S}|\zeta\right) = \mathbb{P}\left\{\bigcap_{m^p} \mathcal{E}(m^p) = 1 \middle| \zeta\right\} = \prod_{m^p} \mathbb{P}\left\{\mathcal{E}(m^p) = 1 \middle| \zeta\right\} \tag{51}$$

Finally, using the normality assumption for the ι_{m^p}'s in Eq. (50), we can write Eq. (51) as Eq. (52) [33]

$$\ln\{\mathbb{P}\left(\mathfrak{S}|\zeta\right)\} = \sum_{m^p} \ln\left(\Phi\left[\frac{\Pi^E\left(B_{m^p}|.\right)}{s}\right] - \Phi\left[\frac{\Pi^E\left(B_{m^p}+1|.\right)}{s}\right]\right) \tag{52}$$

where $\Phi(.)$ is the standard normal cdf. Note here that the model is not scale invariant and, thus, potentially informative about the variance s of locations' profit shocks $\iota_{m^p} \sim_{idd} \mathcal{N}(c_E, s^2)$.

5.2.2 Likelihood of Household Level Data

We now explore the model implications for the household level data, which we will denote by the demand side. In this subsection we extend the methodology in Karaivanov and Townsend (2014), with the key difference that we use the model-implied market shares (in equilibrium) to derive weights for each of the contracts in the likelihood. Intuitively, our model of contracting and the Nash equilibrium in utilities implies a level of utility that each intermediary offers. From this level of utility, we can use the frontier to recover what is the optimal contract. The optimal contract then has implications for the joint distribution of consumption, output and capital for each household.

Before constructing the likelihood, we introduce new notation. Let the results of the model in terms of contracts be given by (53), which are specific to each location (m). As we assume that provinces are independent, we simplify the notation and do not include p as a superscript.

$$\left\{\pi_m(c, q, z|k, u^*_{b\in m})\right\}_m \tag{53}$$

[32] Provinces are always assumed to be independent from each other

[33] For notation purposes, we define $\Pi^E(0|.) = \infty$, since we want $\Phi\left[\frac{\Pi^E(0|.)}{s}\right] = 1$.

Let the cross sectional household level data be given by $\{\hat{y}_j\}_{j=1}^{\mathcal{H}}$, where $j = 1,...,\mathcal{H}$ denotes households. Here, we use capital, income and capital, respectively denoted by $y_j = (c_j, q_j, k_j)$. In other settings, however, one can apply the same estimation method based on a different $\hat{y}_j$ that is the outcome of contracting. To deal with actual measurement error in the data and fitting the data into the discrete grids used in contracting, we assume that the data has a measurement error of the form:

$$\mathcal{N}(0, \gamma_{ME} \cdot \chi^2(X)) \tag{54}$$

where $\chi^2(X)$ denotes the range of the grid $X = C, K, Q$. Given the structural parameters ζ, we can write the density for (c,q) conditional on capital as in Eq. (55).

$$g_v(c,q|k,\zeta) = \sum_u m_v^u(k) \sum_z \pi(c,q,z|k,u) + \left[1 - \sum_u m_v^u(k)\right] \pi^{aut}(c,q,z|k) \tag{55}$$

where $m_v^u(k)$ is the share of agents in village v, capital k that are offered utility u by a FSP - i.e., we must sum the market shares across villages and across intermediaries $b \in B$ to recover the market share of a given level of utility, as in Eq. (56).

$$m_v^u(k) \equiv \sum_{b \in B} \sum_u \mathbb{1}_{u=u^*(b)} \mu_v^b(u,k) \tag{56}$$

The distribution of (c,q,k) in a village is then given by Eq. (57), where we multiply by the distribution of capital in the village, $h_v^k(k)$.

$$f_v(c,q,k|\zeta) = g_v(c,q|k,\zeta) h_v^k(k) \tag{57}$$

Here f_v captures the probability of observing a given tripe (c,q,k) in the data if the model (including the grids) was a perfect representation of the world.

Due to actual measurement error or actual contracting beyond our grids, (c,q,k) data is not limited to be in the small finite grids and, therefore, we associate the probability of observing a triple (c,q,k) given the measurement error as in Eq. (54). Define $\#Y \equiv C \times Q \times K$, i.e., the Cartesian product of the grids. Therefore, for any y implied in the model (pre-measurement error), we have that $y \in \#Y$. Let $l = 1,...,L$ represent the different elements of $\hat{y}_j$ - (c_j, q_j, k_j) here. Then, for a given household j, the likelihood of observing a given $\hat{y}_j$ is given by Eq. (58). In others, for each possible point in the grid, we compute the function $f_v(.)$, the probability that a point is the outcome of contracting. Given a contracting outcome, we sum all probabilities of the actual observed $\hat{y}_j$ given the measurement error (for each element of $\hat{y}_j$), if y was the actual contracting outcome.

$$F_v\left(\hat{y}_j; \zeta, \gamma_{ME}\right) \equiv \sum_{r=1}^{\#Y} f_v(c,q,k|\zeta) \prod_{l=1}^{L} \Phi\left(\hat{y}_j^l \mid y_r^l, \varsigma^l(\gamma_{ME})\right) \tag{58}$$

where $\Phi(.|\mu, \varsigma)$ stands for the CDF of a normal distribution with mean μ and variance ς^2.

We need now to sum F_v over all households in all villages to get the full sample likelihood. Let the village v in province p of household j^p be given by $\hat{v}_j^p$. Denote the demand side data as $\mathfrak{D} = \left\{(\hat{y}_j^p, \hat{v}_j^p)\right\}_{j,p}$, that is, consumption, income, capital and village for each household j, in each province

p. The likelihood of demand $\mathfrak{D}$ *given* the position of banks observed in the data, is given by Eq.(59) (re-introducing province superscripts). To obtain the log-likelihood for the overall sample, we sum the log of F_v for all households in all villages and provinces.

$$\ln\{\mathbb{P}(\mathfrak{D}|\mathfrak{S},\zeta)\} = \sum_{p}\sum_{v^p}\sum_{j^p} \mathbb{1}_{\hat{v}_j^p = v^p} \, \ln F_v^p\left(\hat{y}_j^p, \zeta\right) \tag{59}$$

Eq. (59) is the likelihood of demand $\mathfrak{D}$ *given* supply because we use the actual number of observed banks in each potential location m^p to compute the utilities in equilibrium in Eq.(56). In the next subsection, we show that this is sufficient to combine the likelihoods.

Unobserved Heterogeneity and A reduced Form of Competition. Our structural method is also flexible to deal with unobserved heterogeneity. This can be a relevant state variable that is not observed, such as types θ or initial asset positions, or the equilibrium utilities themselves. For that, assume that we have a set s_0 of unobserved states that are relevant to determine the distribution of s_0.

Let $h_v^{k,s_0}(k, s_0)$ is the joint distribution of capital k and the unobserved state, s_0, in village v. We can write the joint distribution as a function of the marginal distribution of k, observed, and the conditional distribution of s_0 conditional on k, given by $h_v^{s_0|k}$, not observed as in Eq.(60). We assume that this distribution is parametrized by parameters ζ^u, which we can estimate in the likelihood.

$$h_v^{k,s_0}(k, s_0) = h_v^k(k) h_v^{s_0|k}(s_0|k, \zeta^u) \tag{60}$$

We can re-write Eq.(57) as Eq.(61)

$$f_v(c, q, k|\zeta, \zeta^u) = \sum_{s_0} g_v(c, q|k, \zeta) h_v^k(k) h_v^{s_0|k}(s_0|k, \zeta^u) \tag{61}$$

From Eq.(61), we can simply modify the likelihood computation to also include the parameters ζ^u. This methodology is used in Karaivanov and Townsend (2014) to estimate initial asset holdings in simply borrowing/savings contracts.

This likelihood formulation in Eq.(61) also provides a useful result to understand the effects of competition in utilities. Assume that we do not know or don't want to assume the market structure of the model. We can define s_0 to be utilities in equilibrium for level of capital k in a given province. In particular, we can parametrize it as a normal distribution with different means/standard deviation for each level of capital as $s_0 \sim \mathcal{N}(\kappa_k, \sigma_u^k)$. This would allow a researcher to understand the effects of competition *without* a model of how the equilibrium is determined. The downside is that is not possible to conduct counterfactuals in this way, since we do not take a stand on how the equilibrium utilities are determined.

5.2.3 Combining The Two Datasets

We explore now how to combine the likelihood of Section 5.2.1 and Section 5.2.2. We show that we can compute the demand log-likelihood (based on household level data) *given* the observed location of banks and sum with the log-likelihood of the supply (the number of banks and potential locations). We provide also an additional result, [34] which is how to combine the number of banks for all provinces and household level data for the subset of provinces and villages that are in fact observed. This is useful for a researcher that observes the numbers of banks across several provinces, but only has detailed data on households for a few specific provinces.

These two results are mathematicall stated in Lemma 5.1. Intuitively, Lemma 5.1 states that we do not have to re-compute the likelihood of potential deviations from the number of banks that is not observed, which speeds up computation significantly. In other words, we can decompose the competition structure and its results from the demand side data.

Lemma 5.1. *Let ζ be the set of structural parameters and $\mathcal{L}$ to be the log-likelihood of supply and demand data for province p, denoted by, respectively, $\mathfrak{S}_p$, and $\mathfrak{D}_p$. Let P be the provinces we observe both, and $\overline{P}$ the set we observe only $\mathfrak{S}_P$. Then*

$$\mathcal{L}\left(\zeta|\mathfrak{S}_{p\in\overline{P}},\{\mathfrak{D}_p\}_{p\in P}\right)=\sum_{p\in\overline{P}}\ln\left[\mathbb{P}\left(\mathfrak{S}_p|\zeta\right)\right]+\sum_{p\in P}\ln\left[\mathbb{P}\left(\mathfrak{D}_p|\zeta,\mathfrak{S}_p\right)\right] \tag{62}$$

Proof. See Appendix I.1 ■

Moreover, Lemma 5.1 is useful in our numerical method, which we explore in Section 5.2.4. The idea is that there is a subset of structural parameters ζ that is relevant for supply and one that is relative for demand. Some parameters, like spatial costs ψ, are important for both. The measurement error variance γ_{ME}, on the other hand, determines only demand side likelihood. This will prove to be valuable when we discuss the empirical method in Section 5.2.4.

5.2.4 Numerical Method

We explore now the numerical method given the result of Lemma 5.1 and Eqs.(52) and (59). The parameters in our model are

$$\zeta\equiv\{\zeta_S,\psi,\sigma_L,\gamma_{ME},c_E,s\} \tag{63}$$

where ζ_S are the parameters that change the frontier. Our objective is to solve the optimization problem in Eq.(64)

$$\max_{\zeta}\sum_{p\in P}\ln\left[\mathbb{P}\left(\mathfrak{S}_p|\zeta\right)\right]+\ln\left[\mathbb{P}\left(\mathfrak{D}_p|\zeta,\mathfrak{S}_p\right)\right] \tag{64}$$

where $\ln\left[\mathbb{P}\left(\mathfrak{S}_p|\zeta\right)\right]$ is defined in Eq.(52) $\ln\left[\mathbb{P}\left(\mathfrak{D}_p|\zeta,\mathfrak{S}_p\right)\right]$ from Eq.(59). We could simply numerically solve the problem above in Eq.(64). However, Eq.(64) has several characteristics that allow for a more efficient solution. First, not all parameters enter in both terms. Moreover, we can separate the optimization problem in two parts - for any values of ζ_S,ψ,σ_L, we solve for the optimal values of the optimal parameters and then optimize over ζ_S,ψ,σ_L. In particular, we have that the problem in Eq.(64) is equivalent to Eq.(65).

$$\max_{\zeta_S,\psi,\sigma_L}\left(\max_{\gamma_{ME}}\sum_{p\in P}\ln\left[\mathbb{P}\left(\mathfrak{S}_p|\{\zeta_S,\psi,\sigma_L,\gamma_{ME}\}\right)\right]+\max_{c_E,s}\sum_{p\in P}\ln\left[\mathbb{P}\left(\mathfrak{D}_p|\{\zeta_S,\psi,\sigma_L,c_E,s\},\mathfrak{S}_p\right)\right]\right) \tag{65}$$

[34] In this version of this paper we do not use this result, but we do intend to use it in future versions.

First, we explore how to estimate γ_{ME}, c_E and s given ζ_S, ψ and σ_L, that is, the inner maximization problems in Eq.(65). Second, we discuss how to estimate ψ and σ_L and show that it is numerically identified from Monte-Carlo experiments. We do not focus on the estimation of ζ_S on this paper, given that our innovation is on the market structure side given contracting. We do discuss estimation and identification of ζ_S in Appendix J. So we simplify the notation and exclude the dependence from it.

Estimating $\{c_E, s, \gamma_{ME}\}$**.** For now, assume that ψ and σ_L are as if known. We show how to solve inner maximization problems in Eq.(65) given these values.

Note that the variance of the measurement error, denoted by γ_{ME}, does not affect (i) the frontier and (ii) Equilibrium utilities, and (iii) the likelihood of the supply side in Eq.(52). It only changes the likelihood at the computation of F_v in Eq. (58). This does not mean that in maximizing the likelihood there is no interaction between the parameters, but it means that *given* ψ, σ_L and ζ_S we can easily compute the estimator $\hat{\gamma}_{ME}$, the argmax of Eq.(66) as a function of ψ and σ_L , without having to recompute the frontier or the equilibrium.

$$\hat{\gamma}_{ME} \equiv \arg\max_{\gamma_{ME}} \sum_{p \in P} \ln\left[\mathbb{P}\left(\mathfrak{D}_p | \{\gamma_{ME}, \psi, \sigma_L\} \mathfrak{S}_p\right)\right] \tag{66}$$

Furthermore, note that the entry cost, c_E and std. deviation of location specific shocks, s, do not affect (i) the frontier and (ii) Equilibrium utilities, and (iii) the likelihood of the demand side in Eq.(52). Thus, we can easily compute the estimators $\hat{c}_E, \hat{s}$, the argmax of Eq.(67) as a function of ψ and σ_L .

$$(\hat{c}_E, \hat{s}) \equiv \arg\max_{c_E, s} \sum_{p \in P} \ln\left[\mathbb{P}\left(\mathfrak{S}_p | \{c_E, s, \psi, \sigma_L, \zeta_S\}\right)\right] \tag{67}$$

As γ_{ME} is the std. dev. of the normal distribution, F_v in Eq. (58) is differentiable in γ_{ME}. As we know what is the analytical derivative of it, it is straightforward to compute an optimal value for γ_{ME} using a grid search method, which is computationally fast. See Appendix I.2 for specific equations and details.

Analogously, we can take the FOC of Eq.(52) with respect to c_E, s and easily solve for it through a grid search (although now is a FOC system with two equations). See Appendix I.3 for specific equations and details.

Estimating ψ **and** σ_L**.** We have to estimate ψ (spatial cost) and σ_L (logit variance) numerically. For each value used of ψ and σ_L, we use the method above to compute $\{c_E, s, \gamma_{ME}\}$ The numerical optimization of ψ and σ_L is done through a mix of a grid search and the Matlab built-in *patternsearch*. See Appendix I.4 for more details and a pseudo-code.

Numerical Identification. Although we have given indications in Section 3.3 that we can identify the parameters from the micro data, we show that it is the case numerically. The intuition is that the overall levels of utility, which imply consumption, capital and income dynamics, identify σ_L, while the variation between these levels across villages identifies ψ, as seen in Section 3.3. To validate this intuition, we conduct a Monte-Carlo experiment. We generate model simulated data and use it to estimate the parameters in question. We use only data on consumption, production and capital in this exercise. [35]. As we are ultimately interested in estimating the spatial costs ψ and logit variance, σ_L, we mainly focus on the maximization of Eq. (59) on these two parameters. The numerical results show that our method in fact identifies $\{\psi, \sigma_L\}$ from the data. The details and results are in Appendix J.

6 Thai Data and Results

In this section we apply our method to real data. We first describe the data used, which is a combination of the Townsend Thai Data for households and other sources for the distances and travel times from villages to bank branches. We then present our parameter estimates and various counterfactual results. Our results suggest that spatial costs are important for individuals, as an individual would reduce its consumption by 20% to eliminate them. In terms of aggregate welfare, reducing spatial costs by 50% is equivalent to increasing consumption by 4.85%, while reducing σ_L by 50% is equivalent to increasing consumption by 15.36%. One additional bank has limited effects in our results, increasing consumption by only 2.2%.

6.1 Data Description

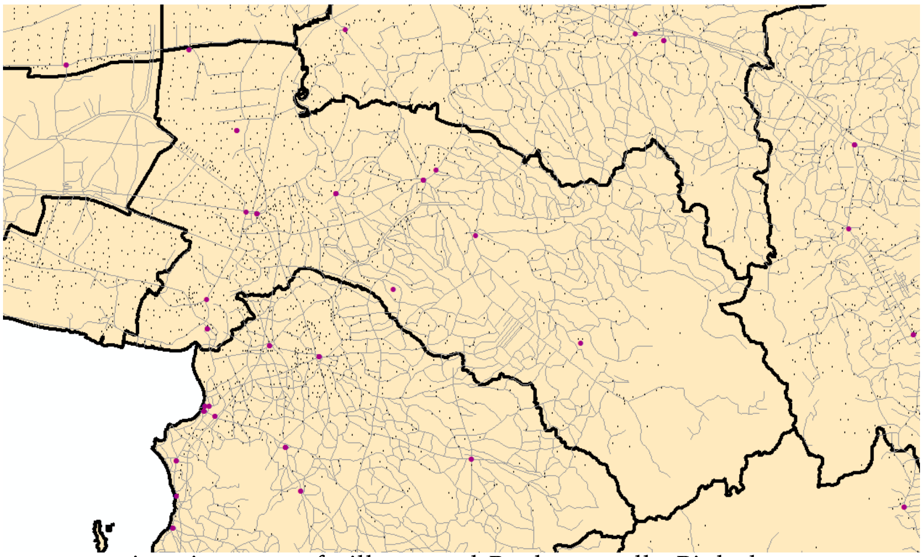
For household level data, we use the Monthly Resurvey of the Townsend Thai Data for the year of 1999, as in Karaivanov and Townsend (2014). For locations of villages in the Townsend Thai Data and banks, we use several data sources. We assume that each bank branch is a different FSP.

Village data is extracted from the Thai Community Development Department (CDD) survey. The information on bank branch location comes the Bank of Thailand, Bank of Agricultural and Agricultural Cooperative, Telephone Authority of Thai land, Community Development Center and several non-traditional financial institutes. We combine these datasets as in Assuncao et al. (2012) to get the open and close date for each bank, as well as bank branch and name. We geo-locate each bank branch and village by the Google Maps API and compute travel time between two points in the map using a GIS platform. We use the road network from the Thailand Environment institute. The data classifies all roads in Thailand among 7 types, with different traveling speeds (e.g., highway vs local road). We use a GIS platform to compute the travel time between any two junctures in the map. As an illustrative example of our spatial data, we plot the position of all villages and FSPs in 1999 for the province of Chacheongsao in Figure 25.

In the Townsend Thai Data, the Monthly Resurvey data consists of data collected for 531 households in 16 villages of 4 provinces. This provinces are: Chacheongsao, Lopburi, Buriram and Sisaket. The provinces of Buriram and Srisaket are located in the North-east region, which is relatively poor and semi-arid. The provinces of Chacheongsao and Lopburi are located near Bangkok and, in part, urban. Consumption expenditures, c, includes expenditures in food, gasoline, education, house and vehicle repairs, clothing, etc. and includes owner-produced consumption. Production, q, is measured on an accrual basis. As we are using annualized data, however, this is close to cash flow. Capital (or

[35]As we assume that the observed data corresponds to an equilibrium in terms of bank entry (as in Bresnahan and Reiss (1991)), we do not know the entry/exit process and what is the dynamics of it, such that it is very challenging to simulate an equilibrium in the position of each bank.

Figure 25: Villages and Banks in Chacheongsao Province



Note: Chacheongsao province in terms of villages and Banks overall. Pink dots represent bank branches, black dots are villages and grey lines are the roads in 1999. Horizontal distance from extremes in the figure corresponds to ≈ 80 miles.

business assets) data, k, includes business and farm equipment and livestock. Financial assets or durable goods are not considered in k. The variables are not converted to per-capita terms, i.e., household size is not brought into consideration. All values are in nominal terms. Table 12 exhibits the summary statistics. As pointed out in Karaivanov and Townsend (2014), an important characteristic of the data is that correlations between income, consumption and capital indicate that there is significant consumption smoothing, but still far away from full insurance. We consider a market as a cluster of bank branches that are at most 30 minutes by car from the nearest village. In our estimation, we assume that banks consider that only the villages in the Monthly Resurvey Sample exist when computing their profits (i.e., the demand is simply given by households on this villages). The error from this assumption comes will enter in our model through the location specific shocks.

Table 12: Summary Statistics

Consumption expenditure, c	
Mean	58,311
Std. Dev.	48,951
Median	43,895
Production, q	
Mean	100,820
Std. Dev.	290,997
Median	42,013
Business Assets, k	
Mean	76,065
Std. Dev.	401,008
Median	10,959

Notes: 1999, Monthly Resurvey of the Townsend Thai Data. Average exchange rate in 1999-2000 was 1 USD = 39 Baht. See text for definitions of consumption, income and business assets.

6.2 Results

Using the method of Section 5.2, we estimate the parameters assuming that FSPs contract with SMEs under MH + LC. We convert all data from Thai currency into 'model units' by dividing all currency values by the 90-th percentile of the assets distribution in the sample (this is approximatelly 180,000 Thai baht). We use the parameters and grids to compute the frontier as in, respectively, Tables 4-5. We do not attempt to estimate the parameters that define the frontier in this paper. We estimate the measurement error parameter γ_{ME}, and the market structure parameters, namely: spatial costs ψ, logit variance σ_L, cost of entry c_E, idiosyncratic location shock variance s

$$\{\psi, \sigma_L, \gamma_{ME}, c_E, s\}$$

We use the functional form for utilities as in Eq.(16). The estimates are Table 13.The estimates for the measurement error γ_{ME} is 21 %. This corresponds to measurement error with standard deviation of 21% of the variables' grid ranges. Moreover, the estimate for s is also low compared to c_E, which indicates the model predicts relatively well the number of FSPs in each potential location.

To understand how relevant spatial costs are, we compute how much a household that pays zero travel costs to every bank would have to receive to be at the median distance. This is the result of Eq.(68). See Appendix K for details. As we use CRRA preferences, this measure depends on the initial level of consumption of this household. For simplicity, we use the average consumption, denoted by $\bar{c}$. Given a $\hat{\psi} = .55$, we have that a household at the average consumption would have to receive a $\psi^u = 19.61\%$ increase to move to the median distance. Note that this is different than the counterfactual exercise on changing spatial costs, as here, we keep the utilities played by FSPs constant, that is, we consider an unilateral move of one household that does not affect the equilibrium.

$$\psi^u(\psi) \equiv \left[\frac{\psi \,\mathrm{med}\left(t(x_v, x_b)\right)}{u(\bar{c})} + 1\right]^{\frac{1}{1-\sigma}} - 1 \tag{68}$$

We move to our counterfactuals. We denote the equilibrium at our parameter estimates as our baseline, and showcase in our results percentage deviations to the equilibrium at the estimated parameters. That is, for any variable X (such as consumption, market shares etc.), we show in the tables the percentage change as in Eq.(69). $X^{'}$ is the value after the change and X_0 the baseline. For welfare, we plot the average utility of household taking into account the spatial costs, that is, we subtract $\psi t(x_v, x_b)$ from equilibrium utilities (weighted by market shares, as in Section 3.3).

$$100\left(\frac{X^{'}}{X_0} - 1\right) \tag{69}$$

To interpret the changes in welfare, we compute how much consumption would have to increase (for certain) to match this change in utility levels. As we use CRRA preferences, this measure depends on the initial level of consumption, which we use the average consumption in our sample. For details, see Appendix K.

Spatial Costs. The results of changing spatial costs ψ are are in Table 14. The averages and standard deviation are computed at the village level (after aggregating for households and different levels of

Table 13: Parameter Estimates

	Estimate	Model
$\hat{\gamma}_{ME}$	.21	Measurement Error
	(0.0139)	
$\hat{\psi}$	.55	Spatial Cost
	(0.0175)	
$\hat{\sigma_L}$	.083	Logit. Var
	(0.005)	
$\hat{c_E}$	1.57	Cost of Entry
	(0.0260)	
$\hat{s}$	0.03	Variance of Location Specific Profit Shock
	(0.0001)	

Parameters estimated by maximizing the likelihood with the 1999 Monthly Resurvey data. γ_{ME}, c_E, s are maximized through the first order conditions. ψ, σ_L are maximized by a grid search followed by the *patternsearch* algorithm in Matlab. Standard error in parenthesis computed using Bootstrap with 200 repetitions. See Section 5.2 and Appedix I for details. All coefficients are significant at 1%.

capital), and averaged for different provinces. [36] The results of the transformation from welfare to consumption are in Table 17, where we repeat the welfare numbers and compute the consumption equivalent change.

By reducing spatial costs by 50%, welfare increases 4.85 % (First Column of Table 17). Note that due to lack of insurance under autarky, average consumption increases when welfare is reduced, just as in our example in Section 2. When we focus on the consumption of intermediated SMEs (i.e., SMEs that used financial services before and after the change), there are no changes with a reduction in spatial costs. The increases in utility come from changes in insurance, effort etc.., without changing consumption, at least for initial reductions. Note also that with lower spatial costs more SMEs use financial services (market shares growing), and, as expected, the standard deviation across villages of the share of SMEs that use financial services is reduced.

Logit Variance. The results of changing the logit variance σ_L are are in Table 15. The aggregation and conversion from welfare and consumption equivalents is made as in the case for spatial costs. By reducing the logit variance by 50%, welfare increases 9.20%, which corresponds to a 15.36% in consumption equivalent terms (last two rows of the first Column of Table 17). Note that in this case, movements in welfare can be understood as changes in the consumption of the intermediated SMEs, although magnitudes are still off. For instance, average effort decreases by more than 20% with the 50% reduction in σ_L.

This is a counterfactual we cannot do with the methodology in Section 5.1. In Section 5.1, we showed how we can use market shares to recover the frontier and thus, conduct welfare of changes of spatial costs and new banks. However, as we cannot identify the scale of utility without a model for utilities, we cannot conduct counterfactuals with respect to σ_L, which is essentially changing the utility scales in our model. To conduct this counterfactual we need to model both building blocks of our model: the frontier and marker structure.

[36]The standard deviation are not those from the parameters, but the standard deviation across the average of different villages.

Table 14: Counterfactual: Percentage changes of outcomes with spatial costs ψ

	$.5\hat{\psi}$	$.75\hat{\psi}$	$1.25\hat{\psi}$	$1.5\hat{\psi}$
Average Welfare	2.7263	0.7919	-0.4456	-1.5471
Std. Dev. Welfare	-6.1825	-1.9299	1.6455	4.8028
Average Consumption	-1.9772	-0.7022	0.7064	2.2048
Std. Dev. Consumption	-8.3253	-2.7837	2.0039	7.4793
Average Market Share	10.3156	3.4218	-2.0866	-8.6046
Std. Dev. Market Share	-1.7316	-0.3950	0.9085	1.1034
Average Consumption of Intermediated	0	0	0.6187	0.7695
Std. Dev. Consumption of Intermediated	0	0	-2.2569	-1.3478

Note: Model outcomes for changes in the spatial cost ψ. Percentage change (Eq. 69) with respect to the baseline of $\hat{\psi}$ and $\hat{\sigma_L}$ in Table 13. Averages and standard deviation computed at the village level (after averaging out households). All results are aggregated across the four provinces we use in our estimation. Contracting is done under MH + FI.

Bank Entry. Our last counterfactual computes changes in model outcomes after bank entry. We compute the average outcome of one bank entry in each of the potential locations in each province. The results are in Table 16. An extra bank increases utilities on average by 2.16%, which translates to a 2.2% equivalent change in consumption. Note that more households do get served (increase in average market share), and those that do go to banks see an increase in their average consumption. Village wide average consumption, however, is still decreasing, since the consumption *level* of intermediated agents is smaller than those in autarky (due to insurance).

Our results indicate that reducing spatial costs, the logit variance and adding extra banks can increase utilities of agents, but in different magnitudes. Although spatial costs are relevant for individual agents, they are less relevant in determining overall welfare than the logit variance (for changes of the same magnitude). Our results suggest that to increase welfare, policy makers should guarantee that markets shares change when utility offerings change. This means that the goal of policy makers should be at financial literacy, platforms where financial products can be compared, bank correspondents, or other policies geared toward making SMEs more likely to choose better financial products, rather than simply increasing the number of FSPs.

Table 15: Counterfactual: Percentage changes of outcomes with logit variance σ_L

	$.5\hat{\sigma}_L$	$.75\hat{\sigma}_L$	$1.25\hat{\sigma}_L$	$1.5\hat{\sigma}_L$
Average Welfare	9.2099	3.6983	-2.4646	-7.7871
Std. Dev. Welfare	13.7923	4.1346	-5.3172	-9.6498
Average Consumption	0.8900	1.0539	0.3388	-1.2744
Std. Dev. Consumption	5.4557	1.4383	-2.6656	-2.8107
Average Market Share	0.7816	0.7479	1.8228	0.6360
Std. Dev. Market Share	23.8761	6.3072	-7.9259	-15.0111
Average Consumption of Intermediated	13.8097	3.9902	-0.5355	-8.3375
Std. Dev. Consumption of Intermediated	26.3510	-2.7190	-1.8731	-11.7399

Note: Model implied outcomes for changes in the logit variance σ_L. Percentage change (Eq. 69) with respect to the baseline of $\hat{\psi}$ and $\hat{\sigma_L}$ in Table 13. Averages and standard deviation computed at the village level (after averaging out households). All results are aggregated across the four provinces we use in our estimation. Contracting is done under MH + FI.

Table 16: Counterfactual: Percentage changes of outcomes with Bank Entry

Average Welfare (Cons. Equivalent)	2.2008
Std. Dev. Welfare	-6.4613
Average Consumption	-2.7193
Std. Dev. Consumption	-12.3310
Average Market Share	15.6661
Std. Dev. Market Share	1.7803
Average Consumption of Intermediated	1.8863
Std. Dev. Consumption of Intermediated	-8.3687

Note: Model implied outcomes for changes in the number of banks. We include an additional bank in each potential location at a time, and compute the averages of all of these counterfactuals to show the results. Results are displayed as percentage change (Eq. 69) with respect to the baseline of $\hat{\psi}$ and $\hat{\sigma_L}$ in Table 13. Averages and standard deviation computed at the village level (after averaging out households). All results are aggregated across the four provinces we use in our estimation. Contracting is done under MH + FI.

Table 17: Counterfactual: From Utilities to Consumption

	$.5\hat{\psi}$	$.75\hat{\psi}$	$1.25\hat{\psi}$	$1.5\hat{\psi}$
Welfare Change (%)	2.7263	0.7919	-0.4456	-1.5471
Consumption Equivalent (%)	4.8523	1.3742	-0.7610	-2.6051
	$.5\hat{\sigma}_L$	$.75\hat{\sigma}_L$	$1.25\hat{\sigma}_L$	$1.5\hat{\sigma}_L$
Welfare Change (%)	9.2099	3.6983	-2.4646	-7.7871
Consumption Equivalent (%)	15.36	5.78	-3.59	- 10.71

Note: Model welfare changes for changes in the spatial cost ψ and logit variance σ_L. We move from welfare to utilities using the equations in Appendix K, Eq. (178). Percentage changes (Eq. 69) with respect to the baseline of $\hat{\psi}$ and $\hat{\sigma_L}$ in Table 13. Averages and standard deviation computed at the village level (after averaging out households). All results are aggregated across the four provinces we use in our estimation. Contracting is done under MH + FI.

7 Conclusion

Given the challenges in interpreting reduced form evidence in settings with contracting and market power in intermediation, we focus on building, solving and estimating a model that allows for frictions (Moral Hazard, Adverse Selection etc.) and different market structures (Monopoly, Oligopoly, Competition). The main insight of our theoretical analysis it to develop a framework in terms of utilities generated by contracts rather than the contracts themselves, and divide the contracting and competition problems in building blocks. This allows us to apply most of the competition toolbox to potentially complex models of competition and contracting.

We focus our analysis in contracting between a entrepreneurs and a set of financial intermediaries for several different financial regimes. Our market structure is on a demand system where entrepreneurs and FSPs are spatially separated and entrepreneurs have idiosyncratic preferences for intermediaries that generate logit market shares.We show that under a few conditions, a unique Nash equilibrium exists and can be computed through iteration of best response functions. Through comparative statics exercises, we show how this method can be applied to understand and quantify the impact of the spatial and technological changes in the banking sector in emerging market countries. For instance, among other results, we show that (i) local competition increases utilities, and it does more under MH + LC than under FI, (ii) reduction in spatial costs can increase or decrease welfare of SMEs, as it can create local monopolies, (iii) if entrepreneurs do not change FSPs based on which contracts they offer (either through regulation, lack of financial literacy etc.), more competition or reduction in spatial costs are not effective to increase welfare.

We provide several ways of taking our framework to the data. With market share data, we show how to recover the contracting frontier from variation in spatial configuration and competition across markets . This allows a researcher to conduct market structure counterfactuals without having to take a stand on which contracting friction is relevant. With household level data, we extend the methodology of Karaivanov and Townsend (2014) which maps unobserved equilibrium utilities to equilibrium contracts and show how to combine this with the entry model of Bresnahan and Reiss (1991). Our results indicate that reducing spatial costs, the logit variance and adding extra banks can increase utilities of agents, but in different magnitudes. Our results suggest that policy makers should focus on mechanisms that guarantee markets shares change when utility offerings change, which could be achieved through policies geared toward making SMEs more likely to choose better financial products (such as financial literacy, lending platforms etc.).

Our larger objective in this paper is to develop a tool kit, an operational empirical framework. In this sense, there are several ways in which our ideas can be naturally extended in future research. First, we believe our methods could be applied to other markets and more developed countries (e.g., health market in the U.S.). Second, we haven't explored the issue of dynamics - both in contracting and competition of FSPs, which may be relevant in other settings. Finally, the case of AdS can still be explored further. There are several implications of our comparative statics exercises (inequality within village, systemic risk etc.) that are not fully understood yet, to which our framework could prove useful, not only in theoretical models, but also in empirical applications.

References

Assuncao, J., S. Mityakov, and R. Townsend (2012). Ownership Matters: the Geographical Dynamics of BAAC and Commercial Banks in Thailand. Working paper.

Azevedo, E. and D. Gottlieb (2018). Perfect competition in markets with adverse selection. *Econometrica 85*(1), 67–105.

Berry, S. (1994). Estimating Discrete-Choice Models of Product Differentiation. *Rand Journal of Economics 25*, 242–262.

Berry, S., J. Levinsohn, and A. Pakes (1995). Automobile prices in market equilibrium. *Econometrica 7*(1), 841–890.

Bresnahan, T. and P. Reiss (1991). Entry and competition in concentrated markets. *Journal of Political Economy 99*, 977–1009.

Brook, Y., R. Hendershott, and D. Lee (1998). The Gains from Takeover Deregulation: Evidence from the End of Interstate Banking Restrictions. *The Journal of Finance 53*(6), 2185–2204.

D'Aspremont, C., J. J. Gabszewicz, and J.-F. Thisse (1979). On Hotelling 'Stability in Competition". *Econometrica 47*(5), 1145–1150.

Demyanyuk, Y., C. Ostergaard, and B. E. Sørensen (2007). U.S. Banking Deregulation, Small Businesses, and Interstate Insurance of Personal Income. *The Journal of Finance 62*(6), 2763–2801.

Doraszelski, U. and A. Pakes (2007). A Framework for Applied Dynamic Analysis in IO. In *Handbook of Industrial Organization*, Volume 3. Elsevier.

Dunn, A. (2009). Does Competition Among Medicare Advantage Plans Matter? An Empirical Analysis of the Effects of Local Competition in a Regulated Environment. DepâĂŹt of Justice Economic Analysis Group Discussion Paper.

Einav, L., A. Finkelstein, and J. Levin (2010). Beyond Testing: Empirical Models of Insurance Markets. *Annu. Rev. Econ. 2*, 311–336.

Einav, L., A. Finkelstein, S. P. Ryan, P. Schrimpf, and M. R. Cullen (2013). Selection on Moral Hazard in Health Insurances. *American Economic Review 103*, 178–219.

Einav, L., M. Jenkins, and J. Levin (2012). Contract pricing in consumer credit markets. *Econometrica 80*, 1387–1432.

Green, E. J. (1987). Lending and the Smoothing of Uninsurable Income. In E. C.Prescott and N. Wallace (Eds.), *Contractual Arrangements for Intertemporal Trade*. University of Minnesota Press, Minneapolis.

Handely, B., I. Hendelz, and M. Whinston. The welfare eÂğects of long-term health insurance contracts, journal = Unpublished manuscript, MIT, year = 2017.

Hotelling, H. (1929). Stability in Competition. *The Economic Journal 39*(153), 41–57.

Karaivanov, A. and R. Townsend (2014). Dynamic Financial Constraints: Distinguishing Mechanism Design from Exogenously Incomplete Regimes. *Econometrica 82*(3), 630–49.

Koijen, R. S. and M. Yogo (2012). The Cost of Financial Frictions for Life Insurers. *NBER Working Paper No. 18321*.

Lester, B., A. Shourideh, V. Venkateswaran, and A. Zetlin-Jones (2018). Screening and Adverse Selection in Frictional Markets. *Journal of Political Economy - Forthcoming*.

Lustig, J. D. (2010). Measuring welfare losses from adverse selection and imperfect competition in privatized medicare. Unpublished manuscript, Boston University.

Martin, A. and F. Taddei (2012). International capital flows and credit market imperfections: A tale of two frictions. *Journal of International Economics*, 441–452.

Moll, B., R. M. Townsend, and V. Zhorin (2017). Economic development, flow of funds, and the equilibrium interaction of financial frictions. *PNAS 114*(24), 6176–6184.

Nguyen, H.-L. Q. (2018). Are credit markets still local? evidence from bank branch closings. *American Economic Journal: Economic Policy - Forthcoming*.

Petersen, M. A. and R. G. Rajan (1995). The Effect of Credit Market Competition on Lending Relationships. *The Quarterly Journal of Economics 110*(2), 407–443.

Prescott, E. C. and M. Visscher (1977). Financial Obstacles and Inter-Regional Flow of Funds. *The Bell Journal of Economics 8*(2), 378–393.

Spear, S. E. and S. Srivastava (1987). On Repeated Moral Hazard with Discounting. *Review of Economic Studies 54*(4), 599–617.

Appendix

A A Model of Risky Production and Intermediation

In this section we discuss the model that generates the consumption-production puzzle (Section 2.1). We first go into the details of the model. Given the model set up, we then provide more details on the experiment and how the results in Tables 1 and 2 were generated.

Consider an economy with a continuum of types of agents indexed by $\theta_i \geq 1$ (to guarantee $\theta_i^2 \geq \theta_i$ later on, which means that intermediation will be profitable for these agents). The share of types in the population is given by $f(\theta_i)$. For each type θ_i, there is a continuum of agents of this type $i \in [0,1]$. An agent i with type θ_i produces a quantity $p_i(\theta_i)$, given by:

$$p_i(\theta_i) = \theta_i\left(1 + \frac{\sigma}{\sqrt{\theta_i}}\varsigma_i\right) \tag{70}$$

where $\varsigma_i \sim_{iid} \mathcal{N}(0,1)$. That is, each agent has a risky production here (and the risk is i.i.d. across agents). Agents with higher θ_i have both higher average payoff (given by θ_i) and higher production risk (the std. of idiosyncratic outcome is $\sqrt{\theta_i}\sigma$). Agents in this economy have the preferences usual risk-return preferences over a risky production process:

$$u(p_i(\theta_i)) = \mathbb{E}\left[p_i(\theta_i)\right] - A\mathbb{V}\left[p_i(\theta_i)\right] \tag{71}$$

where A is a measure of risk aversion, $\mathbb{E}[.]$ denotes the expectation over ε_i and $\mathbb{V}$ denotes the variance. For notatation purposes, we define the utility under autarky of type θ_i to be given by $u_A(\theta_i)$

Autarky. Under autarky, eah agent has to consume its production. Using the process in Eq. (70) and substituting in Eq. (71)

$$u(p_i(\theta_i)) = \theta_i\left(1 - A\sigma^2\right) \ \ (< \theta_i) \tag{72}$$

Here the average and standard deviation in consumption within types θ_i, denoted respectively by $c_A(\theta_i)$ and $s_A(\theta_i)$, is given by

$$c_A(\theta_i) = \theta_i \text{ and } s_A(\theta_i) = \sqrt{\theta_i}\sigma \tag{73}$$

Let $p_A(\theta_i)$ be the expected value of production for type i (which will match the observed for a large enough sample due to our iid assumption) under autarky. We have that $p_A(\theta_i)$ is given by

$$p_A(\theta_i) = \theta_i \tag{74}$$

Financial Intermediation with Full Information. FSPs can provide credit that allows entrepreneurs to increase production and, simultaneously, zero out the production risk (insurance). [37]. That is, they can transform the production process $p_i(\theta_i)$ into an intermediated process $p_I(\theta_i)$, as in Eq. (75), which has no uncertainty.

$$p_I(\theta_i) = \theta_i\lambda_i, \text{ where } \lambda_i > 1 \tag{75}$$

Providing credit is not costless for FSPs (cost of raising deposits, balance sheet constraints etc.). We assume

[37] The full insurance here is just for simplicity. The model extends for cases with partial insurance.

that costs are given by $.5\lambda^2$ to provide financial intermediation. FSPs charge $t(\theta_i)$ for this financial product that combines credit and insurance. In particular, we consider a model of monopolistic competition where each FSP solves Eq. (76)

$$\max_{\lambda_i, t_i} \left[t_i - .5\lambda.^2\right] D(\theta_i, \lambda_i, t_i) \tag{76}$$

where $D(.)$ is a demand with constant elasticity ε given by Eq. (77)

$$D(\theta_i, \lambda_i, t) \equiv (\theta_i \lambda_i - t - u_A(\theta_i))^{\varepsilon} \tag{77}$$

A few comments are in order. We capture imperfect competition in this example in a reduced form way. We assume that the demand is exogenously decreasing in the gap between the implied utility of a contract and the outside option. If the FSP provides a contract (λ_i, t_i) that gives the agent the same level of utility as under autarky, the demand for this contract is zero. The problem of the FSP is then to balance out the tradeoff between offering a low level of utility (by increasing t, for instance), which increases profits, but lowers demand. The elasticity of demand is the parameter in our economy that controls this tradeoff, which will essentially translate into market power from the FSPs.

The problem of the FSPs in Eq. (76) implies the optimal contract and transfers as given by Eq. (78) (See Appendix C for details).

$$\lambda_i \equiv \lambda(\theta_i) = \theta_i \text{ and } t(\theta_i) = \frac{1 + .5\varepsilon}{1 + \varepsilon}\theta_i^2 - \frac{1}{1 + \varepsilon}u_A(\theta_i) \tag{78}$$

and the implied utility for agents given the menu of contracts offered by FSPs, denoted by $u_I(\theta_i, \varepsilon)$, is given by Eq. (79)

$$u_I(\theta_i, \varepsilon) = \frac{\varepsilon}{1 + \varepsilon}\left(.5\theta_i^2\right) + \frac{1}{1 + \varepsilon}u_A(\theta_i) \tag{79}$$

We interpret $\omega \equiv \frac{1}{1+\varepsilon}$ as *market power* in this economy. The total cost of offering $\lambda(\theta_i) = \theta_i$ is given by $.5\theta_i^2$ and, therefore, the total output from intermediation is given by θ_i^2 (total production) minus $.5\theta_i^2$ (cost), which is equal to $.5\theta_i^2$. Market power in our economy determines how this gain is distributed between agents and FSPs. In particular, substituting ε by ω in Eq. (79), we have that

$$u_I(\theta_i, \omega) = (1 - \omega)\left(.5\theta_i^2\right) + \omega u_A(\theta_i) \tag{80}$$

Eq. (80) is a linear combination between autarky utility and the utility under perfect competition (where FSPs would make zero profits). We know that $u_A(\theta_i) \leq .5\theta_i^2$ due to the assumption that $\theta_i \geq 1$. The weights on this combination determines how the intermediation gain is divided is thus the market power of FSPs in this economy. For simplicity, we assume that the FSP provides financial services for all agents at equilibrium contracts even though the demand is downward sloping. [38]

Finally, note that under an intermediation regime, the average and standard deviation in consumption within

[38] For that, we only need to multiply the demand by a scaling factor. This assumption simplifies the analysis and is consistent with a case where the researcher has micro data on who uses financial intermediation. If this data is not available, then the role of market power in the scale of demand is also relevant and a potential source of bias which we are not taking into account in our example.

types θ_i, denoted by $c_I(\theta_i)$ and $s_I(\theta_i)$, respectively, is given by Eq. (81)

$$c_I(\theta_i) = u_I(\theta_i, \omega) \text{ and } s_I(\theta_i) = 0 \tag{81}$$

A.1 The Experiment

In the experiment we discuss in the text, we observe a sample of consumption and production for each household i in each village, denoted by, respectively, $\{c_i^v, p_i^v\}_{v,i=1}^N$. For simplicity, we assume that both samples have the same number of agents of type θ_i, which we denote by N_i, and that it matches the theoretical share of agents, that is $N_i = f(\theta_i)/N$.

Case 1: Autarky to Intermediation. Consider the theoretical difference in average consumption of a given type θ under intermediation with market power ω_1, denoted by c_I (Eq.81), and autarky, denoted by c_A (Eq. 73)

$$c_I(\theta_i) - c_A(\theta_i) = (1-\omega_1)\left(.5\theta_i^2\right) + \omega_1 u_A(\theta_i) - \theta_i = (1-\omega_1)\left(.5\theta_i^2 - \theta_i\right) - \omega_1 \theta_i A \sigma^2 \tag{82}$$

As we are assuming the researcher has a perfect experiment, it is is the case that the sample analogue converges to the theoretical difference in probability (i.e., the problem is not the statistical estimator). The problem in this case is not with the estimation, it is with the *interpretation* of the results. If risk aversion (denoted by A), risk in production (denoted by $\sigma.^2$) or ω are large enough, *average consumption goes down* in a move from autarky to intermediation with market power ω. At the same time, however, as utility is a convex combination of perfect competition and autarky utility, *utility is always increasing with intermediation with respect to autarky.* The intuition behind this result is that average consumption is a mix of three factors: market power, credit and insurance. Insurance can make the agent better off, even if decreases average consumption. If there is enough risk in project or the agent is too risk averse, or, alternatively, if the FSP has enough market power to keep production rents to itself, average consumption potentially is reduced. Overall, the effects on welfare are underestimated (even if in the sample it is true that consumption increases as it is the case with welfare). Moreover, note that types are often not observed by the researcher. In this case, the researcher computes the differences in average consumption across types. Let:

$$c_A \equiv \sum_i c_A(\theta_i) f(\theta_i) \text{ and } c_I \equiv \sum_i c_I(\theta_i) f(\theta_i) \tag{83}$$

From Eq. (82), we have that the difference in average consumption averaged across types in theory is given by

$$c_I - c_A = \sum_i \left\{(1-\omega)\left(.5\theta_i^2\right) + \omega u_A(\theta_i) - \theta\right\} f(\theta_i) \tag{84}$$

$$= (1-\omega)\left(.5\mathbb{V}_\theta + .5\mathbb{E}_{\theta_i}^2 - \mathbb{E}_\theta\right) - \omega \mathbb{E}_\theta A \sigma^2 \tag{85}$$

where $\mathbb{E}_\theta, \mathbb{V}_\theta$ denote, respectively, expectation and variance of θ in the population. Now, not only we confound the parameters of contracting and intermediation (and potentially give a wrong signal for average consumption), but our results are dependent also on the variance of θ due to its heterogeneous effects across agents. Two regions equally productive on average can have different outcomes of financial intermediation simply due to their distribution of productivity and the non-linear effects we see in our model.

One can correctly point out that we could potentially see the other side of this coin, which is that standard deviation of consumption should also fall with the introduction of intermediation. Conditional on types, it is true that the difference between standard deviation under intermediation, denoted by s_I (Eq.81), and autarky, denoted by S_A (Eq. 73) is given by

$$s_I(\theta_i) - s_A(\theta_i) = -\sqrt{\theta_i}\sigma \tag{86}$$

that is, variation in consumption comes down *within type* due to insurance. If types are not observed by the researcher, however, the variance in consumption across the sample can be mostly determined by variation *between types*. Considering the non-linearities introduced by the intermediation, it is possible that standard deviation of consumption between types increases. In the specific case of $\omega = 1$, which is a lower bound for the difference, we can show that

$$s_I(\theta_i) - s_A(\theta_i) = \sqrt{\mathbb{V}_\theta} - \sqrt{\mathbb{V}_\theta + \sigma^2 \mathbb{E}_\theta} \tag{87}$$

if $\mathbb{V}_\theta$ is large with respect to σ, even in this lower bound case, it is possible that the coefficient of variation (mean over standard deviation) increases simultaenously to a decrease in consumption. This can happen if average consumption decreases due to insurance, while variance does not decrease enough to keep the ratio constant, since most ot the variance comes from variation between types and not in production for a given type.

Moreover, note that we can also compute in this case differences in average production between autarky and intermediation with market power ω_1. From Eq. (74) and Eq. (75) we have that

$$p_I(\theta_i) - p_A(\theta_i) = \theta_i(\theta_i - 1) \tag{88}$$

which is positive due to our assumption of $\theta_i \geq 1$. In this case, we have that production is in fact increasing by looking at the micro data and comparing the two samples. In this case, the interpretation of the effects of financial intermediation become even murkier, since consumption is potentially reducing while production is increasing. In particular, note that we can substitute Eq. (88) in Eq.(82) to obtain

$$c_I(\theta_i) - c_A(\theta_i) = (1-\omega_1).5\,(p_I(\theta_i) - p_A(\theta_i)) - \theta_i\left\{.5(1-\omega_1) - \omega_1 A\sigma^2\right\} \tag{89}$$

which means that $(p_I(\theta_i) - p_A(\theta_i))$ is endogenous in Eq.(89). In this case, even with a perfect experiment, changes in production are correlated to changes in consumption through the structure of the model. In this case, it could be perceived that with intermediation agents are producing a larger income (Eq. 88) and the ones that have the bigger leap in income are also the ones to which the consumption decreases by more, which is not true in Eq.(82). This is related to the more empirical version of Eq. (1).

Case 2: Changes in Market Power. Now we focus on the case where the difference between the two samples is the level of market power. In particular, we assume that in one sample the market power is given by ω_1, while in the other it is given by $\omega_2 < \omega_1$. In this case, we can write the difference between consumption

$$c_I(\theta_i;\omega_2) - c_I(\theta_i;\omega_1) = (\omega_1 - \omega_2)\underbrace{\left[.5\theta_i^2 - \theta_i(1-\sigma^2 A)\right]}_{\equiv g_I(\theta_i)} \tag{90}$$

where $g_I(\theta_i)$ corresponds to the total intermediation gains in utility in our model, that is, the difference between the output gains (discount of the cost of intermediation) of credit and the autarky utility of the agent. This is total amount of extra utility this economy is generating through intermediation. Our market power parameter, ω captures how this gain is shared across FSPs and agents. For different levels of ω, changes in consumption are simply a multiplier of this intermediation gains. The reason is that in both of this scenarios there is no risk and same level of credit, so the only difference is the redistribution of gains from intermediation. In a model where competition affects contracts offered, as will be the case we focus paper, we would be back to a problem of multidimensional contracting as seen in moving from autarky to some intermediation. In this case where the two samples differ by market power, if the researcher has a model on gains from financial intermediation, which depend on utility specification and production function, differences in consumption identify differences in market power. If the researcher does not has this model, changes in consumption in the observed sample will pin down changes in market power times the gains from intermediation (which can be small or large, or even different at the market level).

In this case, note that

$$c_I(\theta_i;\omega_2) - c_I(\theta_i;\omega_1) = (\omega_1 - \omega_2)\underbrace{\left[.5\theta_i^2 - \theta_i(1-\sigma^2 A)\right]}_{\equiv g_I(\theta_i)} + p_I(\theta_i, nu_2) - p_I(\theta_i,\omega_1) \tag{91}$$

since $p_I(\theta_i,\omega_2) = p_I(\theta_i,\omega_1)$. Therefore, changes in consumption in this economy have nothing to do with changes in production, since agents are insured against production shocks.

To generate the outcomes in Tables 1 and 2. We use the parameters in Table 18. We compare two potential changes: from autarky to an economy with $\omega_1 = .3$ (Table 1) and from $\omega_1 = .3$ to $\omega = .1$ (Table 2). We assume that the distribution of θ is: $\theta = \min(1, X)$, where $X \sim \mathcal{N}(\mathbb{E}_\theta, \mathbb{V}_\theta)$. In this case, $\mathbb{E}_\theta$ is not the actual average of θ, but this facilitates the computation of the statistics of interest.

Table 18: Parameter Values, Numerical Example

Parameter	Value	Role
σ	1	Variance in Production
A	1	Risk Aversion
$\mathbb{E}_\theta$	2	Mean of Types in Population
V_θ	2	Variance of Types in Population

We use the parameters in Table 18 to simulate 70 households in 100 control and treatment villages, to which we take averages and standard deviations as in Tables 1 and 2. We bootstrap our sample 1,000 times to obtain standard error estimates.

B From Adverse Selection to Full Information

We focus now on the model that is behind the information structure puzzle and Table 3 in Section 2.2. This model is an extension of the model of Section A to unobserved types (AdS).

We assume that there are two types, now unobserved to the FSP, and to the researcher. In the incomplete information case, the problem of the FSP becomes a generalized version of Eq. (76), where we also take into account the truth telling constraints. The problem of an FSP is now given by Eq.(92)

$$\max_{\{\lambda(\theta),t(\theta)\}_\theta} f(\theta_L)D(\theta_L,\lambda_L,t_L)\left[t_L-.5\lambda_L^2\right]+f(\theta_H)D(\theta_H,\lambda_H,t_H)\left[t_H-.5\lambda_H^2\right] \tag{92}$$

s.t. to the *Truth Telling* constraints:

$$\theta_L\lambda_L-t_L\geq\theta_L\lambda_H-t_H \tag{93}$$

$$\theta_H\lambda_H-t_H\geq\theta_H\lambda_L-t_L \tag{94}$$

where $f(\theta)$ is the share of type θ in the population. As usual, only one of the TT constraints potentially bind. Contrary to the textbook case, however, we don't know however which constraint is binding. As the FSP does not have all of the monopoly power, it cannot fully extract rents and the differences in ability to extract production rents and distribution of types in the population will determine which constraint is binding. For simplicity, we assume that

$$\sigma^2A=1$$

which guarantees that autarky utilities of both agents now are zero (See Eq. (72)). The truth telling constraints - Eqs.(93)-(94) - are not binding whenever (See Appendix C.1) :

$$\omega\leq\frac{\theta_H-\theta_L}{\theta_H+\theta_L} \tag{95}$$

Which already starts to provide the relationship between AdS and market power: *the truth telling constraint only binds if there is not enough competition in this model.* With little competition, the transfers for each type are sufficiently different - since they keep most of the surplus of the trade - that no type wants to take the quality-transfer pair of the other. If we experimented with a village at this level of market power, we would observe not effect of a screening system in increasing credit (since AdS is not binding to being with).

From AdS to Full Information. In the previous section we focused on how to understand the effects on welfare from consumption data. Now, we focus on a case where welfare is observed and want to understand the effect of an economy moving from adverse selection to full information, both in terms of credit (λ) and utility (u) for both types of agents at different leves of market power. We use subscripts L,H for credit and utility of each type. For that, we solve the problem of the FSPs of maximizing Eq.(92) subject to the truth telling constraint in Eq. (93)-(94) for various levels of ω, the market power. We use the parameters in Table 19.

We plot the leverage (total credit provided, λ_i) of low and high types chosen by the FSP in Figure 26. The vertical line shows where the constraint is binding - the minimum value of μ such that Eq. (95) is violated. For $\omega\leq\frac{\theta_H-\theta_L}{\theta_H+\theta_L}$, we have that contracts are as in the full information case. However, for $\omega>\frac{\theta_H-\theta_L}{\theta_H+\theta_L}$, the FSP must distort the contracts. Not that when compared with the full information case, the low type can have

Table 19: Outcomes from Different Intermediation Regimes

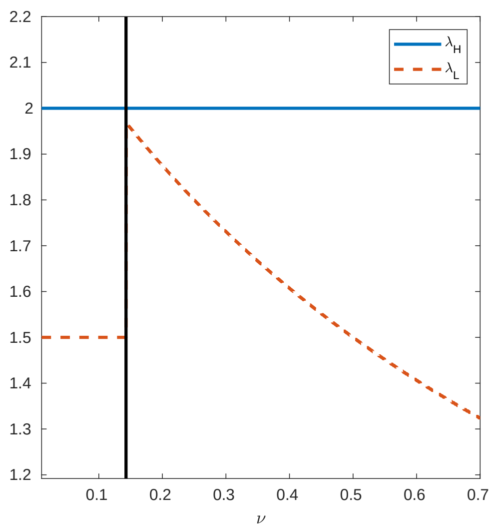
Parameter	Value	Meaning
θ_H	2	High Type
θ_L	1.5	Low Type
f_H	.75	Share of Low Type in Pop.
f_L	.25	Share of High Type in Pop.

more or less leverage under adverse selection. The allocation for the high type, however, is never distorted. This is a result of the fact that in the parameters we use, Eq.(94) is binding. Contrary to the textbook case of adverse selection with two types, however, we do not know *ex-ante* constraint binds in this example. See Appendix C for more details on this.

We plot the difference in utility from an adverse selection to a full information economy for high and low types in Figure 27. Utilities for one or both agents can decrease or increase by moving from AdS to full information. In particular, if the FSPs have enough market power (high ω), *agents are better off in an environment with AdS*. As the FSPs does not know how to differentiate the agents, it cannot extract the rents it would otherwise in a full information case.

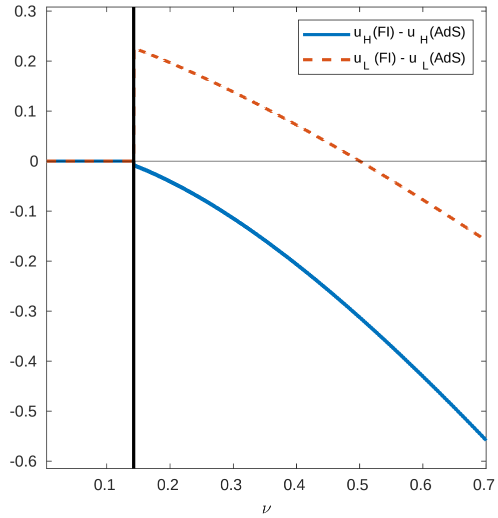
With a well designed experiment and observing utility (with the caveats we discussed in the other example), a researcher can infer what is the effect moving from Adverse Selection to Full Information for a given level of ω. However, we cannot interpret the results more generally, i.e., beyond the studied setting and to study public policy more generally. Furthermore, if financial markets in the economy are heterogeneous in terms of ω, it could be that the study finds no effect as a combination of negative effects of moving from adverse selection to full information in markets where ω is high with the positive effects of when it is low.

Figure 26: Leverage of Low and High Types



Leverage consistent with the solution to the problem of the FSPs of maximizing Eq.(92) subject to the truth telling constraint in Eq. (93)-(94) for various levels of ω, the market power. Types $\theta_H = 1.5$, $\theta_L = 1$, with probability $f_L = .75$ and $f_H = .25$. We use $u_{A,H} = 0$, $u_{A,L} = 0$, consistent with a $\sigma^2 A = 1$. Vertical line represents the point at which truth telling constraint starts to bind - the minimum value of μ such that Eq. (95) is violated.

Figure 27: Utility of Low and High Types



Note: Utility consistent with the solution to the problem of the FSPs of maximizing Eq.(92) subject to the truth telling constraint in Eq. (93)-(94) for various levels of ω, the market power. Types $\theta_H = 1.5$, $\theta_L = 1$, with probability $f_L = .75$ and $f_H = .25$. We use $u_{A,H} = 0$, $u_{A,L} = 0$, consistent with a $\sigma^2 A = 1$. Vertical line represents the point at which truth telling constraint starts to bind - the minimum value of μ such that Eq. (95) is violated.

C Algebra for Section 2

Complete Information. In the complete information case, the problem of the FSP is:

$$\max_{\lambda,t} D(\theta_i,\lambda,t)\left[t - .5\lambda^2\right] \tag{96}$$

taking the FOC:

$$-D'\left[t-\lambda^2\right] + D = 0 \tag{97}$$

$$\theta d'\left[t-\lambda^2\right] - D\lambda = 0 \tag{98}$$

Dividing Eq. (97) by (98

$$\lambda(\theta_i) = \theta_i \tag{99}$$

To compute the transfers, we replace $\lambda(\theta_i) = \theta_i$ in Eq. (97):

$$t - .5\theta_i^2 = D/D' = (\theta_i^2 - t - u_{0,\theta_i})\varepsilon^{-1} \Rightarrow t(\theta_i) = \frac{1+.5\varepsilon}{1+\varepsilon}\theta_i^2 - \frac{1}{1+\varepsilon}u_0(\theta_i) \tag{100}$$

Adverse Selection. The problem of the FSP becomes:

$$\max_{\lambda_L,t_L,\lambda_H,t_H} f_L D(\theta_L,\lambda_L,t_L)\left[t_L - .5\lambda_L^2\right] + f_H D(\theta_H,\lambda_H,t_H)\left[t_H - .5_H^2\right] \tag{101}$$

s.t. to the *Truth Telling* constraints:

$$\theta_L\lambda_L - t_L \geq \theta_L\lambda_H - t_H \tag{102}$$

$$\theta_H\lambda_H - t_H \geq \theta_H\lambda_L - t_L \tag{103}$$

We can transform the problem to become:

$$\max_{\lambda_L,u_L,\lambda_H,u_H} D(u_L)\left[\theta_L\lambda_L - .5\lambda_L^2 - u_L\right] + D(u_H)\left[\theta H\lambda_H - .5_H^2 - u_H\right] \tag{104}$$

s.t.

$$u_L \geq u_H - (\theta_H - \theta_L)\lambda_H \tag{105}$$

$$u_H \geq u_L + (\theta_H - \theta_L)\lambda_L \tag{106}$$

Note that a monotonicity condition ($\lambda_H > \lambda_L$) joint with one truth telling constraint implies the other, that is

$$u_H = u_L + (\theta_H - \theta_L)\lambda_L \Rightarrow u_L = u_H - (\theta_H - \theta_L)\lambda_L > u_H - (\theta_H - \theta_L)\lambda_H \tag{107}$$

and that a binding TT of the low type is not consistent with the TT of the high type:

$$u_L = u_H - (\theta_H - \theta_L)\lambda_H \Rightarrow u_H = u_L + (\theta_H - \theta_L)\lambda_H > u_H + (\theta_H - \theta_L)\lambda_L \tag{108}$$

Therefore, there is only one potentially binding truth telling constraint. Contrary to the textbook case, however, we don't know which constraint is binding. As the FSP does not have full market power, it is not trivial that the high type is the one extracting information rents. This is a function of each type outside option and share that each type appears in the population. To see this, consider a case where $f_H \to 0$. In this case, it is better to keep the allocation to the low type undistorted, while distorting the allocation of the high type to satisfy the truth telling constraints.

Assume without loss of generality that it is the one of the high type. The Lagrangian in the problem in Eq.(104) becomes:

$$\mathcal{L} \equiv f_L D(u_L)\left[\theta_L\lambda_L - .5\lambda_L^2 - u_L\right] + f_H D(u_H)\left[\theta_H\lambda_H - .5\theta_H^2 - u_H\right] + \psi(0 - u_L - (\theta_H - \theta_L)\lambda_L + u_H) \tag{109}$$

The FOC system is:

$$f_L D^{'}(u_L)\left[\theta_L\lambda_L - .5\lambda_L^2 - u_L\right] - f_L D(u_L) - \psi = 0 \quad (u_L) \tag{110}$$

$$f_H D^{'}(u_H)\left[\theta H\lambda_H - .5\lambda_H^2 - u_H\right] - f_H D(u_H) + \psi = 0 \quad (u_H) \tag{111}$$

$$f_L D(u_L)\left[\theta_L - \lambda_L\right] + (\theta_H - \theta_L)\psi = 0 \quad (\lambda_L) \tag{112}$$

$$\lambda_H = \theta_H \quad (\lambda_H) \tag{113}$$

$$u_H = u_L + (\theta_H - \theta_L)\lambda_L \quad (\psi) \tag{114}$$

which shows the no distortion result of Figure 26. In the numerical simulation, we use the following scaling for the demand function:

$$D(u) = \left(\frac{u - u_0}{.5\theta^2 - u0}\right)^{\varepsilon} \tag{115}$$

that is, there is a share between zero and one ($.5\theta^2$ is the perfect competition outcome) that uses intermediation and the curvature is given by the market power.

C.1 Deriving Eq. (95)

Under the assumption that $\sigma^2 A = 1$, we have that autarky utilities are zero for both types, that is $U_{A,H} = u_{A,L} = 0$

$$t_H = \frac{1+.5\varepsilon}{1+\varepsilon}\theta_H^2 \text{ and } t_L = \frac{1+.5\varepsilon}{1+\varepsilon}\theta_L^2$$

Therefore, using the full information solution of t_L, t_H as above and $\lambda_H = \theta_H$ and $\lambda_L = \theta_L$, we can re-write the truth telling constraint of the low type, which is generically given by

$$\theta_L\lambda_L - t_L \geq \theta_L\lambda_H - t_H$$

as

$$\theta_L^2 - \frac{1+.5\varepsilon}{1+\varepsilon}\theta_L^2 \geq \theta_L\theta_H - \frac{1+.5\varepsilon}{1+\varepsilon}\theta_H^2 \Leftrightarrow \theta_L(\theta_L - \theta_H) \geq \frac{1+.5\varepsilon}{1+\varepsilon}(\theta_L - \theta_H)(\theta_L + \theta_H) \tag{116}$$

$$\Leftrightarrow \theta_L(.5\varepsilon) \geq (1+.5\varepsilon)\theta_H \Leftrightarrow \theta_L - \theta_H \geq \frac{2}{\varepsilon}\theta_H \tag{117}$$

Moreover,

$$\omega \equiv \frac{1}{1+\varepsilon} \Rightarrow \varepsilon = \omega^{-1} - 1$$

Substituting and manipulating we arrive at

$$\omega \geq \frac{\theta_L - \theta_H}{\theta_H + \theta_L} \tag{118}$$

which is always satisfied in the first best, so the TT constraint is not binding in the full information contracts. Note, however, that we can simply re-do the analysis for the truth telling constraint on the high type, and in this case get exactly Eq.(95). ■

D From contracts to utilities

The starting point for the theoretical framework is how to take a potentially very complicated object - a financial contract - and simplify it to a tractable concept, *utility*.

Let $\boldsymbol{C}$ represent a contract in a set of contracts $\mathfrak{C}$. $\boldsymbol{C}$ is potentially multidimensional (e.g., interest rate, collateral and cost of default). The set $\mathfrak{C}$ is already constrained by the contracts that satisfy the contracting frictions. The agent in the model has a utility $\mathbb{U}:\mathfrak{C}\to\mathbb{R}$. This utility can represent an expected utility if the contract depends on realizations of stochastic variables. For any contract, we also assume that we can specify the profit of a FSP, $\Pi:\mathfrak{C}\to\mathbb{R}$. Moreover, denote W as the set of utilities generated by any contract, i.e., $W\equiv\{u\mid\exists\,\boldsymbol{C}\in\mathfrak{C}\text{ s.t. }\mathbb{U}(\boldsymbol{C})=u\}$. We assume that our contracting structure is s.t. *Assumption U* holds. *Assumption U* is essentially a limitation in the set $\mathfrak{C}$ beyond the limitations caused by contracting frictions.

Assumption U.

1. *No contract is Pareto Dominated, i.e., for any* $\forall\,\boldsymbol{C}_0,\boldsymbol{C}_1\in\mathfrak{C}$:

$$\Pi(\boldsymbol{C}_0)>\Pi(\boldsymbol{C}_1)\Leftrightarrow\mathbb{U}(\boldsymbol{C}_0)<\mathbb{U}(\boldsymbol{C}_1) \tag{119}$$

2. *There are no different contracts that offer the same utility for agents and profits for FSPs*:

$$\nexists\,\boldsymbol{C}_0,\boldsymbol{C}_1\in\mathfrak{C}\text{ s.t. }\boldsymbol{C}_0\neq\boldsymbol{C}_1,\ \Pi(\boldsymbol{C}_0)=\Pi(\boldsymbol{C}_1)\text{ and }\mathbb{U}(\boldsymbol{C}_0)=\mathbb{U}(\boldsymbol{C}_1) \tag{120}$$

Eq.(119) means that if a contract is more profitable for the FSP, it provides less utility for the agent. What we are ruling in out is that in the set of *feasible* contracts that satisfy all information constraints, there is a contract $\boldsymbol{C}_0$ that $\Pi(\boldsymbol{C}_0)>\Pi(\boldsymbol{C}_1)$ and $\mathbb{U}(\boldsymbol{C}_0)\geq\mathbb{U}(\boldsymbol{C}_1)$ in this case, $\boldsymbol{C}_0$ would be a Pareto improvement over $\boldsymbol{C}_1$ and that there is no reason to play it. In our framework, this assumption is very natural. As it becomes clear later when we discuss limitations, there a few cases where it may not hold. Eq.(120) rules out a contract that is equivalent for agents and FSPs at the same time. This assumption is true in our application, where agents are risk-averse and FSPs risk neutral. If there were two equivalent contracts for the agents, there would be a Pareto improvement of offering the mean contract [39] that would provide a higher utility for the agent and the same profit for the FSP.

What *Assumption U* in fact guarantees is that: $\nexists\,\boldsymbol{C}_0,\boldsymbol{C}_1\in\mathfrak{C}$ s.t. $\boldsymbol{C}_0\neq\boldsymbol{C}_1$ and $\mathbb{U}(\boldsymbol{C}_0)=\mathbb{U}(\boldsymbol{C}_1)$, i.e., no two contracts offer the same utility. Eq. (120) guarantees that they do not have the same profits. Eq. (119) eliminates the possibility that one of them is better for the FSP than the other - i.e., implies that these contracts have the same profit. Therefore, it cannot exists under *Assumption U*. What this means is that the optimization problem in Eq. (121) is well defined:

$$c^*(u)\in\arg\max_{\boldsymbol{C}\in\mathfrak{C}}\pi(\boldsymbol{C})\quad\text{s.t.}\quad\mathbb{U}(\boldsymbol{C})=u \tag{121}$$

and that, $\forall\,\boldsymbol{C}_0\in\mathfrak{C}$, the solution of the Eq. (121) is s.t.:

$$c^*\left(U\left(\boldsymbol{C}_0\right)\right)=\boldsymbol{C}_0$$

[39] Assuming that $\mathfrak{C}$ is convex, i.e., that the mean contract is in the contract space.

i.e., that there is a one-to-one mapping from contracts to utilities implied by this contract. Since utility generated by contract $\boldsymbol{C}_0$ cannot be generated by any other contract, i.e., for the contracts $\boldsymbol{C}_0$ that satisfy this assumption, $\exists!\ \boldsymbol{C} \in \mathfrak{C}$ s.t. $\mathbb{U}(\boldsymbol{C}) = \mathbb{U}(\boldsymbol{C}_0)$. In this case, the constraint $\mathbb{U}(\boldsymbol{C}) = u$ rules out all contracts that are not $\boldsymbol{C}_0$. Thus, Eq.(121) holds under *Assumption U*. Therefore, in this case, the mapping of contracts to utilities is one-to-one and the transformation can be done without loss of generality. In the specifics of our framework, it will be clear that *Assumption U* holds.

E Proof of Lemma 3.1

Proof. The strategy to show that the equilibrium exists and is unique is to show that the vector of best response functions is a contraction. The Nash Equilibrium is then the unique fixed point of the vector of best response functions. This is not only useful theoretically, but also numerically: computing the fixed point of a contraction can be done by an iterative algorithm. The first step of the proof is Lemma E.1, which is a version of Blackwell's sufficient conditions for operators between compact subspaces of $\mathbb{R}^n$, which is our case here.

Lemma E.1. *Let $T : C \to C$, $C \subset \mathbb{R}^n$, C compact. Define $\|x-y\| \equiv \max_i |x_i - y_i|$ and $x \leq y$ if $x_i \leq y_i, i = 1,...,n$. Then, if:*

1. (Monotonicity) $x \leq y \Rightarrow Tx \leq Ty$, $\forall x, y \in C$.

2. (Discount) $T(x + \boldsymbol{e}a) \leq T(x) + \beta \boldsymbol{e}a$, $\forall x \in W, a \in \mathbb{R}_+$, $\boldsymbol{e} = (1,...,1) \in \mathbb{R}^n_+$ and $x + \boldsymbol{e}a \in C$.

T is a contraction with modulus β.

Proof. $\forall x, y$: $x - y \leq \boldsymbol{e}\|x-y\|$. This imples that $x \leq y + \boldsymbol{e}\|x-y\|$ By properties 1 and 2: $Tx \leq Ty + \beta \boldsymbol{e}\|x-y\|$. Also, the same is true for x in place of y: $Ty \leq Tx + \beta \boldsymbol{e}\|x-y\|$. Therefore: $Tx - Ty \leq \beta \boldsymbol{e}\|x-y\|$ and $Ty - Tx \leq \beta \boldsymbol{e}\|x-y\|$ which implies $\|Tx - Ty\| \leq \beta\|x-y\|$. ■

Moreover, we present an auxiliary Lemma E.2 on the argmax of problems of a particular condition - which we then show to hold in our case. This is simply a way to simplify the exposition.

Lemma E.2. *Let $f, g : \mathbb{R} \to \mathbb{R}$ be strictly concave functions, f strictly decreasing in x and g strictly increasing in x. Moreover, Let g be continuously differentiable in x. Let I be a compact interval. Let $\lim_{x \to \max_I} f(x) = -\infty$, and g bounded above. Finally, let δf be the correspondence function of subgradients of f. Define: $x^* \equiv arg\max_{x \in I} f(x) + g(x)$. Then x^* exists, is unique and is s.t.:*

$$(x^* - \min I) \text{ and } g'(x) < -[\max\{\delta f(x)\}], \forall\, x \in I \tag{122}$$

or

$$g'(x^*) \in -\delta f(x^*) \tag{123}$$

Proof. Existence and uniqueness comes from strict concavity and continuity. The optimum is not at max I since $\lim_{x \to \max_I} f(x) = -\infty$. Therefore, the optimum must either be at min I or satisfy the FOC. To derive the FOC, take x_0 s.t. $g'(x_0) \in -\delta f(x_0)$. Due to the strict concavity of f, g, $f(x) - f(x_0) < -g'(x)(x - x_0)$ and $g(x) - g(x_0) < g'(x)(x - x_0), \forall x \neq x_0$, Therefore: $f(x) + g(x) < f(x_0) + g(x_0)$. If the optimum is at min I, however, it must be the case that $g'(x) < -[\max\{\delta f(x_0)\}]$, $\forall\, x \in I$ - otherwise we could find an interior maximum. ■

Back to our original problem, we do the following steps. First, we do a transformation where we include the outside option as a phantom player. Second, we show that the BR function satisfies the condition of the Lemma E.2 and the implications for the specific case of BR functions. Third, we discuss the boundaries on changes of the BR function. Forth, we show that the conditions for Lemma 3.1 are satisfied and conclude the proof.

Step 1. The Phantom Player. We assume that there is a phantom player in the game, the *outside option* player. This player represents the outside option: we assume that it is as if it is another FSP, but it always plays the outside option. We denote it by $b = B_p + 1$, i.e., the extra bank in the province. We do this transformation to facilitate the proof of uniqueness and existence. Define the best response function vector (i.e., of all FSPs) by [40]

$$BR(u_1, u_2, ...u_n) \equiv \left(BR_1\left(u_{-1}\right), BR_2\left(u_{-2}\right), ..., BR_{B_p+1}\left(u_{-B_p+1}\right)\right)$$

BR is a function that maps the cartesian products of the strategy spaces in itself: $BR : W^{B_p+1} \to W^{B_p+1}$ maps a set of strategies of all FSPs $\{u_b\}_b \in W_{B_p}$ in the best response of a FSP given the strategies of all other FSPs.

Step 2. Best Response and Auxiliary Lemma. Note that S is decreasing and concave in u. As S is concave, it is continuous everywhere and differentiable almost everywhere (a.e.) in $u \in W$. Also, note that W is a compact subset of $\mathbb{R}$. Moreover, note that:

$$BR_b\left(\{u_b\}_{b=1}^{B_p+1}\right) = \arg\max_{u \in W} S\left(u\right)\mu\left(u, u_{-b}\right) = \arg\max_{u \in \hat{W}} \ln\left[S\left(u\right)\right] + \ln\left[\mu\left(u, u_{-b}\right)\right] \tag{124}$$

where $\hat{W} \equiv W \cap \{u \mid S(u) \geq 0\}$, which is compact. The idea is that if $\exists u \in W s.t. S(u) > 0$, then no utility in equilibrium is played with $S(u) \leq 0$. This implies that the function

$$f\left(u\right) \equiv \ln\left(S\left(u\right)\right)$$

is: (i) strictly concave, decreasing in u, (ii) $\lim_{u \to \max\ \hat{W}} \Sigma(u) = -\infty$. Therefore, $f(u)$ plays the role of f in the Lemma E.2. Moreover, let

$$g\left(u, u_{-b}\right) \equiv \ln\left(\mu\left(u, u_{-b}\right)\right)$$

which is: (i) strictly concave, increasing in u, (ii) continuously differentiable at $u \in interior(W)$, (iii) bounded above by zero. Therefore, $g(u, u_{-b})$ plays the role of g in Lemma E.2 for any given value of u_{-b}. Define: $\Sigma : \hat{W} \rightrightarrows \mathbb{R}$ as:

$$\Sigma(u) \equiv -\frac{\delta S(u)}{S(u)} \tag{125}$$

and:

$$\Upsilon(u, u_{-b}) \equiv \frac{\partial_{u_b}\mu\left(u, u_{-b}\right)}{\mu\left(u, u_{-b}\right)} \tag{126}$$

to represent the equivalents of $\delta f, g^{'}$ in Lemma E.2, respectively. From Lemma E.2:

$$BR_b\left(\{u_b\}_{b=1}^{B_p+1}\right) = \begin{cases} \min \hat{W} \text{ and } \Sigma\left(u\right) > \Upsilon\left(u, u_{-b}\right), \ \ \forall u \in \hat{W} \\ \text{ or } \Upsilon\left(BR_b\left(\{u_b\}_{b=1}^{B_p+1}\right), u_{-b}\right) \in \Sigma\left(BR_b\left(\{u_b\}_{b=1}^{B_p+1}\right)\right) \end{cases}, \ \forall\ b = 1, ..., B$$

[40]The best responses are also a function of the locations of FSPs and villages, taken as given. We remove it from the notation at this point to facilitate the understanding.

For our phantom player - the outside option:

$$BR_b\left(\{u_b\}_{b=1}^{B_p+1}\right) = u_0$$

Step 3. The BR Boundaries. We start this step with two observations: $\Sigma(u)$ is strictly increasing in u and $\Gamma(u, u_{-b})$ is strictly decreasing in u, strictly increasing in u_{-b}. We know that $-\delta S$ is increasing in u. Moreover, S is strictly decreasing in u. The ratio $\Sigma(u)$ is thus increasing, meaning that $u > \hat{u} \Leftrightarrow x < y, \forall x \in \Sigma(u), y \in \Sigma(\hat{u})$. Second, as μ is log-concave in u, Υ must be decreasing in u and increasing in u_{-b} (since Υ is the first derivative of $\ln(\mu)$).

The fact that $\Sigma(u)$ is strictly increasing in u and $\Gamma(u, u_{-b})$ is strictly decreasing in u guarantees that:

$$\left[BR_b\left(\{u_b\}_{b=1}^{B_p+1}\right) - BR_b\left(\{\hat{u}_b\}_{b=1}^{B_p+1}\right)\right]^2 \leq \left[BR_b^{FOC}\left(\{u_b\}_{b=1}^{B_p+1}\right) - BR_b^{FOC}\left(\{\hat{u}_b\}_{b=1}^{B_p+1}\right)\right]^2 \tag{127}$$

where: $BR_b^{FOC}\left(\{u_b\}_{b=1}^{B_p+1}\right)$ is defined as the point that satisfies the equation:

$$\Upsilon\left(BR_b^{FOC}\left(\{u_b\}_{b=1}^{B_p+1}\right), u_{-b}\right) \in \Sigma\left(BR_b^{FOC}\left(\{u_b\}_{b=1}^{B_p+1}\right)\right)$$

even if $BR_b^{FOC} \notin W$. BR_b^{FOC} is picking the utility that solves the FOC if there is no lower-bound to possible levels of utility that are offered. Given Eq. (127), it is sufficient to show that

$$\left[BR_b^{FOC}\left(\{u_b\}_{b=1}^{B_p+1}\right) - BR_b^{FOC}\left(\{\hat{u}_b\}_{b=1}^{B_p+1}\right)\right]^2 \lesseqq a$$

to guarantee that the same is true for the BR functions. Therefore, we assume in the following step that the condition that the equilibrium utility is in the interior of W never binds. Moreover, note that Eq. (127) is always satisfied for the outside option, since the LHS of Eq. (127) is always zero.

Step 4. The contraction.

Given the conditions on the BR function, we now proceed to show that the two conditions in *Lemma 1* hold for BR_b^{FOCs} and, thus, for BR_b. Focus on the FOC of our problem, that is:

$$\Upsilon\left(u_b^*, u_{-b}\right) \in \Sigma\left(u_b^*\right) \tag{128}$$

where $u_b^* \equiv BR_b\left(\{u_b\}_{b=1}^{B_p+1}\right)$. Then:

1. **Monotonicity.** If $\hat{u}_{-b} \geq u_{-b} \Rightarrow \hat{u}_b^* \geq u_b^*$. Assume by contradiction that $\hat{u}_{-b} \geq u_{-b}$ and $\hat{u}_b^* < u_b^*$. We know that:
$$\Upsilon\left(\hat{u}_b^*, \hat{u}_{-b}\right) < \Upsilon\left(\hat{u}_b^*, u_{-b}\right) < \Upsilon\left(u_b^*, u_{-b}\right) \in \Sigma\left(u_b^*\right) < \Sigma\left(\hat{u}_b^*\right) \Rightarrow \Upsilon\left(\hat{u}_b^*, \hat{u}_{-b}\right) < \Sigma\left(\hat{u}_b^*\right) \tag{129}$$
which cannot happen at an interior solution.

2. **Discounting.** If $\tilde{u}_{-b} = u_{-b} + ea$, $a > 0, e = (1, ..., 1), u_{-b} + ea \in \hat{W}^B \subset \mathbb{R}^B \Rightarrow \tilde{u}_b^* \in (u_b^*, u_b^* + a)$. We know from monotonicity $\tilde{u}_b^* > u_b^*$. Assume by contradiction that $\tilde{u}_b^* \geq u_b^* + a$.

$$\Upsilon\left(\tilde{u}_b^*, \tilde{u}_{-b}\right) > \Upsilon\left(u_b^* + a, \tilde{u}_{-b}\right) = \Upsilon\left(u_b^*, u_{-b}\right) \in \Sigma\left(u_b^*\right) > \Sigma\left(\tilde{u}_b^*\right) \Rightarrow \Upsilon\left(\tilde{u}_b^*, \tilde{u}_{-b}\right) > \Sigma\left(\tilde{u}_b^*\right) \tag{130}$$

which cannot happen at an interior solution. Note that this is where we use the condition of Eq. (128). The above reasoning guarantees that, $\forall a \in \mathbb{R}_+$, $\exists \beta_b(a)$ s.t.:

$$\tilde{u}_b^* \leq u_b^* + \beta_b(a) \tag{131}$$

As we know that: $\beta_b(a) < 1$, take the β of the contraction as: $\beta \equiv \max_b \max_a \beta_b(a)$. Note that, as $\hat{W}$ is compact, $\beta_b(a) < 1 \Rightarrow \beta < 1$.

Conclusion. Note that, as the Steps 1-4 above are true for all FSPs and the phantom bank, it must be true that the BR of Eq. (124) is a contraction. The Nash Equilibrium is the unique fixed point of the best responses and, thus, can be found through an iterative procedure (see). For a more intuitive approach to the proof, see Figure (6) in the main text. ■

F Proof of Lemma 3.2

Proof. **Step 1. Bounded Away from Zero.** With the logit formulation, the minimum market share in a given village is s.t.

$$\mu_v(\varphi_b) > \frac{e^{\sigma_L^{-1}[\mathbb{V}(u_{min}, x_b, x_v) - u_0]}}{1 + \sum_{\hat{b}=1}^{B} e^{\sigma_L^{-1}[\mathbb{V}(u_{max}, x_{\hat{b}}, x_v) - u_0]}} > 0$$

where the inequality comes from replacing u_{max} also for b at the denominator. As $\sigma_L > 0$, it is the case that the RHS is larger than zero.

Step 2. Log-concave in u. Taking the derivative of Υ (which corresponds to the second derivative of log-market share)

$$\Upsilon(u_b, u_{-b}) = \frac{\partial_{u_b}\mu(\varphi_b)}{\mu(\varphi_b)} = \frac{\sum_{v=1}^{V} N_v \mu_v(\varphi_b)[1 - \mu_v(\varphi_b)]}{\sum_{v=1}^{V} N_v \mu_v(\varphi_b)} = 1 - \frac{\sum_{v=1}^{V} N_v \mu_v(\varphi_b)^2}{\sum_{v=1}^{V} N_v \mu_v(\varphi_b)} \tag{132}$$

Note that in the case with spatial cost $\psi = 0$ (or a single market, i.e., $V = 1$), the above condition reads as $\Upsilon(u_b, u_{-b}) = 1 - \mu(\varphi_b)$, which is trivially strictly decreasing in u_b. In our problem with $\psi > 0, V > 1$, however, we need to do some additional steps. Taking the derivative of Eq. (132) and simplifying the notation of $\mu(\varphi_b)$ to μ:

$$\partial_{u_b}\Upsilon(u_b, u_{-b}) = -\frac{2\left[\sum_{v=1}^{V} N_v \mu_v^2 (1 - \mu_v)\right]\left[\sum_{\hat{v}=1}^{V} N_{\hat{v}} \mu_{\hat{v}}\right] - \left[\sum_{v=1}^{V} N_v \mu_v (1 - \mu_v)\right]\left[\sum_{\hat{v}=1}^{V} N_{\hat{v}} \mu_{\hat{v}}^2\right]}{\left[\sum_{v=1}^{V} N_v \mu_v\right]^2} \tag{133}$$

Selecting the terms on top for any pair $v, \hat{v}$, we recover the equation in Lemma 3.2.

Step 3. Υ increasing in u_{-b}. Taking the derivative of Υ w.r.t. $u_{\hat{b}}$ (which corresponds to the cross derivative of log-market share)

$$\partial_{u_{\hat{b}}}\Upsilon(u_b, u_{-b}) = \frac{2\left[\sum_{v=1}^{V} N_v \left(\mu_v^b\right)^2 \mu_v^{\hat{b}}\right]\left[\sum_{\hat{v}=1}^{V} N_{\hat{v}} \mu_{\hat{v}}^b\right] - \left[\sum_{v=1}^{V} N_v \mu_v^b \mu_v^{\hat{b}}\right]\left[\sum_{\hat{v}=1}^{V} N_{\hat{v}} \left(\mu_{\hat{v}}^b\right)^2\right]}{\left[\sum_{v=1}^{V} N_v \mu_v^b\right]^2} \tag{134}$$

Selecting the terms on top for any pair $v, \hat{v}$, we recover the equation in Lemma 3.2.

Step 4. The Sufficiency of $\overline{\psi}$. We show here that it guarantees the log-supermodularity condition, but the proof is the same for the log-concavity. As utility in equilibrium is bounded below (since consumption is greater than the lower bound of the grid), whenever there is a level of utility that the bank can offer and make a positive profit:

$$\mu_v^b(\varphi_b) \geq \underline{\mu} > 0$$

On the other hand, as there is the outside option:

$$\mu_v^b(\varphi_b) \leq \overline{\mu} < 1$$

Therefore: $\mu_v \in [\underline{\mu}, \overline{\mu}]$. Note that the signal of the $\partial_{u_{\hat{b}}} \Upsilon(u_b, u_{-b})$ is the same as

$$\begin{aligned}
&\sum_v \sum_{\hat{v}} N_v N_{\hat{v}} \mu_v^b \mu_{\hat{v}}^b \mu_{\hat{v}}^{\hat{b}} \left[2\mu_{\hat{v}}^b - \mu_v^b\right] \\
&= \sum_v N_v^2 \left[\mu_v^b\right]^3 \mu_{\hat{v}}^b + \sum_v \sum_{\hat{v}} N_v N_{\hat{v}} \left\{\mu_v^b \mu_{\hat{v}}^b \mu_{\hat{v}}^{\hat{b}} \left[2\mu_{\hat{v}}^b - \mu_v^b\right] + \mu_{\hat{v}}^b \mu_v^b \mu_v^{\hat{b}} \left[2\mu_v^b - \mu_{\hat{v}}^b\right]\right\} \\
&= \sum_v N_v^2 \left[\mu_v^b\right]^3 \mu_{\hat{v}}^b + \sum_v \sum_{\hat{v}} N_v N_{\hat{v}} \mu_v^b \mu_{\hat{v}}^b \left\{\mu_{\hat{v}}^{\hat{b}} \left[2\mu_{\hat{v}}^b - \mu_v^b\right] + \mu_v^{\hat{b}} \left[2\mu_v^b - \mu_{\hat{v}}^b\right]\right\}
\end{aligned} \tag{135}$$

Assume that there is a gap of $\mu_{u_{max}} = A + \mu_{u_{min}}$ given a level of utility. In this case:

$$\begin{aligned}
\mu_{\hat{v}}^{\hat{b}} \left[2\mu_{\hat{v}}^b - \mu_v^b\right] + \mu_v^{\hat{b}} \left[2\mu_v^b - \mu_{\hat{v}}^b\right] &\geq \mu_{u_{min}} \left[2\mu_{u_{max}} - \mu_{u_{min}}\right] + \mu_{u_{max}} \left[2\mu_{u_{min}} - \mu_{u_{max}}\right] \\
&= 4\mu_{u_{min}} \mu_{u_{max}} - \mu_{u_{min}}^2 - \mu_{u_{max}}^2 \\
&= 4A\mu_{u_{min}} + 4\mu_{u_{min}}^2 - A^2 - \mu_{u_{min}}^2 + 2A\mu_{u_{min}} - \mu_{u_{min}}^2 \\
&= 6A\mu_{u_{min}} + 2\mu_{u_{min}}^2 - A^2 \\
&\geq 0 \Leftrightarrow \frac{A}{\mu_{u_{min}}} \in \left[3 - \sqrt{11}, 3 + \sqrt{11}\right] \Leftrightarrow \mu_{u_{max}} \in \left[0, \mu_{u_{min}}\left(4 + \sqrt{11}\right)\right] \\
&\Leftrightarrow \frac{\mu_{u_{max}}}{\mu_{u_{min}}} < 7.31 (\approx)
\end{aligned} \tag{136}$$

$$\mu_{u_{max}} = \frac{e^{u_b - \psi \min \|x_b - x_v\|}}{e^{u_b - \psi \min \|x_b - x_v\|} + \sum_{\beta \in B/b} e^{u_\beta - \psi \|x_\beta - x_v\|} + e^{u0}}$$

$$\mu_{u_{min}} = \frac{e^{u_b - \psi \max \|x_b - x_v\|}}{e^{u_b - \psi \max \|x_b - x_v\|} + \sum_{\beta \in B/b} e^{u_\beta - \psi \|x_\beta - x_v\|} + e^{u0}}$$

Note then that:

$$\frac{\mu_{u_{max}}}{\mu_{u_{min}}} < e^{[\psi \max_{b,v} \|x_b - x_v\| - \psi \min_{b,v} \|x_b - x_v\|]\sigma_L^{-1}}$$

A sufficient condition for log-supermodularity of the game is that:

$$\psi \leq \frac{\log(4 + \sqrt{11})\sigma_L}{[\max_{b,v} \|x_b - x_v\| - \min_{b,v} \|x_b - x_v\|]}$$

and $\log(4+\sqrt{11}) \approx 2$, which is easily verifiable. $\blacksquare$

G Proof of Lemma 4.1

Proof. The proof is comprised of two steps. The first establishes that we can simplify the set of truth telling constraints - Eq. (34) - to neighboring types only and a monotonicity condition. Second, we show that the proof of Lemma 3.1 can still be applied using the equivalence of the first step.

Neighboring Truth Telling Constraints. Before proceeding, we define some extra notation. Let π^z be the marginal distrubution of a contract on effort, z (i.e., summing over $q,z,k^{'}$). Moreover, define the following dot notation:

$$\pi(\theta)\cdot U(\theta) \equiv \sum_{c,q,z,k^{'}} \pi^{\theta}(c,q,z,k^{'})\mathbb{U}(c,z|\theta) \tag{137}$$

$$\pi(\theta)\cdot U(\theta)\cdot P(\theta,\hat{\theta}) \equiv \sum_{c,q,z,k^{'}} \pi^{\theta}(c,q,z,k^{'})\mathbb{U}(c,z|\theta)\frac{P(q|k^{'},\theta,z)}{P(q|k^{'},\hat{\theta},z)} \tag{138}$$

Finally, denote the truth telling condition between two types as

$$TT(\theta,\hat{\theta}) \equiv \sum_{c,q,z,k^{'}} \pi^{\theta}(c,q,z,k^{'})\mathbb{U}(c,z|\theta) - \sum_{c,q,z,k^{'}} \pi^{\hat{\theta}}(c,q,z,k^{'})\mathbb{U}(c,z|\theta)$$

Our claim in this step is that iff $\pi^z(\theta)$ is increasing (in a first order stochastic dominance) in z and $TT(\theta_i,\theta_{i-1}) \geq 0$, then $TT(\theta_i,\theta_{i-j}) \geq 0, \forall j$.

⇒. Let $\hat{\pi}$ be a contract worse to θ_{i-1} than $\pi(\theta)$, that is

$$\pi(\theta_i)\cdot U(\theta_i) \geq \pi(\theta_{i-1})\cdot U(\theta_i) \tag{139}$$

$$\pi(\theta_{i-1})\cdot U(\theta_{i-1}) \geq \hat{\pi}\cdot U(\theta_{i-1}) \tag{140}$$

Then, we can write

$$\begin{aligned}
\pi(\theta_i)\cdot U(\theta_i) &\geq \pi(\theta_{i-1})\cdot U(\theta_i) && (141)\\
&= \pi(\theta_{i-1})\cdot[U(\theta_i)-U(\theta_{i-1})] + \pi(\theta_{i-1})\cdot U(\theta_{i-1}) && (142)\\
&= [\pi(\theta_{i-1})-\hat{\pi}]\cdot[U(\theta_i)-U(\theta_{i-1})] + \pi(\theta_{i-1})\cdot U(\theta_{i-1}) + \hat{\pi}\cdot[U(\theta_i)-U(\theta_{i-1})] && (143)\\
&= \hat{\pi}\cdot U(\theta_i) + [\pi(\theta_{i-1})-\hat{\pi}]\cdot[U(\theta_i)-U(\theta_{i-1})] + [\pi(\theta_{i-1})-\hat{\pi}]\cdot U(\theta_{i-1}) && (144)
\end{aligned}$$

Therefore:

$$[\pi(\theta_i)-\hat{\pi}]\cdot U(\theta_i) \geq \underbrace{[\pi(\theta_{i-1})-\hat{\pi}]\cdot[U(\theta_i)-U(\theta_{i-1})]}_{\equiv I} + \underbrace{[\pi(\theta_{i-1})-\hat{\pi}]\cdot U(\theta_{i-1})}_{\equiv II} \tag{145}$$

We know that $II > 0$ since $\hat{\pi}$ is not preferred by θ_{i-1} (Eq. (140)). Moreover, we can rewrite I as

$$[\pi(\theta_{i-1})-\hat{\pi}]\cdot[U(\theta_i)-U(\theta_{i-1})] = (\theta_i-\theta_{i-1})\,[\pi^z(\theta_{i-1})-\hat{\pi}^z]\cdot v(z) \geq 0$$

where the inequality comes from $\pi^z(\theta)$ is increasing (in a first order stochastic dominance) in z. Therefore, if θ_{i-1} prefers a contract to other, so does θ_i. Thefore,

$$\begin{aligned}&\pi(\theta_i)\cdot U(\theta_i) \geq \pi(\theta_{i-1})\cdot U(\theta_i) \text{ and } \pi(\theta_{i-1})\cdot U(\theta_{i-1}) \geq \pi(\theta_{i-2})\cdot U(\theta_{i-1})\\ &\Rightarrow \pi(\theta_i)\cdot U(\theta_i) \geq \pi(\theta_{i-j})\cdot U(\theta_i),\ j>0\end{aligned} \tag{146}$$

$\Leftarrow$. Trivially, all truth telling conditions imply the neighboring ones. We focus on the monotonicty condition of π^z. Subtracting the truth telling constraints for types θ_i, θ_{i-1}, we can write

$$0 \leq [\pi(\theta_i)-\pi(\theta_{i-1})]\cdot[U(\theta_i)-U(\theta_{i-1})] = (\theta_i-\theta_{i-1})\,[\pi^z(\theta_{i-1})-\hat{\pi}^z]\cdot v(z) \tag{147}$$

Therefore, the monotonicity condition of π^z must be satisfied.

Step 3. Extension of Lemma 3.1 proof. We assume here that all constraints actually do bind. This simplifies the notation, but can easily be relaxed. [41]. Moreover, we focus on the proof assuming that S,μ are differentiable. For the technicalities if S is piece-wise linear, see the proof of Lemma 3.1.

Given that all constraints bind, choosing the utility of θ_H pins down the utility of all types through the TT. Let u_H be this utility, and define $\mathcal{U}(\theta_i\,|u_H)$ as this mapping. The FOC of a FSP is

$$\sum_i \Big\{\partial_{u_b(\theta_i)} S(u_b(\theta_i))\mu(u_b(\theta_i),u_{-b}(\theta_i)) + S(u_b(\theta_i))\partial_{u_b(\theta_i)}\mu(u_b(\theta_i),u_{-b}(\theta_i))\Big\} f(\theta_i)\partial_{u_H}\mathcal{U}(\theta_i\,|u_H) = 0 \tag{148}$$

For notation purposes, define the FOC w.r.t. $\partial_{u_b(\theta_i)}$ as $\mathcal{F}(u_b(\theta_i),u_{-b}(\theta_i))$ s.t. we can rewrite Eq.(148) as Eq. (149)

$$\sum_i \mathcal{F}(u_b(\theta_i),u_{-b}(\theta_i))\partial_{u_H}\mathcal{U}(\theta_i\,|u_H) = 0 \tag{149}$$

Note that: $\partial_{u_H}\mathcal{U}(\theta_i\,|u_H) > 1$. To see that, assume that we change the contract of type θ_{i-1} to $\hat{\pi}$ such that its utility increases by a. In Eq. (145)

$$[\pi(\theta_i)-\hat{\pi}]\cdot U(\theta_i) \geq [\pi(\theta_{i-1})-\hat{\pi}]\cdot[U(\theta_i)-U(\theta_{i-1})] + [\pi(\theta_{i-1})-\hat{\pi}]\cdot U(\theta_{i-1}) > a \tag{150}$$

where the inequality comes from Eq. (147). This is where the Step 1 is relevant. It shows that solving the problem with the TTs is equivalent to the neighboring TTs and the monotonicity condition, which has implications for how two types see new contracts. In particular, if a bad type prefers a given contract between two, so does the good type - by even more.

Since $\partial_{u_{-b}(\theta_i)}\mathcal{F}(u_b(\theta_i),u_{-b}(\theta_i)) > 0$ (given our assumptions on μ in Lemma 3.1) and the fact that $\partial_{u_H}\mathcal{U}(\theta_i\,|u_H) > 1$, we have that if competitors raise all of their offers to $\hat{u}_{-b}(\theta_i) \geq u_{-b}(\theta_i)$:

$$\sum_i \mathcal{F}(u_b(\theta_i),\hat{u}_{-b}(\theta_i))\partial_{u_H}\mathcal{U}(\theta_i\,|u_H) > 0 \tag{151}$$

thus, the FSP increases u_H. Moreover, if competitors raise all of their offers to $\hat{u}_{-b}(\theta_i) = a + u_{-b}(\theta_i)$:

$$\sum_i \mathcal{F}(u_b(\theta_i)+\hat{a}_i, a+u_{-b}(\theta_i))\partial_{u_H}\mathcal{U}(\theta_i\,|u_H+a) < 0 \tag{152}$$

[41] Without this assumption, we would have to consider all sequences of constraints the bind.

since $\hat{a}_i > a$ (given $\partial_{u_H}\mathcal{U}(\theta_i|u_H) > 1$). Thus, the best response of u_H is still between $(0,1)$, and can still apply the contraction argument of Lemma 3.1. ∎

H Distance to Nash

Here we propose the following conservative technique to order by rank all possible strategies with metric we call "distance to Nash". We use a more general notation, since this algorithm can be used in other general settings. For simplicity, we illustrate the algorithm with two players: 1 and 2. Let G be a set of strategies by both players and $P_1(G)$ and $P_2(G)$ their payoffs. In the case of adverse selection, this corresponds to Eq.(33). Let G_1 be strategies of player 2 - that is, all that is necessary for 1 to compute its best response (and equivalently G_2).

We can define and compute for any of those deviating strategies the following metrics

$$d(G,G_1) = \max(P_1(G_1) - P_1(G), 0)$$

$$d(G,G_2) = \max(P_2(G_2) - P_2(G), 0)$$

Thus, in the first step of procedure we compute $P_1(G)$ and $P_2(G)$ for a trial strategy set of G. Then, in the second stage we solve

$$\max_{G1} d(G,G_1) \text{ subject to } P_1(G) > 0, \ \forall G_1. \tag{153}$$

$$\max_{G2} d(G,G_2) \text{ subject to } P_2(G) > 0, \ \forall G_2. \tag{154}$$

Let us denote the solution of those maximization problems as $\overline{d(G,G_1)}$ and $\overline{d(G,G_2)}$. Then we compute distance to Nash as

$$d(G,G_1,G_2) = \overline{d(G,G_1)} + \overline{d(G,G_2)}$$

And in the final stage we solve

$$\min_{G} d(G), \ \forall \{G,G_1,G_2\}. \tag{155}$$

At true Nash equilibrium G_{Nash} the solution of this two-step optimization problem

$$\overline{d(G_{Nash})} = 0.$$

At all other strategies this function is strictly positive and well-defined. All possible strategies can be rank-ordered by their "distance from Nash" even if no true Nash equilibrium exists.

H.0.1 Numerical accuracy of distance to Nash algorithm

When distance to Nash is Lipschitz bounded[42] $d(G) < \lambda * P_{1,2}(G)$ we accept the outcome as an instance of Nash equilibrium. We conduct the same accuracy checks for each case of simultaneous Nash equilibrium we study. Although we don't provide proofs of existence and sufficiency conditions here, those checks serve to filter numerically well-bounded constructively obtained equilibria from outcomes where Nash equilibrium might

[42] In this case Lipschitz constant λ specifies a stopping criteria for optimization algorithm with distance to Nash $d(G)$ to act as a "measure" of Nash-closeness in the space of strategies with respect to profit level.

not exist. The Lipschitz condition λ is set at 10^{-6} value for Nash equilibrium to be considered well-resolved in our numerical examples.

I Numerical Method

I.1 Proof of Lemma 5.1

Proof. Note that

$$\prod_{p\in P}\mathbb{P}\left(\mathfrak{S}_p,\mathfrak{D}_p|\zeta\right)=\prod_{p\in P}\mathbb{P}\left(\mathfrak{S}_p,\mathfrak{D}_p|\zeta\right)\prod_{p\in\overline{P}-P}\mathbb{P}\left(\mathfrak{S}_p,\mathfrak{D}_p|\zeta\right)$$
$$=\prod_{p\in P}\mathbb{P}\left(\mathfrak{S}_p,\mathfrak{D}_p|\zeta\right)\prod_{p\in\overline{P}-P}\mathbb{P}\left(\mathfrak{S}_p|\zeta\right)\mathbb{P}\left(\mathfrak{D}_p|\mathfrak{S}_p,\zeta\right) \tag{156}$$
$$\propto\prod_{p\in P}\mathbb{P}\left(\mathfrak{S}_p,\mathfrak{D}_p|\zeta\right)\prod_{p\in\overline{P}-P}\mathbb{P}\left(\mathfrak{S}_p|\zeta\right)\prod_{p\in P}\mathbb{P}\left(\mathfrak{D}_p|\mathfrak{S}_p,\zeta\right)\prod_{p\in\overline{P}}\mathbb{P}\left(\mathfrak{S}_p|\zeta\right) \tag{157}$$

Taking logs and re-arranging delivers the expected result.

■

I.2 Estimator for (γ_{ME})

The partial derivative of the likelihood of demand given supply in Eq. (59) to γ_{ME} is given by

$$\begin{aligned}\frac{\partial\ln\{\mathbb{P}\left(\mathfrak{D}|\mathfrak{S},\zeta\right)\}}{\partial\gamma_{ME}}&=-L\sum_{p}\sum_{v^p}\sum_{j^p}\mathbb{1}_{\hat{v}_j^p=v^p}\frac{\sum_{r=1}^{\#Y}f_v(c,q,k|\zeta)\gamma_{ME}^{-L-1}\exp\left\{\sum_{l=1}^{L}-\frac{\left(\hat{y}_j^l-y_r^l\right)^2}{2\chi_l^2\gamma_{ME}^2}\right\}}{F(\hat{y}_j,\zeta)}\\&+\sum_{p}\sum_{v^p}\sum_{j^p}\mathbb{1}_{\hat{v}_j^p=v^p}\frac{\sum_{r=1}^{\#Y}f_v(c,q,k|\zeta)\gamma_{ME}^{-L}\exp\left\{\sum_{l=1}^{L}-\frac{\left(\hat{y}_j^l-y_r^l\right)^2}{2\chi_l^2\gamma_{ME}^2}\right\}\left[\sum_{l=1}^{L}\frac{\left(\hat{y}_j^l-y_r^l\right)^2}{\chi_l^2\gamma_{ME}^3}\right]}{F(\hat{y}_j,\zeta)}\\&=\gamma_{ME}^{-1}\sum_{p}\sum_{v^p}\sum_{j^p}\mathbb{1}_{\hat{v}_j^p=v^p}\mathbb{1}_{j\in v}\frac{\sum_{r=1}^{\#Y}f_v(c,q,k|\zeta)\exp\left\{\sum_{l=1}^{L}-\frac{\left(\hat{y}_j^l-y_r^l\right)^2}{2\chi_l^2\gamma_{ME}^2}\right\}\left[\sum_{l=1}^{L}\frac{\left(\hat{y}_j^l-y_r^l\right)^2}{\chi_l^2\gamma_{ME}^2}-L\right]}{\sum_{r=1}^{\#Y}f_v(c,q,k|\zeta)\exp\left\{\sum_{l=1}^{L}-\frac{\left(\hat{y}_j^l-y_r^l\right)^2}{2\chi_l^2\gamma_{ME}^2}\right\}}\end{aligned} \tag{158}$$

From Eq.(158), we have that

$$\lim_{\gamma_{ME}\to 0}\frac{\partial\ln\{\mathbb{P}\left(\mathfrak{D}|\mathfrak{S},\zeta\right)\}}{\partial\gamma_{ME}}>0\quad\text{and}\quad\lim_{\gamma_{ME}\to 1}\frac{\partial\ln\{\mathbb{P}\left(\mathfrak{D}|\mathfrak{S},\zeta\right)\}}{\partial\gamma_{ME}}<0 \tag{159}$$

In the optimum, $\hat{\gamma}_{ME}$:

$$\hat{\gamma}_{ME}^2=L^{-1}\sum_{p}\sum_{v^p}\sum_{j^p}\mathbb{1}_{\hat{v}_j^p=v^p}\frac{\sum_{r=1}^{\#Y}f_v(c,q,k|\zeta)\exp\left\{\sum_{l=1}^{L}-\frac{\left(\hat{y}_j^l-y_r^l\right)^2}{2\chi_l^2\hat{\gamma}_{ME}^2}\right\}\left[\sum_{l=1}^{L}\frac{\left(\hat{y}_j^l-y_r^l\right)^2}{\chi_l^2}\right]}{\sum_{r=1}^{\#Y}f_v(c,q,k|\zeta)\exp\left\{\sum_{l=1}^{L}-\frac{\left(\hat{y}_j^l-y_r^l\right)^2}{2\chi_l^2\hat{\gamma}_{ME}^2}\right\}} \tag{160}$$

We can re-write it as

$$\sum_{p}\sum_{v^p}\sum_{j^p}\mathbb{1}_{\hat{v}_j^p=v^p}\frac{\sum_{r=1}^{\#Y}f_v(c,q,k|\zeta)\exp\left\{\sum_{l=1}^{L}-\frac{\left(\hat{y}_j^l-y_r^l\right)^2}{2\chi_l^2\hat{\gamma}_{ME}^2}\right\}\left[\sum_{l=1}^{L}\frac{\left(\hat{y}_j^l-y_r^l\right)^2}{\chi_l^2\hat{\gamma}_{ME}^2}\right]}{\sum_{r=1}^{\#Y}f_v(c,q,k|\zeta)\exp\left\{\sum_{l=1}^{L}-\frac{\left(\hat{y}_j^l-y_r^l\right)^2}{2\chi_l^2\hat{\gamma}_{ME}^2}\right\}}=1+L \tag{161}$$

the LHS of Eq.(161) is constant and the RHS is a weighted average. When γ_{ME} increases, we increase the relative weight of high $\left(\hat{y}_j^l-y_r^l\right)^2$ terms and decrease all terms, therefore, it is not trivial to state if the LHS is decreasing or increasing. Therefore, there is no general proof that the function is concave [43], but we know from Eq. (159) that a zero partial derivative is a necessary condition, which translate to $\hat{\gamma}_{ME}$ satisfying Eq. (160).

I.3 Estimators for c_E, s

As $\{c_E, s\}$ are the mean and std. dev. of the normal distributions, $\ln\{\mathbb{P}(\mathfrak{S}|\zeta)\}$ in Eq. (52) is differentiable in $\{c_E, s\}$. As we know what is the analytical derivative of it, it is straightforward to compute an optimal value for $\{c_E, s\}$ using a grid search method, which is computationally fast. In particular, analogous to what did in the Section I.2, $\{\hat{c}_E, \hat{s}\}$ are the solution to the non-linear system in Eq. (162)-(163) (as we show later on).

$$\sum_{m^p}\frac{\phi\left[\frac{\Pi^E(B_{m^p}+1|.)}{\hat{s}}\right]-\phi\left[\frac{\Pi^E(B_{m^p}|.)}{\hat{s}}\right]}{\Phi\left[\frac{\Pi^E(B_{m^p}|.)}{\hat{s}}\right]-\Phi\left[\frac{\Pi^E(B_{m^p}+1|.)}{\hat{s}}\right]}=0 \tag{162}$$

$$\sum_{m^p}\frac{\phi\left[\frac{\Pi^E(B_{m^p}+1|.)}{\hat{s}}\right]\Pi^E\left(B_{m^p}+1|.\right)-\phi\left[\frac{\Pi^E(B_{m^p}|.)}{\hat{s}}\right]\Pi^E\left(B_{m^p}|.\right)}{\Phi\left[\frac{\Pi^E(B_{m^p}|.)}{\hat{s}}\right]-\Phi\left[\frac{\Pi^E(B_{m^p}+1|.)}{\hat{s}}\right]}=0 \tag{163}$$

where the dependence of the system of c_E comes implicitly from its effect on profits, i.e.: profit $=$ revenue$-c_E$ in Eq. (50). We now move to show that Eq. (162)-(163) determine the optimal value of $\{\hat{c}_E, \hat{s}\}$.

Fixed Cost. the partial derivative of the supply likelihood (Eq. 52) w.r.t. to the fixed cost c_E is given by Eq. (164).

$$\frac{\partial\ln\{\mathbb{P}(\mathfrak{S}|\zeta)\}}{\partial c_E}=s^{-1}\sum_{m^p}\frac{\phi\left[\frac{\Pi^E(B_{m^p}+1|.)}{s}\right]-\phi\left[\frac{\Pi^E(B_{m^p}|.)}{s}\right]}{\Phi\left[\frac{\Pi^E(B_{m^p}|.)}{s}\right]-\Phi\left[\frac{\Pi^E(B_{m^p}+1|.)}{s}\right]} \tag{164}$$

Note that at $c_E \leq 0$, as we have that the profit Π^E is non-increasing in the number of intermediaries in a location m^p and $\Pi^E \geq 0$ (since there is no cost and the state space is limited to points with points with

[43] In all numerical runs, the likelihood was concave in γ_{ME}, although, as shown above, it is not trivial to guarantee this analytically

positive frontier without loss of generality), we have that:

$$\phi\left[\frac{\Pi^{E}\left(B_{m^{p}}|.\right)}{s}\right]\leq\phi\left[\frac{\Pi^{E}\left(B_{m^{p}}+1|.\right)}{s}\right] \tag{165}$$

and:

$$\phi\left[\frac{\Pi^{E}\left(1|.\right)}{s}\right]>0 \tag{166}$$

Together, (165)-(166) imply that (164) is positive at $c_E \leq 0$, i.e.:

$$\left.\frac{\partial \ln\{\mathbb{P}\left(\mathfrak{S}|\zeta\right)\}}{\partial c_{E}}\right|_{c_{E}\leq 0}>0 \tag{167}$$

Moreover, at $c_E \to \infty$:

$$\begin{aligned}
\sum_{m^{p}}\mathbb{1}_{m^{p}>0}\lim_{c_{E}\to\infty}\frac{\phi\left[\frac{\Pi^{E}(B_{m^{p}}+1|.)}{s}\right]-\phi\left[\frac{\Pi^{E}(B_{m^{p}}|.)}{s}\right]}{\Phi\left[\frac{\Pi^{E}(B_{m^{p}}|.)}{s}\right]-\Phi\left[\frac{\Pi^{E}(B_{m^{p}}+1|.)}{s}\right]} &\underset{L'Hospital}{=}\sum_{m^{p}}\mathbb{1}_{m^{p}>0}\lim_{c_{E}\to\infty}\frac{\phi'\left[\frac{\Pi^{E}(B_{m^{p}}|.)}{s}\right]-\phi'\left[\frac{\Pi^{E}(B_{m^{p}}|.)}{s}\right]}{\phi\left[\frac{\Pi^{E}(B_{m^{p}}|.)}{s}\right]-\phi\left[\frac{\Pi^{E}(B_{m^{p}}|.)}{s}\right]}\\
&=\sum_{m^{p}}\mathbb{1}_{m^{p}>0}\lim_{c_{E}\to\infty}\frac{\phi'\left[\frac{\Pi^{E}(B_{m^{p}}|.)}{s}\right]-\phi'\left[\frac{\Pi^{E}(B_{m^{p}}+1|.)}{s}\right]}{\phi\left[\frac{\Pi^{E}(B_{m^{p}}+1|.)}{s}\right]-\phi\left[\frac{\Pi^{E}(B_{m^{p}}|.)}{s}\right]}\\
&=\sum_{m^{p}}\mathbb{1}_{m^{p}>0}s^{-1}\lim_{c_{E}\to\infty}\frac{\Pi^{E}\left(B_{m^{p}}+1|.\right)\phi\left[\frac{\Pi^{BR}(N_{m}+1|.)}{s}\right]-\Pi^{E}\left(B_{m^{p}}|.\right)\phi\left[\frac{\Pi^{E}(B_{m^{p}}|.)}{s}\right]}{\phi\left[\frac{\Pi^{E}(B_{m^{p}}+1|.)}{s}\right]-\phi\left[\frac{\Pi^{E}(B_{m^{p}}|.)}{s}\right]}\\
&=-\infty
\end{aligned} \tag{168}$$

and:

$$\lim_{c_{E}\to\infty}\frac{\phi\left[\frac{\Pi^{E}(1|.)}{s}\right]}{1-\Phi\left[\frac{\Pi^{E}(1|.)}{s}\right]}=0 \tag{169}$$

Therefore, it must be the case that:

$$\partial_{c_{E}}\left.\frac{\partial \ln\{\mathbb{P}\left(\mathfrak{S}|\zeta\right)\}}{\partial c_{E}}\right|_{\hat{c}_{E}}=0 \tag{170}$$

where $\hat{c}_E$ is the argmax of the likelihood given the data (i.e., the estimator). The idea here is that as the function is differentiable, increasing at zero and decreasing at $c_E \to \infty$, there is a global max and it must satisfy the necessary condition in Eq. (170).

Standard Deviation s. the partial derivative of the supply likelihood (Eq. 52) w.r.t. to the variance of location specific profit shocks s is given by Eq. (171).

$$\frac{\partial \ln\{\mathbb{P}(\mathfrak{S}|\zeta)\}}{\partial s} = s^{-2}\sum_{m^p} \mathbb{1}_{m^p>0} \frac{\phi\left[\frac{\Pi^E(B_{m^p}+1|.)}{s}\right]\Pi^E\left(B_{m^p}+1|.\right) - \phi\left[\frac{\Pi^E(B_{m^p}|.)}{s}\right]\Pi^E\left(B_{m^p}|.\right)}{\Phi\left[\frac{\Pi^E(B_{m^p}|.)}{s}\right] - \Phi\left[\frac{\Pi^E(B_{m^p}+1|.)}{s}\right]}$$
$$+ s^{-2}\sum_{m^p} \mathbb{1}_{m^p=0} \frac{\phi\left[\frac{\Pi^E(1|.)}{s}\right]\Pi^{BR}(1|.)}{1-\Phi\left[\frac{\Pi^E(1|.)}{s}\right]} \tag{171}$$

Using the same arguments as in Eq. (168), one can show that

$$\lim_{s\to 0}\frac{\partial \ln\{\mathbb{P}(\mathfrak{S}|\zeta)\}}{\partial s} > 0 \text{ and } \lim_{s\to\infty}\frac{\partial \ln\{\mathbb{P}(\mathfrak{S}|\zeta)\}}{\partial s} < 0 \tag{172}$$

Therefore, it must be the case that:

$$\left.\frac{\partial \ln\{\mathbb{P}(\mathfrak{S}|\zeta)\}}{\partial s}\right|_{\hat{s}} = 0 \tag{173}$$

where $\hat{s}$ is the argmax of the likelihood given the data (i.e., the estimator). The idea here is that as the function is differentiable, increasing at zero and decreasing at $s \to \infty$, there is a global max and it must satisfy the necessary condition in Eq. (173).

I.4 Details on Numerical Maximization

First, we discuss the pseudo-code we use for numerical maximization. We discuss first how we compute the likelihood for fixed values of $\{\psi, \sigma_L\}$.

One likelihood computation. Given $\{\psi, \sigma_L\}$.

1. As a function of $\{\sigma, \theta\}$ only, compute the frontier of Section 3.1, S .
   - Don't have to redo this step if we are calibrating $\{\sigma, \theta\}$ instead of estimating it (as we are here).
   - Use the LP formulation with the *gurobi* [44] linear solver.
2. Given Step 1 (the frontier), compute the equilibrium utilities and the resulting contracts.
   - uses iterative procedure based on supermodularity of Lemma 3.1, Section .
3. Likelihood:
   - Demand Given Supply. Given Steps 1 and 2, compute the likelihood of the demand using the adaptation of Karaivanov and Townsend (2014) method presented (Eq. 59).
     - Use a grid search to find $\hat{\gamma}_{ME}$ that satisfies Eq. (160) and compute Eq. (59) already at the optimum $\hat{\gamma}_{ME}$.
   - Supply. Given Steps 1 and 2 - i.e., it can be compute parallelly to the demand - , compute the likelihood of the supply using the entry model and the normality assumption (Eq. 52) .
     - Use a grid search to find $\{\hat{c}_E, \hat{s}\}$ that satisfie Eqs.(170) -(173) and compute Eq. (52) already at the optimum $\{\hat{c}_E, \hat{s}\}$.

[44] Available for free for academic use at:

- Sum Demand Given Supply and Supply. as in Eq.()62).

Global Optimum. We optimize over ψ, σ_L by first doing a grid search and then use the *patternsearch* command in Matlab from the optimal point in the grid search. We guess $\psi = 1$, $\sigma_L = .33$ and use 25 point grids between .1 and 5 times the original values for both of them.

J Identification

J.1 Numerical Identification

Although we have given indications in Section 3.3 that we can identify the parameters from the micro data, we now show that it is the case numerically. For that, we conduct a Monte-Carlo experiment. We generate model simulated data and use it to estimate the parameters in question. We use only data on consumption, production and capital in this exercise. As we simply assume that the observed data corresponds to an equilibrium in terms of bank entry (as in Bresnahan and Reiss (1991)), our methodology makes it much harder to discuss identification from it. [45]. As we are ultimately interested in estimating the spatial costs ψ and logit variance, σ_L, we focus on this section on the maximization of Eq. (59) on these two parameters, but present results for risk aversion σ and ost of exerting effort φ.

To be closer to our actual application, we use a spatial configuration here with two dimensions, in a 'Manhattan' style as in in Figure 28. We assume each intersection has a random number of FSPs (which can be zero) and villages are uniformly distributed within roads. [46]

We simulate the data for four provinces, with 10 villages in each road, each of them with $N = 75$ households each, with the same parameters of the comparative statics exercises (Section 3.3). [47]. In particular, a pseudo-code for this identification experiments is as follows:

Pseudo-Code for identification:

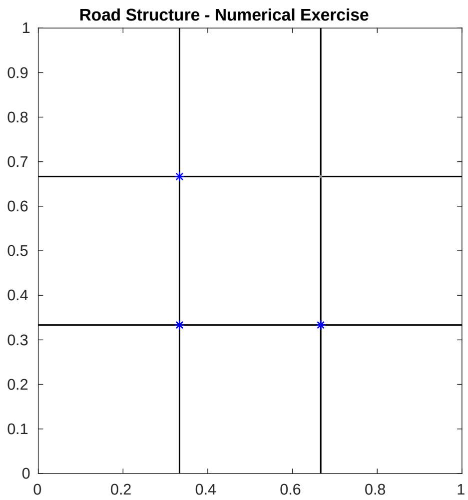
1. Draw a number of banks for each position as in Figure (28)).
2. Given the map configuration of Step 1, solve for the equilibrium in utilities and compute the optimal contracts.
3. Produce probability distributions of (c, q, k) assuming that the distribution of k in the simulated data is the same as in the sample.
4. Draw N observations from Step 3 and then add the measurement errors to (c, q, k) obtained (belonging to the grid) to generate the dataset.
5. Fix all of the other parameters and vary the parameters we are plotting.

[45] We are assuming that we do not know the entry/exit process and what is the dynamics of it, such that it is very challenging to simulate an equilibrium in the position of each bank.

[46] We have the equivalent results using the actual spatial configuration in the Townsend Thai Data upon requrest. We prefer to showcase this version since we can understand better the dynamics of competition and contracting.

[47] Frontier parameters given by Table 4 and the true value of ψ and σ_L as in Table 7

Figure 28: Simplified Map 'Manhattan' Style



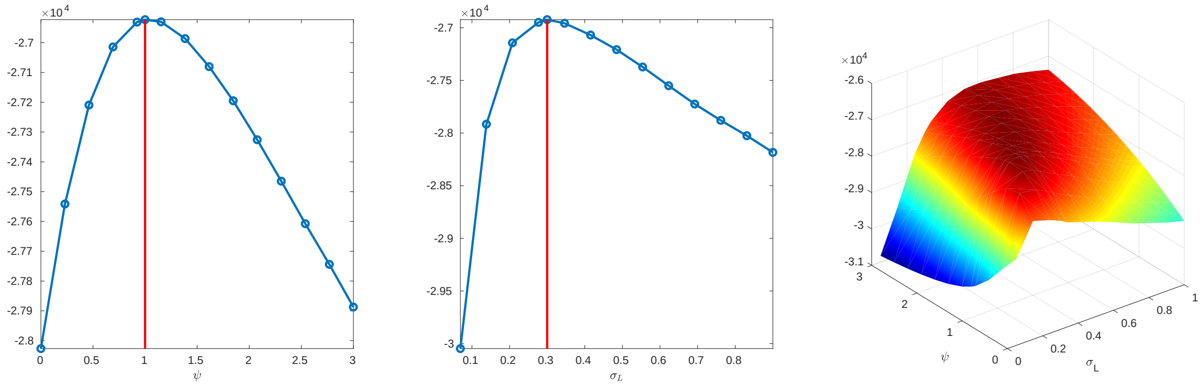
A circular dot in one of the intersection means that there is one bank there and a star means that there are two. Villages are uniformly distributed throughout the roads.

Results are in Figure 29, where we can see our method is successful in identifying the market structure parameters: spatial cost ψ and logit variance σ_L. We first vary one parameter at a time (panels (a) and (b)), and both parameters jointly. The overall levels of utility, which imply consumption, capital and income dynamics, identify σ_L, while the variation between these levels across villages identifies ψ, as seen in Section 3.3. Note that although ψ and σ_L are jointly identified, the higher differences from the likelihood appear for different ratios of ψ/σ_L, which is the overall spatial cost in terms of utilities.

We repeat the experiment above for risk aversion σ and disutility of effort φ. The results are in Figure 30 Again, our results show that the model is successful in identfying this parameters. The identify them, to model is using the joint variation in consumption/income and the implicit distribution of all variables given the utilities and effort levels implied by competition.

Figure 29: Log-Likelihood of Household Level Data as a Function of Spatial Cost (ψ) and Logit Var (σ_L) for Simulated data

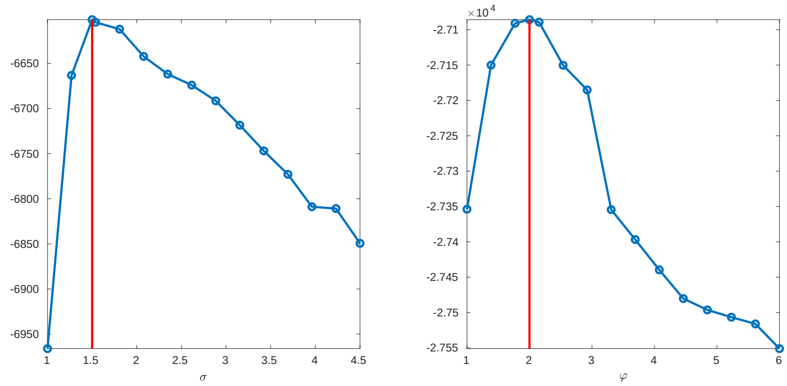
(a) Spatial Cost, ψ (b) Logit Var, σ_L (c) Joint



Note: Likelihood of household level data (Eq. 59) using . Red line is the true value, dotted blue line is log-likelihood. Map in here is 'Manhattan' style (Figure 28). Data simulated for four provinces, with 10 villages in each road, each of them with $N = 75$ households. Frontier parameters given by Table 4 and the true value of ψ and σ_L as in Table 7.

Figure 30: Log-Likelihood of Household Level Data as a Function of Risk Aversion (σ) and Disutility of Effort (φ) for Simulated data

(a) Risk Aversion (σ) (b) Disutility of Effort (φ)



Note: Likelihood of household level data (Eq. 59) using . Red line is the true value, dotted blue line is log-likelihood. Map in here is 'Manhattan' style (Figure 28). Data simulated for four provinces, with 10 villages in each road, each of them with $N = 75$ households. Frontier parameters given by Table 4 and the true value of ψ and σ_L as in Table 7.

K From Utilities to Consumption

Consumption and Distance. With the utility given by Eq.(16), we want to solve for Δ in Eq.(174). The value Δ is the % growth in consumption that corresponds to moving from zero to the median distance of intermediaries and villages in the sample. In this case:

$$u((1+\Delta)c) - z^{\varphi} - \psi \,\text{med}\,(t(x_v, x_b)) = u(c) - -z^{\varphi} \tag{174}$$

Eq.(174) implies Eq.(175)

$$\left\{(1+\Delta)^{1-\sigma} - 1\right\} u(c) = \psi \,\text{med}\,(t(x_v, x_b)) \tag{175}$$

and Eq.(175) yields:

$$\Delta = \left[\psi \,\frac{\text{med}\,(t(x_v, x_b))}{u(c)} + 1\right]^{\frac{1}{1-\sigma}} - 1 \tag{176}$$

which delivers Eq. (68).

Utilities and Consumption. With the utility given by Eq.(16), we now want to solve for Δ in Eq.(176). The value $\tilde{\Delta}$ is the % growth in consumption that corresponds to moving Δu. In this case:

$$\Delta u = u((1+\tilde{\Delta})c) - z^{\varphi} - \psi t(x_v, x_b) - [u(c) - z^{\varphi} - \psi t(x_v, x_b)] = (1-\sigma)^{-1}\left[(1+\tilde{\Delta})^{1-\sigma} - 1\right] c^{1-\sigma} \tag{177}$$

For $\sigma = 2$

$$\tilde{\Delta} = \left[\frac{\Delta u}{u(c)} + 1\right]^{\frac{1}{1-\sigma}} - 1 \tag{178}$$