Optimal Contracting, Spatial Competition among Financial Service Providers, and the Impact of Digital Lending

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

September 24, 2025

Abstract

We present a contract-based model of industrial organization for financial markets characterized by both informational and other frictions (Moral Hazard, Adverse Selection, Limited Commitment, etc.) and different market structures (Monopoly, Oligopoly, Competition). We show that considering both contracting frictions and market power simultaneously is qualitatively and quantitatively relevant. We apply our model to the entry of new providers, including both legacy and FinTech firms, and to the rise of new technologies that can mitigate access costs and contractual obstacles to trade, such as digital lending. Our model is micro-founded and thus suitable for a broad set of counterfactuals. We derive a likelihood estimator for the structural parameters determining both contracting frictions and market structure and apply it to the Townsend Thai data on small and medium enterprises and GIS data on bank locations. Finally, we assess which remedies for enhancing competition are most effective for improving welfare.

[*]Federal Reserve Bank of Boston, e-mail: gustavo.joaquim@bos.frb.org

[†]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, Yale, ASSA and SAET conferences for very useful comments. We gratefully acknowledge research support from the Thailand Research Fund, the Bank of Thailand, Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD) (grant number R01 HD027638), the John Templeton Foundation (grant number 12470), the Consortium on Financial Systems and Poverty at the University of Chicago (funded by Bill & Melinda Gates Foundation under grant number 51935), the Centre for Economic Policy Research (CEPR) and the Department for International Development (DFID) under grant MRG002_1255, and the BFI through the Macro Financial Modeling Fellowship. The views expressed herein are those of the author and do not indicate concurrence by the Federal Reserve Bank of Boston, the principals of the Board of Governors, or the Federal Reserve System.

1 Introduction

1.1 Motivation: The Phenomena

Innovation in financial intermediation is evident in a myriad of settings. Typically, market power is a consideration. In China, one can study the geography of branch deregulation. Changed regulations allowed joint-equity banks to enter new markets, with local impact varying with a measure of incumbent bank concentration (Gao et al., 2019). In the US, branches were closed with the merger of large national banks, with an impact that varied by product and client, i.e., low impact for mortgage origination, likely due to alternative internet service providers, but higher for small and medium enterprises, likely due to relationship lending (Nguyen, 2018). Related, innovation beyond branches is increasingly relevant. Recent financial technology (FinTech) sector advancements have allowed peer-to-peer borrowing/ lending e-platforms to enter the traditional financial intermediation space rapidly. According to Wolfe and Yoo (2018), a substantial fraction (26.7%) of peer-to-peer loan volume in the US substitutes for small commercial bank personal loan volume. Finally, FinTechs in China, Argentina, Indonesia, and other countries innovate, at first with e-commerce and electronic payments services, but then with credit and insurance products, using the acquired payments data to mitigate traditional obstacles to trade. Thus, digital financial service providers, FSPs, have the potential to increase financial access by making the geography of branches less relevant and lowering information obstacles in lending – moral hazard within contracts and adverse selection across contacts. There are regulatory concerns, however, as FinTechs and FSPs disrupt existing paradigms. Their emergence may come with market concentration and a lesser role for commercial banks if the latter do not adjust. Finally, the transmission of signals through cell towers remains imperfect, measured in Ridhwan et al. (2024). The closure of legacy bank branches can still create service voids when internet services are not available or certain customers require physical access (Bennett, 2020).

1.2 Brief Summary of Goals, Model Experiments, and Empirical Application

Yet, we have relatively few frameworks for the analysis of these empirical phenomena. This paper develops a framework to solve, estimate, and do policy analysis in models that deal simultaneously with imperfect spatial competition among providers and contracting with underlying obstacles to trade. We apply this framework to example settings to address conceptual issues.

To summarize briefly, in one setting we experiment with the entry of more banks and also in the same setting with better information technologies, thinking of digital data. The impact depends on the interaction between spatial competition and contract aspects. The welfare of households running small and medium enterprises (SMEs) can be reduced under a new contract technology if there is little competition. In another setting, we consider competition between a legacy national-level bank with many branches, offering high financial access yet relatively uninformed about clients, versus a local service lender with few branches but with better information on clients in the area of service. In this setting, a digital lender with both an information and financial access advantage can become quite large, though dealing with all risk types. However, we show via an example that this can hinge on the ability of such a digital lender to lower both frictions. Improved information while retaining access frictions can create selection into safer types, so that risky types move to the national bank, hence a regulatory concern. More generally, the impact on the ground is heterogeneous, depending on the distance of a household/ SME to a bank/ provider and the degree of competition in bank clusters.

Still summarizing briefly, we take the framework to an actual application for which we have data, a legacy setting with SMEs, measured in survey data and physical bank branch location data from geospatial maps. We estimate the parameters of the model and then consider various counterfactuals in both legacy and digital lending: more entry of legacy banks to enhance spatial competition, lower artificial product differentiation (through platform competition), and lower spatial access costs. Again, this is done at empirically estimated obstacles to trade and parameter values. We thus can answer the quantitative question of which counterfactual change would have the highest impact on welfare.

1.3 Overview of the Model

More generally, 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, limited commitment) and a variety of market structures (monopoly, oligopoly where the number of providers is varied and potentially endogenous). The model has implications for profits and market shares of financial service providers as contract providers, for FSP portfolio risk, and for the distribution of consumption, income, and capital of agents in the economy. In this context, we consider the impact of new technologies.

Specifically, we analyze contracting between a group of entrepreneurs and a set of financial intermediaries as service providers for credit and insurance. Entrepreneurs in the model are risk-averse households running small and medium enterprises in need of external credit and insurance. Financial intermediaries are banks, financial services providers, FinTechs, or platforms that provide contracts to customers and compete with each other. The financial/ information regimes we consider: full information (complete insurance/perfect credit), unobserved effort (moral hazard) but with complete information on types, unobserved types thus creating adverse selection, and limited enforcement of repayment of loans (limited commitment). The market structure we consider is based on a demand system where SMEs have idiosyncratic preferences for intermediaries that generate logit market shares. Traditionally, financial intermediaries and SMEs are spatially separated, and SMEs must pay a travel cost to go to a branch to obtain a contract from an intermediary. We run experiments motivated by digitization where this cost can be set to zero. A given intermediary in equilibrium typically attracts a set of SMEs and can offer contracts that provide funding and pool risks among these clients, subject to the information and commitment restrictions.

1.4 The Model Experiments in More Detail

In the first of a set of conceptual experiments, we imagine we are able to increase the financial service providers into areas that had none (hence new financial access) and into areas where FSPs were already present (hence more competition). In a second experiment, we imagine digital data as a better technology that gives all FSPs better information on otherwise unobserved types of potential borrower clients.

In the first experiment using empirical difference-in-difference (DID) methods, there is an expenditure puzzle –consumption, investment, and other expenditures drop when there is new financial access but not when competition among providers is enhanced. The puzzle is resolved in a model that delineates how contracts change with new financial access, from autarky to a new contract form, and this change with unobserved heterogeneity makes DID inappropriate. We show how to use the structural model to quantify large welfare gains. In contrast, enhanced competition leaves the form of contracts intact but lowers market power, the premium paid by clients for financial services, thus enhancing welfare in a different way.

The order of magnitude of impact from entry experiments depends on the underlying financial/ information regime, that is, the nature and extent of obstacles to trade, and on the degree of local competition, number of FSPs, all which we can vary. We can also vary the extent of monopolistic competition through intermediation platforms, that is, lowering the logit variance, both on its own and in conjunction with lower spatial access costs that come with digitization. We find these are complements. Thus, we want both lower access costs (through digitization) and lower barriers that create artificial walled gardens (otherwise high logit variance).

In the second experiment above, the elimination of adverse selection through better, digitized data can lower client welfare relative to a comparison group with no new FSP yet enhance client welfare if there are FSPs in the comparison group. The puzzle is resolved with the recognition that better information allows imperfectly competitive FSPs to extract rent, but with sufficient competition, and the possibility of switching to another provider, the equilibrium elasticity of demand is high, thus allowing clients to benefit from the improved information technology.

In another set of conceptual experiments, financial service providers are heterogeneous in outreach and technologies. There is a large national-level bank using a legacy non-digital system, thus having relatively little information on clients but offering easier access due to its dense branch network (captured in the model, as an approximation, by zero spatial access costs). There is also a geographically service provider (e.g., cooperative) with better information on clients but with nontrivial spatial costs (access to it is limited, hence, its market is local). In an equilibrium, we can consider which institution is attracting low and high-risk clients, their market shares, and their profits. In one experiment, we set spatial access costs of the local informed cooperative to zero as well, and then it turns into a big-tech digital lender. For given parameter values, at zero spatial costs, the local informed cooperative has an advantage in profits and has higher market shares of both risky and safe types. In another experiment, which can differentiate the impact of digitization on access versus the impact of better information, we impose identical spatial costs on both types of institutions, legacy and local, leveling the playing field on that dimension of technology while retaining information differences, that is, with heterogeneity in information technologies, legacy vs digital. The better-informed digital provider retains a profit advantage throughout the entire range of spatial costs, but as the better-informed provider can distinguish high-risk from low-risk clients, it drops the former relatively quickly at higher spatial access costs. The relatively uninformed national bank thus attracts in equilibrium relative more of the high risk types, which would be a concern to regulators. But again, BigTechs with both low access costs and better information do not necessarily generate this problem, as the numerical example demonstrates.

1.5 Explicit Empirical Applications

Finally, we show how to apply our general method not just to model experiments but to real data. We break this into two parts. In one part, the data concerns the market shares of FSP. In the second part, the data concerns the geo-locations of banks and villages, hence road travel times, a legacy context.

For the second part, we go one step further and take the model to actual data, with bank locations and Townsend Thai micro-level panel data on the SMEs - consumption, income, and capital. We estimate the structural parameters of the model using maximum likelihood methods. This is the baseline scenario. Then, in a set of counterfactuals, we compute the effects of three innovations on the supply side of financial services: an increase in the number of branches of FSPs, a reduction in spatial access costs as with digitization, and lower variance of idiosyncratic preference shocks as with platforms with standardization. We show that depending on the baseline competition level, each policy has a different effect on SMEs.

Overall, to increase welfare, policymakers need to guarantee that market shares change when product offerings change. This means policymakers should pursue policies that make SMEs more likely to choose better financial products, such as through financial literacy, or insist on platforms where financial services can be easily compared and at lower access costs. At estimated parameters, simply increasing the number of FSPs without changes in technology has a positive but relatively low impact. Digitization that lowers spatial access costs has a greater impact. This is surpassed at estimated parameters by the reduction of artificial barriers that enhance monopolistic competition.

1.6 Key Contract-Theoretic Approach and the Building Blocks

Our theoretical framework focuses on utilities generated by contracts. This draws on a theory literature that uses promised utilities as a key state variable (Green, 1987; Spear and Srivastava, 1987). Though client outcomes are specified as utilities, much of the usual toolbox of competition remains at our disposal. The overall framework is then divided into two building blocks: one, the utilities and profits frontier, and two, the market structure. The frontier, as in Karaivanov and Townsend (2014), represents the profits of a contract service provider for a given level of utility of clients. 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; it also makes economic sense: changes in contracting frictions only alter the initial frontier block, and all strategic interactions are contained in the market structure block.

1.7 Estimation Method

An equilibrium among financial service providers makes the distribution of promised utilities faced by SMEs endogenous, along with profits and market shares for the FSPs. The structure can thus be estimated using full information maximum likelihood techniques, comparing both information regimes and market structures. Using Bank of Thailand data on FSP locations and Townsend Thai project data on SME consumption, income, and capital, we develop a numerical method based on a synthesis of Bresnahan and Reiss (1991) and Karaivanov and Townsend (2014). We use the model of entry from Bresnahan and Reiss (1991) with information on the number of banks in each location. We extend the methodology of Karaivanov and Townsend (2014), which maps unobserved equilibrium utilities of the model to equilibrium contracts by assuming that financial service providers provide optimal contracts on an efficiency frontier (subject to the frictions). These imply a joint distribution for consumption, capital, and income. Thus, we derive a likelihood function that maps the model-generated data to observed data.

1.8 Related Literature

Our work is related to three strands of research in economics. First, we contribute to the literature analyzing markets with both market power in intermediation and asymmetric information. Lester et al. (2018) incorporate a search model of imperfect competition into an adverse selection model with specific structural specifications for each. Mahoney and Weyl (2014) use a graphical price theoretic model that parameterizes both the degree of market power and selection, using areas under curves as a measure of welfare. In both these papers, more competition or reduced information asymmetry *can* increase or be detrimental to welfare. From a more empirical perspective, Crawford et al. (2018) shows how to estimate a standard industrial organization model with these two ingredients to quantify the role of information asymmetries, captured by the error structure in a discrete choice model. The advantage of such empirical models is that they can flexibly identify the degree of moral hazard and adverse selection from contract data. The approach is data-based and can be used for counterfactuals changing the degree of contracting frictions. Our framework is model-based and attempts to be general on multiple dimensions, granted with some limitations. We can deal with situations where a counterfactual affects at the same time both contracting and the degree of importance of each contracting friction *endogenously*. In sum, we expand on this literature by bridging the gap between the theoretical approach, based on micro-foundations of contracting, and the more reduced form approach of Crawford et al. (2018).

We also contribute to the broad literature on contracting frictions in a variety of fields. Though our paper is specifically aimed at developing methods that could be applied to financial institutions, we believe these same methods could be applied to other markets. 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, and Einav et al. (2013) on behavior responses under exogenously introduced variation in medical contracts.

We emphasize that our framework is quite operational in that it can include any obstacle to trade, with contracts endogenous to those frictions, and the framework can be taken to the data in several different forms. Though we display calibrated and numerical examples, as in the literature, Azevedo and Gottlieb (2018) and others, we also move beyond those to the estimation of key parameters as with fully specified structural models. We illustrate this in the Townsend Thai Data using a numerically efficient method. At the same time, we clarify what is needed, that depending on the policy counterfactual of interest, one does not necessarily have to deal with both blocks of the model, the contracting block and the competition block. This is useful, of course, in framing policy guidance coming from empirical applications.

Third, we contribute to the literature on the effects of credit policies and reforms in the credit market from a macro perspective. The macro-development literature (e.g., Buera and Shin (2013), Itskhoki and Moll (2019)) focuses largely on a simple collateral constraint as a financial friction or a more elaborate version of a contracting problem (Cole et al., 2016; Buera et al., 2011). Our paper contributes to the literature by providing a flexible intermediation block for macroeconomists to study different phenomena under information asymmetries and market power that can be directly linked to the microdata.

From a broader perspective, the 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 the geography of branches, expansion over time of some providers, market shares, the actual loan/insurance contracts that are offered, and then the impact of digitization and new technologies. On the normative side, we seek to answer policy questions such as regulation given 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; Nguyen, 2018; Gao et al., 2019); and the welfare and distributional consequences of market structures and their potential change, different obstacles to trade (information, trade costs) (Koijen and Yogo, 2012; Martin and Taddei, 2012) and their alleviation, and the interaction of obstacles with market structure.

1.9 Organization of the Paper

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 in Section 4 to the specific case of adverse selection and discuss the model's implications for relationship lending. We show in Section 5 how to take the model to various kinds of data. One subsection uses market shares, and we show how to identify the contract frontier. In another subsection, we show how to construct the likelihood function for the number and location of banks and SME outcomes using household-level data. We discuss which moments are key for identification. Finally, we use these likelihood functions for the Townsend Thai Data in Section 6 and provide parameter estimates and counterfactual experiments. Section 7 concludes the paper and points to directions for future research. Appendices A to E are available in the current draft, while appendices F to O are available as Supplemental Material.

2 Motivation: Puzzles from the Perspective of RCT Experiments, Resolved through the Lens of the Model

Before presenting our framework explicitly, 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 regions with varying degrees of competition. Second, we study the effects of the introduction of a screening system, such as better data through digital transactions, that essentially eliminates the selection of unobservables.

2.1 Introduction of FSP in Villages, enhanced financial access through branching

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 received it are the treatment group. We are interested in the following question: what are the effects of this introduction on the welfare of villages?

We consider two different subsamples. First, we focus on an area where, before the intervention, there was no FSP reaching a village - both in the treatment and control 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 expenditure (consumption + investment) and output. The results are in Table 1.

When we compare the treatment group, that is, the villages that now have an FSP, with the control group of no FSPs, we find that average expenditure *decreases*, an unexpected result, as it might appear households are worse off. We also examine the standard deviation of expenditure. Yet, the standard deviation in expenditure *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 output increases. A potential conclusion could be that the introduction of the FSP was not welfare-enhancing for these villages. Apart from behavioral explanations, this is, of course, a puzzle. Entrepreneurs had the *option* of using intermediation services or not. There are no general equilibrium effects on prices or spillovers. Thus, household welfare should not decrease. We refer to this puzzle as the *cash expenditure-production puzzle*.

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

	Control	Treatment	Difference
Avg. cash expenditure	2.2089	2.1344	-0.0745***
	(.0223)	(0.0257)	(0.0216)
Std. Dev. of cash expenditure	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 a 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 expenditure *conditional* with changes in production output by running a regression, as in Eq. (1) at the household level. For a household h, in village v, we compute changes in expenditure 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 cash expenditure, 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 expenditure is negative, and expenditure 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 reaching villages 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. Cash expenditure increases on average, and the changes in its standard deviation are simply because cash expenditure is larger on average, [1] What can explain the

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

differences between Table 1 and Table 2?

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

	Control	Treatment	Difference
Avg. Cash Expenditure($C+I$)	2.1344	2.7442	0.6098***
	(.0257)	(0.0331)	(0.0074)
Std. Dev. of Cash Expenditure	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 the 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 .

Cash Expenditures-Production puzzle. Although the movements in cash expenditure, production, and 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 who 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 cash expenditure is reduced in Table 1 as entrepreneurs prefer to insure cash expenditure - and pay for it. Variation in cash expenditure, however, increases because most of the variability in cash expenditure 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 cash expenditure does drop dramatically for each type. The details of the model are in Appendix A. In particular, the cash expenditure equivalent gains in welfare in treatment villages with respect to control villages due to the intervention are 91.45% - a large effect yet still not successfully estimated from cash expenditure. This explains the *cash expenditure-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 [2], 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 cash expenditure 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 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.

2.2 Introduction of Screening Technologies Through Better Data, One Aspect of Digital Lending

We are interested in answering the following question now: What is the welfare effect of introducing a screening system that all FSPs here have access to that virtually eliminates adverse selection?

In our subsample of 300 villages that originally had an FSP operating, we first selected a group of 150 to receive a new FSP. We sequentially select a random set of villages to receive the screening system, with the FSP using digital data, as it were. We now have four types of villages: those that randomly were assigned to receive a new FSP (or not) and those that were assigned to receive the new digital 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 cash expenditure and production dynamics, we put some structure into the problem and calculate welfare through inputs on cash expenditure 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 a new FSP was introduced. In Column 2, we report the difference in the case where we compare villages where no new FSP was 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 Eq. (2), while in Column

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

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 cash expenditure 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. However, when there is no introduction of new FSPs, 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 whether 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 are computed through 1,000 bootstrap repetitions. *** denotes 1 % significance .

Information Structure puzzle. The data on the *information structure puzzle* are also model generated. There are two types of entrepreneurs in each village, with varying productivity, $\theta_L < \theta_H$, now unobserved by the FSP and the econometrician. This creates a potential adverse selection problem. 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.

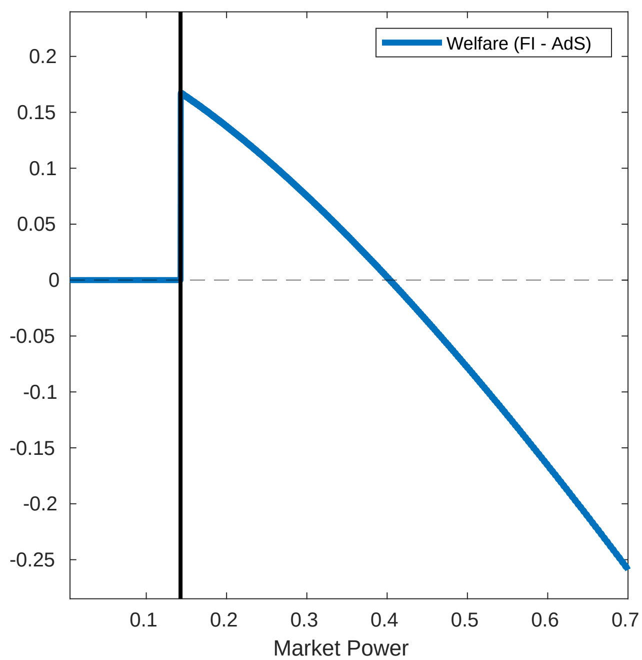
The difference between the results in Table 3 comes from differences in market power in the underlying economies. With new FSPs, all villages (treatment and control) have a relatively more competitive intermediation sector (average of $\omega = .3$), while with no new FSPs, all villages are in an economy 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, they cannot extract the rents of the full information case.

We show the difference between welfare [3] in the AdS selection case versus the Full infor-

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

mation in Figure 1 for various levels of market power. The vertical line is the minimum level of market share, such that the adverse 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: Welfare differences between adverse selection and full information for economies with contracting and competition. 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. 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 by the FSP in the AdS case. Welfare shown here is the average welfare of entrepreneurs in the economy. See Appendix B for details.

Our analysis suggests that new information technologies could make entrepreneurs worse off if concentration ω is high. A potential example concerns FinTechs that use digitally acquired data to mitigate information problems. The impact, beneficial or adverse, depends, according to the model, on the extent of competition with other providers, which, however, may be higher than stereotypical stories imply and on the objective function of the provider under regulation.

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 an 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 the utility that it is offering. Profit for an intermediary is a multiplication of the two building blocks: profits it would have contracted 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 the 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 (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 rates, covenants, etc., while utility is a uni-dimensional object. As a uni-dimensional object, a representation in utility space allows for most IO tools designed for price (also an uni-dimensional 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 reach of applications to insurance markets, healthcare, or any sector characterized by incomplete markets.

The separation of the model into building blocks is not only pedagogical but also economic in meaning and relates to 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 between the framework and 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 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 an 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 per unit is simply price minus average cost (times quantities). In a model of contracting, however, the price minus the 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 an FSP are represented by a point on the Pareto frontier: the profit is the maximum profit that can be generated conditional on offering a specified level of utility. The actual contract can then be recovered from the argmax of the optimization problem. From this point forward, we thus refer to the profit function as the *Frontier*. That is, a certain level of utility will imply a level of profit.[4]

We consider an economy populated by output-producing households running small and medium enterprises (SMEs)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 determine the probability of output q being observed conditional on capital k and effort z [5]. Type θ is potentially a vector, and both preferences and the production function can depend on it. We assume output and capital are observable and, thus, the contract can be made conditional on them[6]. 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 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. (this rate is fixed, as with a small open economy; we do not consider the savings side of intermediation, though we come back to this in the conclusion). 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

[4]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 an 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 C for a mathematical formulation of these ideas. From now on, we focus on utilities.

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

[6]Extensions in kt allow unobserved but reported output and unobserved capital. This would not change the conceptual framework here.

section, we discuss dynamic extensions. In notation, here is the surplus (profit) maximization problem

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

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 alternatively as 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). [7]

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

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

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

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

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

and the contracting frictions [8]:

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

[7]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 - the ability to offer lotteries - and the numerical approximation of the discrete grid if the true model has continuous supports for variables.

[8]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.

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. [9]

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:

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

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

Eq. (12) is the *Incentive Compatibility Constraint*, which guarantees that, when the effort is not observed, it is optimal for the agent to execute the effort recommended by the FSP. In reverse, when effort is observed or inferred through data, the moral hazard constraint is weakened. Eq.(13) simply states that if the FSP can recover $(1-\rho)$ of the output, the utility offered is 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. Note that this is one of the many possible ways of writing a LC constraint.[10]. (We will however assume that household participation can be monitored so that relationships are exclusive.) 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 (2). 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 in which geographic costs

[9]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.

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

can be mitigated with digital lending, in ways which are clearly delineated. Still, one could apply the model to several other contracting problems or frictions, i.e., other S's. Better data though digitization can also change those. 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 and maximal effort. The maximum feasible utility in FI, MH and LC the utility with maximum consumption and minimum effort. To summarize 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 parameterize 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 Table 4. 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 4 summarizes the problem and constraints. We leave the detailed discussion of computation for later, when we discuss the numerical method.

Table 4: Parameter Values, Grids and Constraints for Frontier Construction

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

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

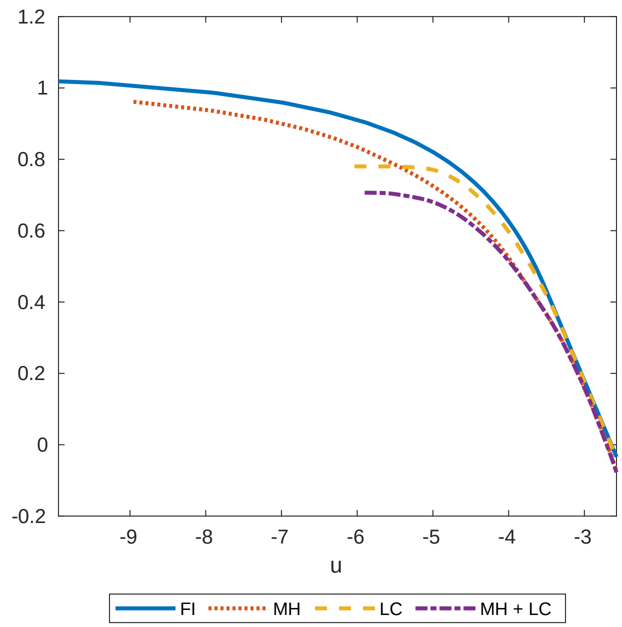
Friction	Constraint(s)
Full Information (FI)	-
Moral Hazard (MH)	Eq. (12)
Limited Commitment (FI)	Eq. (13)
MH + LC	Eqs. (12) and (13)

Note: parameters for a utility function given by: $\mathbb{U}(c,z \mid \theta) = \frac{c^{1-\sigma}}{1-\sigma} - \theta z^{\psi}$.Grids and grid sizes based on the Townsend Thai Data (Section 6) and as in Karaivanov and Townsend (2014). Parameter values for risk aversion and effor aversion are within the range of the various estimated parameters reported there. The value for THETA is identical. The grid for consumption has enough points to guarantee that the frontier is smooth.Linear programming problems solved with the Gurobi Linear Solver for Matlab for 150 utility levels equally spaced between $[u_{min}, u_{max}]$.

For a given level of capital, Figure 2 displays the frontier for different contracting frictions. In all cases, the higher the level of utility that must be offered to an agent, the lower the level of profit for a bank/ SP. (But take note that in the market structure section utility will be endogenously determined by the degree competition and FSPs.) Moreover, as the agent has a strictly concave utility function and the FSPs are effectively risk-neutral after pooling away risk across SMEs, higher levels of utility imply marginally lower levels in profits. 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 the frontiers for given frictions cross. 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 simply avoid paying back.

Figure 3 displays expected levels and standard deviation of consumption, effort and capital. In panel (c), for instance, we can see how the behavior of capital allows us to differentiate the behavior of capital between MH and non-MH models once utility u is determined. Note in Figure 3 how MH also induces the FSP to increase the standard deviation of consumption (panel (d)). For MH, the FSP needs to create risk to incentivize effort.[11]

Figure 2: The Profit Function of FSPs as a Pareto Frontier: FI, MH, LC and MH + LC



Note: FSP profits 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.

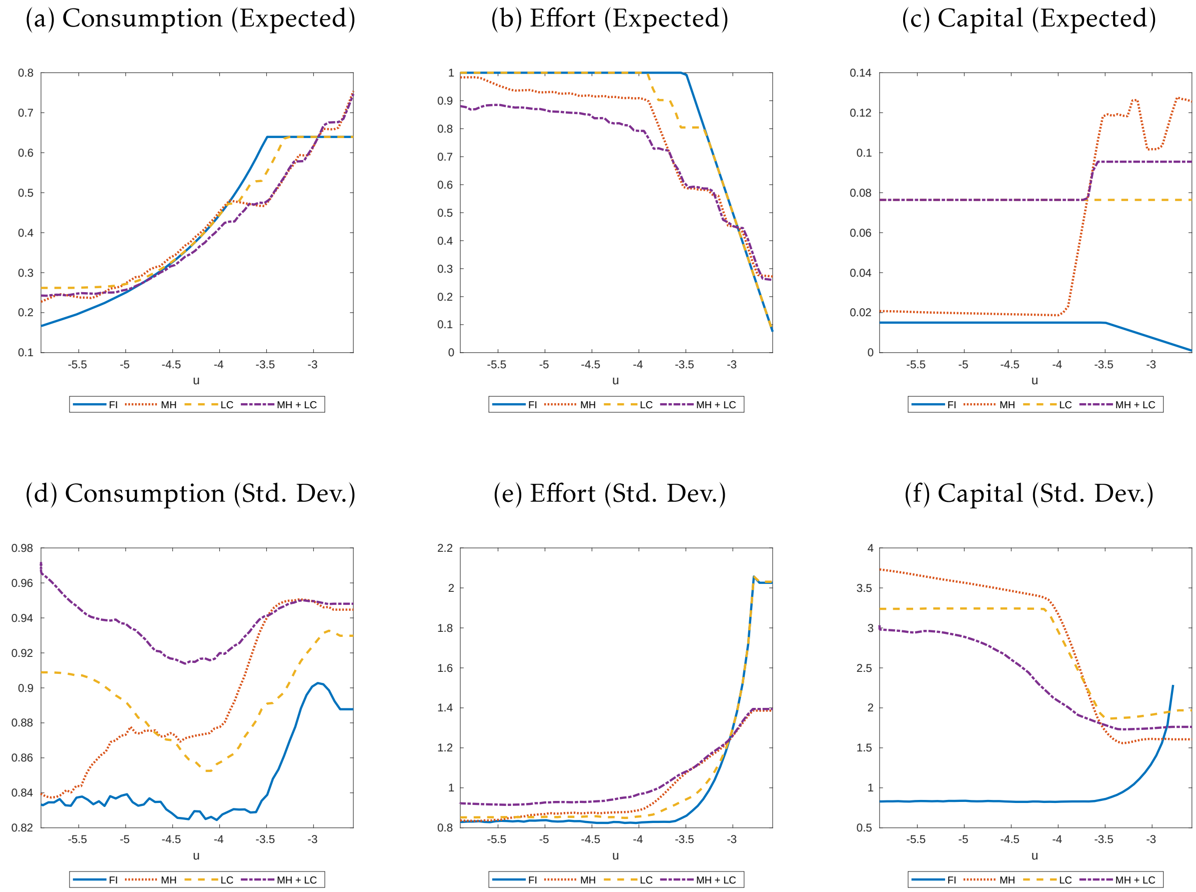
3.1.2 Extensions and Limitations

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 FSPs). We describe here the full commitment version, but the problem can be re-adapted for no-commitment contracting. In a full-commitment case, we follow

[11] 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: Contracts: Expected level and Standard Deviation in consumption (c), effort (z) and capital (k) for varying levels of utility



Note: The expected levels and 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.

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 for this variable in the next period. Including w' as a choice variable is consistent with choosing lotteries over it to represent future promises. With dynamics we can assume that θ now explicitly follows a Markov Process as in most applications. This allows us to write the problem recursively, as in Karaivanov and Townsend (2014).

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 to a specific household contract 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. (Alternatively, commitment could be maintained in a dynamic setting by punishments for leaving.) Overall, the issue of common agency and renegotiation is still very model dependent (e.g., Handely et al. (2017)) and one would have to proceed case by case.

3.2 Market Structure

In this subsection we focus on the legacy market structure where FSPs compete with each other. We also discuss how the model can be altered to accomodate digital lending.

Our model features a logit demand system and in the baseline legacy version the model, 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 $p = 1, \dots, P$, there are B_p FSPs, located at a position $x_b \in \mathbb{R}^2$, $b = 1, \dots, B_p$. Households reside in villages, denoted by $v = 1, \dots, V_p$ in each market. Each village v has a population of N_v. We can display all of this on a map, and indeed as short-hand we denote an individual markets as a *Map*. 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.[12]

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. 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 in the baseline version of the model that there is a linear

[12] 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.

spatial cost ψ. In particular, the value a households located at location x_v attributes to a contract that offers u_b offered by bank b is given by Eq. 17

$$\mathbb{V}(u_b, x_b, x_v) = u_b - \psi t(x_b, x_v) \tag{17}$$

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. 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:

$$\varphi_b \equiv \{u_b, u_{-b}, u_0, x_b, x_{-b}, \{x_v\}_v\} \tag{18}$$

as the vector of relevant variables in the profit of a financial service provider b, where the subscript $-b$ denotes the 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. (19)

$$\Pi(\varphi_b) \equiv S(u_b)\mu(\varphi_b) \tag{19}$$

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

In some of the counterfactual policy changes considered below, we introduce digital lending. This changes the financial access map. As a first pass, a digital lender could potentially reach all customers at zero spatial cost ψ.[13] However, a salient aspect of reality that we retain is that internet access is imperfect. In these settings, one could think of a financial service provider as facing a choice between being a digital lender/FinTech or establishing legacy branches. As a digital lender, with no branches, its own location is indeterminate. However, it does not reach all geographical locations equally. A key determinant is the distance between a broadcast station with coordinates x and the receiver with coordinate y (Adegboyega et al., 2014). Thus, the higher the product of parameter ψ with $t(x,y)$, the lower the access is. But note low service areas mean that a client is far from a given configuration of cell towers rather than far from a branch of the branch of an FSP [14]

Indeed, potentially, a financial service provider could choose to be a legacy provider with physical branches, with the previous interpretation of spatial distance, competing, however, with digital banks that face varying signal strength depending on client location. The choice

[13]In practice, digital banking has been incremental, with physical ATMs substituting for some transactions within branches. Relatedly, some transactions could be digital but others still required a trip to the branch. We abstract from this history.

[14].

is non-trivial in the sense that the business models and operating systems of digital service providers vs legacy banks are very different. Note also that digital lenders may also have superior information by virtue of better data, hence a change in underlying obstacles to trade on the contracting side of the model. That is, digital lending entails both access and contracting obstacles. Below, we stick to the traditional spatial interpretation of access unless otherwise noted.

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 $v = 1, \ldots, V$ where households reside [15]

$$\mu(\varphi_b) \equiv \sum_{v=1}^{V} N_v \mu_v(\varphi_b) \tag{20}$$

The functional form we use for μ_v in the benchmark specification is given by Eq.(21)

$$\mu_v(\varphi_b) = \frac{e^{\sigma_L^{-1}[\mathbb{V}(u_b, x_b, x_v) - u_0]}}{1 + \sum_{\hat{b}=1}^{B} e^{\sigma_L^{-1}\left[\mathbb{V}(u_{\hat{b}}, x_{\hat{b}}, x_v) - u_0\right]}} \tag{21}$$

This is the textbook logit model where u_0 denotes the utility of autarky. It can be microfounded with extreme type 1 idiosyncratic preference shocks to agents to contract with each FSP, drawn with mean 0 and variance σ_L at the household level. (Heterogeneous preferences also incorporate tastes for operating systems of the various providers). 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. (21) for three main reasons. First, it speaks directly to the data. Without idiosyncratic preference shocks, a single FSP would, except for ties, dominate the market of a given village, which we typically do not see. Second, an interpretation of the logit preference shocks is that this comes in part from spurious product differentiation, that is artificial heterogeneity. Third, parameter σ_L allows us to summarize dimensions of the model that we are not considering that eventually affect the elasticity of demand. Fourth, it smooths the demand functions and guarantees existence and uniqueness of 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.:

$$u_b^* = \arg\max_{u \in W} \; \Pi\left(u_b, u_{-b}^*, u_0, x_b, \{x_v\}_v, x_{-b}\right), \;\; \forall b \tag{22}$$

[15] 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.

note that the equilibrium is at the 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 below, with an intuitive explanation in appendix C, characterizes the equilibrium properties. It shows that under the assumption that μ is log-concave in u_b and an additional equation holds (which we show to be true in the case of the logits, as in Eq. (21)), the equilibrium exists, is unique and can be computed by an iterative algorithm.

Lemma 3.1. *Let the demand μ, as in Eq.(20), be log-concave in u_b, log-supermodular in (u_b, u_{-b}), bounded away from zero and satisfy Eq. (23) $\forall\ a \in \mathbb{R}$:*

$$\mu(\varphi_b(a)) = \mu(\varphi_b(0)) \tag{23}$$

then $\exists!\{u_b^*\}_b$ that satisfies Eq. (22). Moreover, $\{u_b^*\}_b$ can be computed by an iteration of best responses starting at any strategy.

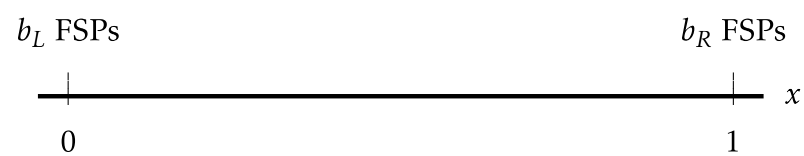
Comparative Statics. We conduct several counterfactual exercises in our model to understand the terms of loans offered and how each contracting friction and the market structure affect the equilibrium contracts and real outcomes. We use a spatial confirguation/ map as Hotelling (1929), D'Aspremont et al. (1979), Prescott and Visscher (1977), which is familiar and interpretable, not a black box, and avoid for now the complications of the reality of geospatial configurations of real maps (though we return to this with likelihood estimation later.) We provide several numerical results. These are consistent with the results of the first two experiments mentioned in our leading initial section and also, again, the conceptual and empirical applications in subsequent sections below.

First, we illustrate how asymmetry in the number of FSP at distinct locations changes the equilibrium utilities and hence change 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, an increase in spatial costs can increase or decrease welfare of SMEs. For example it can creates local monopolies, which are able to charge more for financial services, but at high spatial costs providers have to provide higher utility and hence lower profit in order to form reasonable markets. Fourth, 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 in increasing welfare or reducing financial costs.

These exercises are not only useful to understand the inner workings of the model. They illustrate important mechanisms in reality. One can think of advancements in the financial lending 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 can 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 also how this depends on imperfect competition and the underlying financial frictions. Again we go beyond Hotelling and quantify these in an actual geospatial environment, the Thai application in section 6 below.

Throughout this section, we use the median level of capital observed in the Townsend Thai Data (more details on Section 6). The frontier we use for each contracting regime is the one in Figure 2. 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 4). We denote b_L as the number of FSPs on the left at $x = 0$ and b_R as the number of FSPs on the right 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.

Figure 4: 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.

The parameters for the frontier are as in Table 4, while the baseline parameters in market structure are as in Table 5 below. To allow for easier comparison between experiments, we re-scale utilities such that a zero utility represents the autarky level and 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 5: 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 $x = 1$

3.2.1 Heterogeneity Across Villages

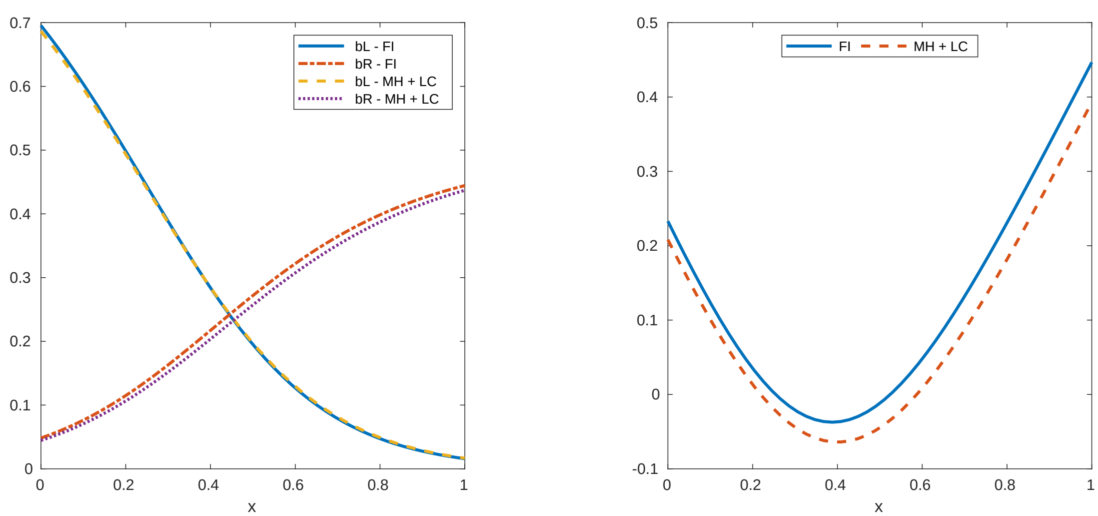
We first show the equilibrium implications for each village in $[0,1]$, that is, for fixed spatial costs and fixed number of providers, but asymmetric across endpoints, and compare across regimes with different obstacles. Using the parameters in Table 5, we solve for the equilibrium in utilities (with Lemma 3.1) and recover the implied equilibrium contracts. Note that there is one FSP at the left point, $x = 0$, and two at 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.(24).[16] The results are in Figure 5. [17]

$$W_v(\psi, \sigma_L) \equiv b_L \mu_{v,x=0} \left[u_{v,x=0} - \psi t(x_v, 0)\right] + b_R \mu_{v,x=1} \left[u_{v,x=1} - \psi t(x_v, 1)\right] \tag{24}$$

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.

Figure 5: Market Share and Welfare by villages located in $x \in [0,1]$ for FI and MH + LC

(a) Market Share (b) Welfare



Note: Market shares by village in position x (as in Eq. (21)) and welfare (as in Eq. (24)) in the equilibrium with spatial configuration of Figure 4 and parameters of Table 5.

First, see in panel (a) of Figure 5 the effects of spatial access costs in which FSP provides

[16] 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.(24) .

[17] Note that Eq.(24) 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 .

more services for each village. As expected, villages closer to $x = 0$ mostly contract with the FSP in $x = 0$. That said, the heterogeneity generated by logic plays a non-trivial role. The key model implication is how this curve decays as distance grows. In our baseline calibration, the market share of the FSP at $x = 0$ decays from .7 to .02 from the closest to the furthest village. Second, in panel (b), one can see the effects of *local* competition. The FSPs offer higher utilities in $x = 1$ since we use $b_R = 2$ and $b_L = 1$ as our baseline. Third, the utility difference 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. Likewise, market shares of a given village at x are equal to the left of the midpoint .5, as the right FSPs are more attractive.

Given this difference in utility levels between regimes and their differential spatial effect, we can see that the average and standard deviation of consumption and effort in villages will also differ due to different contracts. Note in panel (a) of Figure 6 that average consumption is always larger under full information contracting, but the difference is reduced closer to where there is more competition (at $x = 1$, where there are two FSPs). The opposite is true for the standard deviation in consumption. These results are a combination of different utility levels implying different contracts (Figure 3) weighted by different market shares at each region (Figure 5).

Note that average consumption behaves the exact opposite of utilities. This is a consequence of insurance in the model. In the location between providers, a relatively larger fraction of households choose to stay in autarky and, facing more risk, choose higher effort, resulting in higher expected consumption; note also how the standard deviation in consumption is also decreasing closer to the FSPs. Also, for effort, note that under full information, the higher the competition, the greater leisure (lower effort) is, but the MH/LX regime overturns this effect.

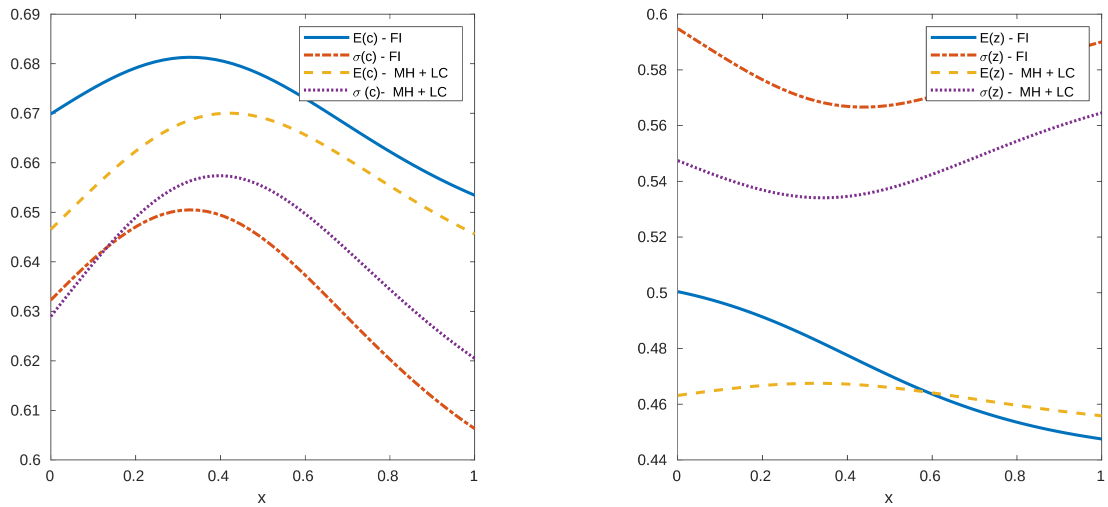
The levels and heterogeneous behavior of consumption and effort across villages will allow us to identify the structural parameters later. A model that ignores this spatial variation, as was the case of the model behind the experiment of Section 2, will mistakenly use this cross-sectional variation in location as consumption variation, which was the source behind the consumption-production puzzle.

3.2.2 Spatial Costs and Local Competition

We change gears now to display how the equilibrium changes with changes in spatial costs, denoted by ψ, applied uniformly to all FSPs. For simplicity, we focus on an economy with

Figure 6: 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 4 and parameters of Table 5.

obstacles to trade, where there is MH + LC in contracting. [18]

Increasing spatial access costs in a legacy branching environment can increase or decrease the overall level of profits of FSPs depending on local competition (Figure 7, panel (a)). When spatial costs are zero, profits of the left and right banks are the same, as there are no access frictions. 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 offers higher levels of utility to increase its market. Likewise, for high values of ψ, profits increase as the market becomes more segmented, that is, close to a local monopoly. In contrast, at $x = 1$, where there are two FSPs, for low values of ψ, the local monopoly effect is reduced with that local competition. Then, increasing the spatial costs makes it harder for banks to make profits on the extensive margin with larger markets. Thus, local competition becomes more intense, and profits decrease.

Increasing spatial costs also have non-monotone and heterogeneous effects across villages in terms of welfare. To illustrate this result, we compute the welfare of villages situated at three locations, $x \in \{0, .5, 1\}$ as defined in Eq. (24). Without spatial costs, all villages have the same access to the three FSPs, and thus, all have the same welfare (panel (b) of Figure 7). In a comparison of legacy economies with increasing spatial access costs, a resident of village $x = 0$ not only has to potentially pay higher costs if it wants to visit FSPs at $x = 1$ but the

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

utility being offered by FSPs at $x = 0$ is reduced (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 otherwise reduce market share, and this 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).

In the appendix we conduct more experiments, changing the logit variance and increasing the number of FSPs at x=1. A high logit variance makes the impact of lower spatial access costs lower. The impact of digitization reducing moral hazard through better data enhances the gains from having more bank branches.

In an environment with digital lending, we assume cell towers exist at both endpoints, but unlike the number of banks, towers are the same at both. The signal strength from towers declines with distance, and consumers access the nearest tower. Thus, the disutility of clients is lowest in the middle region at x=.5. If the signal reaches zero, there would be a cell service void around x-.5, so that the population reached by digital lenders would be zero, limiting the overall size of their markets. (Potentially, a legacy bank might enter there). Unlike the legacy model, the profits of digital lenders do not depend on their location, and welfare for households would be symmetric in distance. In a conceptual experiment to capture the impact of increased digitization, one could improve the signal communication system, allowing the markets to be larger, thus impacting utility offerings of digital lenders as they compete for customers [19].

4 Adverse Selection

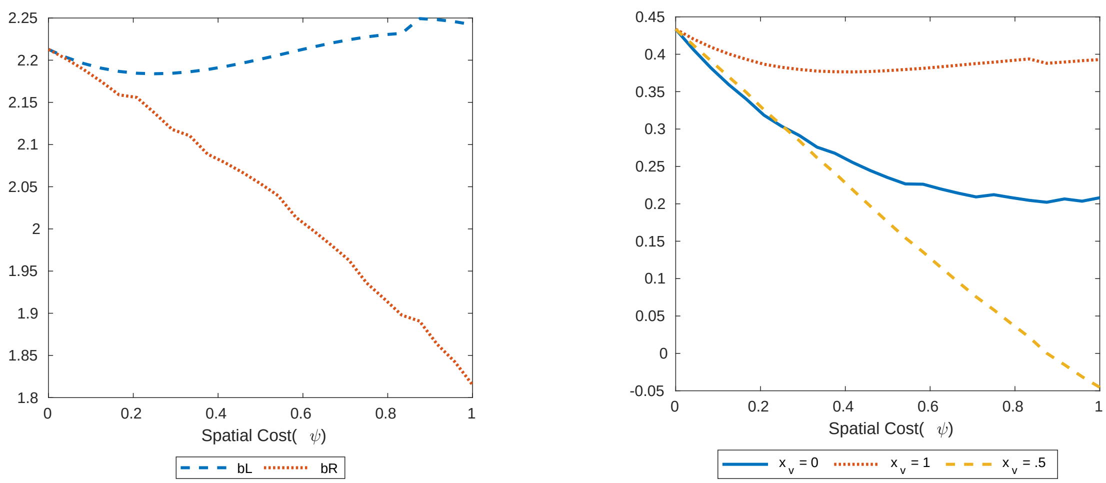
We now explore the case of Adverse Selection, where FSPs do not observe the type θ of households. Adverse selection is more complex than when types are observed because the frontier is now a function of the *contract menu* offered for all types, and not simply the utility offered for one given type θ.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

As an application, we consider a specific case of adverse selection with two types that differ only in their cost of exerting effort. We focus on a case where one FSP can contract under FI but is subject to spatial costs (a Local Cooperative), and the other FSP is subject to MH + AdS but, due to a larger network structure, is not subject to spatial costs (A National Bank).

[19] Note that for a single digital provider, it can reach more customers, as it maximizes its own profits, as per SEE EQUATION 28 BELOW. The overall level of utility of customers would be higher the greater the number of providers through competition.

Figure 7: 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.(24) for three villages - the ones located in $x \in \{0, .5, 1\}$. Equilibrium with spatial configuration of Figure 4 and parameters of Table 5, changing the spatial cost parameter, denoted by ψ. Contracting frictions are MH + LC.

To preview, our results indicate that the local cooperative will always offer higher utilities (better contracts from credit cooperatives, for instance). Moreover, as spatial costs rise, the share of good types on the credit market falls significantly (with the rest now producing under autarky), while the share of bad types remains practically the same (a lemon problem).

We also consider in Appendix J a case where both FSPs are subject to the same spatial costs, but one contracting under FI and the other under MH + AdS. Relationship lending generating full information can significantly increase within-region inequality between those who have access to credit through previous relationships and those who don't. This effect is larger for higher spatial costs, where markets are more isolated. We also show that with a small logit variance (fiercest competition), there is a selection mechanism where the informed local bank retains mostly good customers, while the national bank has most of the bad customers. This result has key implications for macroprudential regulation regulation if the policy maker is concerned with bank-level risk.

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 an FSP for a given capital level k is given by Eq.(25)

$$S^{AdS}\left(\{u^{\theta}\}_{\theta\in\Theta} \mid k\right) \equiv \max_{\{\pi^{\theta}(c,z,q,k')\}_{\theta\in\Theta}} \sum_{\theta\in\Theta} \left\{ \sum_{c,z,q,k'} \pi^{\theta}(c,q,z,k') \left\{q - c + (1+r)\left[(1-\delta)k - k'\right]\right\} \right\} f(\theta) \tag{25}$$

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

$$\sum_{c,q,z,k'} \pi^{\theta}(c,q,z,k')\mathbb{U}(c,z|\theta) \geq \sum_{c,q,z,k'} \pi^{\hat{\theta}}(c,q,z,k') \frac{P(q|k',\theta,z)}{P(q|k',\hat{\theta},z)} \mathbb{U}(c,z|\theta), \ \forall \hat{\theta},\theta \in \Theta \tag{26}$$

and, potentially, the other contracting frictions (MH, LC, etc.). The difference is the added truth-telling constraints (Eq. 26). 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 an FSP, contract choices cannot be made 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. (26) complicates the problem significantly since we cannot separate the contracting problem for different types. Without the constraint in Eq. (26), we could separate the sum in θ in independent problems (the case of Section 3.1).

Simplified Case: Ordered Types and Binding Constraints. If we assume that (i) utility is separable in consumption and effort and SMEs only differ in cost of effort θ, as Eq.(27) [20] and (ii) the only truth-telling constraints that are binding are those of a lower cost of effort θ taking the contract of a higher one, we have that the result of Lemma 3.1 still applies in this case. See Appendix E for details. Although this result is powerful, since it is hard to guarantee existence and uniqueness in models of AdS, it still relies on two strong assumptions. In particular, the second assumption is not innocuous. Different from the textbook case, 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. [21]. As a result of these complications, we move on to a more general, numerical method.

General Case: To solve for the equilibrium in utilities with the Frontier as in Eq.(25), we use a *distance-to-Nash* algorithm (See Appendix I for details). The idea behind the algorithm is to write the Nash Equilibrium as an optimization problem (instead of a fixed point one).

[20] Can be generalized for heterogeneity in any dimension (as long as its only one) and satisfies a concavity condition. See Appendix E for details.

[21] 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 .

In this case, we do not have proofs of existence or uniqueness for the equilibrium, 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: Local vs National FSPs

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.(71)

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

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 D, we use a Hotelling line as our spatial configuration with villages uniformly distributed over it. However, we place only one FSP at $x = 0$ in the FI regime and one at $x = 1$ in the MH/LC regime (instead of two at $x = 1$). Thus, the asymmetry now comes from the information that each FSP has. We consider the case of local vs national FSPs. The local cooperative has an information advantage - not subject to MH or AdS. The national bank has a spatial advantage: as an approximation, with a dense network of branches, SMEs do not have to pay travel costs to visit that national bank. We focus on how the equilibrium changes with spatial costs.

We use the same parameters as in Section 3.1 to generate the frontier (Table 4) with the addition now of the high type, $\theta_H = 2 > 1 = \theta_L$. The market structure parameters are as in Table 6. 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 5 and Table 6 comes from the fact that now utility scales are naturally different for different types.

The results of equilibrium utilities and market shares and profits with varying spatial costs are displayed in Figures 8-9, respectively. If spatial costs ψ are at or near zero, the informed local cooperative has a higher market share in both good and bad types since it can offer higher values of utility (no constraints between the two in the frontier). In fact, note that local FSPs always offer higher utilities but end up with lower market shares and profits if ψ, the spatial costs, are nontrivial. The two providers can co-exist in our model since each of them will have some advantage (informational vs spatial). When spatial costs ψ increase in a cross-sectional comparison (and are small to begin with), we observe that the local FSP increases its utility offerings to both types 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 facing the national bank has decreased as a consequence

Table 6: 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 $x = 1$

of this increase. If spatial costs ψ are large in the baseline, we have that both FSPs reduce their utility offerings.

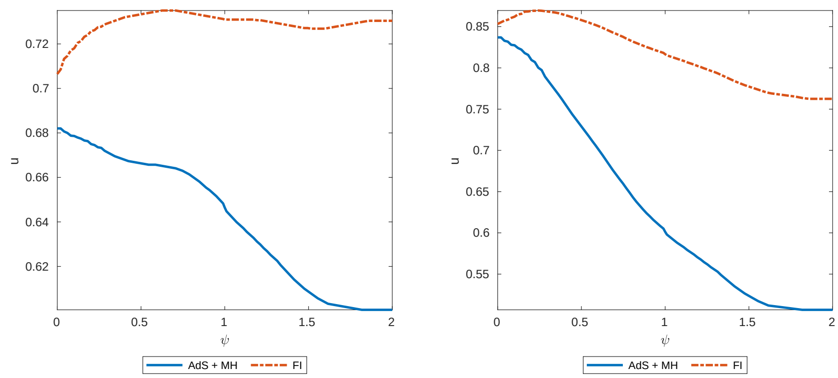
In Figure 9, we see the consequences for market shares. The total participation of good types is reduced from 75% to 37% as spatial costs increase (with the rest of the population producing through autarky), while the total market share of bad types simply transfers from the local cooperative to the national bank. This is a version of the lemons problem. Competition of an informed (local cooperative) with an uninformed FSP (national bank) can lead to a reduction in relative participation of good types (θ_L) as spatial costs increase.

Going the other way, at zero spatial costs, the fully informed FSP and the otherwise local FSP have no disadvantage, turning into a BigTech digital lender, securing a higher share of both types and higher profits.

In Appendix J, we separate the two features of digitization: lower access costs and better information. The information regimes are the same, with the local cooperative having better information, but spatial costs are symmetric across the two providers, i.e., the national bank suffers the same impediments to attracting customers as the local cooperative. The impact of changing spatial costs is then significantly different. As is evident from the figures, profits of the informed digital lender lie well above those of the legacy national bank for all spatial costs. Profit curves more in parallel with spatial costs. However, market shares of bad types are higher for the uninformed national bank, hence a concern for regulators. But again, this does not happen when the digital lender has both positive attributes, easier access, and lower obstacles to trade.

Figure 8: 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 no 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 6. We solve the equilibrium using the distance to Nash algorithm (Appendix I). The x-axis, ψ, denotes spatial costs. Utilities are normalized such that zero is the autarky and one is the FI, the perfect competition level for the bad type.

5 Taking the Model to Data

This section explores how to take the model to actual data. For simplicity, we assume that types are observed here (i.e., no AdS). [22] Our ultimate goal is to develop an empirical toolkit for models of competition and contracting that other researchers can use. We can discuss model implications for the data under two different assumptions for what data is available.

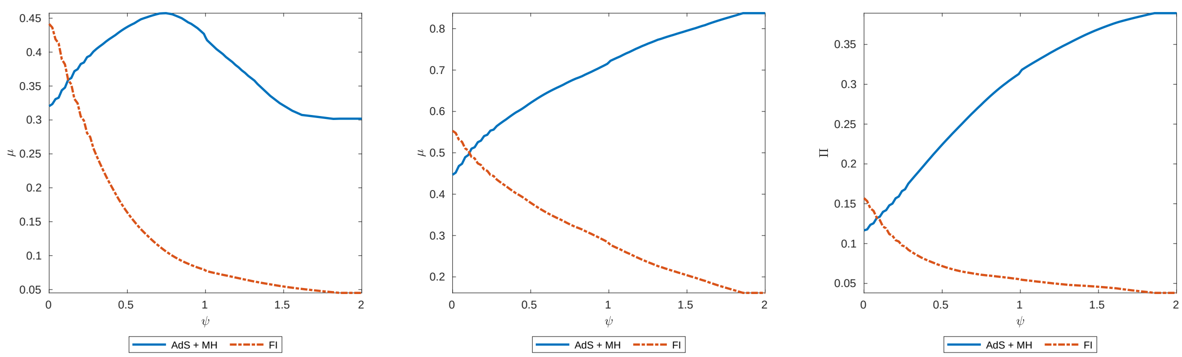
First, we discuss how to use the theoretical framework when there is market share data. We show how to identify the frontier from market share data, which in turn provides a way of computing some, but not all, counterfactuals *without* a model of contracts, even in the presence of unobserved heterogeneity.

Second, we discuss how to use the theoretical model when there is data on the number of intermediaries in a given location (as in Bresnahan and Reiss, 1991) and data on households' outcomes to construct a likelihood function that maps the model to the distribution of consumption, income, and capital of households, 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, we are interested in the implications for household-level outcomes. Therefore, a method that maps the 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,

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

Figure 9: Local vs National Banks and Spatial Costs: 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, but SMEs pay no 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 6. We solve the equilibrium using the distance to Nash algorithm (Appendix I). The x-axis, ψ, denotes spatial costs.

which we also discuss in this section. Using this numerical method, we provide Monte-Carlo evidence that we can identify the parameters of interest.

To preserve the brevity of the main text, we present the requisite notation in Appendix 5.1: How to Use Market Share Data, 5.2: Identification of the Frontier, 5.3: The Bank Model and Interaction with SME Data, and 5.4: Likelihood of Households.

5.1 Market Share data

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

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 control for this in our estimation method. For simplicity in the exposition here, we assume that villages are homogeneous apart from their market structure in terms of 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

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

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 the data perfectly, we could add an error to bank-village level market shares and write an empirical version of Eq. (21) as Eq. (27) [24]

$$\ln\left(\mu^p_{v,b}\right) - \ln\left(\mu^p_{0,v}\right) = \sigma_L^{-1}\left[u^p_b(\zeta_S) - \psi t(x^p_b, x^p_v) - u_0(\zeta_S)\right] + \vartheta^p_{b,v} \tag{27}$$

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. (27), we could estimate the structural parameters by 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, how can we use market share data to allow for the identification of the frontier without having to define which contracting frictions are relevant in the data? This is the topic of the next subsection.

Estimating the frontier is related but significantly different from what is generally done in the IO literature. In the classic IO literature, as Berry et al. (1995), the researcher observes product characteristics, prices, and market shares (or individual choices) and tries to estimate how these characteristics affect the utility function (and thus the decision process). In this case, the profit function of firms is known: it is given by prices minus cost times quantities. What we propose here is different. Due to the contracting frictions, we do not know *ex-ante* what is the shape of the profit function. We use the market share data with a first-order condition of FSPs to recover how this profit function is shaped.

5.2 Identification of the Frontier with Market Share Data

Instead of assuming that the utility is parametrized by ζ_S, 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

[24]By taking the log at Eq.(21) and the fact that market shares must sum to one (including the outside option). See Appendix K for details

[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 set of parameters, we can solve for the frontier and equilibrium, 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.

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. Imagine that both μ and 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. (28)

$$-\frac{\partial_{u_b} S(u_b^*)}{S(u_b^*)} = \frac{\partial_{u_b} \mu\Big(u_b^*, u_{-b}, u_0, x_b, x_{-b}, \{x_v\}_v\Big)}{\mu\Big(u_b^*, u_{-b}, u_0, x_b, x_{-b}, \{x_v\}_v\Big)} \tag{28}$$

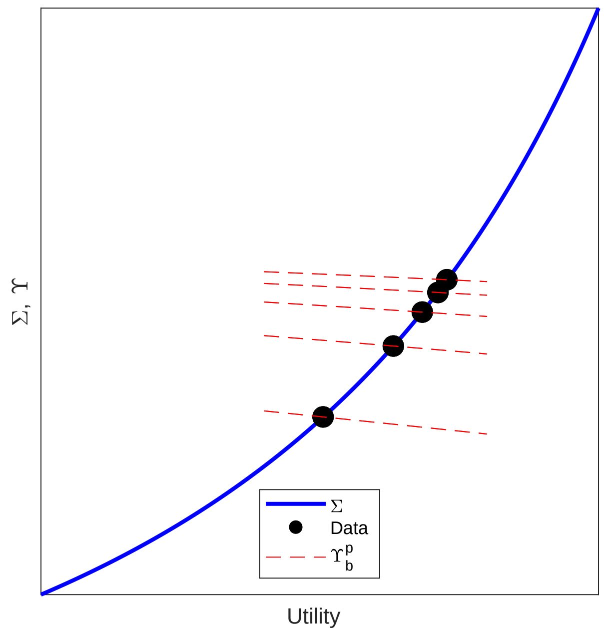
As in Eq.(28), we can represent the solution to the problem of each FSP b as the intersection of the marginal loss of offering a higher utility as determined by the frontier with its market share gains. For FSPs in different provinces, however, the marginal gains in market share, the RHS of Eq.(70), 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, assumed here to be common over all FSPs, 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. (29).

$$\Sigma(u) \equiv -\frac{\partial_{u_b} S(u_b)}{S(u_b)} \text{ and } \Upsilon_b^p(u) \equiv \frac{\partial_{u_b} \mu\Big(u_b, u_{-b^p}^{*,p}, .\Big)}{\mu\Big(u, u_{-b^p}^{*,p}, .\Big)} \tag{29}$$

With the model's structure, competition variation allows us to estimate the curvature of the frontier, as seen in Figure 10. This intuition here is the same as that of instrumental variable estimation. We have variation in the market structure that does not directly affect the frontier; it 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, we would observe the circular dots in Figure 10. 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 10, we focus first on what we *cannot* identify. We cannot identify the scale of utilities, as 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 together. 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 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

Figure 10: First Order Condition of an FSP in Different Provinces



Note: visual Representation of FOC of an FSP b in market p, i.e., Eq. (70) 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.

details on identification of the frontier Section M.0.1.)

5.3 Data on the Location of Banks and Household Level Outcomes

We break this section into a number of pieces, first the bank model and then household-level SME data (details of the interaction with the bank model and household outcomes are in the Appendix).

For the bank model, we develop the likelihood of FSP location data as in a model of entry (Bresnahan and Reiss, 1991). The data that is informative about the model's parameters is the number of banks in each potential entry location in each province. In particular, the next bank to enter any potential location in province p would have negative profits, which 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 our model. Let m^p be a potential location for a new entrant

FSP in province p. [27]. 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. (30).

$$\Pi^E\left(B_{m^p}|m^p,\{x_b^p\}_{b\notin m^p}\right) \equiv S(u_b^*)\mu\left(u_b^*, u_{-b}^*, x_b^p, x_{-b}^p, u_0\right) \tag{30}$$

Note that the right-hand side depends on utility offerings, but these are the equilibrium offerings. If the model perfectly replicates reality, inequalities (31) 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. The marginal bank, which would imply $B_{m^p}+1$ banks, should imply negative profits for all banks in a given location (otherwise, it would have entered). [28]

$$\Pi^E\left(B_{m^p}|m^p,\{x_b^p\}_{b\notin m^p}\right) \geq 0 \;\; \cap \;\; \Pi^E\left(B_{m^p}+1|m^p,\{x_b^p\}_{b\notin m^p}\right) < 0, \forall \;\; m^p, p \tag{31}$$

For simplicity, we define the indicator variable $\mathcal{E}(m^p) = 1$ if Eq. (31) 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.

$$\Pi^F\left(B_{m^p}|m^p,\{x_b^p\}_{b\notin m^p}\right) \equiv \Pi^E\left(B_{m^p}|m^p,\{x_b^p\}_{b\notin m^p}\right) + \iota_{m^p}, \;\; \iota_{m^p} \sim \mathcal{N}(c_E, s) \tag{32}$$

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 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). Given the stochastic term on the entry costs, we can generate a likelihood function for observed FSP locations.

More details are in Appendix M.1.1.

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

[27] 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 are any location with an FSP between 1997-2011

[28] Note here that we assume 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 the entry game, the number of deviations - and of equilibrium calculations to compute the likelihood - grows exponentially.

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 the optimal contract. The optimal contract then has implications for each household's joint distribution of consumption, output, and capital.

Before constructing the likelihood, we introduce new notation. Let the results of the model in terms of contracts be given by (33), which are specific to each location (m). Assuming that provinces are independent, we simplify the notation and do not include p as a superscript.

$$\left\{\pi_m(c,q,z|k,u^*_{b\in m})\right\}_m \tag{33}$$

Let the cross sectional household level data be given by $\{\hat{y}_j\}_{j=1}^{\mathcal{H}}$, where $j = 1,\ldots,\mathcal{H}$ denotes households. Here, we use consumption, 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:

$$\mathcal{N}(0, \gamma_{ME} \cdot \chi^2(X)) \tag{34}$$

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. (35).

$$g_v(c,q|k,\zeta) = \sum_u m_v^u(k) \sum_z \pi(c,q,z|k,u) + \left[1 - \sum_u m_v^u(k)\right] \sum_z \pi^{aut}(c,q,z|k) \tag{35}$$

where $m_v^u(k)$ is the share of agents in village v, capital k that are offered utility u by an 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. (36).

$$m_v^u(k) \equiv \sum_{b\in B} \sum_u \mathbb{1}_{u=u^*(b)} \mu_v^b(u,k) \tag{36}$$

The distribution of (c,q,k) in a village is then given by Eq. (37), where we multiply by the distribution of capital in the village, $h_v^k(k)$.

$$f_v(c,q,k|\zeta) = g_v(c,q|k,\zeta) h_v^k(k) \tag{37}$$

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.

More details are in Appendix M.1.2.

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.

We feature various counterfactual results, namely, what if, on top of the baseline with legacy infrastructure, there were changes associated with digitization and the cost of trips to banks or banks competing on a common platform? Our results suggest that digital banks, conversely, spatial costs, are important for individuals, as an individual would reduce their consumption by 20% to eliminate such costs. In terms of aggregate welfare, reducing spatial costs by 50% is equivalent to increasing consumption by 4.85%. Reducing the variance of idiosyncratic logit shock, lessening the importance of spurious product differentiation, σ_L by 50% is equivalent to increasing consumption by 15.36%. One additional legacy bank in the baseline setting has limited effects on our results, increasing consumption by only 2.2%. In short, technology is more important for welfare gains than competition.

6.1 Data Description

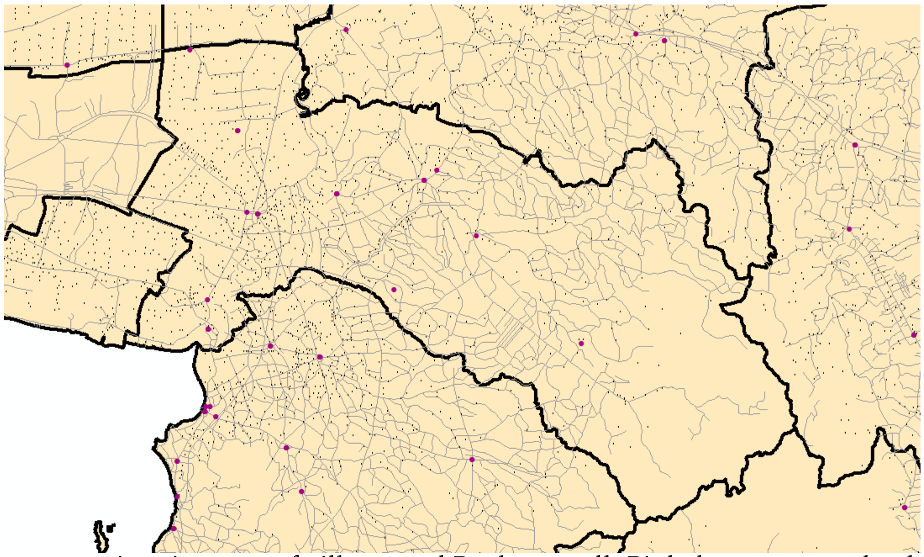
For household-level data, we use the Monthly Resurvey of the Townsend Thai Data for the year 1999, as in Karaivanov and Townsend (2014). We use several data sources for locations of villages in the Townsend Thai Data and banks. We assume that each bank branch is a different FSP. [29]

Village data is extracted from the Thai Community Development Department (CDD) survey. The information on bank branch locations comes from the Bank of Thailand, Bank of Agricultural and Agricultural Cooperative, Telephone Authority of Thailand, Community Development Center, and several non-traditional financial institutes. We combine these datasets as in Assuncao et al. (2012) to get each bank's open and close date, as well as the bank branch and name. We geo-locate each bank branch and village using the Google Maps API and compute travel time between two points on the map using a GIS platform. We use the road network from the Thailand Environment Institute. The data classifies all roads in Thailand among seven 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

[29] This is equivalent to assuming branches of the same bank are competing with each other. In practice, what branches of the same bank can do may be limited by headquarters, a different model, and one that limits the impact of more branches.

illustrative example of our spatial data, we plot the position of all villages and FSPs in 1999 for the province of Chacheongsao in Figure 11.

Figure 11: Villages and Banks in Chacheongsao Province



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

In the Townsend Thai Data, the Monthly Resurvey data consists of data collected for 531 households in 16 villages of 4 provinces. These provinces are Chacheongsao, Lopburi, Buriram, and Sisaket. The provinces of Buriram and Sisaket 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 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 considered. All values are in nominal terms. Table 7 exhibits the summary statistics. As pointed out by Karaivanov and Townsend (2014), an important characteristic of the data is that correlations between income, consumption, and capital indicate significant consumption smoothing but are still far 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 in these villages). The error from this assumption will enter our model through the location-specific shocks.

Table 7: 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. The 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 M.1, we estimate the parameters assuming that FSPs contract with SMEs under both obstacles, MH + LC. We convert all data from Thai currency into 'model units' by dividing all currency values by the 90th percentile of the assets distribution in the sample (approximately 180,000 Thai baht). We use the parameters and grids to compute the frontier as in Table 4. 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 8. The estimate for the measurement error γ_{ME} is 21 %. This corresponds to measurement error with a 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 the number of FSPs in each potential location relatively well.

To understand how relevant spatial costs are, we compute how much a household that pays zero travel costs to every bank, as if complete digital lending, would have to receive to be at the median distance when spatial costs are still imposed. This is the result of Eq.(38). See Appendix O for details. As we use CRRA preferences, this measure depends on this household's initial consumption level. 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, a relatively large number.

Note that this is different than the counterfactual exercise on changing spatial costs below, as here, we keep the utilities played by FSPs constant; that is, we consider a unilateral move of one household that does not impact the equilibrium.

$$\psi^u(\psi) \equiv \left[\frac{\psi \, \text{med}\,(t(x_v, x_b))}{u(\bar{c})} + 1\right]^{\frac{1}{1-\sigma}} - 1 \tag{38}$$

Table 8: 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 M.1 and Appedix M for details. All coefficients are significant at 1% .

We take the legacy baseline structure as given and feature counterfactuals associated with new technology or policy changes. We denote the equilibrium at our parameter estimates as our baseline and showcase in our results percentage deviations from 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.(39). X' is the value after the change and X_0 the baseline. For welfare, we plot the average utility of the 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 D).

$$100\left(\frac{X'}{X_0} - 1\right) \tag{39}$$

To interpret the changes in welfare, in addition to changes in the levels of utility, we have a consumption equivalent measure; that is, 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 as the average consumption in our sample, as above. For details, see Appendix O.

Decreasing Spatial Costs The results of changing spatial costs ψ are are in Table 9. The averages and standard deviation are computed at the village level (after aggregating for households and different levels of capital) and averaged for different provinces. [30] The results of the transformation from welfare to consumption are in Table 12, where we repeat the welfare numbers and compute the consumption equivalent change.

By reducing spatial costs by 50%, welfare increases by 4.85% (First Column of Table 12). Due to the 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 utility increases 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 are growing), and, as expected, the standard deviation across villages of the share of SMEs that use financial services is reduced.

Table 9: 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. 39) with respect to the baseline of $\hat{\psi}$ and $\hat{\sigma_L}$ in Table 8. Averages and standard deviations are 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 .

Reducing the Logit Variance/ A Platform to Facilitate Comparison Across Providers. The results of changing the logit variance σ_L are are in Table 10. The aggregation and conversion

[30]The standard deviation is not those from the parameters but the standard deviation across the average of different villages.

from welfare and consumption equivalents are made in the case of spatial costs. By reducing the logit variance by 50%, welfare increases by 9.20%, which corresponds to 15.36% in consumption equivalent terms (last two rows of the first Column of Table 12). 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.

Table 10: 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. 39) with respect to the baseline of $\hat{\psi}$ and $\hat{\sigma}_L$ in Table 8. Averages and standard deviations are 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 in the Legacy Baseline. Our last counterfactual computes changes in model outcomes after bank entry. We compute the average outcome of one bank entry in each potential location in each province. The results are in Table 11. An extra bank increases utilities on average by 2.16%, which translates to a 2.2% equivalent change in consumption. For those previously intermediated, consumption increases. Note also that more households do get served (increase in average market share). 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 agents' utilities, but in different magnitudes, the number declines across these experiments. Adding another legacy bank while keeping technology the same has the least

Table 11: 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 a percentage change (Eq. 39) with respect to the baseline of $\hat{\psi}$ and $\hat{\sigma_L}$ in Table 8. Averages and standard deviations are 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 .

impact. Although any reduction in spatial costs (as associated with increased access coming from digitization) is more relevant for individual agents, this, in turn, is less relevant in determining overall welfare relative to reducing the logit variance (for changes of the same magnitude). Our results suggest that to increase welfare; policymakers should guarantee that market shares change when utility offerings change. This means that the goal of policymakers should be 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.

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 is to develop a framework in terms of utilities generated by contracts rather than the contracts themselves and divide the contracting and competition problems into building blocks. This allows

Table 12: 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 O, Eq. (185). Percentage changes (Eq. 39) with respect to the baseline of $\hat{\psi}$ and $\hat{\sigma}_L$ in Table 8. Averages and standard deviations are 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.

us to apply most of the competition toolbox to potentially complex models of competition and contracting.

We focus our analysis on contracting between 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 the 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 variations in spatial configuration and competition across markets. This allows a researcher to conduct market structure counterfactuals without taking 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 agents' utilities but at different magnitudes. Our results suggest that policymakers should focus on mechanisms that guarantee market shares change when utility offerings change. These could be achieved through policies that make 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 aim to extend the model to include the savings side of intermediation, in which households and SMEs could choose to offer investment products. This also introduces several complications. Regulators typically do not allow payment providers to intermediate directly; that is, offered savings and investment accounts cannot include the loans of their own portfolio. Regulators also tend to protect incumbent commercial banks that provide deposit accounts and tend to be less innovative in their business models. Second, we believe our methods could be applied to other markets and more developed countries (e.g., the health market in the U.S.). Third, we haven't explored the issue of dynamics - both in contracting and competition of FSPs- here, which may be relevant in various settings. Fourth, the case of AdS can still be explored further. Several implications of our comparative statics exercises (within-village inequality, systemic risk, etc.) are not fully understood yet, so our framework could prove useful in both theoretical models and empirical applications.

References

Adegboyega, G. A., O. FapohundaKofoworola, and J. O. Famoriji (2014). Impact of transmission distance on the strength of received signals within the vicinity of four base stations. *American Journal of Engineering Research (AJER) 3*(1), 272–279.

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.

Bennett, O. (2020). Bank branches: why are they closing and what is the impact? Briefing Paper CBP 8740, House of Commons Library.

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.

Buera, F. J., J. P. Kaboski, and Y. Shin (2011). Finance and development: A tale of two sectors. *American economic review 101*(5), 1964–2002.

Buera, F. J. and Y. Shin (2013). Financial frictions and the persistence of history: A quantitative exploration. *Journal of Political Economy 121*(2), 221–272.

Cole, H. L., J. Greenwood, and J. M. Sanchez (2016). Why doesn't technology flow from rich to poor countries? *Econometrica 84*(4), 1477–1521.

Crawford, G. S., N. Pavanini, and F. Schivardi (2018, July). Asymmetric information and imperfect competition in lending markets. *American Economic Review 108*(7), 1659–1701.

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.

Dunn, A. (2009). Does Competition Among Medicare Advantage Plans Matter? An Empirical Analysis of the Effects of Local Competition in a Regulated Environment. Dep't of Justice Economic Analysis Group Discussion Paper.

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.

Gao, H., H. Ru, R. M. Townsend, and X. Yang (2019). Rise of bank competition: Evidence from banking deregulation in china. *Available at SSRN 3087081*.

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 (2017). The welfare effects of long-term health insurance contracts. *Unpublished manuscript, MIT*.

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

Itskhoki, O. and B. Moll (2019). Optimal development policies with financial frictions. *Econometrica 87*(1), 139–173.

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*.

Mahoney, N. and E. G. Weyl (2014). Imperfect Competition in Selection Markets. SSRN 2372661.

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

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.

Ridhwan, Hidayat, Ismail, and A. Suryahadi (2024, October). How financial digitalization affects business performance and business innovation among ultra-micro, micro and small enterprises: Evidence from indonesia. Technical report, BI Institute and SMERU Research Institute.

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

Wolfe, B. and W. Yoo (2018). Crowding out banks: credit substitution by peer-to-peer lending.

Appendix

A A Model of Risky Production and Intermediation

This section discusses the model that generates the cash expenditure-production puzzle (Section 2.1). We first go into the details of the model. Given the model setup, 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:

$$p_i(\theta_i) = \theta_i\left(1 + \frac{\sigma}{\sqrt{\theta_i}}\varsigma_i\right) \tag{40}$$

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:

$$u(p_i(\theta_i)) = \mathbb{E}\left[p_i(\theta_i)\right] - A\mathbb{V}\left[p_i(\theta_i)\right] \tag{41}$$

where A is a measure of risk aversion, $\mathbb{E}[.]$ denotes the expectation over ε_i and $\mathbb{V}$ denotes the variance. For notation purposes, we define the utility under autarky of type θ_i to be given by $u_A(\theta_i)$. Our model is static, so cash expenditure (consumption + investment) is equivalent to consumption.

Autarky. Under autarky, each agent has to consume its production. Using the process in Eq. (40) and substituting in Eq. (41)

$$u(p_i(\theta_i)) = \theta_i\left(1 - A\sigma^2\right) \ \ (< \theta_i) \tag{42}$$

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

$$c_A(\theta_i) = \theta_i \text{ and } s_A(\theta_i) = \sqrt{\theta_i}\sigma \tag{43}$$

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

$$p_A(\theta_i) = \theta_i \tag{44}$$

Financial Intermediation with Full Information. FSPs can provide credit that allows entrepreneurs to increase production and simultaneously zero out the production risk (insurance). [31]. That is, they can transform

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

the production process $p_i(\theta_i)$ into an intermediated process $p_I(\theta_i)$, as in Eq. (45), which has no uncertainty.

$$p_I(\theta_i) = \theta_i \lambda_i, \text{ where } \lambda_i > 1 \tag{45}$$

Providing credit is not costless for FSPs (cost of raising deposits, balance sheet constraints, etc.). We assume 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. (46)

$$\max_{\lambda_i, t_i} \left[t_i - .5\lambda.^2 \right] D(\theta_i, \lambda_i, t_i) \tag{46}$$

where $D(.)$ is a demand with constant elasticity ε given by Eq. (47)

$$D(\theta_i, \lambda_i, t) \equiv (\theta_i \lambda_i - t - u_A(\theta_i))^{\varepsilon} \tag{47}$$

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 to balance 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 translate into market power from the FSPs.

The problem of the FSPs in Eq. (46) implies the optimal contract and transfers as given by Eq. (48) (See Appendix F for details).

$$\lambda_i \equiv \lambda(\theta_i) = \theta_i \text{ and } t(\theta_i) = \frac{1 + .5\varepsilon}{1 + \varepsilon}\theta_i^2 - \frac{1}{1 + \varepsilon}u_A(\theta_i) \tag{48}$$

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. (49)

$$u_I(\theta_i, \varepsilon) = \frac{\varepsilon}{1 + \varepsilon}\left(.5\theta_i^2\right) + \frac{1}{1 + \varepsilon}u_A(\theta_i) \tag{49}$$

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. (49), we have that

$$u_I(\theta_i, \omega) = (1 - \omega)\left(.5\theta_i^2\right) + \omega u_A(\theta_i) \tag{50}$$

Eq. (50) 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 determine how the intermediation gain is divided and, 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. [32]

Finally, note that under an intermediation regime, the average and standard deviation in consumption within types θ_i, denoted by $c_I(\theta_i)$ and $s_I(\theta_i)$, respectively, is given by Eq. (51)

$$c_I(\theta_i) = u_I(\theta_i, \omega) \text{ and } s_I(\theta_i) = 0 \tag{51}$$

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.51), and autarky, denoted by c_A (Eq. 43)

$$c_I(\theta_i) - c_A(\theta_i) = (1-\omega_1)\left(.5\theta_i^2\right) + \omega_1 u_A(\theta_i) - \theta_i = (1-\omega_1)\left(.5\theta_i^2 - \theta_i\right) - \omega_1 \theta_i A \sigma^2 \tag{52}$$

As we are assuming the researcher has a perfect experiment, it is the case that the sample analogue converges to the theoretical difference in probability (i.e., the problem is not the statistical estimator). The issue 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, 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 it decreases average consumption. If there is enough risk in a project, the agent is too risk averse, or the FSP has enough market power to keep production rents to itself, average consumption potentially is reduced. Overall, the effects of welfare are underestimated (even if, in the sample, it is true that consumption increases as 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:

$$c_A \equiv \sum_i c_A(\theta_i) f(\theta_i) \text{ and } c_I \equiv \sum_i c_I(\theta_i) f(\theta_i) \tag{53}$$

[32]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 microdata 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.

From Eq. (52), we have that the difference in average consumption averaged across types in theory is given by

$$c_I - c_A = \sum_i \Big\{(1-\omega)\left(.5\theta_i^2\right) + \omega u_A(\theta_i) - \theta\Big\} f(\theta_i) \tag{54}$$

$$= (1-\omega)\left(.5\mathbb{V}_\theta + .5\mathbb{E}_{\theta_i}^2 - \mathbb{E}_\theta\right) - \omega \mathbb{E}_\theta A\sigma^2 \tag{55}$$

where $\mathbb{E}_\theta, \mathbb{V}_\theta$ denote, respectively, expectation and variance of θ in the population. Now, not only do we confound the parameters of contracting and intermediation (and potentially give a wrong signal for average consumption), but our results also depend on the variance of θ due to its heterogeneous effects across agents. Two regions that are 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 the 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.51), and autarky, denoted by S_A (Eq. 43) is given by

$$s_I(\theta_i) - s_A(\theta_i) = -\sqrt{\theta_i}\sigma \tag{56}$$

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 the 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

$$s_I(\theta_i) - s_A(\theta_i) = \sqrt{\mathbb{V}_\theta} - \sqrt{\mathbb{V}_\theta + \sigma^2 \mathbb{E}_\theta} \tag{57}$$

if $\mathbb{V}_\theta$ is large with respect to σ, even in this lower bound case, the coefficient of variation (mean over standard deviation) may increase simultaneously 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 of 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. (44) and Eq. (45) we have that

$$p_I(\theta_i) - p_A(\theta_i) = \theta_i(\theta_i - 1) \tag{58}$$

which is positive due to our assumption of $\theta_i \geq 1$. In this case, production is increasing by looking at the microdata and comparing the two samples. In this case, the interpretation of the effects of financial intermediation becomes even murkier since consumption is potentially reducing while production is increasing. In particular, note that we can substitute Eq. (58) in Eq.(52) to obtain

$$c_I(\theta_i) - c_A(\theta_i) = (1-\omega_1).5\left(p_I(\theta_i) - p_A(\theta_i)\right) - \theta_i\Big\{.5(1-\omega_1) - \omega_1 A\sigma^2\Big\} \tag{59}$$

which means that $(p_I(\theta_i) - p_A(\theta_i))$ is endogenous in Eq.(59). 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 intermediation agents are producing a larger income (Eq. 58), 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.(52). 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

$$c_I(\theta_i;\omega_2) - c_I(\theta_i;\omega_1) = (\omega_1 - \omega_2)\underbrace{\left[.5\theta_i^2 - \theta_i(1-\sigma^2 A)\right]}_{\equiv g_I(\theta_i)} \tag{60}$$

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 the total amount of extra utility this economy generates 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 these intermediation gains. The reason is that there is no risk and the same level of credit in both scenarios, 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 on in this 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 depends on utility specification and production function, consumption differences identify differences in market power. Suppose the researcher does not have this model. In that case, 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

$$c_I(\theta_i;\omega_2) - c_I(\theta_i;\omega_1) = (\omega_1 - \omega_2)\underbrace{\left[.5\theta_i^2 - \theta_i(1-\sigma^2 A)\right]}_{\equiv g_I(\theta_i)} + p_I(\theta_i, nu_2) - p_I(\theta_i, \omega_1) \tag{61}$$

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 13. 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.

We use the parameters in Table 13 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.

Table 13: 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

B From Adverse Selection to Full Information

We focus now on the model 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 there are two types, which are now unobserved by the FSP and the researcher. In the incomplete information case, the problem of the FSP becomes a generalized version of Eq. (46), where we also take into account the truth-telling constraints. The problem of an FSP is now given by Eq.(62)

$$\max_{\{\lambda(\theta),t(\theta)\}_\theta} f(\theta_L)D(\theta_L,\lambda_L,t_L)\left[t_L - .5\lambda_L^2\right] + f(\theta_H)D(\theta_H,\lambda_H,t_H)\left[t_H - .5\lambda_H^2\right] \tag{62}$$

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

$$\theta_L\lambda_L - t_L \geq \theta_L\lambda_H - t_H \tag{63}$$

$$\theta_H\lambda_H - t_H \geq \theta_H\lambda_L - t_L \tag{64}$$

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

$$\sigma^2 A = 1$$

which guarantees that autarky utilities of both agents now are zero (See Eq. (42)). The truth-telling constraints - Eqs.(63)-(64) - are not binding whenever (See Appendix F.1) :

$$\omega \leq \frac{\theta_H - \theta_L}{\theta_H + \theta_L} \tag{65}$$

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 trade surplus - 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 no 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 understanding 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 levels of market power. We use subscripts L,H for the credit and utility of each type. For that, we solve the problem of the FSPs of maximizing Eq.(62) subject to the truth-telling constraint in Eq. (63)-(64) for various levels of ω, the market power. We use the parameters in Table 14.

Table 14: 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.

We plot the leverage (total credit provided, λ_i) of low and high types chosen by the FSP in Figure 12. The vertical line shows where the constraint is binding - the minimum value of μ such that Eq. (65) 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 more or less leverage under adverse selection. The allocation for the high type, however, is never distorted. This results from the fact that in the parameters we use, Eq.(64) 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 F 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 13. 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.

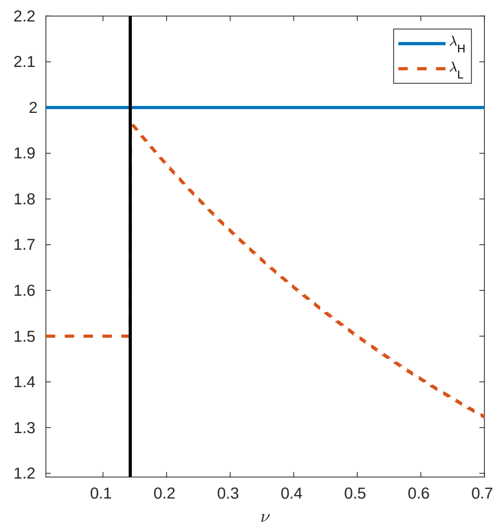
With a well designed experiment and observing utility (with the caveats we discussed in the other example), a researcher can infer what is the effect moving from Adverse Selection to Full Information for a given level of ω. However, we cannot interpret the results more generally, i.e., beyond the studied 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.

C 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

Figure 12: Leverage of Low and High Types



Leverage consistent with the solution to the problem of the FSPs of maximizing Eq.(62) subject to the truth-telling constraint in Eq. (63)-(64) 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 $\sigma^2 A = 1$. The vertical line represents the point at which truth-telling constraint starts to bind - the minimum value of μ such that Eq. (65) is violated .

contract depends on realizations of stochastic variables. For any contract, we also assume that we can specify the profit of an 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 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* $\forall\, \boldsymbol{C}_0, \boldsymbol{C}_1 \in \mathfrak{C}$:

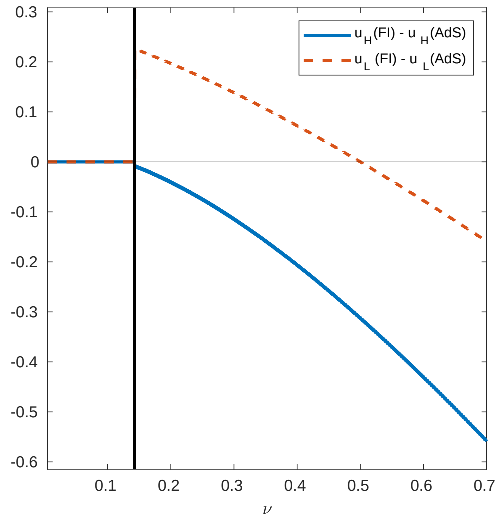
$$\Pi(\boldsymbol{C}_0) > \Pi(\boldsymbol{C}_1) \Leftrightarrow \mathbb{U}(\boldsymbol{C}_0) < \mathbb{U}(\boldsymbol{C}_1) \tag{66}$$

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

$$\nexists\, \boldsymbol{C}_0, \boldsymbol{C}_1 \in \mathfrak{C} \text{ s.t . } \boldsymbol{C}_0 \neq \boldsymbol{C}_1,\ \Pi(\boldsymbol{C}_0) = \Pi(\boldsymbol{C}_1) \text{ and } \mathbb{U}(\boldsymbol{C}_0) = \mathbb{U}(\boldsymbol{C}_1) \tag{67}$$

Eq.(66) means that if a contract is more profitable for the FSP, it provides less utility for the agent. What we are ruling 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, a few cases may not hold. Eq.(67) rules out a contract equivalent for agents and FSPs simultaneously. This assumption is true in our application, where agents are risk-averse, and FSPs are risk-neutral. If there were two equivalent contracts for the agents, there would be a Pareto improvement of

Figure 13: Utility of Low and High Types



Note: Utility consistent with the solution to the problem of the FSPs of maximizing Eq.(62) subject to the truth-telling constraint in Eq. (63)-(64) 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 $\sigma^2 A = 1$. The vertical line represents the point at which truth-telling constraint starts to bind - the minimum value of μ such that Eq. (65) is violated .

offering the mean contract [33] 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. (67) guarantees they do not have the same profits. Eq. (66) 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 exist under *Assumption U*. What this means is that the optimization problem in Eq. (68) is well defined:

$$c^*(u) \in \arg\max_{\boldsymbol{C}\in\mathfrak{C}} \pi(\boldsymbol{C}) \quad \text{s.t . } \mathbb{U}(\boldsymbol{C}) = u \tag{68}$$

and that, $\forall\ \boldsymbol{C}_0 \in \mathfrak{C}$, the solution of the Eq. (68) is s.t.:

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

i.e., a one-to-one mapping from contracts to utilities is 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.(68) 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.

Lemma 3.1 and Nash Equilibrium

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

We provide an intuitive explanation. 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. (69)

$$\varphi_b(a) \equiv \{u_b - a, u_{-b} - a, u_0 - a, x_b, x_{-b}, \{x_v\}_v\} \tag{69}$$

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

The idea behind Lemma 3.1 can be represented pictorially. Imagine that both μ and 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. (70)

$$-\frac{\partial_{u_b} S(u_b^*)}{S(u_b^*)} = \frac{\partial_{u_b}\mu\left(u_b^*, u_{-b}, u_0, x_b, x_{-b}, \{x_v\}_v\right)}{\mu\left(u_b^*, u_{-b}, u_0, x_b, x_{-b}, \{x_v\}_v\right)} \tag{70}$$

In Eq. (70), 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.(70) is strictly decreasing, while the concavity of S guarantees that the LHS is increasing. One can see this trade-off in Figure 14.

Consider that we are in an equilibrium $\{u_b^*\}_b$. Let's 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. (70) 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.(23) 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.

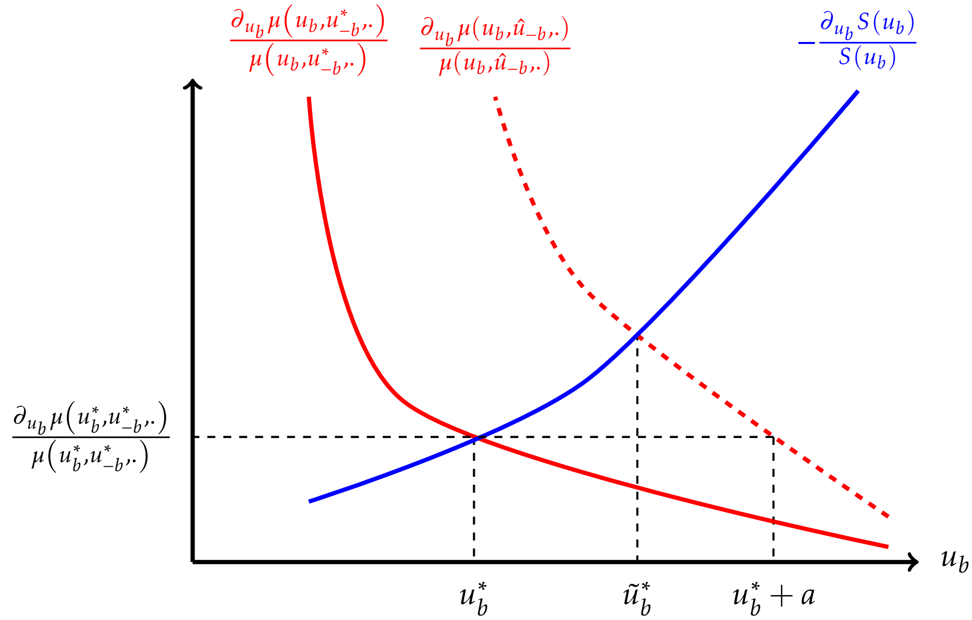
Jointly, the monotonicity and discounting guarantee a unique equilibrium that can be computed through an iteration of best response functions.

The application of Lemma 3.1 is not straightforward. For example, the logit itself is log-concave, but the sum of logits in Eq. (20) may not be if there is enough variation in market shares across villages in a given market. This means we have to bind either 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). In Appendix H, we provide a sufficient condition that guarantees the log-concavity even with the spatial costs $\psi > 0$. In practice, all parametric values we tested satisfy this condition.

D Comparative Statics

In Appendix L.1, we show the equivalent results for changes in the logit variance, σ_L. 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 are higher. Contrary to what we see with spatial costs, this effect

Figure 14: Nash Equilibrium: Monotonicity and Discounting

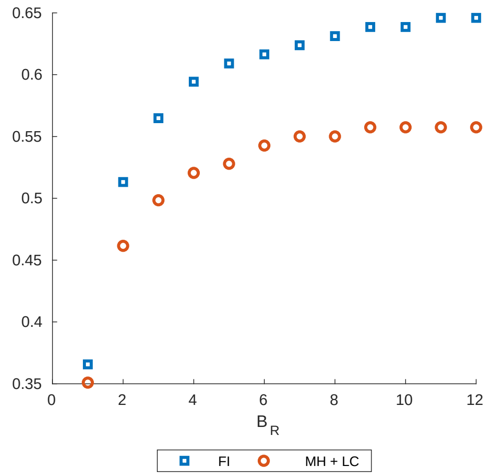


Note: Pictorial representation of maximization of one FSP abstracting from technical details (non-differentiability, corner solutions, etc.) as in Eq. (70). The equilibrium, u^*, is where the marginal benefit is equal to the marginal cost (for all FSPs, although the picture denotes only one). u^* is the baseline equilibrium, and $\tilde{u}_b = u^*_b + a$ is the best response to the deviation if all other FSPs increase their utilities played by a.

is homogeneous across all villages. For larger values of σ_L, we observe a smaller utility difference across all villages (which we know how to match to consumption data, for instance, as in Figure 3), whereas, in contrast, high spatial costs should lead to dispersion in utilities across villages in a given market. In Appendix L.2, we simultaneously vary both spatial costs ψ and the logit variance σ_L to understand if they are complements or substitutes (and how this changes with the level of competition among FSPs). The effects of reducing spatial costs are more pronounced with lower values of σ_L. Reduction in spatial costs is passed through more to consumers when competition is higher. The difference is at its highest when either ψ or σ_L are sufficiently small.

Local Competition and Information Structure. 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 $b_R = 1, \ldots, 8$. We do this under FI and and MH + LC to highlight the interaction of local competition and information structure. The results are in Figure 15. Not only is the *level* of utility higher under full information so that clients benefit from the alleviation of obstacles to trade but the *gains* from an increasing number of providers is larger. Thus, digital data, which can reduce moral hazard, increases the added impact of competition among FSPs.

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.(22) for FSPs located in $x = 1$. Equilibrium with the spatial configuration of Figure 4 and parameters of Table 5, changing the number of FSPs in $x = 1$, denoted by b_R. Contracting frictions are FI (blue curve) and MH + LC (red curve).

E Simple Model of Adverse Selection

We consider here 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: the cost of exerting effort, etc. In particular, we focus on the utility function in Eq. (71) [34]

$$\mathbb{U}(c,z|\theta) \equiv u(c) - \theta v(z), \;\; \theta \in \Theta \tag{71}$$

Eq.(71) 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 [35]

$$\sum_{c,q,z,k'} \pi^{\theta}(c,q,z,k')\mathbb{U}(c,z|\theta) \geq \sum_{c,q,z,k'} \pi^{\hat{\theta}}(c,q,z,k')\mathbb{U}(c,z|\theta), \;\; \forall \theta, \;\; \forall \hat{\theta} > \theta \tag{72}$$

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 model's parameters (as, for instance, the share of

[34] 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 .

[35] 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. (71) .

each type in the population) determine the incentives of FSPs to distort the allocation across types. For more details, 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 E.1, a version of Lemma 3.1 for the case of AdS. The intuition behind Lemma E.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 simultaneously, the FSP ends up doing this. Thus, the equilibrium still satisfies monotonicity. Analogous reasoning shows discounting.

Lemma E.1. *Assume that all conditions for Lemma 3.1 are true. Additionally, assume that the utility function is as in Eq.(71). Finally, assume that only the truth-telling constraints that potentially bind are those in Eq.(72). Then, a Nash Equilibrium in utilities exists, is unique, and can be computed iteratively.*

Proof. The proof is comprised of two steps. The first establishes that we can simplify the set of truth-telling constraints - Eq. (26) - 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:

$$\pi(\theta) \cdot U(\theta) \equiv \sum_{c,q,z,k'} \pi^{\theta}(c,q,z,k^{'})\mathbb{U}(c,z|\theta) \tag{73}$$

$$\pi(\theta) \cdot U(\theta) \cdot P(\theta,\hat{\theta}) \equiv \sum_{c,q,z,k'} \pi^{\theta}(c,q,z,k^{'})\mathbb{U}(c,z|\theta)\frac{P(q|k^{'},\theta,z)}{P(q|k^{'},\hat{\theta},z)} \tag{74}$$

Finally, denote the truth-telling condition between the 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

$$\pi(\theta_i) \cdot U(\theta_i) \geq \pi(\theta_{i-1}) \cdot U(\theta_i) \tag{75}$$

$$\pi(\theta_{i-1}) \cdot U(\theta_{i-1}) \geq \hat{\pi} \cdot U(\theta_{i-1}) \tag{76}$$

Then, we can write

$$\pi(\theta_i)\cdot U(\theta_i) \geq \pi(\theta_{i-1})\cdot U(\theta_i) \tag{77}$$
$$= \pi(\theta_{i-1})\cdot[U(\theta_i)-U(\theta_{i-1})]+\pi(\theta_{i-1})\cdot U(\theta_{i-1}) \tag{78}$$
$$= [\pi(\theta_{i-1})-\hat{\pi}]\cdot[U(\theta_i)-U(\theta_{i-1})]+\pi(\theta_{i-1})\cdot U(\theta_{i-1})+\hat{\pi}\cdot[U(\theta_i)-U(\theta_{i-1})] \tag{79}$$
$$= \hat{\pi}\cdot U(\theta_i)+[\pi(\theta_{i-1})-\hat{\pi}]\cdot[U(\theta_i)-U(\theta_{i-1})]+[\pi(\theta_{i-1})-\hat{\pi}]\cdot U(\theta_{i-1}) \tag{80}$$

Therefore:

$$[\pi(\theta_i)-\hat{\pi}]\cdot U(\theta_i) \geq \underbrace{[\pi(\theta_{i-1})-\hat{\pi}]\cdot[U(\theta_i)-U(\theta_{i-1})]}_{\equiv I}+\underbrace{[\pi(\theta_{i-1})-\hat{\pi}]\cdot U(\theta_{i-1})}_{\equiv II} \tag{81}$$

We know that $II > 0$ since $\hat{\pi}$ is not preferred by θ_{i-1} (Eq. (76)). 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. Therefore,

$$\pi(\theta_i)\cdot U(\theta_i) \geq \pi(\theta_{i-1})\cdot U(\theta_i) \text{ and } \pi(\theta_{i-1})\cdot U(\theta_{i-1}) \geq \pi(\theta_{i-2})\cdot U(\theta_{i-1})$$
$$\Rightarrow \pi(\theta_i)\cdot U(\theta_i) \geq \pi(\theta_{i-j})\cdot U(\theta_i),\ j>0 \tag{82}$$

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

$$0 \leq [\pi(\theta_i)-\pi(\theta_{i-1})]\cdot[U(\theta_i)-U(\theta_{i-1})] = (\theta_i-\theta_{i-1})\,[\pi^z(\theta_{i-1})-\hat{\pi}^z]\cdot v(z) \tag{83}$$

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. [36]. 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 an FSP is

$$\sum_i \Big\{\partial_{u_b(\theta_i)}S(u_b(\theta_i))\mu(u_b(\theta_i),u_{-b}(\theta_i))+S(u_b(\theta_i))\partial_{u_b(\theta_i)}\mu(u_b(\theta_i),u_{-b}(\theta_i))\Big\}f(\theta_i)\partial_{u_H}\mathcal{U}(\theta_i\,|u_H)=0 \tag{84}$$

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.(84) as Eq. (85)

$$\sum_i \mathcal{F}(u_b(\theta_i),u_{-b}(\theta_i))\partial_{u_H}\mathcal{U}(\theta_i\,|u_H)=0 \tag{85}$$

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

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. (81)

$$[\pi(\theta_i) - \hat{\pi}] \cdot U(\theta_i) \geq [\pi(\theta_{i-1}) - \hat{\pi}] \cdot [U(\theta_i) - U(\theta_{i-1})] + [\pi(\theta_{i-1}) - \hat{\pi}] \cdot U(\theta_{i-1}) > a \tag{86}$$

where the inequality comes from Eq. (83). 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 $\hat{u}_{-b}(\theta_i) \geq u_{-b}(\theta_i)$:

$$\sum_i \mathcal{F}(u_b(\theta_i), \hat{u}_{-b}(\theta_i))\partial_{u_H}\mathcal{U}(\theta_i|u_H) > 0 \tag{87}$$

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

$$\sum_i \mathcal{F}(u_b(\theta_i) + \hat{a}_i, a + u_{-b}(\theta_i))\partial_{u_H}\mathcal{U}(\theta_i|u_H + a) < 0 \tag{88}$$

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. ■