# Spatial competition among financial service providers and optimal contract design

Robert M. Townsend and Victor V. Zhorin

May 17, 2013

PRELIMINARY AND INCOMPLETE

**Abstract**

We present a contract-based model of industrial organization that allows us to consider in unified way both different information frictions (moral hazard, adverse selection) and a variety of market structures (monopoly, perfect and imperfect competition, strategic interaction). We show how this method can be applied to banking and insurance industries. Transitions from extractive local monopoly organizations to globally competitive inclusive environments are discussed alongside with frictions affecting the outcome of such transitions.

## 1    Introduction

We have in mind the local environment of a county or province in a developing country, such as Thailand, but what we observe is typical of emerging market countries as in Brazil and especially low income countries such as Bangladesh with a well-developed micro finance industry. Typically banks in developing countries are sparsely located on the ground, with relatively few branches and few banks operating in a given area. Travel to branches is non trivial in terms of time and repeat customer visits as well as visits of credit officers to the field. We thus focus on the absence of centralized markets and focus on bank lending and competition among relatively few banks. For that matter the actual structure of observed bank contracts (credit and insurance arrangements for households and small and medium, SME businesses) is not simple, i.e. does not fit the stylized contracts of theory, of borrowing at interest with collateral and fixed term payments, with presumed repayment but allowing for default. Rather typical contracts offered by banks represent a blend of credit and insurance, e.g, loans are rolled over, some interest is forgiven, and indeed there are well known and explicit contingencies under which an effective indemnity is paid and some or all of principal is written off (as if paid with the indemnity). More complicated dynamic environments allow periodic and randomized audits triggered by signals of borrower performance with loans terms and credit lines a function of both the audit and publicly observed index of sector wide performance.

Our motivation for this research is both positive and normative. On the positive side we seek to understand better the industrial organization of financial service providers in terms

of both the geography of branches and expansion over time as well as in terms of the actual loan/insurance contracts which are offered[1]. On the normative side, we seek to answers policy questions such as the coexistence of local and national banks and the role of information and competition (Petersen and Rajan (1995)): the impact of deregulation which alleviates artificial geographic or policy/segmentation boundaries (Brook et al. (1998), Demyanyuk et al. (2007)); the interplay between competition among banks and branches and financial stability (Nicoló et al. (2004), Nicoló and Boyd (2005), Martinez-Miera and Repullo (2010)); and the welfare and distributional consequences of different market structures, different obstacles to trade (information, trade costs) (Koijen and Yogo (2012), Martin and Taddei (2012)) and the interaction of these obstacles with market structure.

The setting and two previous strands of the literature are coming together here in this paper. One line of research of Karaivanov and Townsend (2012) shows how to estimate financial/information regimes for SME's, distinguishing moral hazard constrained lending and insurance, as in urban areas, versus more limited contracts, buffer stock savings with bounds on borrowing, as in rural areas, using Townsend Thai project data on consumption, income, investment, and capital stock, at a point in time and over time as in the panel.

A second line of research of Assuncao et al. (2012) uses data on the timing and location of the opening of new branches for both the commercial banking sector and government banks (in the same setting, Thailand). When there are only a few branches around, households would need to travel relatively long, time consuming distances to get to a branch or choose to not participate in the (formal) financial system. As new banks/branches enter, the market catchment areas effectively evolve. The key point is that a "market" is not a fixed object with heterogeneous characteristics and the environment is not modeled as being in a steady state.

Here we report on work to bring these these two strands together with both the location of bank branches and the contracts they offer as endogenous (though our framework allows for regulatory restrictions if we choose to further restrict the environment exogenously), to match the contracts we see in reality and allow for those we do not see out of equilibrium[2]. Specifically this paper is devoted to developing methods that could potentially be applied, not only in Thailand, but in other countries as well.

---

[1]Agarwal and Hauswald (2010) study the effects of physical distance on the acquisition and use of private information in credit markets. Rajan and Petersen (2002) document that the distance between small firms and lenders is increasing. Alessandrini et al. (2009b) show show that greater functional distance stiffened financing constraints, especially for small firms. Butler (2008) suggests that investment banks with a local presence are better able to assess private information and place difficult bond issues. Degryse and Ongena (2005) report the comprehensive evidence on the occurrence of spatial price discrimination in bank lending. Alessandrini et al. (2009a) show that small and medium enterprises (SMEs) located in provinces where the local banking system is functionally distant are less inclined to introduce process and product innovations, while the market share of large banks is only slightly correlated with firms propensity to introduce new products.

[2]Our work is close in spirit to the work of Einav et al. (2010), Einav et al. (2013) on health care and Einav et al. (2012) on auto loans except that we try to make few restrictions on contracts to see how far we can get and we add in the supply, competition in financial services

# 2 Micro foundations: contracts for household enterprise

We consider an economy populated by spatially distributed SME's, output-producing agents and financial intermediaries. By default all agents are in an autarky regime, as a reservation strategy. They produce an output using a stochastic technology and that output is fully consumed by the agents.

The agents also can choose to contract with an intermediary. Intermediated agents can augment their effort by borrowing capital $k$ from intermediaries[3]
Agents have preferences

$$u(c, a|\theta),$$

where $c$ is consumption and $a$ is an action that can be either observable or hidden. Utility is strictly concave featuring risk aversion hence optimal contracts offer insurance, not simply credit. In principal we can use any parameterized utility function, such as CARA below. We do not have to rule out wealth effects. The bank keeps all surplus left from the output $q$. The output is fully observable and the contract can be made conditional on output. There is a set of stochastic production technologies available to intermediated agents

$$P(q = high|k, a, \theta) = p(q = high|a)f(\theta)k^\alpha; P(q = low|k, a, \theta) = 1 - P(q = high|k, a, \theta) \quad (1)$$

where $P(q|k, a, \theta)$ is a probability to reach the output $q$ that depends on agent's type $\theta$ and the effort $a$ exercised by an agent. Here $\theta$ stands for observed (and potentially unobserved) characteristics of the household/SME. We can also use non parametrically estimated production function if relevant empirical data is available. Depending on informational frictions we consider, thus either the type or the effort could be unobservable leading to either moral hazard or adverse selection. We can allow correlation of $\theta$-types in preferences and in production, more specific interpretation of $\theta$-heterogeneity will be discussed later, types can be either observable or not. Heterogeneity in wealth endowments is possible as well.

We begin with out basic building block, as if there is one lender. The optimal contract offered to the agents maximizes bank surplus extracted from each agent:

$$S^{\overline{\omega(\theta)}} := \underset{\pi(q,c,k,a|\theta)}{\text{maximize}} \left[ \sum_{q,c,k,a,\theta} \pi(q, c, k, a|\theta)[q - c - k] \right] \quad (2)$$

where $\pi(q, \underline{c}, k, a|\theta)$ is a probability distribution[4] over the vector $(q, c, k, a)$ given the agent's type $\theta$ and $\overline{\omega(\theta)}$ is specific utility offered to the agents by the bank.

---

[3]To give intermediation an extra advantage, output is higher in the intermediated sector. This is not essential but it does make the utility gain over autarky larger, making it easier to display solutions. There is also risk-sharing available through the intermediaries mitigating potential downside risk of failure. However, the output is effectively under control of the intermediaries who structure the contract in a way to maximize their own profit.

[4]We use Prescott-Townsend lotteries over discrete grids to find solution, it is also possible to write down the problem in general integral form typically used in contract theory

We put the following constraints in place.

- Participation Constraints (participation must be voluntary):
  $\forall\, \theta \in \Theta$

$$\sum_{Q,C,K,A} \pi\left(q,c,k,a|\theta\right) u\left(c,a|\theta\right) \geq u_0\left(\theta\right)$$

  where $u_0\left(\theta\right)$ is the autarky utility.

- Utility Assignment Constraints (UAC):
  $\forall\, \theta \in \Theta$

$$\sum_{q,c,k,a} \pi(q,c,k,a|\theta)u(c,a|\theta) = \overline{\omega(\theta)} \tag{3}$$

- Mother Nature/Technology Constraints:
  $\forall\, \left\{\overline{q},\overline{k},\overline{a}\right\} \in Q \times K \times A$ and $\forall\, \theta \in \Theta$

$$\sum_{c} \pi(\overline{q},c,\overline{k},\overline{a}|\theta) = P\left(\overline{q}|\overline{k},\overline{a},\theta\right) \sum_{q,c} \pi(q,c,\overline{k},\overline{a}|\theta) \tag{4}$$

- Incentive Compatibility Constraints (ICC) for action variables[5]
  (Moral Hazard problem on unobserved effort):
  $\forall\, a,\hat{a} \in A \times A$ and $\forall\, k \in K$ and $\forall\, \theta \in \Theta$:

$$\sum_{q,c} \pi\left(q,c,k,a|\theta\right) u\left(c,a|\theta\right) \geq \sum_{q,c} \pi\left(q,c,k,a|\theta\right) \frac{P(q,|k,\hat{a},\theta)}{P(q,|k,a,\theta)} u(c,\hat{a}|\theta) \tag{5}$$

- Truth Telling Constraints (TTC) for unobservable types (Adverse Selection problem: type $\theta$ must be prevented from pretending to be of type $\theta', \theta \neq \theta'$) :
  $\forall\, \theta,\theta' \in \Theta \times \Theta$ :

$$\sum_{q,c,k,a} \pi\left(q,c,k,a|\theta\right) u\left(c,a|\theta\right) \geq \sum_{q,c,k,a} \pi\left(q,c,k,a|\theta'\right) \frac{P(q,|k,a,\theta)}{P(q,|k,a,\theta')} u(c,a|\theta) \tag{6}$$

With optimal contract $\overline{\pi(q,c,k,a|\theta)}$ as solution of optimization problem (2) we get bank surplus $S^{\overline{\omega(\theta)}}$, where

$$\overline{\omega(\theta)} = \sum_{q,c,k,a} \overline{\pi(q,c,k,a|\theta)} u(c,a|\theta)$$

We chose the following standard functional form for utility

$$u(c,a|\theta) = \frac{c^{(1-\sigma(\theta))}}{(1-\sigma(\theta))} + \chi(\theta)\frac{(1-a)^{\gamma(\theta)}}{\gamma(\theta)}$$

---

[5]We act as if either there is moral hazard or, below, adverse selection

In our numerical example (see specifications in Table 1 that are similar to Phelan and Townsend (1991)) we get the following graph (Figure 1)[6] for Pareto frontier for first-best (full insurance) and moral hazard case. Here we assume only one type is present. There is the usual trade off between profits of the lender and utility of the borrower. In many settings, as with a seller, the price the seller gets is what the buyer surrenders. Our set up is similar but with risk aversion the frontier is concave.

The properties of the optimal contract are shown at Figures 2-3. Notice that the borrowing is higher in moral hazard regime with lower utility offerings compared to first-best regime. The borrowing is complementary to the effort in this setup. The conditional reward structure for the agent's compensation contract is typical for moral hazard problems. When effort goes down to zero both full-insurance and moral hazard solutions coincide. The first-best compensation contract is not conditional on output at all utility offers. When utility goes up in full insurance case the agent first gets more leisure then consumption starts to rise[7]. Specifics of adverse selection contracts will be discussed in a separate section.

It can be seen from Figures 1-2 that at $\omega > 2.3$ there is full consumption insurance for moral hazard contract at zero effort level. The contract and surplus for moral hazard, however, are still different from those for full information. Finally, at $\omega > 2.6$ both moral hazard and full information become indistinguishable even on contracts since incentive compatibility constraints do not bind at utilities higher than the level at which the full information effort goes to zero.
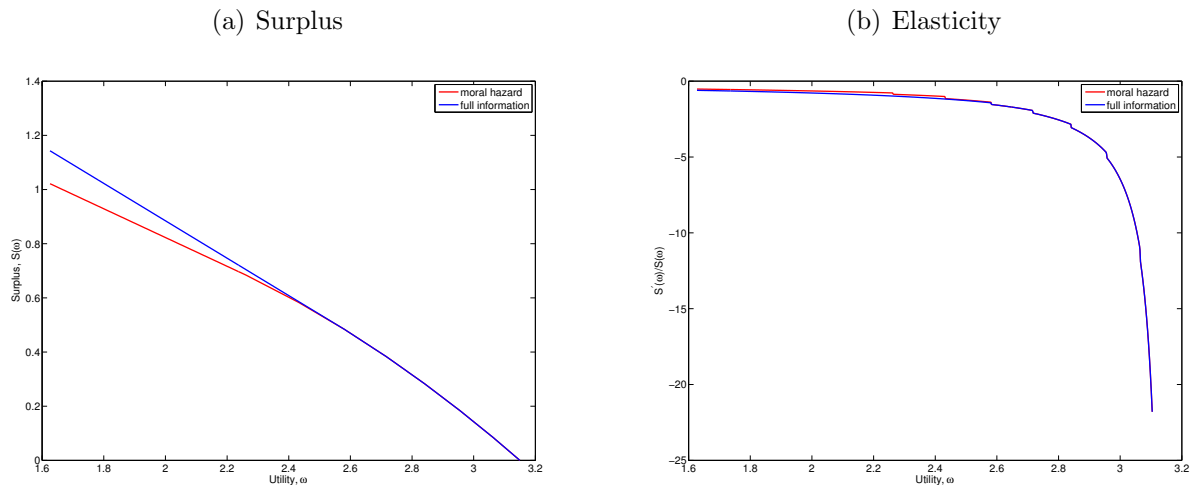
(a) Surplus

(b) Elasticity



Figure 1: Surplus frontier $S(\omega)$ and elasticity $\epsilon_S = S'(\omega)/S(\omega)$

---

[6]We have generated many such figures for different parameters values. What we present in this paper is illustrative.

[7]Financial service provider minimizes the loss in surplus for a given positive change in utility. The impact of consumption change on utility is the marginal utility of consumption times consumption necessary and cost in surplus is one to one to consumption change. The same applies to cost measured in leisure but the impact in surplus is through the production function. Either consumption or leisure is a corner solution depending on specifications costs leading to different preferences for a provider to purchase an additional unit of input from agents either by giving more leisure (cutting the work time at the same wage paid) or by paying more in consumption (increasing wages at the same work time).
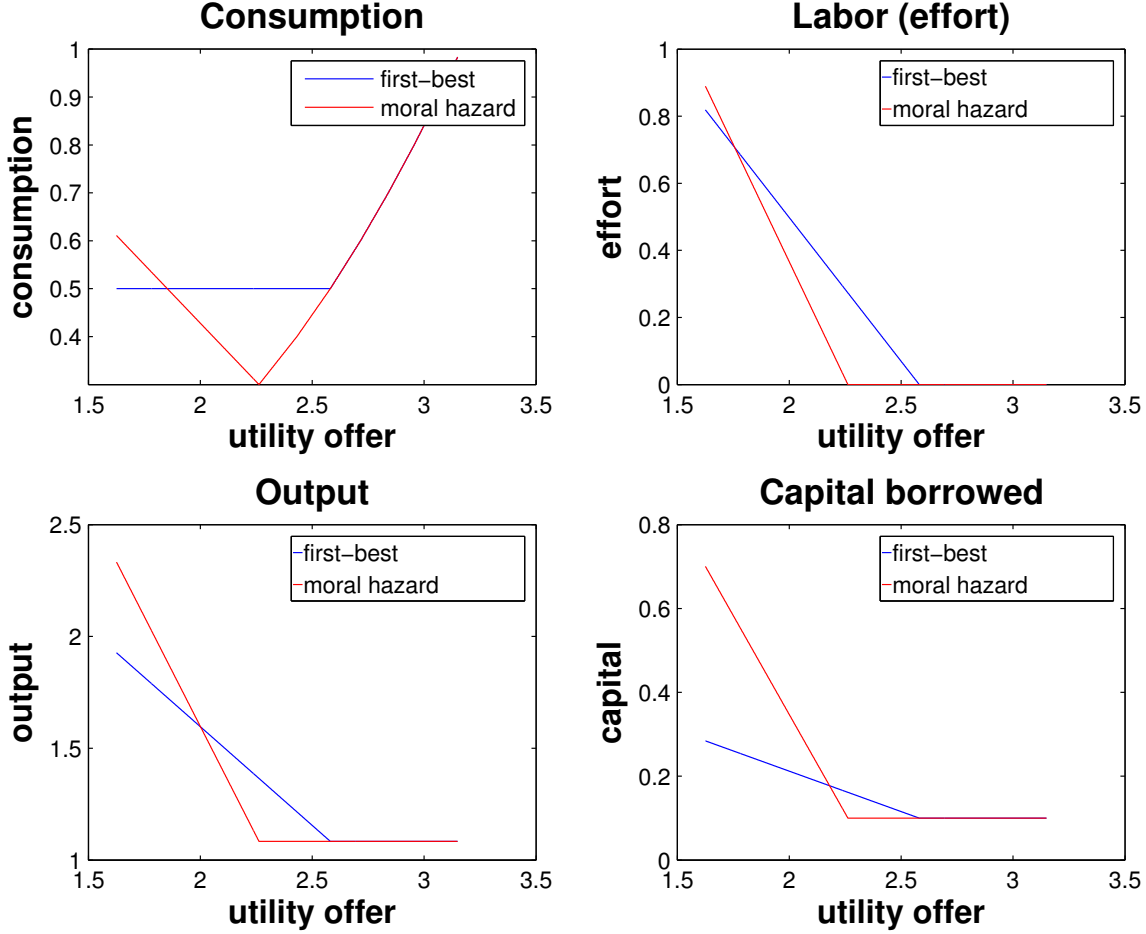
Figure 2: Contract characteristics, expected values

The surplus frontier starts at autarky utility that bounds utility offer space from below. The first point on all surplus frontiers pictured is always an autarky utility, the last one is the utility at zero surplus for the lender.

The effort grid is set in [0 1] range. The outcome space for autarky lies in [0, 2] range and for bank-mediated sector it lies in [1, 4] range. Thus the bank-mediated sector is in first-order stochastic dominance over the autarky. Type dependent technology function is set to be linear: $f(\theta) = \theta$. The surplus maximization problem is a standard linear program.

To conclude, in this section we introduced the space of contracts defined by a range of profit-optimal utility offers from financial intermediaries that depend on agents type and underlying information frictions. We can amend this structure to allow only a fraction of the surplus as written to enter as profits of the bank, the rest covering real intermediation costs. Likewise we can add shocks as random variables into the surplus function, even arguably as a function of locations that we describe below, but though that would move us toward more standard industrial organization setups, it would only cloud the picture here as we try to focus on basics, first.

In the next section we consider several types of market structure that define Pareto optimal contracts with specific utility offers that depend on competition (or lack of such) among financial
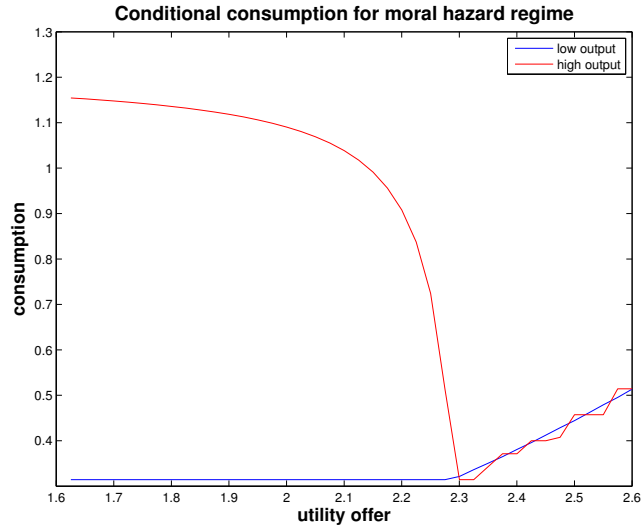
Figure 3: Conditional consumption offered by the contract in moral hazard regime

| structural parameter | $\sigma$ | $\chi$ | $\gamma$ | $\alpha$ |
|---|---|---|---|---|
| value | 0.5 | 1.4 | 1.2 | 1/3 |

| a | $p(q = low)$ | $p(q = high)$ |
|---|---|---|
| 0 | .9 | .1 |
| .2 | .75 | .25 |
| .4 | .6 | .4 |
| .6 | .4 | .6 |
| .8 | .25 | .75 |
| 1 | .1 | .9 |

Table 1: Specifications and technology relating action to probability of each output

service providers and their strategic positioning.

# 3 Financial market structure and equilibrium contracts

## 3.1 Demand side

The agents are distributed spatially inside a unit ball (for simplicity here we limit to linear Hotelling competition case in $\mathbb{R}^1 : [0; 1]$ [8]. Total market mass is set to one. Location $l$ of the agent defines cost to access financial services at location $l_i$. Let's assume that cost is equal to $\overline{L} * |l - l_i|$ with a scale factor $\overline{L}$. Different local environments can have specific values of cost parameter $\overline{L}$, representing a variety of regional or historical features. We imagine we see the entire cross section of economies and, for now, that cost $\overline{L}$ is fully observable.

The agents at location $l$ choose to go to bank 1 if the offer of utility from bank 1 satisfies participation constraint

$$\omega_1(\theta) - \overline{L} * |l - l_1| \geq \hat{\omega}_0(\theta)$$

where $\hat{\omega}_0(\theta)$ is reservation utility from agent's problem solution (autarky value $u_0(\theta)$). Hence, we additively separate the utility (contract) as such from the costs, as opposed to selection into same contract which ex post has different incentives given unobserved heterogeneity and nonseparability.

The agents at location $l$ choose to go to bank 2 at location $l_2$ if the offer of utility from bank 2 satisfies both the participation constraint and bank 2 strong domination constraint against bank 1

$$\omega_2(\theta) - \overline{L} * |l - l_2| \geq \hat{\omega}_0(\theta)$$
$$\omega_2(\theta) - \overline{L} * |l - l_2| > \omega_1(\theta) - \overline{L} * |l - l_1|$$

when $\omega_2(\theta)$ is offered by bank 2.

The spatial distribution of agents can be thought of literally in geographical sense as in our preferred application. However one can also imagine that households have preferences for a certain bank, following a smooth distribution as in discrete choice. These considerations might help smooth out reaction functions, depending on the market structure under consideration (as it stands without this modification, we do have some non-existence problems in a subset of problems we identify below).

## 3.2 Supply side

### 3.2.1 Single branch monopoly

Let's consider a single branch monopoly that selects location $\bar{l}$ and utility offer $\overline{\omega}$ simultaneously. We start with the simplification of one type, thus dropping dependence on $\theta$ temporarily.

The share of market captured is

$$\mu(\overline{\omega}, \bar{l})$$

attractive all agents at locations $l$ where the following is satisfied

$$\overline{\omega} - \overline{L} * |l - \bar{l}| \geq \hat{\omega}_0$$

---

[8]Restriction to finite nodes as potential bank locations is a trivial simplification. Allowing location choices in unrestricted $\mathbb{R}^2$ space is feasible.

That market share at level of surplus provided by the utility offer $\overline{\omega}$ determines a total profit for monopoly to maximize

$$P(\overline{\omega}, \overline{l}) := \underset{\{\overline{\omega}, \overline{l}\}}{\text{maximize}} \left[ S^{\overline{\omega}} \mu(\overline{\omega}, \overline{l}) \right]$$

**Definition** A contract with utility $\overline{\omega}$ offered by financial service provider at location $\overline{l}$ is *Pareto optimal* if there is no other feasible contract $\omega^*$ and location $l^*$ such that $P(\omega^*, l^*) \geq P(\overline{\omega}, \overline{l})$ and $\omega^* - \overline{L} * |l - l^*| \geq \overline{\omega} - \overline{L} * |l - \overline{l}|$ for the agents at locations $l$ that joined financial system under contract $\overline{\omega}$ offered at location $\overline{l}$

**Proposition 3.1.** *Let $S(\omega) * \mu(\omega)$ be the total profit for one branch monopoly at location $l$. Suppose that monopoly profit is at maximum and that monopoly doesn't cover the whole market. Then any increase in utility offer $\omega$ will decrease monopoly profit.*

*Proof.* Since the monopoly does not cover the entire market, there are numerous locations that maximize profit. For example, location $l = 1/2$ is fine without loss of generality. Hence there is only one argument to choose, w. From FOC wrt to utility promise

$$\frac{\partial \left[ S(\omega) * \mu(\omega) \right]}{\partial \omega} = S'(\omega) * \mu(\omega) + S(\omega) * \mu'(\omega) = 0$$

From here we derive useful equality for surplus and market share elasticities

$$\frac{S'(\omega)}{S(\omega)} = -\frac{\mu'(\omega)}{\mu(\omega)} \tag{7}$$

The second-order condition is written as

$$S''(\omega)\mu(\omega) + 2S'(\omega)\mu'(\omega) + S(\omega)\mu''(\omega)$$

Let's consider terms in this expression

$$S''(\omega) \leq 0 \wedge \mu(\omega) \geq 0 \Rightarrow S''(\omega)\mu(\omega) \leq 0$$

since surplus function is concave and market share is non-negative by definition.

$$S'(\omega) \leq 0 \wedge \mu'(\omega) \geq 0 \Rightarrow 2S'(\omega)\mu'(\omega) \leq 0$$

since surplus function is monotonically decreasing and market share is a linearly increasing function of $\omega$ (follows from definition of market share).

$$\mu''(\omega) = 0 \Rightarrow S(\omega)\mu''(\omega) = 0$$

Thus second-order condition is negative. □

The surplus elasticity $\epsilon_S(\omega) = \frac{S'(\omega)}{S(\omega)}$ is downward sloping (see Fig. 1 (b)), it asymptotically goes to $-\infty$ at $S(\omega) \to 0$. The market share elasticity $\epsilon_\mu(\omega) = -\frac{\mu'(\omega)}{\mu(\omega)}$ is upward sloping or flat at zero level (when utility offer is high enough to cover the whole market, $\mu'(\omega) = 0$), it asymptotically goes to $-\infty$ at $\omega \to \hat{\omega}_0$. Thus, equation (7) is a condition where surplus elasticity function intersects with market share elasticity function at optimal (profit maximizing) utility level. The surplus elasticity is independent of spatial costs $\overline{L}$, so the increase in $\overline{L}$ moves market share elasticity to the right as will be shown in examples later. To anticipate, as $\overline{L}$ moves we trace out the market share elasticity function for this monopoly problem.

9

**Proposition 3.2.** *If monopoly covers the whole market and monopoly profit is at maximum. Then no change in location l or utility $\omega$ is possible without hurting optimal monopoly profit at maximal monopoly utility. However, when the spatial cost $\overline{L}$ rises, if the monopolist is to maintain 100% of the market, then utility will have to be increased to retain this marginal customer at his autarky value.*

*Proof.* Since in this case monopoly still wants the marginal customer who is located at maximum distance $l = 0.5$ from monopoly location $\overline{l} = 0.5$ (without loss of generality) it offers minimum utility $\omega = \hat{\omega}_0$ to attract that borderline customer. Any higher offer would hurt monopoly surplus without increasing market share any further.

$\square$

**Corollary 3.3.** *(Pareto Optimality of single-branch monopoly contract)*
*Let $S(\omega) * \mu(\omega, l)$ be a total profit for one branch monopoly. Then monopoly equilibrium is Pareto optimal.*

*Proof.* Immediately follows from Propositions 3.1 - 3.2

$\square$

A monopolist could potentially offer a spatially spatially discriminating contract. In this case the monopolist without loss of generality locates at $\overline{l} = 1/2$ and offers a location specific utility $\omega(l) = \hat{\omega}_0 + \overline{L} * |l - \overline{l}|$ to leave each and every household at $l$ indifferent to autarky while extracting maximum surplus and profit. The market is covered up to the point $l = \hat{l}$ such as $S(\omega(\hat{l})) \to 0$.

### 3.2.2 Two branch monopoly (collusion among banks that are supposed to be competitors)

Let's consider a monopoly that selects locations $\overline{l_1}$ and $\overline{l_2}$ simultaneously for two branches or two banks that collude to share total surplus.

This monopoly also offers utilities $\overline{\omega_1}, \overline{\omega_2}$ simultaneously at each of its branches.

Thus for each branch we can write the the share of market captured as

$$\mu_1(\overline{\omega_1}, \overline{l_1}, \overline{\omega_2}, \overline{l_2}),$$
$$\mu_2(\overline{\omega_1}, \overline{l_1}, \overline{\omega_2}, \overline{l_2})$$

with agents at all locations $l$ where the following is satisfied for those locations constituting market share of branch 1,

$$\overline{\omega_1} - \overline{L} * |l - \overline{l_1}| \geq \hat{\omega}_0$$

and for those locations constituting markets share of branch 2

$$\overline{\omega_2} - \overline{L} * |l - \overline{l_2}| > \overline{\omega_1} - \overline{L} * |l - \overline{l_1}|$$

Those market shares for the two branches provides a total profit for monopoly to maximize defined as

$$\underset{\{\overline{\omega_2}, \overline{l_2}, \overline{\omega_1}, \overline{l_1}\}}{\text{maximize}} [S^{\overline{\omega_1}} \mu_1(\overline{\omega_2}, \overline{l_2}, \overline{\omega_1}, \overline{l_1}) +$$
$$S^{\overline{\omega_2}} \mu_2(\overline{\omega_2}, \overline{l_2}, \overline{\omega_1}, \overline{l_1})]$$

**Proposition 3.4.** *Let $S_1(\overline{\omega_1}) * \mu_1(\overline{\omega_1}, \overline{\omega_2}, \overline{l_1}, \overline{l_2}) + S_2(\overline{\omega_2}) * \mu_2(\overline{\omega_1}, \overline{\omega_2}, \overline{l_1}, \overline{l_2})$ be a total profit for two branch monopoly with branches at locations $\overline{l_1}$ and $\overline{l_2}$ offering utility $\overline{\omega_1}$ and $\overline{\omega_2}$ correspondingly. Suppose that monopoly profit is at maximum. Then monopoly equilibrium is Pareto optimal.*

*Proof.* Follows as in Propositions 3.1 - 3.2 earlier, with one branch, since this problem is strictly concave. Intuitively, without loss of generality monopoly branches are located at $\overline{l_1} = 1/2, \overline{l_2} = 3/4$. Each of those monopolists offers a contract that is surplus maximizing according to Propositions 3.1 - 3.2 for its half of the market. See the discussion which follows. $\square$

### 3.2.3  Real value for households from financial contracts

Here we introduce the concept of real value for a household at location $l$ from financial contracts defined as

$$V^{\overline{\omega_1},\overline{\omega_2},\overline{l_1},\overline{l_2}}(l) = \max(\overline{\omega_1} - L * |l - \overline{l_1}|, \overline{\omega_2} - L * |l - \overline{l_2}|, \hat{\omega}_0), \tag{8}$$

where $\overline{\omega_1}, \overline{\omega_2}, \overline{l_1}, \overline{l_2}$ are utility promises and again locations for two competing banks or two branches of monopoly or a central planner choices, $\hat{\omega}_0$ is the level of utility the household gets from staying in autarky and not incurring any costs to join financial system.

The real value from contracts to households is plotted on Fig.4
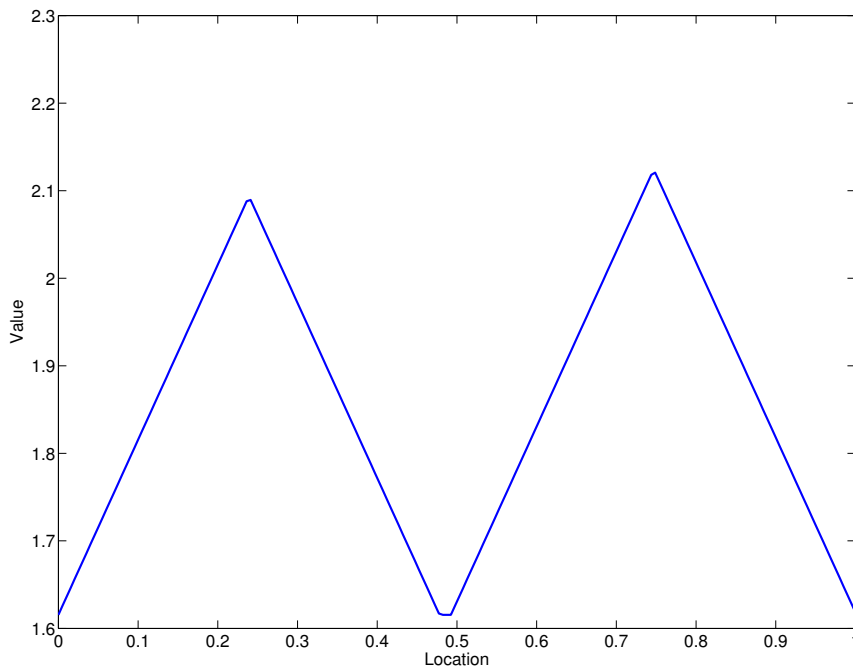


Figure 4: Real value for household from financial contracts, spatial cost $\overline{L} = 2$, two-branch monopoly

Those Tipi-shaped graphs represent net benefit for households. The area above an autarky utility value horizontal line is a gain for households who join financial system.

The apex of each Tipi corresponds to financial service provider branch location and utility offer (location on the $x$ axis and promise $\omega$ on the $y$ axis). Households that are located in the immediate vicinity get the largest value since utility offered is not spatially discriminating. The marginal customer in this case is the customer who gets exactly the autarky value from contract when spatial costs are subtracted from utility offered. There is a little island of autarky in the middle. Spatial costs are high enough so as to make total monopoly profit optimal while servicing $< 100\%$ of the market. High surplus extracted from agents at relatively lower utility offer overweights the benefits of additional agents attracted by higher utility offers.

### 3.2.4 Monopoly: comparative statics as spatial cost $L$ is varied

Fig.5 shows how monopoly utility offers change as we scan the range of spatial costs from $\overline{L} = 5$ to $\overline{L} = 0$.

(a) Profit
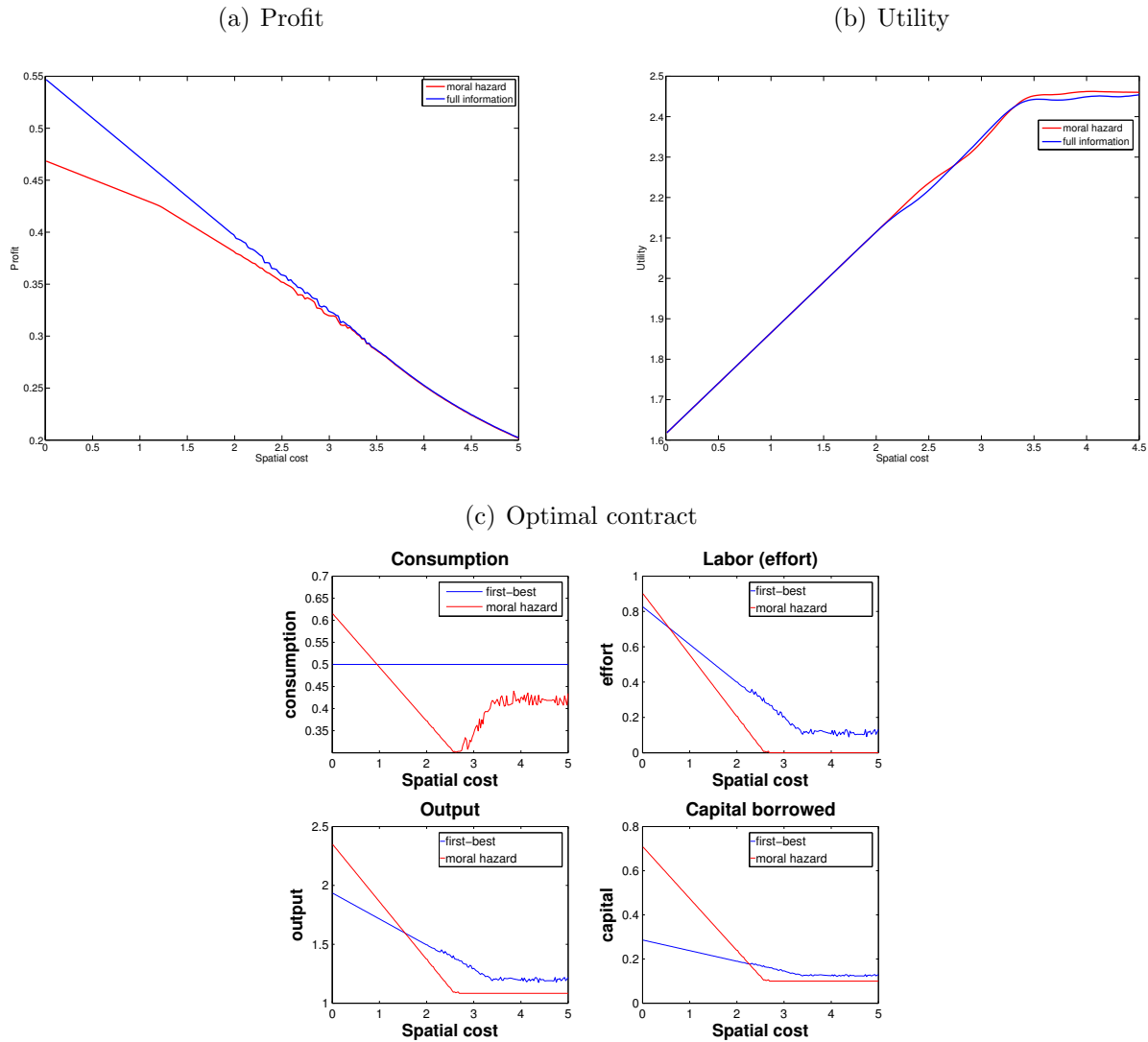
(b) Utility



(c) Optimal contract



Figure 5: Two-branch monopoly (collusion)

Utility choice is fixed by Tipi-shaped real household values from intermediation, see Fig.4. In case of both full information and moral hazard, the offer from banks has to beat autarky at the margin. Since the autarky value is the same for full information and moral hazard, we have identical utility for both regimes when market is fully covered. The contracts are different in full information compared to moral hazard regime throughout the whole range scanned.

The value offered to intermediated agents (whose share in total market goes to zero at infinite spatial costs) at high spatial costs $\overline{L} > 3.5$ is both relatively large and approximately constant at the apex of Tipi. Now, lets move to the left as we lower spatial costs to allow for profitable profit extraction by monopoly from larger share of the market. The autarky islands shrink, more people get access to financial services and utility offer monotonically goes down

till zero spatial costs where monopoly simply offers minimum utility that is slightly larger than autarky utility $\omega \approx \hat{\omega}_0$. This can be thought of as transition from local banking with most population left in autarky to a banking system with whole population intermediated under government restricted competition (state-charted banks protected from out-of-state banks by interstate banking legal limitations).

At Fig.6 we plot surplus and market share elasticity for two-branch monopoly to illustrate how the functions and their intersection point move as we scan the range of spatial costs.
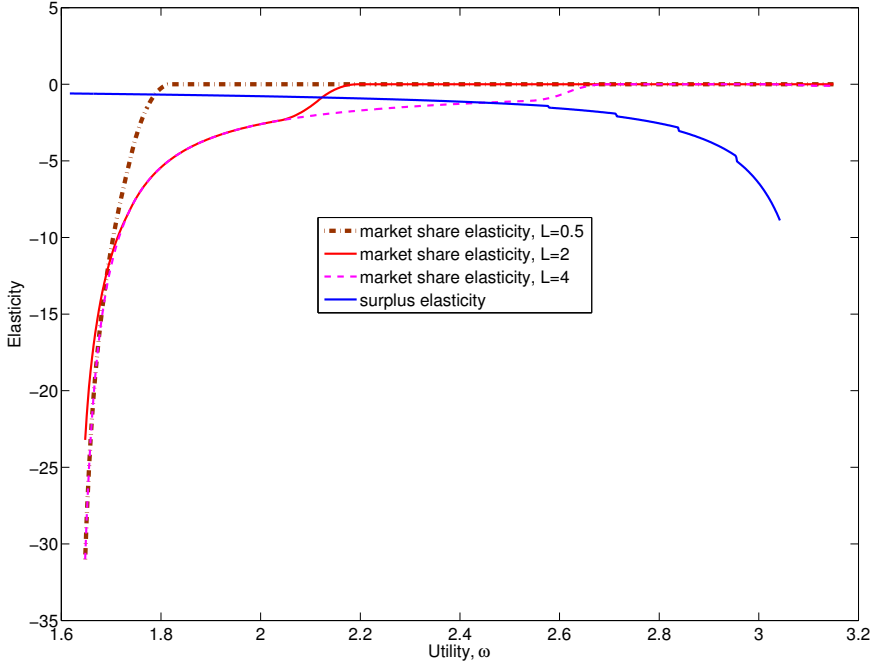


Figure 6: Surplus and market share elasticity, two-branch monopoly (collusion)

14

### 3.2.5 Competition: Sequential Nash Equilibrium (SNE) with full commitment on location choice and contracts

Let's consider a sequential game with first bank coming to location $l_1$. This bank offers utility $\omega_1$ that provides a surplus $S^{\omega_1}$. The first bank anticipates the entry of the second bank and it chooses its location and offer with respect to the best possible response by the second bank. The second entrant chooses optimal location $l_2$ and utility offer $\omega_2$ for any choice of the first entrant taken as given. In principle, we can incorporate shocks that impact profits just prior to entry, so that the first entrant gets a shock, centered at zero but may make it want to move left or right, and more to the point, the shocks for the second entrant makes it such that the first entrant cannot anticipate entirely what the second entrant will do, hence taking expectations.

Conditional on bank 1 choice of location $\overline{l_1}$ and offer of utility $\overline{\omega_1}$ the second bank gets a share $\mu_2(\omega_2(\overline{\omega_1},\overline{l_1}),l_2(\overline{\omega_1},\overline{l_1}),\overline{\omega_1},\overline{l_1})$ of the market with agents at all locations $l$ where the following is satisfied

$$\omega_2(\overline{\omega_1},\overline{l_1}) - \overline{L}*|l - l_2(\overline{\omega_1},\overline{l_1})| \geq \hat{\omega}_0$$
$$\omega_2(\overline{\omega_1},\overline{l_1}) - \overline{L}*|l - l_2(\overline{\omega_1},\overline{l_1})| > \overline{\omega_1} - \overline{L}*|l - \overline{l_1}|$$

Second bank profit to maximize is

$$P_2(\overline{\omega_2}(\overline{\omega_1},\overline{l_1}),\overline{l_2}(\overline{\omega_1},\overline{l_1}),\overline{\omega_1},\overline{l_1}) := \underset{\{\omega_2(\overline{\omega_1},\overline{l_1}),l_2(\overline{\omega_1},\overline{l_1})\}}{\text{maximize}} S^{\omega_2(\overline{\omega_1},\overline{l_1})}\mu_2(\omega_2(\overline{\omega_1},\overline{l_1}),l_2(\overline{\omega_1},\overline{l_1}),\overline{\omega_1},\overline{l_1})$$

After second entrant makes the offer the first entrant gets market share $\mu_1(\omega_2(\omega_1,l_1),l_2(\omega_1,l_1),\omega_1,l_1)$ with agents at all locations $l$ where the following is satisfied

$$\omega_1 - \overline{L}*|l - l_1| \geq \hat{\omega}_0$$
$$\omega_1 - \overline{L}*|l - l_1| > \omega_2(\omega_1,l_1) - \overline{L}*|l - l_2(\omega_1,l_1)|$$

First bank profit to maximize is

$$P_1(\omega_2(\overline{\omega_1},\overline{l_1}),l_2(\overline{\omega_1},\overline{l_1}),\overline{\omega_1},\overline{l_1}) := \underset{\{\omega_1,l_1\}}{\text{maximize}} S^{\omega_1}\mu_1(\omega_2(\omega_1,l_1),l_2(\omega_1,l_1),\omega_1,l_1)$$

**Definition** A strategy $G = \{\overline{\omega_2}(\overline{\omega_1},\overline{l_1}),\overline{l_2}(\overline{\omega_1},\overline{l_1}),\overline{\omega_1},\overline{l_1}\}$ constitutes sequential Nash eqilibrium (SNE) if for for all $\omega_1,\omega_2 \in \Omega$ where $\Omega$ is a feasible contract space that is surplus optimal and for all $l_1,l_2 \in \{0,1\}$

$$P_1(\omega_2(\overline{\omega_1},\overline{l_1}),l_2(\overline{\omega_1},\overline{l_1}),\overline{\omega_1},\overline{l_1}) \geq P_1(\omega_2(\omega_1,l_1),l_2(\omega_1,l_1),\omega_1,l_1)$$

s.t.

$$P_2(\overline{\omega_2}(\overline{\omega_1},\overline{l_1}),\overline{l_2}(\overline{\omega_1},\overline{l_1}),\overline{\omega_1},\overline{l_1}) \geq P_2(\omega_2(\overline{\omega_1},\overline{l_1}),l_2(\overline{\omega_1},\overline{l_1}),\overline{\omega_1},\overline{l_1})$$
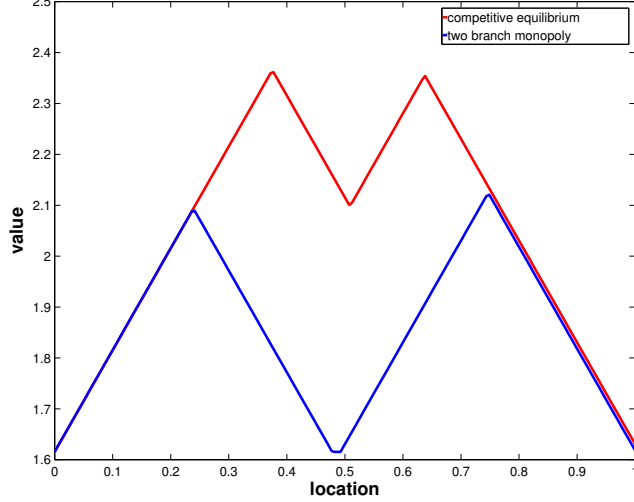
Figure 7: Real value for households, spatial cost $\overline{L} = 2$, SNE with full commitment

The example of real value to households under SNE at $\overline{L} = 2$ is plotted on Fig.7 together with comparable graph for two-branch monopoly. Although Tipi-shaped values now are strictly higher, meaning that households are strictly better off under competitive environment compared to monopoly, those improvements come at the expense of bank profits. No strict Pareto improvement is found, the effect is redistributive in nature. The island of autarky is gone under competitive environment and banks are competing intensely in the market segment between their chosen location providing additional benefits for their agents. The marginal customer is at the left and right extreme borders of the spatial interval where nearest bank always has relative price advantage over its distant competitor. Note therefore that changes in promises $\omega$ have different consequences for how market shares move with $\omega$.

We also solve a social planner problem that attempts to find a strategy

$$G^{SP} = \{\overline{\omega_2}^{SP}, \overline{l_2}^{SP}, \overline{\omega_1}^{SP}, \overline{l_1}^{SP}\}$$

such as to improve real value for households compared with optimal SNE strategy G:

$$V^{\overline{\omega_2}^{SP}, \overline{l_2}^{SP}, \overline{\omega_1}^{SP}, \overline{l_1}^{SP}}(l) \geq V^{\overline{\omega_2}(\overline{\omega_1}, \overline{l_1}), \overline{l_2}(\overline{\omega_1}, \overline{l_1}), \overline{\omega_1}, \overline{l_1}}(l), \forall l \in [0, 1]$$

s.t.

$$P_1(\overline{\omega_2}^{SP}, \overline{l_2}^{SP}, \overline{\omega_1}^{SP}, \overline{l_1}^{SP}) \geq P_1(\overline{\omega_2}(\overline{\omega_1}, \overline{l_1}), \overline{l_2}(\overline{\omega_1}, \overline{l_1}), \overline{\omega_1}, \overline{l_1})$$

$$P_2(\overline{\omega_2}^{SP}, \overline{l_2}^{SP}, \overline{\omega_1}^{SP}, \overline{l_1}^{SP}) \geq P_2(\overline{\omega_2}(\overline{\omega_1}, \overline{l_1}), \overline{l_2}(\overline{\omega_1}, \overline{l_1}), \overline{\omega_1}, \overline{l_1})$$

No strategy $G^{SP}$ that allows to obtain Pareto improvement over optimal SNE strategy G is found in our numerical experiments.
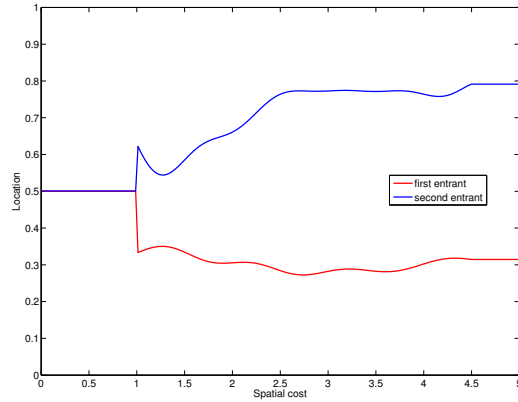
16

### 3.2.6 Competition: comparative statics as spatial cost $L$ is varied

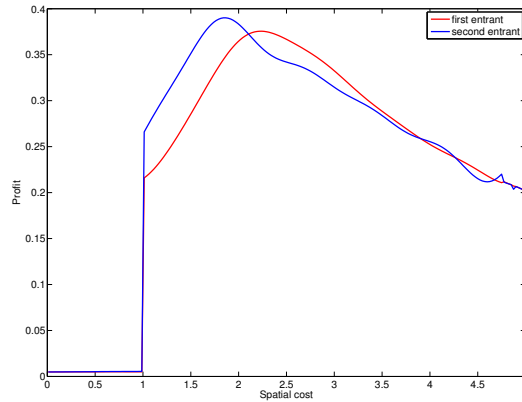Figure 8: Competition: full insurance, full commitment

(a) Utility offer choice
(b) Location choice
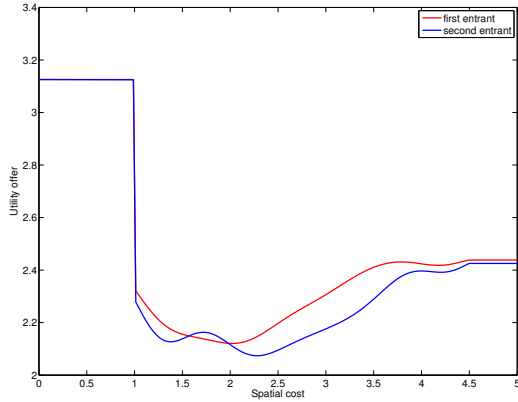


(c) Profit



To illustrate the effect of spatial costs on SNE competitive outcome in this full commitment regime we solve for SNE over the range of spatial costs from $L = 0$ to $L = 5$. Results are shown in Fig 8. We successfully reproduce classical Bertrand result for lower spatial costs, i.e. both banks locate at $l = 1/2$. There is non-differentiable transition from perfect competition utility offer $\omega(L) = \omega_{max}$ to competitive utility offer $\omega(L) << \omega_{max}$ starting at $L \approx 1$. Below that point the first entrant was under threat of being completely and profitably eliminated by a competitor at any strategy different from Bertrand. With $L > 1$ the first entrant can choose a strategy of spatial differentiation that leaves it with positive profit while making second entrant strictly better off if it chooses to move in opposite direction. The profit for the second entrant from eliminating its competitor becomes strictly worse than by cooperating with the first one by limiting intense competition only to the central part of the market.

The process of transition from perfectly competitive case to local monopoly at very high spatial costs is not monotone. At $1 < L < 2$ the contract is more extractive with agents getting relatively smaller value as spatial costs go up and banks solidify their stance as local

17

monopolists and their profits rise. Afterwards, attracting marginal agents from autarky at high cost becomes the dominant force and profits start to drop.

The first entrant, the incumbent, has a profit disadvantage up to cost $L \approx 2$ and after that the second entrant suffers. Note also that location choices are not exactly symmetric.

We also solve for SNE with full commitment under moral hazard. As in the earlier figure for monopoly, it would appear that locations and promises are similar even though they are not exactly the same. In this case it helps to identify the obstacle to go back to the characteristics and outcomes of the actual contracts (see Fig. 2). Note that in SNE not all utilities are covered as cost $L$ is varied, due to this jump, empirical identification of surplus function has to be modified to take this into account.

### 3.2.7 Competition: partial commitment
### (SNE on location choice, simultaneous Nash on contracts)

Let's consider a sequential game with first bank coming to location $\overline{l}_1$ and second bank coming later to location $\overline{l}_2$.

The first bank anticipates the entry of the second bank and it chooses its location with respect to the best possible response by the second bank. Both banks anticipate subsequent simultaneous Nash competition in contracts.

Simultaneous Nash equilibrium (ex-post competition in utilities for arbitrary locations) in case of two market entrants would be defined by $G^N = \{\omega_1^*(\overline{l}_1, \overline{l}_2)), \omega_2^*(\overline{l}_1, \overline{l}_2)\}$ that satisfy

$$P_2(G^N) = S^{\omega_2^*(\overline{l}_1, \overline{l}_2)}\mu_2(\omega_2^*(\overline{l}_1, \overline{l}_2), \overline{l}_2, \omega_1^*(\overline{l}_1, \overline{l}_2), \overline{l}_1) \geq S^{\omega_2(\overline{l}_1, \overline{l}_2)}\mu_2(\omega_2(\overline{l}_1, \overline{l}_2), \overline{l}_2, \omega_1^*(\overline{l}_1, \overline{l}_2)), \overline{l}_1) \quad (9)$$

$$P_1(G^N) = S^{\omega_1^*(\overline{l}_1, \overline{l}_2)}\mu_1(\omega_2^*(\overline{l}_1, \overline{l}_2), \overline{l}_2, \omega_1^*(\overline{l}_1, \overline{l}_2), \overline{l}_1) \geq S^{\omega_1(\overline{l}_1, \overline{l}_2)}\mu_1(\omega_2^*(\overline{l}_1, \overline{l}_2), \overline{l}_2, \omega_1(\overline{l}_1, \overline{l}_2), \overline{l}_1) \quad (10)$$

$$\forall\{\omega_1(\overline{l}_1, \overline{l}_2), \omega_2(\overline{l}_1, \overline{l}_2)\}$$

Second bank chooses location $\overline{l}_2(\overline{l}_1)$ so as to maximize

$$P_2(\omega_2^*(\overline{l}_1, \overline{l}_2(\overline{l}_1)), \overline{l}_2(\overline{l}_1), \omega_1^*(\overline{l}_1, \overline{l}_2(\overline{l}_1)), \overline{l}_1) := \underset{\{l_2(\overline{l}_1)\}}{\text{maximize}}\ S^{\omega_2^*(\overline{l}_1, l_2(\overline{l}_1))}\mu_2(\omega_2^*(\overline{l}_1, l_2(\overline{l}_1)), l_2(\overline{l}_1), \omega_1^*(\overline{l}_1, l_2(\overline{l}_1)), \overline{l}_1)$$

The first bank chooses location $\overline{l}_1$ so as to maximize

$$P_1(\omega_2^*(\overline{l}_1, \overline{l}_2(\overline{l}_1)), \overline{l}_2(\overline{l}_1), \omega_1^*(\overline{l}_1, \overline{l}_2(\overline{l}_1)), \overline{l}_1) := \underset{\{l_1\}}{\text{maximize}}\ S^{\omega_1^*(l_1, \overline{l}_2(l_1))}\mu_2(\omega_2^*(l_1, \overline{l}_2(l_1)), \overline{l}_2(l_1), \omega_1^*(l_1, \overline{l}_2(l_1)), l_1)$$

**Definition** A strategy $G^{SNE} \otimes G^N = \{\overline{l}_1, \overline{l}_2(\overline{l}_1)\} \otimes \{\omega_1^*(\overline{l}_1, \overline{l}_2)), \omega_2^*(\overline{l}_1, \overline{l}_2)\}$ in continuos utilities and locations constitutes sequential Nash eqilibrium (SNE) on location with ex-post simultaneous Nash competition in utilities if for for all $\omega_1, \omega_2 \in \Omega$ where $\Omega$ is a feasible contract space that is surplus optimal and for all $l_1, l_2 \in \{0, 1\}$

$$P_1(G^{SNE} \otimes G^N) \geq P_1(\{l_1, l_2(l_1)\} \otimes G^N)$$

s.t.

$$P_2(G^{SNE} \otimes G^N) \geq P_2(\{\overline{l}_1, l_2(\overline{l}_1)\} \otimes G^N)$$

and Eq.(9)-(10)

This model is illustrated with results of numerical experiments in Table 2.

The first entrant at costs tries to use his first mover advantage by choosing more central location location in the middle and the second entrant tries to get his market share by staying at the margin $\overline{l}_2 \approx 1$. Thus, for the first entrant the optimal business strategy is to get larger market share even at relatively smaller surplus per agent utilizing better central location while the second entrant compensates worse location choice by playing at the surplus margin. But note how the separating equilibrium for banks here with ex post competition in utility promises

| Spatial Cost $L = .5$ | | | | |
|---|---|---|---|---|
| bank | location | utility offer | market share | profit |
| 1 | 0.44 | 3.15 | 84% | 0.12 |
| 2 | 0.93 | 2.99 | 16% | 0.08 |
| Spatial Cost $L = 1$ | | | | |
| bank | location | utility offer | market share | profit |
| 1 | 0.43 | 2.13 | 56% | 0.42 |
| 2 | 0.94 | 2.38 | 44% | 0.25 |
| Spatial Cost $L = 2$ | | | | |
| bank | location | utility offer | market share | profit |
| 1 | 0.45 | 2.44 | 60% | 0.36 |
| 2 | 0.71 | 2.21 | 36% | 0.27 |

Table 2: SNE on location choice, simultaneous Nash on contracts, full information

is different from pooling eqilibrium for banks under full commitment in contracts (see Figure 8)[9]. Obvious but worth repeating, if utilities are different, contracts are different.

We can also consider simplified version when bank locations $\{\overline{l_1}, \overline{l_2}\}$ are fixed (pre-determined by historic or other reasons) and we compute Nash eqilibrium on contracts only. Generally, the second entrant now can not undercut the first entrant on price as in full commitment case. That in turn prevents price war escalation as well as clustering of financial service providers in the middle of the market space. The choice of location becomes relatively more important compared to the case of full commitment. The zero total profit is no longer Nash optimal strategy at $\overline{L} > 0$.

We'll discuss this simplified model of competition in more details in later sections.

---

[9]A similar effect of ex-post differentiation on Bertrand competition is discussed in Moscarini and Ottaviani (2001) where competitive pressure on prices is also sharply reduced with revelation of private information. The sellers become local monopolists and make high profits by fully extracting the customer's surplus.

| Spatial Cost $L = 0.5$ | | | | |
|---|---|---|---|---|
| bank | location | utility offer | market share | profit |
| 1 | 0.5 | 3.15 | 50% | 0.0 |
| 2 | 0.5 | 3.15 | 50% | 0.0 |
| Spatial Cost $L = 2$ | | | | |
| bank | location | utility offer | market share | profit |
| 1 | 0.36 | 2.33 | 58% | 0.39 |
| 2 | 0.72 | 2.14 | 39% | 0.32 |

Table 3: No commitment, full information

### 3.2.8 Competition: no commitment (simultaneous Nash on location choice and contracts)

Nash equilibrium in case of two market entrants is defined by $\{\omega_1^*, l_1^*, \omega_2^*, l_2^*\}$ that satisfy

$$S^{\omega_2^*}\mu_2(\omega_2^*, l_2^*, \omega_1^*, l_1^*) \geq S^{\omega_2}\mu_2(\omega_2, l_2, \omega_1^*, l_1^*) \tag{11}$$
$$S^{\omega_1^*}\mu_1(\omega_2^*, l_2^*, \omega_1^*, l_1^*) \geq S^{\omega_1}\mu_1(\omega_2^*, l_2^*, \omega_1, l_1) \tag{12}$$

$\forall\{\omega_1, l_1, \omega_2, l_2\}$

This model with results of numerical experiments in Table 3 can be compared with partial commitment case in Table 2.

In this competitive structure at low spatial costs both banks choose middle location and they both raise utility offers up to the zero surplus level. This is close to full commitment case so far and it is very different from niche specialization strategy in partial commitment case. Both banks here are left with zero profits at $L < 1$. With spatial costs rising $L > 1$ banks try to find spatially separate locations (converging at $\overline{l_1} = 0.25$ and $\overline{l_2} = 0.75$ on unit interval as in local monopoly case at very high spatial cost) while lowering utility offers to increase profits. During this transition from a pooling to separating eqilibrium, our numerically constructed equilibrium is highly unstable and it might not exist at $1 < L < 2$. Sufficient condition for Nash equilibria is that the reaction functions be continuous in competitor offer given other players offer. In fact there can be jumps and one Tipi can at a critical point completely cover the other player's. While at full commitment case the first entrant can send a signal to competitor with commitment to specific location to avoid price war escalation, such mechanism doesn't exit in no commitment case (see Moscarini and Ottaviani (2001)). With higher spatial costs $L > 2$ a separating Nash equilibrium becomes stable and utility offers start to rise in correlation with increasing costs as banks try to keep their market shares attracting marginal customers left in autarky. The autarky islands start to appear, eventually banks converge at their local monopoly positions with each bank Tipi safely spatially separated from the other one by an autarky island.

More graphs for this model of competition together with a metric for Nash equilibrium stability are given in the Appendix.

This model of competition can represent, for example, the result of competing banks expanding nationally (or globally) into the same area that was previously left without any financial

intermediation. Each nationwide (or global) bank tries to outcompete by selecting both location and contracts at the same time as its competitor. There is a possibility that outcome of such competition becomes highly uncertain at specific spatial costs and full identification thus might not be possible with empirical analysis impacted strongly by high level of random noise.

# 4 Heterogeneous Agents

## 4.1 Fully observed types

Now we explore the effects of heterogeneity among the agents with banks offering a menu of type depending contracts. There is a single location $l_i$ for each bank $i$ with utility offer vector $\{\omega_i(\theta_1), \omega_i(\theta_2), ..., \omega_i(\theta_n)\}$ for types $\theta_j, j = \{1, n\}$ resulting to total profit $P_i = \sum_j P_i(\theta_j)$.

First, lets assume there are two types of agents in population $\{\theta_1, \theta_2\}$. The low-risk type $\theta_2$ corresponds to our earlier baseline results and discussions. It has stricter larger expected output produced at all effort levels than the type $\theta_1$, the riskier type. With higher effort the riskier type $\theta_1$ can overcome the natural ability limitations to achieve similar expected output at larger variance compared to safe type. That is, if type $\theta_2$ shirks from effort $a(\theta_2) < a(\theta_1)$ we can get $E[q(\theta_2)|a(\theta_2)] \approx E[q(\theta_1)|a(\theta_1)]$, $\sigma(q(\theta_2)|a(\theta_2)) < \sigma(q(\theta_1)|a(\theta_1))$ (see Fig.9, where expected output values and one standard deviation bars are plotted at zero investment level).
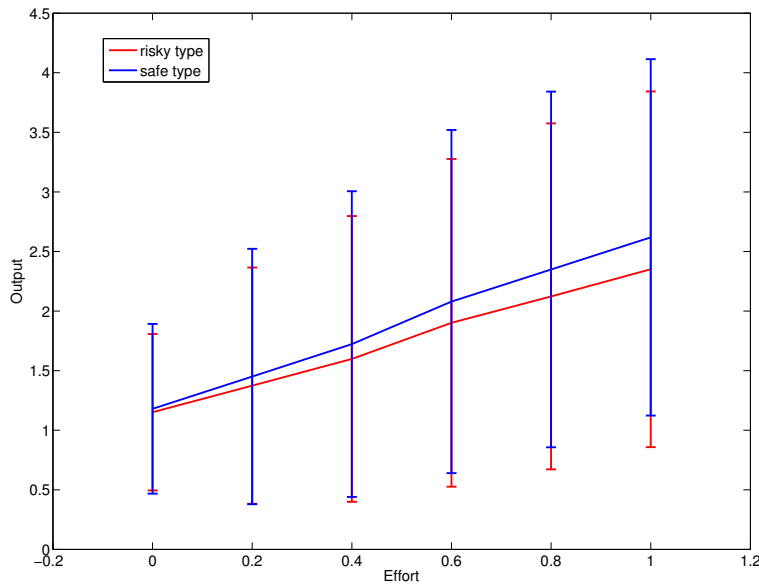


Figure 9: Type dependence of stochastic output $q$, $E[q] \pm \sigma$

The shares of both agents in population for two types case without loss of generality are assumed to be equal. Results for any mix of types can be easily obtained by scaling type-specific surplus correspondingly so that only the total profit level is different for arbitrary mixes.

We plot surplus frontiers at Fig. 10 (a) for types that we will use further. For comparison we also provide results for wider range of heterogeneity in population at Fig. 10 (b). Autarky utilities are different for types, and that causes surplus frontier to start at different minimum utility offers since participation constraints cut the curve at different values. The slope of surplus for agents carrying high risk is lower in absolute value than for safer types because the effort of those types doesn't have as much influence on the output (relatively more of output realizations is due to luck) as it does for safer types, hence there is relatively less difference
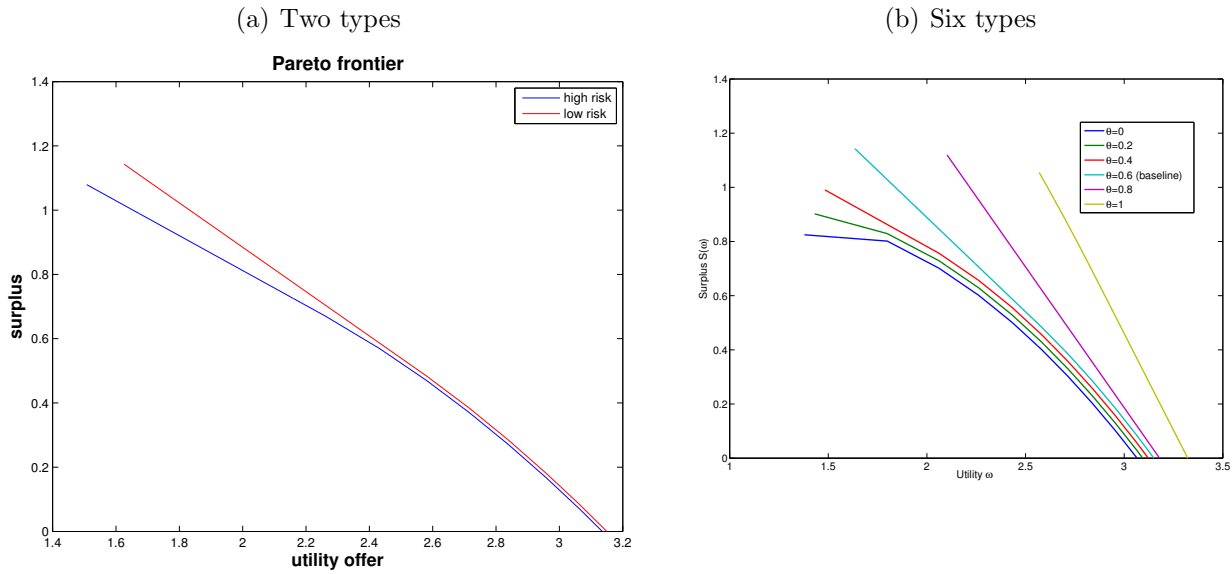
| (a) Two types | (b) Six types |

Figure 10: Surplus, Full Insurance

in optimal output/consumption/effort schedules with contracts varying in utility. That means financial service providers could raise utility offer, getting a larger market share without losing most of surplus. It potentially can make risky types relatively more attractive compared to safer types when competition among financial service providers is most intense in full insurance regime. However, the absolute value of surplus per risker type is lower, attracting safer types even at larger utility offer compared to risky types can be optimal for banks.

In general, two most important factors defining optimal level of utility offers chosen by financial service providers are the level and the slope of surplus but in combination. The surplus elasticity $S'(\omega(\theta))/S(\omega(\theta))$, plotted at Fig 11 defines how much of surplus is lost when utility is raised to get extra market share and that in turn interacts with the market share elasticity $-\mu'(\omega(\theta))/\mu(\omega(\theta))$.

Of considerable interest, the absolute value of elasticity for high risk types is lower at low utility promises than at high promised utilities. That means relatively less surplus is lost at low utility promises for risky types when utility is raised at the margin than for safe types and vice versa for high promised utilities. So the attractive type so to speak varies with the form of overall competition, as that determines the range of utilities in play.
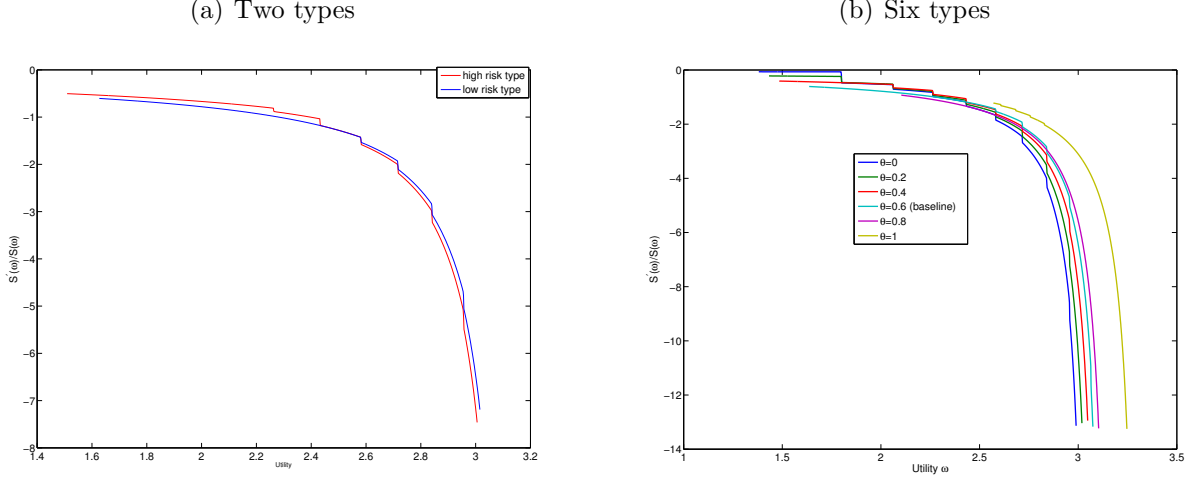
Figure 11: Elasticity, Full Insurance

## 4.2 Optimality for menus of contracts in adverse selection problem

In the adverse selection case with menu of contacts we have type specific utility constraints.

Utility Assignment Constraints (UAC) - type dependent offer from bank is set at utility level $\overline{\omega(\theta)}$ :

$\forall\ \theta \in \Theta$

$$\sum_{q,c,k,a} \pi(q,c,k,a|\theta)u(c,a|\theta) = \overline{\omega(\theta)} \tag{13}$$

For two types

$$\theta \in \{\theta_1, \theta_2\}$$

truth-telling constraints (TTC) can be combined with UAC (given offer of utility per type) to result in explicit bounds for TTC:

$$\overline{\omega(\theta_2)} \geq \sum_{q,c,k,a} \left[ \pi\left(q,c,k,a|\theta_1\right) \frac{P(q,|k,a,\theta_2)}{P(q,|k,a,\theta_1)} u(c,a) \right] \tag{14}$$

$$\overline{\omega(\theta_1)} \geq \sum_{q,c,k,a} \left[ \pi\left(q,c,k,a|\theta_2\right) \frac{P(q,|k,a,\theta_1)}{P(q,|k,a,\theta_2)} u(c,a) \right] \tag{15}$$

**Proposition 4.1.** *When both constraints (14) and (15) are not active the solution of adverse selection problem is identical to solution of full-information problem*

**Proposition 4.2.** *The larger the gap between utility offers $\overline{\omega(\theta_1)}$ and $\overline{\omega(\theta_2)}$ the stronger truth-telling binds affecting maximum surplus possible under feasible contracts*

**Proposition 4.3.** *When one of the constraints (14) and (15) is active (binding with non-zero Lagrange multiplier) the solution of adverse selection problem is separated into solution of full-information problem for one type utility offer that is binding the opposite type contracts plus full-information problem for the second type with one additional strict equality constraint*

**Proposition 4.4.** *In adverse selection if $\overline{\omega(\theta_2)} < \overline{\omega(\theta_1)}$ then low ability type gets a constrained contract and high ability type gets a full-information contract or both types get full-information contracts. It is also true that if $\overline{\omega(\theta_1)} < \overline{\omega(\theta_2)}$ then either high ability type gets a constrained contract and low ability type gets full-information contract or both types get full-information contracts.*

Using those insights we can reduce the space of infeasible contracts under adverse selection while solving for equilibrium in any of market structures considered by fixing (assigning) the utility of one type and then maximizing the surplus by choice of utility of the other subject to truth telling constraints. Then, we can search over all promises for the first, initial type while getting the other type contract linked through truth-telling constraints.

In our numerical examples we have two unobserved types as described above, with a safe type set at the level of talent of our baseline scenario $\theta = 0.6$ and a risky type at level of talent $\theta = 0.5$. In adverse selection case both types are not separable anymore, the utility offer for the opposite type changes the structure of contract for a given type. To illustrate a changed structure of contracts in such intertwined problem Figure 12 shows Pareto frontier and intertwined contract menus. Financial service provider can no longer offer a menu of independent contracts for types.

In the zero total profit limit there is small loss on risky types that got subsidized by safe types. It is not optimal to drop intermediating risky types since resulting loss of good types through intertwined contracts is not going to provide a better solution. As long as TTC are in place the optimal contract menu at zero total profit limit must attract a mix of both types. The contracts of course are still different in consumption/effort/borrowing schedules for types even at close utility levels.
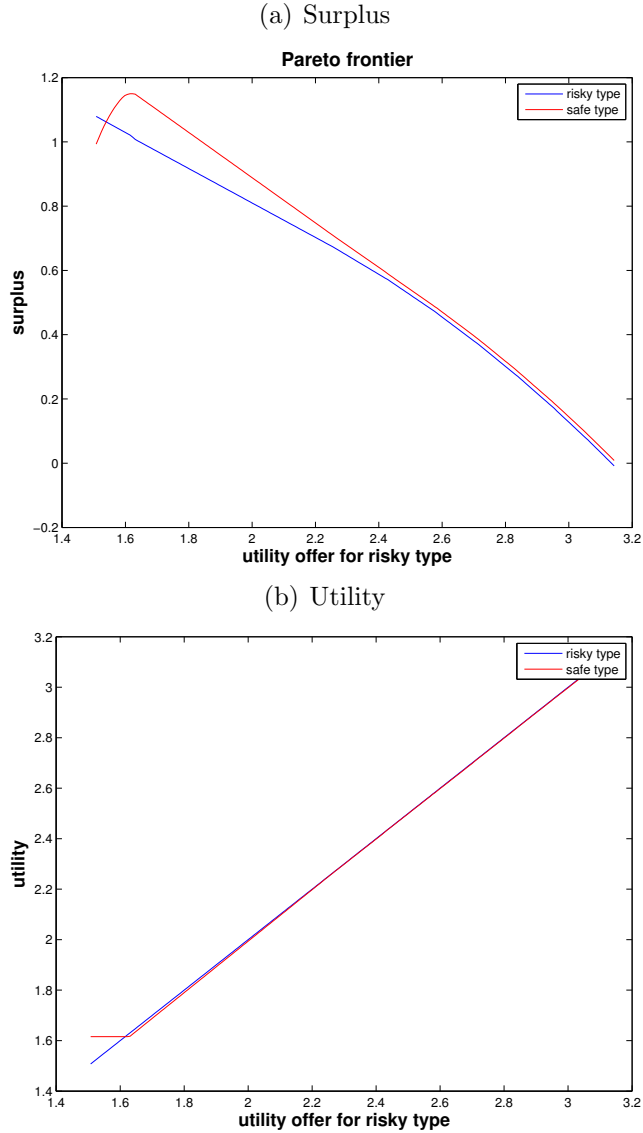
(a) Surplus



(b) Utility



Figure 12: Adverse Selection

# 5 Full insurance versus adverse selection

## 5.1 Profits under type dependent contract and partial commitment with exogenous locations

In this section we consider a special case of partial commitment on contracts with banks location fixed at $\overline{l_1} = 0.25$ and $\overline{l_2} = 0.75$. We compare the effects of heterogeneity in types for full insurance and adverse selection regimes using varying spatial costs. Since locations are completely symmetric we use this also as a test of our numerical procedure that indeed computes the same policy choices for both banks.

The first notable difference in profit dependence (Fig.13(a)) is flat zero profit area characteristic for perfect competition in full information regime observed under the full commitment

market structure previously for single type of agents. Both types under full information get strictly perfectly competitive utility (Fig.13(b)) with a zero profit outcome on those types in full information up till spatial cost $L = 1$. Non zero profit in adverse selection comes from a mix of risky and safe types, both of which get lower utility compared with full information case.

In the adverse selection regime, the truth-telling constraints prevent banks from offering higher utility offer for safer types. The contract for risky types in adverse selection is a TTC constrained contract, the contract for safe types in adverse selection solves full information optimization problem for that type, but at lower level of utility on surplus Pareto frontier compared to market outcome for safe types when full information regime is in place for both types. The presence of worse types drives the utility of good types down in adverse selection competition outcome compared to full information competition outcome.

Total profits in adverse selection are higher than in full information regime at low spatial cost ($L < 1$). Yet profits can be lower at higher spatial costs. The TTC prevent banks from capturing full market share at household value extractive utilities as efficiently as under full-information regime. In adverse selection banks are forced to offer utility that is at safe type value in full-information regime for both risky and safe types (see (Fig.13(b)). Riskier types in adverse selection get higher (safe type) utility that is damaging for bank surplus. The presence of better types, thus, at those spatial costs, drives utility of worse types up while their labor does not produce as much output as it does at full-information contracts for better types.

(a) Profits                                    (b) Nash optimal contract
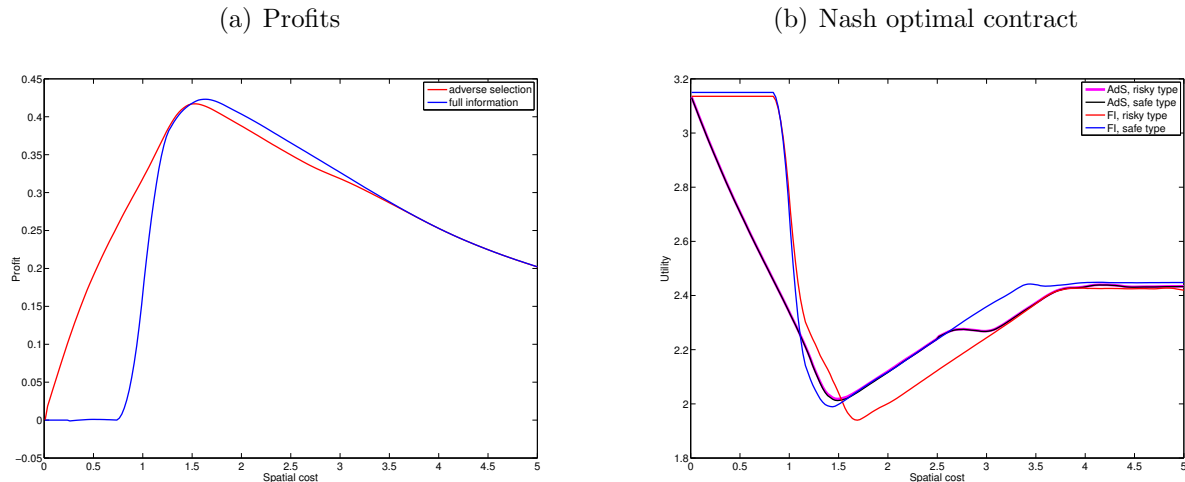


Figure 13: Adverse Selection versus Full Insurance, two types, fixed locations

Fig. 14 illustrates phenomenon we call "relative type depletion" that appears with increasing spatial costs. Best types drop out of coverage first since they have better outside option, banks are left with relatively larger share of high-risk types under Nash optimality conditions even in full information regime. However in adverse selection the depletion of the safe type happens even more rapidly.

Fig. 15 shows spatial dependence of consumption and investment schedules in contract as a reminder that that the various contracts have real consequences on observables. We can interpret it as transition from low investment economy at high spatial cost $\overline{L} > 3$ to high investment economy at intermediate spatial costs $1.5 < \overline{L} < 3$. Finally there is low bound level of investments for all types at low spatial costs.

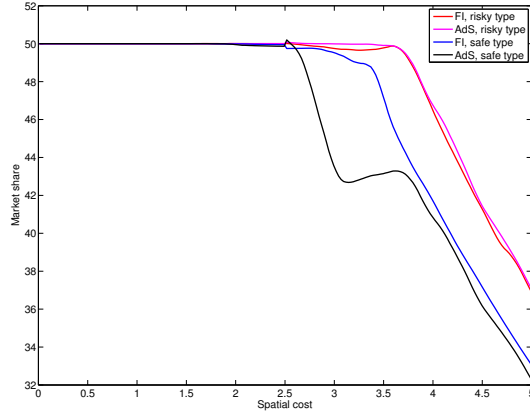28

Figure 14: Adverse Selection versus Full Insurance, market share

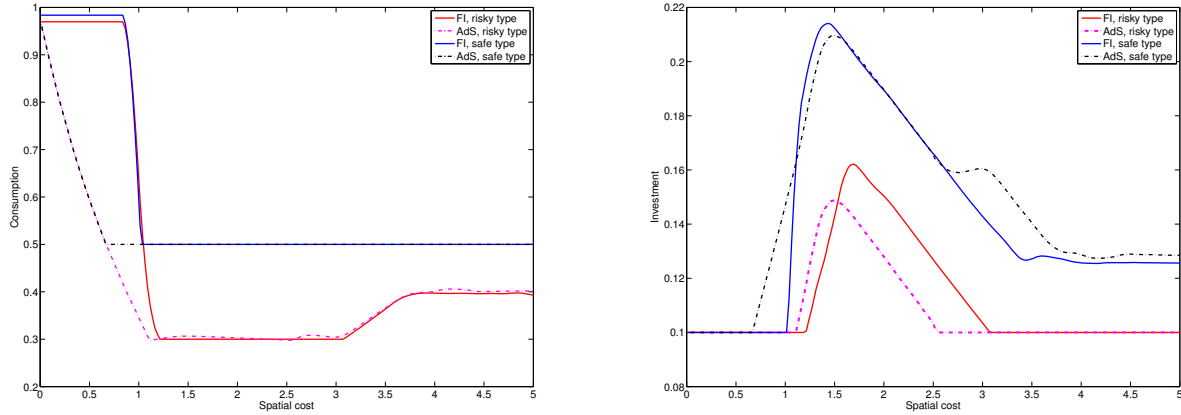(a) Consumption                                      (b) Investment



Figure 15: Adverse Selection versus Full Insurance, contract schedules

## 5.2 Spatial dependence of Nash optimal contract with endogenous location choice

In this section we consider a full blown case of partial commitment on contracts and full commitment on endogenous location choice determined from anticipatory sequential Nash game with subsequent simultaneous Nash game on utility offers. We concentrate on adverse selection case here to illustrate the concept of niche positioning for banks under such market structure.

It can be seen from Fig. 16 (a) that the first entrant indeed has first mover advantage. The first bank at low spatial costs occupies central location (Fig. 16 (c)) and he makes a lower utility offer (Fig. 16 (b)) thus obtaining a larger profit.

The second bank at low spatial costs prefers a marginal niche location (Fig. 16 (c)) at the border and he makes higher utility offer (Fig. 16 (b)) thus obtaining smaller but positive profit. With larger spatial costs both banks separate further from each other, their profits go up, the first entrant still keeps marginal its first mover advantage.
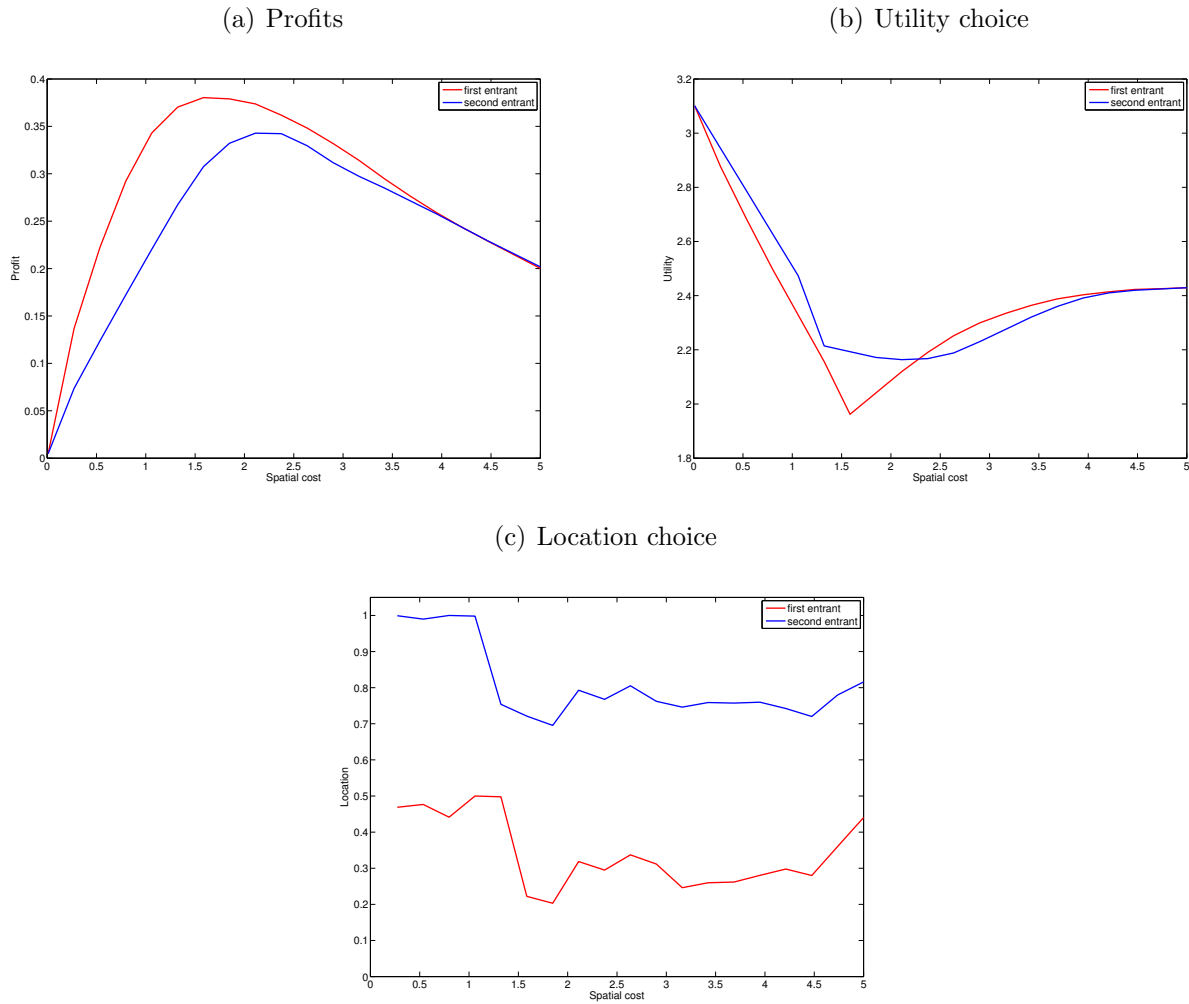
(a) Profits

(b) Utility choice

(c) Location choice

Figure 16: Adverse Selection, endogenous choice of locations

# 6 Relationship banking and origin of subprime lending

Here we consider another case of full commitment on contracts and locations choice with banks operating in different information regimes. There are two types (riskier and safer) among the agents that are exactly the same types described in adverse selection section of this paper.

The first entrant (a local bank or an incumbent) comes to the area and it studies his customers long enough to gain full information about their true types establishing relationships in the process. The first entrant anticipates that there will another player in the area (a global bank or a challenger) and it commits to location and contract menu based on established relationships trying to prevent a challenger from taking over his market share. The global bank doesn't have information advantage that the first entrant possesses and he operates in adverse selection regime using truth-telling constraints to offer the optimal menu of contracts for types. The global bank has the second mover advantage in a sense that he already knows both location and contracts offered by the local bank, so potentially it is capable to undercut the competitor if such strategy is profit maximizing.
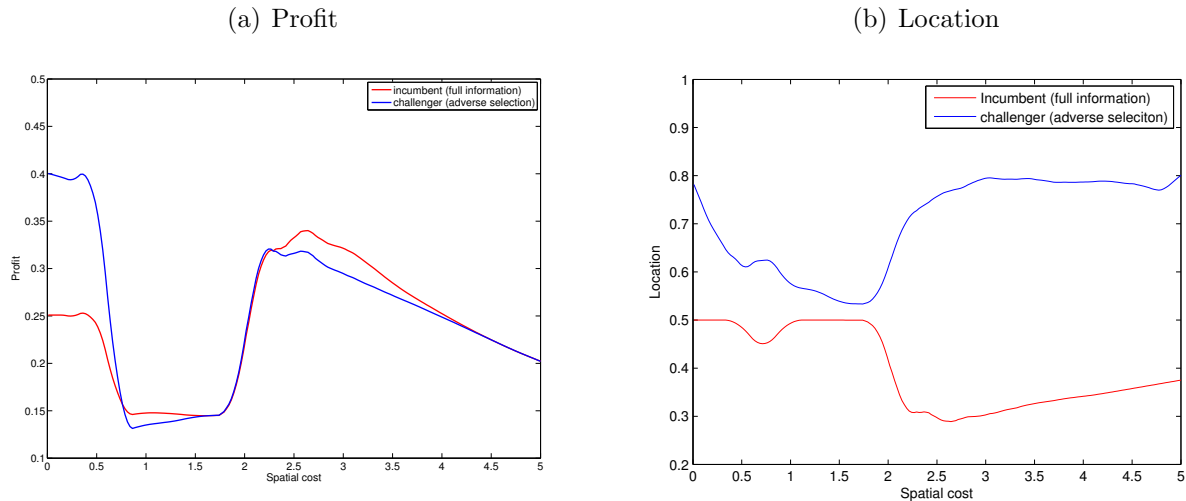
Figure 17: Information Advantage, profit and location choice

From Fig. 17(a) we see that both local and global bank are left with non-zero profit even at zero spatial costs. Overall, at most spatial costs the local bank makes larger profit and he keeps close to central location as shown in Fig. 17(b).

To understand both banks strategies we need to look at contracts and resulting market shares shown in Fig. 18 and Fig. 19 respectively.

We observe remarkable result that at low spatial costs the incumbent gets exactly 100% of good (safer) types and the challenger gets exactly 100% of bad (riskier) types. The global bank specializes in what can be called "subprime lending", the local bank keeps relationships established with the better clients with no subprime activity on the books. Why does it happen in such a dramatic way?
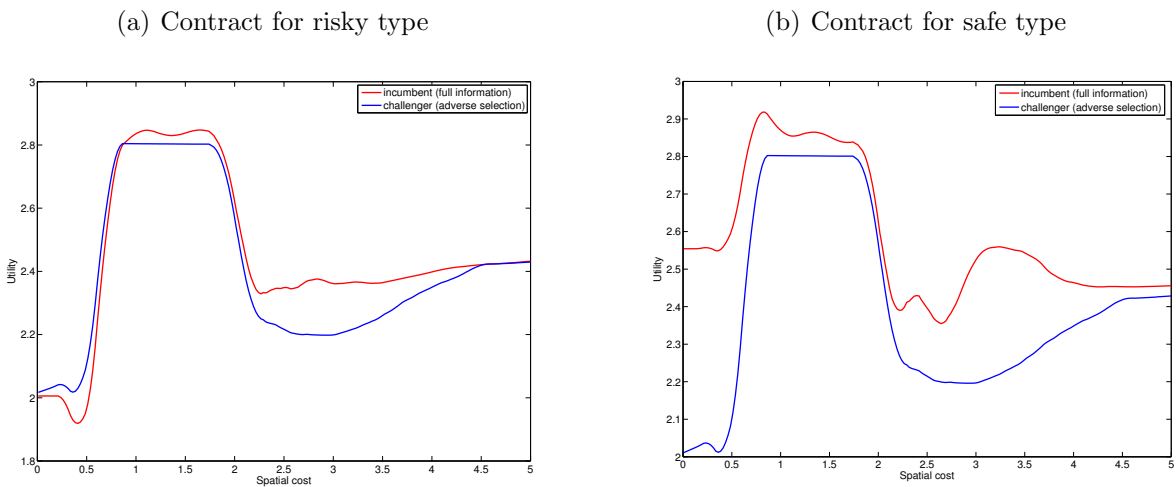
(a) Contract for risky type

(b) Contract for safe type



Figure 18: Information Advantage, menu of contracts offerred

When spatial costs are low enough the local bank sees the opportunity and it suddenly increases the gap between offers for riskier and safer types. The local bank stops caring about

attracting bad types and raises utility offer for good clients in such way so as to make global bank incapable of taking over by offering high utility for good clients. Since the global bank is constrained by truth-telling conditions it can't customize the offers in a way that full information bank is capable of. As a result, the global bank outcompetes local bank on riskier types by offering utility for riskier types just above the one posted by first entrant. It, however, loses completely on good (safer) types. The global bank is more than compensated by intermediating riskier types with higher profits resulting from local banks refusing to fight

This strategy changes, however, when spatial costs reach the level $\bar{l} \approx 1$ where attracting marginal customers from autarky becomes dominant problem. At such costs both banks raise utility offers for both types and they compete most fiercely with substantial drop in profit. The local bank has an advantage since it can structure a menu of contracts in such way so as to attract customers from autarky having different type dependent autarky utilities more efficiently compared to global bank. The global bank is forced by truth-telling conditions to underpay good types more, the local bank picks more of good types from autarky in result. Then, both banks completely separate spatially, utility offers merge and eventually each bank becomes a local monopolist at large spatial costs.
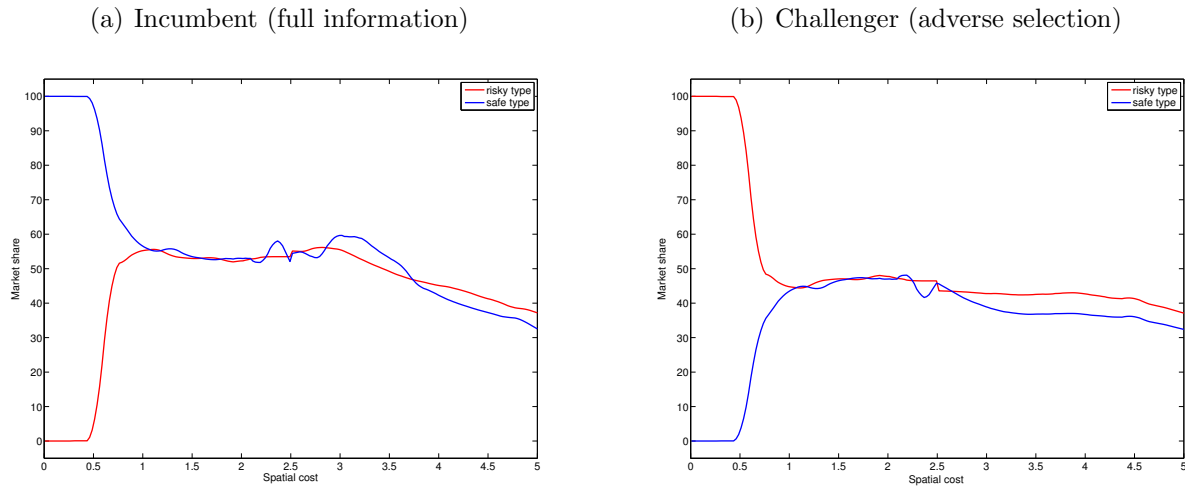
(a) Incumbent (full information)                    (b) Challenger (adverse selection)



Figure 19: Information Advantage, market shares

32

# 7   Dynamics and efficiency considerations

Virtually all of this paper could be extended to include explicit dynamics, i.e. multi-period contracts for borrowers, not simply dynamic IO considerations. On the household side, one need only put next period's promised utility $\omega'$ as an additional control variable. Karaivanov and Townsend (2012) is about likelihood estimation with this structure. Again equilibrium outcome will be information-constrained Pareto efficient, if there is perfect costless long term commitment on all sides. This allows both moral hazard and a kind of adverse selection (hidden types) simultaneously. However, in the context of sequential competition among a local incumbent and new entrant, it seems more likely that households cannot commit to stay with the local bank especially if there is a limit in the extent to which bank can front load the contract. This would be true either if the new entrant comes in as a surprise move, unanticipated, or if the incumbent anticipates. Under the latter, the outcome would almost surely not be information constrained efficient, as the first bank cannot take full advantage of long term contracts wich are otherwise beneficial. This is the first instance then in the paper where the IO equilibrium can be inefficient and may require some regulation, fine tuned to deal with this problem.

# 8   Estimation and Empirical Methods

We begin this discussion as if there were only one lender and exclude heterogeneity for borrowers, but pick this up below.

1. Given the economic environment, that is the parameters of preferences and technology, and given one or several obstacles to trade (full information and commitment as baseline, moral hazard, or exogenously incomplete contracts, etc) the optimized contract for a given promised utility to the borrower will generate a specific profit for a lender. As promised utility is varied we trace out the maximized surplus frontier. Though its level determines the overall profits at a given utility promise, a key characteristic is its elasticity, derivative divided by level.

Much of the discussion below on market outcomes is conditional on this surplus, and to the extent the surplus were the same despite variation in preference and technology parameters and underlying obstacles, then these are not identified and the market outcome will be the same. On the other hand, variation in levels of a surplus, even with the same elasticities, may help identify regimes.

2. For a given contract generating a particular level of expected utility, the same for all potential borrowers gross of disutility transportation costs and not varying with the location of borrower, that same contract also generates the size of the market, those who come to bank. A key characteristic is how this market size varies with promised utility, an elasticity, derivative divided by levels.

3. Thus observed variations in transportation costs and travel times (of going to branches of banks, given road networks) determine variation in that market elasticity.

4. A first order condition of a lender equates surplus/profit elasticities with market elasticities (assuming no boundary condition is binding and less than 100% of market is reached). Thus, we can effectively trace out the downward sloping surplus/profit function of the bank as a function of promised utility as transportation/travel times are varied. This is much like

varying demand and tracing out supply. The surplus function is a key to understanding both market structure and data points on underlying contracts through promised utilities. Ideally we would have a monopoly environment for a given spatial cost (assuming parameters of preferences and technology do not vary with the environment as well). This argument is weakened when spatial costs are observed with error, but travel times from villages to location are accurately measured on GIS systems. (when the 100% of the market taken and the lender is at a boundary condition, variation in promised utility with travel costs is still necessary, in order to retain customers at the extensive, autarky margin. Thus we can continue to trace out the surplus function as promised utility is varied, despite the corner.
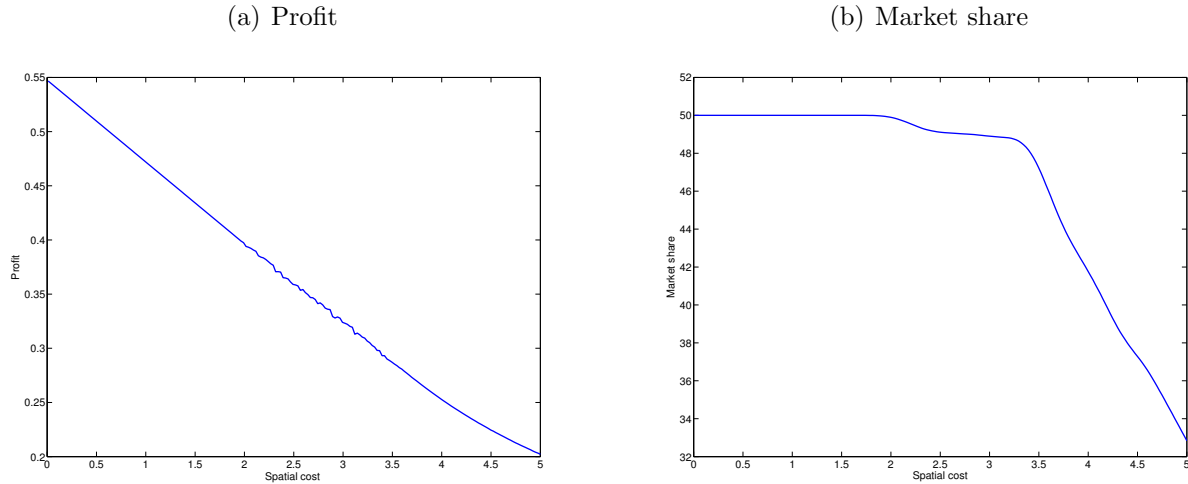
(a) Profit

(b) Market share



Figure 20: Monopoly, full information

5. Promised utility is not observed. We do see profits but with measurement error. Still, through the lens of the model and above considerations, we infer true underlying promised utilities with error. That is, the observed profit, and hence corresponding utility, may be observed without error, but that same profit may be observed from true profit levels nearby that are contaminated with error. For example, with a parameterized error structure we back out the distribution of true profits, given observed profit, simply the probability of error that generates observed minus actual. Obviously, the smaller is measurement effort the tighter the distribution.

6. For given surplus/profit function, again regardless of how it is generated from underlying contract obstacles and preference/technology parameters, different market structures will lead to different equilibrium outcomes in terms of branch locations and promised utilities especially as we consider the entire range of outcomes with variation in spatial costs. (One caveat, there can be jumps or discontinuities in profits for some market structures so if have only data from those markets we will not have data on part of the surplus frontier). On the other hand, and the more general point here, variation in observed outcomes can allow us to infer which market structure is in place. We consider collusion (effectively, multi branch monopoly), sequential Nash equilibrium in which banks enter one at a time with full commitment to both location and contract, version of simultaneous Nash equilibria in which banks compete ex post in contracts, and potentially in initial location as well. Subject to measurement error in

profits, and measurement error in location, data on location and profits (hence utilities) allow market structure to be inferred (with exceptions, some of which are created by occasional non existence problems). Actually, this is where observing location and costs without error can cause problems, if certain location decisions happen with probability zero in the model. To remedy this one can add shocks to location preferences of banks, as was anticipated in the text.

7. In principal we can put additional variables into household demand side over and above travel to branches. If there were proportional (marginal) costs for banks for generating a given surplus and these varied in an observable way, then vertical shifts in the surplus/profit function as this tax/subsidy is varied would trace out the the market elasticity on the household side, but note that variation in an outside interest rate is not an instrument in this sense as the cost of funds enters the consumers problem. With invertibility and other regularity conditions, we could likely back out something about the distribution of the additional factor.

8. We can let location of the borrower enter the contract, and this impacts equilibrium outcomes, but then we do need another factor which varies smoothly across agents, so mitigate jumps in reaction functions. Otherwise, current methods apply.

9. Related, we can introduce observed heterogeneity in some factor(s) impacting preferences and/or technology. Then a bank in practice, as in the model, would offer a menu of contracts , one for each observed type. We would see this variation in the data. Thus, most of the above arguments still apply in terms of identification of the curves and market structure (though location of a branch would be in common, serving multiple types).

10. If the type of a customer is unobserved, as in adverse selection and different risk types, or with different disutility of effort (with and without moral hazard,) then a bank in practice, as in the model, would still offer a menu of contracts, likely separating across types. The overall equilibrium outcome would however be different from the fully observable case, so conditioned on observables in a relevant application, one could postulate a distribution of unobservables and back out restriction on market outcomes. To confirm that there is a selection effect does require taking a stand on the source of unobserved heterogeneity, as the model with endogenous contracts has implications for default rates, induced effort and hence observed productivity, and so on.

11. Hence there is a strong interaction between obstacles (i.e. full information vs adverse selection) and the impact of a given market structure on observables such as profits, market share, locations.

12. For a given surplus/profit and for a given market size function, the equilibrium for a given market structure determines the promised utilities. Again the latter are contaminated with measurement error as profit is mismeasured. But conditional on a promised utility as if the true promise, (and the true surplus of a branch), the underlying contracting problem for the household determines the likelihood of observables such as consumption, output, investment. The fraction of households at various observables comes from the solution to the linear program, i.e. from histograms. Observables are given and the number of households at underlying grid points is a function of underlying parameters of technology and preferences. This generates the likelihood function, which is maximized. The key state variable of the model is promised utility (along with observables such as the capital stock). One can postulate these promises follow some distribution in the population, and estimate key parameters of that distribution. Here in this paper more structure from the supply side can be brought into play, in that the distribution of promised utilities is an endogenous object, and varies with the identity

of competitor, market structure, and equilibrium IO outcomes. Previous research establishes that underlying parameters and obstacles can be jointly identified from these data, though the evidence is not analytic, i.e. there is no proof, but rather through massive Monte Carlo runs. Some structure is needed but observed histograms on outputs and inputs can be inserted so as to make this estimation partially non parametric.

13 The most power will be gained by consideration of the joint equilibrium determination of contracts and branch locations if the model is correct, or at least a good approximation to reality.

# References

Agarwal, S. and R. Hauswald (2010). Distance and private information in lending. *Review of Financial Studies 23*, 2757–2788.

Alessandrini, P., F. F. Presbitero, and A. Zazzaro (2009a). Bank size or distance: what hampers innovation adoption by SMEs? *Journal of Economic Geography 10*, 845–881.

Alessandrini, P., F. F. Presbitero, and A. Zazzaro (2009b). Banks, distances and firms financing constraints. *Review of Finance 13*, 261–307.

Assuncao, J., S. Mityakov, and R. Townsend (2012). Ownership matters: the geographical dynamics of BAAC and commercial banks in Thailand. Working paper.

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), 21852204.

Butler, A. W. (2008). Distance still matters: Evidence from municipal bond underwriting. *Review of Financial Studies 21*(1).

Degryse, H. and S. Ongena (2005). Distance , lending relationships , and competition. *Journal of Finance 60*, 2757–2788.

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), 27632801.

Einav, L., A. Finkelstein, and J. Levin (2010). Beyond testing: Empirical models of insurance markets. *Annu. Rev. Econ. 2*, 311–336.

Einav, L., A. Finkelstein, S. P. Ryan, P. Schrimpf, and M. R. Cullen (2013). Selection on moral hazard in health insurances. *American Economic Review 103*, 178–219.

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

Karaivanov, A. and R. Townsend (2012). Dynamic financial constraints: Distinguishing mechanism design from exogenously incomplete regimes. *Working Paper, MIT*.

Koijen, R. S. and M. Yogo (2012). The cost of financial frictions for life insurers. *NBER Working Paper No. 18321*.

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

Martinez-Miera, D. and R. Repullo (2010). Does competition reduce the risk of bank failure? *Review of Financial Studies 23*(10), 3638–3664.

Moscarini, G. and M. Ottaviani (2001). Price competition for an informed buyer. *Journal of Economic Theory 101*, 457–493.

Nicoló, G. D. and J. Boyd (2005). The theory of bank risk-taking and competition revisited. *Journal of Finance 60*(3), 1329–1343.

Nicoló, G. D., J. Boyd, and B. Smith (2004). Crises in competitive versus monopolistic banking systems. *Journal of Money, Credit, and Banking 36*(3), 487–506.

Petersen, M. A. and R. G. Rajan (1995). The effect of credit market competition on lending relationships. *The Quarterly Journal of Economics 110*(2).

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

Rajan, R. G. and M. A. Petersen (2002). Does distance still matter? The information revolution in small business lending. *Journal of Finance 57*(6).

# Appendix

# A    Distance to Nash

Here we propose the following technique to order by rank all possible strategies by metric we call "distance to Nash". Both banks enter the market simultaneously and use a strategy set of $G = \{\omega_1^*, \omega_2^*, l_1^*, l_2^*\}$ with payoff $P_1(G) = S^{\omega_1^*}\mu_1(\omega_2^*, l_2^*, \omega_1^*, l_1^*)$ and $P_2(G) = S^{\omega_2^*}\mu_2(\omega_2^*, l_2^*, \omega_1^*, l_1^*)$.

If a strategy set $G$ is a true Nash equilibrium then there exist no deviating strategy sets $G_1 = \{\omega_1, \omega_2^*, l_1, l_2^*\}$, $G_2 = \{\omega_1^*, \omega_2, l_1^*, l_2\}$ that would provide higher profit for any of the correspondingly deviating banks taking other bank strategy as given. We can define and compute for any of those deviating strategies the following metrics

$$d(G, G_1) = \max(P_1(G_1) - P_1(G), 0), P_1(G_1) = S^{\omega_1}\mu_1(\omega_2^*, l_2^*, \omega_1, l_1)$$

$$d(G, G_2) = \max(P_2(G_2) - P_2(G), 0), P_2(G_2) = S^{\omega_2}\mu_2(\omega_2, l_2, \omega_1^*, l_1^*)$$

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

$$\underset{G1}{\text{maximize }} d(G, G_1) \text{ subject to } P_1(G) > 0, \ \forall G_1. \tag{16}$$

$$\underset{G2}{\text{maximize }} d(G, G_2) \text{ subject to } P_2(G) > 0, \ \forall G_2. \tag{17}$$

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

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

And in the final stage we solve

$$\underset{G}{\text{minimize }} d(G), \ \forall\{G, G_1, G_2\}. \tag{18}$$

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

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

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

### A.0.1    Numerical accuracy of distance to Nash algorithm

We report distance to Nash values in the units of bank profits on Fig. 21.

The distance to Nash is often exactly zero and when it is bounded $d(G) < \lambda * P(G)$ with $\lambda \approx 0.01$ or better we accept the outcome as an instance of Nash eqilibrium. We conduct the same accuracy checks for each case of simultaneous Nash equilibrium we study. Although we don't provide proofs of existence and sufficiency conditions here, those checks serve to filter numerically well-bounded constructively obtained equilibria from outcomes where we can't claim that Nash eqilibrium is found.

For example, at Fig. 22 there exist a range of spatial costs $1 < \overline{L} < 2$ where distance to Nash is comparable to profit level when measured in the same units meaning that Nash eqilibrium might not exist.
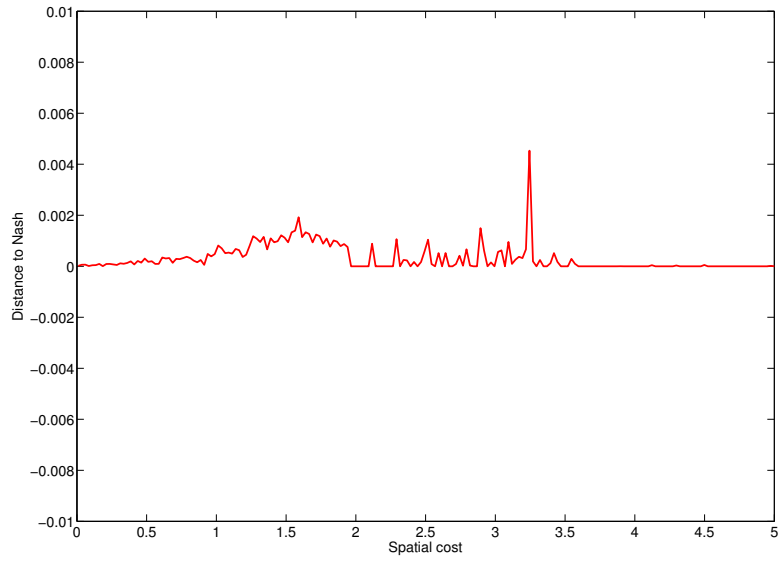
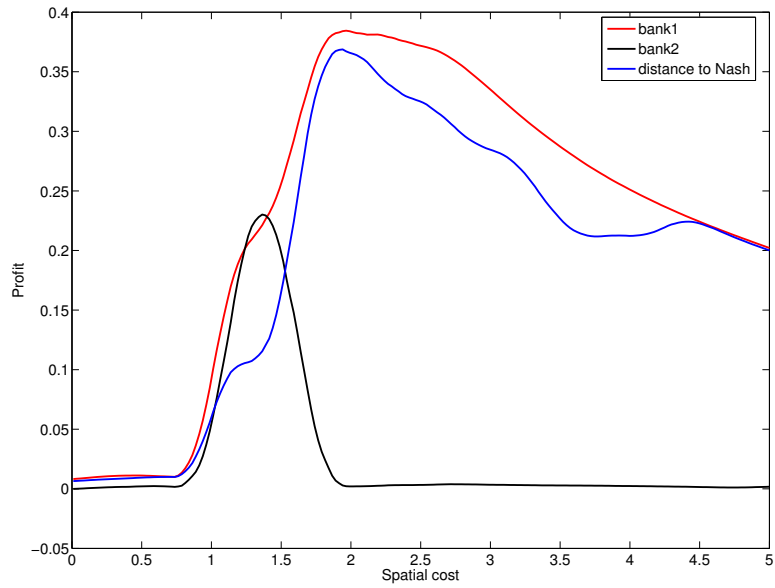Figure 21: Distance to Nash, Full Insurance, two types, fixed locations



Figure 22: Profits and distance to Nash, full information, no commitment