# Essays on Information and Competition in Banking and Service Industries

Autor: Silvio Petriconi

TESI DOCTORAL UPF / ANY 2014

DIRECTORS DE LA TESI

Prof. Jaume Ventura
Prof. Vasco Carvalho

(Centre de Recerca en Economia Internacional, CREi)

upf. Universitat Pompeu Fabra Barcelona

To my daughter, Laura.

# Acknowledgments

It is a pleasure to thank the many people who have supported my research in many different ways.

First and foremost, I am deeply indebted to my two advisors, Jaume Ventura and Vasco Carvalho. Not all doctoral students are in the fortunate position to continuously receive valuable comments, feedback and guidance from their advisor. In this aspect, I always thought of myself as blessed twice: I have had the rare luck to have not only one but even two wonderful advisors for the entire time of my dissertation work. Their continued feedback and support has been crucial for me. I wish to extend my sincere gratitude to both of them.

For the banking-related part of my research, I have enjoyed many interesting conversations with Xavier Freixas. I thank him for his continued interest in my work, for his guidance and for many valuable suggestions. Many thanks also go to Doh-Shin Jeon with whom I have co-authored the paper which forms the last chapter of this thesis. He has taught me many important things about research in general and about the theory of Industrial Organization in particular, for which I am deeply grateful.

Further important feedback has emerged from many conversations with other faculty members, visitors and students at Universitat Pompeu Fabra and CREI. I especially wish to thank Giovanni Dell'Arriccia, Joan de Martí, Robert Marquez, Alberto Martin, Kristoffer Nimark and the participants of the CREI International Lunch Seminar for many helpful comments.

In all the administrative tasks ranging from thesis registration and scholarship renewals to the mailing of my job market materials, the GPEFM secretaries Marta Araque and Laura Augusti have been nothing short of magicians. Their competent and efficient handling of all those urgent matters has allowed me to keep my mind focused on my research while they fought bureaucratic monsters on my behalf. I thank them for their tremendous support.

I am grateful for the funding sources that enabled me to finance my graduate school studies and to cover my expenses when presenting at conferences: the Studienstiftung-Hölderlin scholarship, the Generalitat de Catalunya FI grant, and the EBES travel grant of Universitat Pompeu Fabra.

Finally, I wouldn't be who I am today without my family and friends who have supported me so much throughout my entire life. In particular, there are no words how much I owe to my wife Sofia. Without her love, encouragement and understanding I would never have completed this thesis.

# Abstract

This thesis investigates the relation between information and competition in two particular settings. In the first chapter, I ask which screening and lending choices banks make if their loan offers can be observed by uninformed competitors. I show that banks grant excessive credit and underinvest in screening in order to deter borrower poaching; more competition exacerbates this problem. In a dynamic setting, the model reproduces many stylized facts of lending booms.

The second chapter (jointly with Doh-Shin Jeon) investigates the incentives behind commercial open source software. We find that the production of open source software by profit-maximizing firms can be understood as a form of optimal information sharing: by releasing source code, service firms create their own competition which mitigates customer concerns about hold-up problems. We show that the resulting increases in customer investment and collaboration can yield higher profits for the service firm.

# Resum

Aquesta tesi investiga la relació entre la informació i la competència en dues situacions. En el primer capítol, s'analitzen les opcions de cribratge i de préstecs que els bancs fan si les seves ofertes de préstecs poden ser observats pels competidors desinformats. Mostro que els bancs concedeixen préstecs excessius i inverteixen poc en el cribratge per tal de dissuadir la caça furtiva prestatari; més competència exacerba aquest problema. En un entorn dinàmic, el model reprodueix molts fets estilitzats dels auges de crèdit.

El segon capítol (en col·laboració amb Doh-Shin Jeon) investiga els incentius darrere de programari de codi obert comercial. Ens trobem que la producció de programari de codi obert per les empreses pot ser entesa com una forma de compartir la informació de manera òptima: alliberant el codi font, les empreses de serveis creen la seva pròpia competència, que mitiga les preocupacions dels clients sobre problemes de captivitat. Es demostra que els augments resultants de la inversió del client i la col·laboració poden rendir més beneficis per a l'empresa de serveis.

# Foreword

This thesis comprises of two essays which study the strategic role of information for competition in banking and knowledge-based service industries.

In the first essay, I ask how banks make their loan approval decisions and select their screening activity if their loan offers can be observed by uninformed rival lenders. I find that the impact of such an informational imperfection on screening incentives and lending standards can be rather dramatic: in order to prevent uninformed outsiders from poaching their customers, informed banks will intentionally make "more erratic" lending decisions and will include too many borrowers of poor quality in their loan portfolio. Moreover, they will intentionally underinvest in screening and accept that their screening decisions remain quite noisy. Both effects aggravate the adverse selection problem that an uninformed outsider faces which helps informed banks to protect themselves from competition. Most interestingly, these distortions in credit are not constant in time but vary substantially as the expected cash flows of projects change: I find that especially in times of low interest rates and high collateral values banks are likely to lend too much and to screen too little. I argue that the predictions of the model are well in line with many known stylized facts about credit booms. Finally, I explore a dynamic version of the model in which banks' choices of screening precision exhibit some sort of rigidity. This extended model generates boom-bust cycles in credit that are increasing in the amount of bank competition: it features both excessive credit in boom times and inefficiently severe lending contractions in recessions.

The second essay, jointly with Doh-Shin Jeon, sheds more light on the incentives for profit-maximizing IT firms to engage in open source software development. Using a fairly general dynamic model of knowledge-based service exchanges we show that open source software can be understood as a specific example of optimal information sharing in knowledge-based service industry: our finding is that it may be optimal for knowledge-based service firms to share some of their knowledge stock with rival firms at no charge. This surprising result arises because most knowledge-based service firms depend heavily on working in close collaboration with their customers, yet virtually all relevant dimensions of this collaboration (e.g., delivered service value; customer collaboration effort) are not contractible. Since an incumbent service firm can not commit not to hold up customers for their participation, customer participation remains inefficiently low. By releasing its software as open source and thus strengthening competitors, an incumbent service firm can commit to leave adequate surplus to its customers, which mitigates hold-up concerns and induces stronger customer participation. Our theory also sheds light on how the openness of a firm's open source software development process affects competition in knowledge-based IT service industry.

# Contents

# Chapter 1

# BANK COMPETITION, INFORMATION CHOICE AND INEFFICIENT LENDING BOOMS

## 1.1 Introduction

Over the past two decades, many countries worldwide have undertaken a process of deregulation[1] of their financial sectors that has turned tightly regulated oligopolistic banking landscapes into much more competitive industries. It is widely believed that these reforms have provided credit-constrained firms and self-employed workers with better access to bank credit. However, the increasing occurrence of credit-driven boom-bust cycles[2] raises the question whether this increase in competition may have come at a cost: anecdotal evidence suggests that in episodes of strong economic outlook, credit may actually have become *too easy* to obtain. In good times, so the narrative goes, competitive pressure drives banks to reduce their screening effort and lower their lending standards, resulting in the build-up of large positions of poorly screened assets in their balance sheets that lay the foundations for the next financial crisis.

In this chapter, I develop a theory of excessive lending that is driven by banking competition. In a model of borrower screening, I obtain two main results: first, more competition can incentivate informed incumbent banks to knowingly

[1]e.g. in the U.S. the 1994 Riegle-Neal Act eliminated previous interstate baadvisornking and branching restrictions, and the 1999 Gramm-Leach-Bliley Act repealed the separation between investment and commercial banking. For information on banking liberalization across countries, see Abiad et al. (2010).

[2]For stylized facts and more discussion of credit-driven boom-bust cycles, see Borio and Lowe (2002), Tornell and Westermann (2002), Eichengreen and Mitchener (2003), Mendoza and Terrones (2008, 2012), Hume and Sentance (2009) and Schularick and Taylor (2009).

take bad projects into their loan portfolio. This is optimal because it protects them against the entry of uninformed lenders who could poach their customers. Instead of raising welfare by reducing markups, more bank competition exerts detrimental downward pressure on lending standards. Since in equilibrium banks offer credit to some borrowers despite of a negative screening evaluation, they also fail to make best use of their information which leads to my second main result: if banks choose their screening precision endogenously, I find that competition reduces borrower screening to inefficiently low levels. The combination of these two results provides me with a powerful theory of the credit cycle.

The key to my findings is to acknowledge that just like all other information centric industries, banks may face difficulties in protecting their private information when assessing credit. Specifically, I assume in my model that the lending decisions of privately informed banks can be observed by uninformed competitors. This could for example be the case if banks are legally required to present their loan offers in writing. Loan-approved borrowers could then use the written offer to credibly signal their positive evaluation to an uninformed outside lender in the attempt to receive a better offer.

When actions are observable by competitors, an informed incumbent bank must strike a careful balance between optimizing her portfolio quality and protecting herself from competition: the more wisely she chooses her loan portfolio, the more positive information will be conveyed by every loan approval, and the more profitable it becomes for uninformed outside lenders to enter the incumbent's market and poach loan-approved customers. If the informed incumbent "poisons the well" by making less prudent choices for her own loan portfolio, she will clearly suffer from more defaulting loans but may be rewarded with the successful deterrence of entry.

I show that the distortions in screening and lending decisions that arise from banks' optimal response to informational free-riding problems reproduce the known stylized facts about credit procyclicality remarkably well. When the economy emerges from a recent recession, the composition of the borrower pool is sufficiently poor that competitive threats barely matter at all. Banks then screen very precisely to mitigate adverse selection, and the average quality of loans that are originated in such times is high. As economic conditions improve, competitive threats begin to distort banks' lending decisions: lending standards fall, and credit grows at an increasing pace because banks start to admit some negative net present value projects into their loan pool in order to keep competition away. At the same time, they reduce the precision of borrower screening because they can no longer appropriate the full returns to their costly information acquisition effort. In a booming economy, the average borrower becomes credit-worthy. The reduced level of screening precision then contributes further to the expansion of inefficient lending because the combination of high prior belief and imprecise screening information leaves many poor projects in the portfolio that is financed. In line with empirical evidence, my model predicts that the worst loans are made right at the peak of the boom, and that low policy rates can exacerbate inefficient lending. Once the boom goes bust, the situation reverses. If screening precision requires time to build, the bust phase will be characterized by a credit crunch and a flight to quality: as screening remains impaired from the effects of competition during the preceding boom phase, only exceptionally good borrowers succeed in generating a sufficiently positive signal to receive credit.

Finally, my model also emphasizes the important role of project liquidation values for the lending cycle. In recessions, expected liquidation values are low. This makes banks hesitant to lend because the option of repossessing defaulted projects offers them little protection against risk. The opposite is true when liquidation values are high: not only will those high values reduce banks' incentives to screen thoroughly (Manove et al., 2001), but in the presence of competition they can also exacerbate the inefficiencies in bank lending: this is because diminishing risk encourages uninformed outside banks to compete more aggressively, and incumbents respond to this threat of competition by making more bad loans.

My work is related to several strands of literature. First, it contributes to the literature on credit screening. Imperfect credit-worthiness testing in competitive markets has been analyzed, among others, by Broecker (1990) and Riordan (1993). However, they focus on externalities between banks that arise from credit rejection decisions which exacerbate adverse selection within the applicant pool. In sharp contrast, I analyze here a setting of contestable markets where banks' lending decisions have no impact on the composition of competitors' applicant pools, but where informational spillovers limit the ability of lenders to fully appropriate the returns to costly private information gathering. Lending standards are softened in this type of competition as a result of cross-subsidization between supra- and submarginal projects, which is somewhat similar to de Meza and Webb (1987). Ruckes (2004) and Direr (2008) analyze credit screening models based on Broecker (1990) in which the choice of information is endogenous, just as it is here. The scope of their analysis, however, is different from mine: they are interested in the screening activity of banks in the *absence* of any informational spillovers whereas I study the very consequences of such an imperfection.

To the best of my knowledge, the endogenous screening choice in the presence of informational spillovers between competing lenders has so far only been analyzed by Ogura (2006). In his model, lenders use observational learning to update their beliefs regarding the credit-worthiness of a borrower from a competitor's lending decision in the previous period which leads to reduced screening. There are, however, substantial differences: my model here is a story of informed inside banks who adjust the composition of their loans in the face of competing uninformed outside lenders, whereas his model analyzes the lending decisions of equally informed transactional lenders that are confined to financing only at most one single customer each. This assumption precludes the use of cross-subsidization strategies to deter entry from informational free-riders, an idea that lies at the heart of this paper. Thus, he finds that increasing competition reduces the credit risk taken by every individual bank, which is the opposite of my conclusions.

The idea that adverse selection can constitute an effective barrier to bank entry goes back to Dell'Ariccia et al. (1999). They show that the rejection decisions from Bertrand competition between two informed lenders can deteriorate the applicant pool for a third entrant so much that entry is blockaded.

This paper is also related to the literature on the impact of bank competition on information and lending standards. Dell'Ariccia and Marquez (2006) present a model of a credit boom driven by increases in competition, but their booms, whilst possibly increasing financial fragility, are surplus-increasing and thus suggest substantially different policy than the inefficient booms which I discuss here. Both Marquez (2002) and Hauswald and Marquez (2006) address information in a banking system under increasing levels of competition, where the latter present the idea that increases in competition reduce returns to information; however, in their model information acquisition nevertheless remains socially excessive, which is not the case here.

Finally, my model also adds to the literature that explores the economic mechanisms behind lending booms. Rajan (1994) develops the idea that in the presence of thriving competitor banks, short-horizon reputational concerns of bank managers can promote inefficiently lax lending standards. In Lorenzoni (2008), lenders' failure to internalize the general equilibrium price impact of their liquidation of collateral gives rise to excessive lending. Both of these theories are complimentary to the idea of informational problems that I pursue here.

## 1.2 Model

### 1.2.1 Setup

The economy in my model comprises of two islands, with a unit mass of entrepreneurs and a financial intermediary on each of these islands.

Every entrepreneur $i \in [0, 1]$ is endowed with a risky project. The project requires one unit of investment, and may either succeed and yield a perfectly verifiable payoff of $R > 1$, or fail and yield its liquidation value, $r < 1$:

$$X_i = \begin{cases} R & \text{with probability } p_i \\ r & \text{with probability } 1 - p_i \end{cases} \tag{1.1}$$

The quality of projects is heterogeneous in the sense that some projects are more likely to succeed than others: the success probability $p_i$ of entrepreneur $i$'s project is drawn independently from a uniform probability distribution with mean $\bar{p}$:

$$p_i \sim U\left(\bar{p} - \frac{\varepsilon}{2}, \bar{p} + \frac{\varepsilon}{2}\right) \tag{1.2}$$

I assume that this distribution is the same for both islands and is publicly known whereas an individual project's success probability $p_i$ is only known to the entrepreneur. Entrepreneurs have zero initial wealth, so they need to borrow an amount of $1$ against the state-contingent promise of repayment of $(D_i, d_i)$ from a bank in order to develop their project. I assume that all entrepreneurs apply for a loan irrespective of their type[3], and that they do not have any signaling devices such as collateralization or self-financing at their disposal. Banks are risk-neutral and can access an unlimited amount of financing through the interbank market at a cost of $\rho$. I will interpret $\rho$ directly as the monetary policy rate, and assume that $r < \rho < R$. Before coming to a decision on a loan, banks can screen every *domestic* project (i.e. every project on the same island) by using all soft and hard information available to them to generate an informative but imperfect private signal about its idiosyncratic probability of success.

To this end, they resort to an imperfect credit-worthiness test. The precision of this test depends on a parameter $\lambda \in [0, 1]$ which captures the extent to which the lender's institutional setup and business strategies are well-aligned with the purpose of generating reliable credit-relevant information and channeling it to the decision-making loan officers. In the case $\lambda = 0$ the test does not generate any useful information at all, whereas for $\lambda \to 1$ the test reveals every borrower's type perfectly. By making its choice of $\lambda$, a bank endogenously selects its optimal degree of exposure to asymmetric information problems in its lending activities. The cost associated with holding a given level $\lambda$ of precision in credit testing are primarily fixed cost which are independent of the number of screened customers and are described by a strictly convex cost function $c(\lambda)$ that satisfies $c' > 0$, $c'' > 0$, $\lim_{\lambda \to 0} c(\lambda) = 0$ and $\lim_{\lambda \to 1} c(\lambda) = \infty$; I assume that once these fixed cost are paid, the screening of an individual domestic applicant is costless and yields an imperfect signal of his true type $p_i$. I refer to the signal's random variable as $\sigma_i$, and denote its realization as $s_i$. Active screening only works with domestic entrepreneurs; I assume that the screening of an entrepreneur from a different island always results in a completely uninformative signal.

Whilst the key results in this paper generalize to a wide class of screening technologies I postulate for now that the following specific test technology is in

[3]This could, for example, be ensured by postulating that entrepreneurs derive an additional non-verifiable rent from running their projects.

place: with probability $\lambda$, the credit-worthiness test generates a signal realization $s_i = p_i$ that is identical to the project's true success probability whereas with probability $1-\lambda$ it yields a totally uninformative value $s_i$ that is randomly drawn from $p_i$'s prior distribution (1.2). In sharp contrast to the signal structure analyzed by Ruckes (2004), the bank does not know whether a given signal realization represents a "real" or an uninformative draw. I show in the appendix that Bayesian updating of beliefs given the observation of a signal realization of $s_i$ results in a posterior expectation of

$$E[p_i|\sigma_i = s_i] = \lambda s_i + (1-\lambda)\bar{p} \tag{1.3}$$

Note that the posterior expected value of $p_i$ is simply a convex combination of the observed signal realization and the prior mean $\bar{p}$ whereby the test's precision determines the relative weight given to each of them. A more precise test will therefore move beliefs about $p_i$ further away from the prior mean and create more dispersion in posterior expectations. As we shall see later, this is the first key ingredient for understanding the central properties of the model.

### 1.2.2 Competition

The second key ingredient to the model is competition between asymmetrically informed lenders where an imperfection of the market gives rise to informational spillovers. Specifically, I assume that legislation prohibits charging fees to a loan applicant for the review of a loan application before actually closing the loan[4] and that any loan offer must be made in writing. Entrepreneurs with a favorable loan evaluation by their domestic bank can then use the written offer to credibly signal their positive domestic evaluation to the uninformed outside lender on the other island. In this way, the uninformed bank on the other island can attempt to poach customers which the domestic bank considers credit-worthy. However, a cost disadvantage limits the ability of the outside bank to compete for customers that do not reside on its island: I assume that for every loan made to an off-shore customer, the lender incurs a cost of $\gamma > 0$. Whilst there are several possible interpretations[5] to this charge, I prefer to think of it as the extra cost from monitoring the execution of a project that does not reside on the same island. In my analysis, I use $\gamma$ to parametrize changes in the strength of competition: $\gamma \to \infty$ will result in monopoly, whereas $\gamma \to 0$ generates the highest possible level of competition.

[4]For example, this is the case under the German law.

[5]I discuss alternative interpretations of $\gamma$ at the end of this chapter.

### 1.2.3 Timing

I now let the two banks compete with each other. In particular, I assume the following structure and timing of the game:

1. every bank $j \in \{1, 2\}$ chooses its publicly observable level of screening effort $\lambda^j$, pays screening cost $c(\lambda^j)$ and obtains private signals $\sigma_{i,\lambda}$ for all domestic projects $i \in [0, 1]$.

2. both banks then make their domestic loan offers: they choose a domestic *loan portfolio* comprising of a set $\mathcal{P}_j$ of projects to be offered a loan, and agreed repayment terms $(D_i, d_i)$ for every project $i \in \mathcal{P}_j$ in case of success or failure, respectively.

3. each bank observes the domestic loan offers made on the other island $j' \neq j$ and decides whether and under which state-contingent repayment terms $(O_i^{j\prime}, o_i^{j\prime})$ it offers outside credit to the other islands' loan-approved entrepreneurs.[6]

4. entrepreneurs choose the loan offer with the lowest expected repayment terms; when two offers leave them indifferent, they stay with their domestic bank.

I solve the game by backward induction. Before doing so, I will however make a quick detour and analyze the information precision and credit allocation that a social planner who is constrained to using the same screening technology would choose.

## 1.3 The Social Planner's Solution

### 1.3.1 Constrained Optimal Portfolio Choice

Since I solve backwards, I first take $\lambda$ as exogenously given. The constrained optimal portfolio is easily found by observing that the expected surplus $\pi_i$ from a single project is monotonically increasing in its expected probability of success and reads

$$E[\pi_i|\sigma_{i,\lambda} = s_i] = E[p_i|s_i]R + (1 - E[p_i|s_i])r - \rho \tag{1.4}$$

[6]Note that only domestically loan-approved entrepreneurs have incentive to apply for a loan from the outside bank, and that the outside lender will never find it profitable to offer credit to entrepreneurs without domestic loan approval.

Equating this expression to zero yields that a project with an expected success probability of

$$q \equiv \frac{\rho - r}{R - r} \tag{1.5}$$

will contribute exactly zero surplus, and every project with a higher expected success probability will generate positive surplus. The constrained optimal portfolio must therefore include all projects that have an expected success probability above the cut-off threshold $q$, and deny financing to all those projects that fall below it. As we shall see, all important expressions in the paper can be written in terms of this cut-off threshold $q$, so it deserves a brief discussion at this point: holding $R$ and $r$ constant, it is obvious that $q$ lies between $0$ and $1$ when the interbank rate $\rho$ moves within its assumed boundaries, $r < \rho < R$. Taking $R$ and $\rho$ constant and varying $r$ it is also easy to see that higher liquidation values $r$ decrease the probability of success $q$ that is required to break even.

With the cut-off threshold $q$ at hand, it is straightforward to calculate the volume and average success probability of the constrained optimal portfolio. It is useful to first derive the cumulative distribution function of $E[p_i|\sigma_{i,\lambda}]$,

$$\mathcal{E}_\lambda(q) \equiv \text{Prob}\left(E[p_i|\sigma_{i,\lambda}] \leq q\right) \tag{1.6}$$

which can be obtained using equation (1.3):

$$\mathcal{E}_\lambda(q) = \text{Prob}\left(\sigma_{i,\lambda} \leq \frac{q-(1-\lambda)\bar{p}}{\lambda}\right) = \begin{cases} 0 & \text{if } q \leq \bar{p} - \frac{\lambda\varepsilon}{2} \\ \frac{q-\bar{p}}{\lambda\varepsilon} + \frac{1}{2} & \text{if } \bar{p} - \frac{\lambda\varepsilon}{2} < q < \bar{p} + \frac{\lambda\varepsilon}{2} \\ 1 & \text{if } q \geq \bar{p} + \frac{\lambda\varepsilon}{2} \end{cases} \tag{1.7}$$

The constrained efficient portfolio's mass of credit is then simply $m_\lambda^{SB} = 1 - \mathcal{E}_\lambda(q)$, whilst the portfolio's average success probability can be calculated as

$$\mathcal{A}_\lambda^{SB} = \frac{\int_q^1 p \,\mathrm{d}\mathcal{E}_\lambda(p)}{\int_q^1 \mathrm{d}\mathcal{E}_\lambda(p)} \tag{1.8}$$

The results are conveniently summarized in the following proposition:

**Proposition 1.** *The constrained efficient portfolio contains all projects with expected success probability above the cut-off* $q = \frac{\rho - r}{R - r}$ *and none below. The optimally financed mass of projects reads*

$$m_\lambda^{SB} = \begin{cases} 1 & \textit{if } q \leq \bar{p} - \frac{\lambda\varepsilon}{2} \\ \frac{1}{2} - \frac{q-\bar{p}}{\lambda\varepsilon} & \textit{if } \bar{p} - \frac{\lambda\varepsilon}{2} < q < \bar{p} + \frac{\lambda\varepsilon}{2} \\ 0 & \textit{if } q \geq \bar{p} + \frac{\lambda\varepsilon}{2} \end{cases} \tag{1.9}$$

*and the projects in the portfolio attain an average success probability of*

$$\mathcal{A}_\lambda^{SB} = \bar{p} + \frac{\lambda\varepsilon(1 - m_\lambda^{SB})}{2} \tag{1.10}$$

*Proof.* see appendix. □

Remembering that the limit $\lambda \to 1$ yields the perfect information case, we can now directly compare the second-best choice with imperfect testing precision $\lambda < 1$ to the outcomes of a (clearly unattainable) first-best world in which project types are perfectly observable:

**Corollary 1.1.** *The second-best allocation when project types are noisily observed compares to the allocation when types are perfectly known as follows:*

- *If $\bar{p} < q$, strictly less projects are developed than in the first-best with perfect information. Moreover, the size of the constrained efficient portfolio is increasing in $\lambda$.*

- *If $\bar{p} > q$, the amount of projects that are developed exceeds the perfect information case. The size of the constrained efficient portfolio is decreasing in $\lambda$.*

In other words, imprecise information drives a wedge between the second- and the first-best amounts of financing whereby the sign of the wedge depends on whether the average project is credit-worthy: in exceptionally good times (i.e. $q < \bar{p}$), noisy information results in the financing of *more* projects than under perfect observability of types, whilst in normal times (i.e. $\bar{p} < q$), noisy information leads to *less* investment.

### 1.3.2 Constrained Optimal Screening Precision

Having determined the planner's optimal project portfolio decision, the next step is to ask how much screening precision $\lambda$ the social planner would optimally choose to acquire. For this purpose I first calculate the gross surplus as a function of $q$ and $\lambda$:

$$\begin{aligned}
\Pi_\lambda^{SB} &= \int_q^1 pR + (1-p)r - \rho \,\mathrm{d}\mathcal{E}_\lambda(p) \\
&= (R - r)\int_q^1 \left(p - \tfrac{\rho - r}{R - r}\right) \mathrm{d}\mathcal{E}_\lambda(p) \\
&= m_\lambda^{SB}\,(R - r)\left(\mathcal{A}_\lambda^{SB} - q\right) \qquad (1.11)
\end{aligned}$$

Note the very intuitive structure of this expression: for every financed project, the planner obtains a surplus that is equal to the risky part $R - r$ of the project's payoffs multiplied by the margin by which the portfolio's average success probability exceeds the zero-surplus cut-off $q$. Substituting previous results leads to

$$\Pi_\lambda^{SB} = \begin{cases} (R-r)(\bar{p}-q) & \text{if } q \leq \bar{p} - \frac{\lambda\varepsilon}{2} \\ (R-r)\frac{(2\bar{p}-2q+\lambda\varepsilon)^2}{8\lambda\varepsilon} & \text{if } \bar{p} - \frac{\lambda\varepsilon}{2} < q < \bar{p} + \frac{\lambda\varepsilon}{2} \\ 0 & \text{if } q \geq \bar{p} + \frac{\lambda\varepsilon}{2} \end{cases} \tag{1.12}$$

The three regions in this equation are the same as for the constrained-efficient project mass in equation (1.9): the first region of the equation describes the case in which projects are so good and information is so imprecise that the size of the constrained optimal portfolio is $1$. Region two stands for the intermediate case in which $0 < m_\lambda^{SB} < 1$; finally, the third region is relevant whenever screening precision is so low and average projects are so bad that it is impossible to identify any project with positive gross surplus at all.

The constrained efficient information choice $\lambda_{SB}^*$ maximizes surplus net of information cost:

$$\begin{aligned} \max_{\lambda} \quad & \Pi_\lambda^{SB} - c(\lambda) \\ \text{s.t.} \quad & \lambda \geq 0 \end{aligned} \tag{1.13}$$

Although the exact shape of the solution to this problem depends on the specific functional form of $c(\lambda)$ about which I have not made any assumption so far, its main properties are independent of this choice:

**Proposition 2.** *If the planner's information choice attains a nonzero level of information $\lambda_{SB}^*$,*

- *$\lambda_{SB}^*$ is strictly increasing in the risky part $R-r$ of the projects' payoffs, and $\lim_{r \to R} \lambda_{SB}^* = 0$.*
- *$\lambda_{SB}^*$ as a function of $q$ follows an inverse-U shape. It reaches its maximum when the cut-off quality $q$ coincides with the average quality of the pool: $q = \bar{p}$.*

*Proof.* see appendix. □

These results are quite intuitive: screening is only useful in the presence of risk, and the acquisition of a very precise signal pays off most when it resolves a maximal amount of ambiguity. This corroborates the observation of Ruckes (2004) that there is low incentive to screen when the average quality of the project pool is either very good or very bad because in either case the signal obtained by screening will change investment decisions only for a very small percentage of the project portfolio.

I provide a graphical representation of the results in this section in Figure 1.1. Let me go over this figure in detail because it not only illustrates the planner's solution but also paves the way for the subsequent discussion of the competitive equilibrium. We have found that when the average borrower is credit-worthy and information is relatively imprecise (i.e., $\bar{p} - q \geq \frac{\lambda\varepsilon}{2}$), the volume of lending grows until every project is financed. Let me draw the set of all $(q, \lambda)$ pairs for which this applies in the diagram and label it as area (1). Since within this parameter region every entrepreneur receives a loan, screening is useless there. We can therefore conclude that the curve of optimal information choice $\lambda^*_{SB}(q)$ will either remain outside this region or drop to zero wherever it overlaps with it. The same holds for the opposite case where information is relatively imprecise and the average project makes an expected loss: no projects will be financed, the mass of credit is $0$ and there are no returns to information acquisition within this region either. I mark the corresponding area as (2).

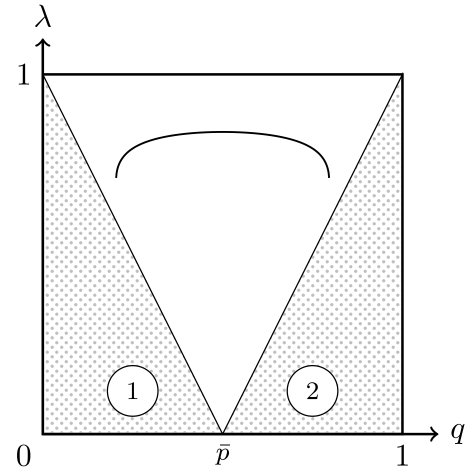


Figure 1.1: The planner's constrained optimal information choice.
The parameters of the plot are $\varepsilon = 1$, $\bar{p} = \frac{1}{2}$, $r = \frac{3}{4}$, $R = 2$ and $c(\lambda) = c_0 \frac{\lambda}{1-\lambda}$ with $c_0 = \frac{1}{250}$. Areas 1 and 2 which are shown in gray shading correspond to the cases handled in the first and third line of equation (1.9), respectively, where screening has no value because everybody, or nobody, is financed.

Wherever the constrained-optimal screening precision $\lambda^*_{SB}(q)$ is not zero, its graph must thus lie inside the white triangular area: these are the only combinations of economic state $q$ and information precision $\lambda$ for which screening can deliver positive value. Whilst the findings of the previous propositions are independent of the functional form of the cost function $c(\lambda)$, we need to make such choice in order to draw a graph. I use here the specification $c(\lambda) = c_0 \frac{\lambda}{1-\lambda}$ which yields closed-form solutions[7] for both the planner's choice and the competitive case.[8]

Looking at the graph of $\lambda^*(q)$ we can see that for the chosen parameters the interior solution indeed covers a wide range of $q$, but as $q$ moves further away from $\bar{p}$, the payoff to screening diminishes so much that it eventually becomes unprofitable to screen.

## 1.4 Competitive Equilibrium

### 1.4.1 Monopoly

Before discussing the full equilibrium under competition, it is instructive to look at the special case of monopoly. We can model this case within the existing setup by assuming that the remote monitoring cost $\gamma$ are sufficiently high that no bank can ever profitably offer outside credit to projects outside its own island. This is the case for all possible values of $q$ whenever $\gamma > (R - r)\varepsilon/2$.

It is easy to see that a profit-maximizing monopolistic lender will replicate the planner's choices perfectly: monopoly power enables the bank to extract the full surplus from all entrepreneurs by lending to them against a state-contingent repayment of $(R, r)$. In order to maximize profits, the monopolist will choose the surplus-maximizing level $\lambda^*_{SB}$ of testing precision and (assuming that the lender observes the same signal realizations as the planner) give loans to exactly the same set of projects that the planner would choose to develop. Despite the fact that all surplus is given to the monopolistic lender, the allocation is therefore constrained Pareto optimal. It is noteworthy that this optimality result depends crucially on the monopolist's ability to extract all project surplus. The assumed exogenous nature of project size plays here an important role because it enables the monopolistic lender to perfectly know each entrepreneur's individual ability of repayment. If entrepreneurs would instead choose the size of their projects endogenously while facing decreasing returns to scale, monopolistic pricing would result in inefficiently small projects that are no longer constrained efficient. Similarly, if a regulator were to limit the monopolist's surplus extraction by curbing the maximum

[7] See the appendix for details.

[8] Note that this specific cost function implies that marginal cost of information acquisition are strictly positive everywhere, even for $\lambda \to 0$: $c'(0) = c_0$. Thus, if $c_0$ is sufficiently large, holding any positive amount of information will be suboptimal and the surplus maximization problem (1.13) yields the corner solution $\lambda^*_{SB} = 0$ for all values of $q$. I avoid this situation by assuming that $c_0$ sufficiently low, i.e. $c_0 < \frac{\varepsilon}{8}$, that there exists an interior solution $\lambda^*_{SB} > 0$ for some neighborhood of $q = \bar{p}$.

repayment on each unit of funds to some value $D_{\max} < R$, the resulting portfolio of financed projects would remain inefficiently small since the corresponding cut-off $\hat{q} = \frac{\rho - r}{D_{\max} - r}$ up to which the monopolist provides funding would lie higher than $q$.

### 1.4.2 Portfolio Allocations under Competition

I now study the optimal portfolio choices of the two lenders when they compete against each other. Under competition, lenders must be careful in how they use the private information that they have obtained in the screening process because their decisions will become known to their uninformed competitor. If their loan approval conveys sufficiently good news about the quality of a borrower, the uninformed rival bank can enter the market as an outside lender, undercut the incumbent's loan offer and poach some or all of its loan-approved customers. However, the outside lender's cost disadvantage of $\gamma$ when lending to the other island puts a limit to this poaching activity: entry is not profitable if a poached borrower earns the outside lender in expectation less than this extra cost $\gamma$. This motivates the following definition:

**Definition** Denote with $\Pi[\mathcal{P}]$ the gross profit that accrues to a domestic lender in the absence of competition from holding a given domestic loan portfolio $\mathcal{P}$, and denote the volume of loans in the portfolio as $|\mathcal{P}|$. Then, a loan portfolio $\mathcal{P}$ shall be called *noncontestable* if its gross surplus per loan is less or equal to $\gamma$: $\Pi[\mathcal{P}] \leq \gamma|\mathcal{P}|$.

In equilibrium, every bank will choose a noncontestable portfolio: otherwise, the competing bank could profitably poach the entire set of customers by undercutting the domestic bank's offer for every domestically loan-approved entrepreneur by a small $\delta > 0$: $(O_i^j, o_i^j) = (D_i - \delta, d_i)$. The domestic lender's profits would drop to zero and be strictly lower than if it had chosen some noncontestable portfolio which yields lower but positive gross profits.[9] It must be pointed out that noncontestability is a necessary, but not a sufficient condition to guarantee that a loan portfolio does not attract entry: it only secures that an entrant financing the *entire* loan portfolio will suffer losses. If the variation in repayment terms $(D_i, d_i)$ of individual borrowers carries any information on borrower quality, the entrant might still be able to profitably target a subset of projects that are identified as more lucrative than others. Such scenario is no concern if the incumbent's portfolio is by chance not only noncontestable but also *symmetric*, i.e. every loan offer in the portfolio has the same repayment terms $(D_i, d_i) = (D, d)$. Then there exists no profitable possibility of entry, and the incumbent wins the full surplus of the

[9] We will see that such profitable noncontestable portfolio exists for every $\gamma > 0$.

portfolio. As the following lemma shows, choosing a symmetric portfolio does not need to be inferior to choosing a non-symmetric one:

**Lemma 3.** *Let $\mathcal{P}$ be a (not necessarily symmetric) noncontestable loan portfolio. Then there always exists a symmetric noncontestable portfolio $\mathcal{Q}$ for which the incumbent's payoff in the competition game is at least as high as when choosing $\mathcal{P}$.*

*Proof.* see appendix. □

A direct consequence of the lemma is that if we manage to find within the class of all symmetric noncontestable portfolios some portfolio $\mathcal{Q}^*$ that is profit-maximizing in the sense that no other symmetric noncontestable portfolio generates higher surplus, it must be an equilibrium of the game: to see this, observe that if any arbitrary portfolio $\mathcal{R}$ could generate a strictly higher payoff for the incumbent, such portfolio would be noncontestable for sure (otherwise, as we know, payoff would be zero). But then the lemma assures us that a symmetric noncontestable portfolio that generates at least the same payoff as $\mathcal{R}$ exists, thus generating an immediate contradiction to the maximality of $\mathcal{Q}^*$. Therefore, $\mathcal{R}$ can not have existed in first place, and we have

**Corollary 3.1.** *Any symmetric portfolio $\mathcal{Q}^*$ that maximizes gross surplus subject to the constraint of noncontestability,*

$$\begin{aligned} \max_{\mathcal{Q}=(\mathcal{S},(D,d))} \quad & \Pi[\mathcal{Q}] \\ \textit{s.t.} \quad & \Pi[\mathcal{Q}] \leq \gamma|\mathcal{Q}| \end{aligned} \tag{1.14}$$

*is an equilibrium allocation of credit under competition.*

The maximality of $\Pi[\mathcal{Q}^*]$ also makes it clear that even though the equilibrium portfolio $\mathcal{Q}^*$ is not necessarily unique, the equilibrium payoffs are always uniquely determined. The same is true for credit mass and average success probabilities: those two quantities are uniquely pinned down by the constraint whenever it is binding, whereas any possible indeterminacy that can arise if the constraint does not bind are eliminated by the previously made assumption that no project with zero expected surplus receives funding.

As a solution to the maximization problem for symmetric portfolios I find the following equilibrium allocations:

**Proposition 4.** *Let $\mathcal{F}$ denote the portfolio of the full unit measure of projects financed against state-contingent repayment promise of $(R, r)$. Then, under competition the following portfolio allocations constitute an equilibrium:*

- *If the monopolistic portfolio is noncontestable, it is an equilibrium. This is the case if information is sufficiently imprecise,* $\lambda \leq \frac{2(q-\bar{p})}{\varepsilon} + \frac{4\gamma}{\varepsilon(R-r)}$, *and the threshold* $q$ *is sufficiently high that* $\mathcal{F}$ *is noncontestable,* $q > \bar{p} - \frac{\gamma}{R-r}$. *The bank then acts as a monopolist and reproduces the constrained optimal allocation by lending to entrepreneurs whose expected success probability exceeds* $q = \frac{\rho-r}{R-r}$ *against state-contingent repayment of* $(R, r)$.

- *If information is sufficiently precise that a monopolistic portfolio choice would attract competition, i.e.* $\lambda > \frac{2(q-\bar{p})}{\varepsilon} + \frac{4\gamma}{\varepsilon(R-r)}$, *but the full portfolio* $\mathcal{F}$ *remains noncontestable,* $q > \bar{p} - \frac{\gamma}{R-r}$, *it is an equilibrium to finance all positive net present value (NPV) projects and some projects with negative NPV by offering credit at terms* $(R, r)$ *to projects whose expected success probability lies above* $\hat{q} = \bar{p} - \frac{\lambda\varepsilon}{2} + 2\left(q - \bar{p} + \frac{\gamma}{R-r}\right)$, *whereby* $\hat{q} < q$.

- *If both the monopolistic portfolio and the full portfolio* $\mathcal{F}$ *are contestable, i.e.* $q \leq \bar{p} - \frac{\gamma}{R-r}$, *it is an equilibrium to offer credit to every domestic entrepreneur against a state-contingent repayment of* $(D, r)$, *with* $D = \frac{\gamma+\rho-r(1-\bar{p})}{\bar{p}}$.

*Proof.* see appendix. □

This important result deserves several comments:

First, the surprising finding is that more competition has virtually zero impact on the cost of borrowing until every available project receives a loan. Before this point, an increase in competition (i.e. fall in $\gamma$) only lowers banks' lending standards and prompts them to take projects with negative net present value into their loan portfolios. The central force behind this result is that it is more profitable for the bank to deter entry by reinforcing the entrant's adverse selection problem than to do so by lowering its terms of repayment: imagine for a moment that the bank had instead reduced its repayment terms for its mass $m$ of borrowers to some $(D, r)$ with $D < R$ but had not lowered its lending standards. Its portfolio must then satisfy the noncontestability constraint with equality which bounds profits to $\gamma|m|$. This is no equilibrium because the bank can do better: it can extract the full surplus $(R, r)$ of those $m$ projects and use the proceeds to finance exactly as many additional negative NPV projects $n$ as to restore noncontestability, $\Pi[m \cup n] \leq \gamma(|m| + |n|)$. The larger size of the project portfolio relaxes the noncontestability constraint and thereby allows the bank to raise its profits by $\gamma|n|$. This mechanism works until the bank runs out of bad borrowers that could be used to further poison its pool: only then, it optimally reacts to competition by offering lower repayment terms. Note that my assumption of fixed refinancing cost contributes to the *complete* absence of any price response: if we had considered a richer model in which careless lenders incur higher refinancing cost, banks might choose an optimal mix of more aggressive pricing and lower lending standards in order to defend themselves against the entry of competitors.

Another important observation is that the game's competitive equilibrium is essentially a contestable market: the incumbent's informational advantage and the cost disadvantage of the entrant make it impossible for the entrant to actually ever poach any customers from the incumbent. It is the mere threat of entry that drives all the changes in allocations although in equilibrium entry never occurs.

Finally, the fact that the game has an equilibrium in pure strategies stands in sharp contrast to most credit screening games in the literature for which the nonexistence of pure strategy equilibria has been established (see e.g. Broecker, 1990; von Thadden, 2004). This is, however, not really unexpected given my substantially different setup: the sequential nature of bidding together with the observability of actions ensure that the entrant can correctly infer the incumbent's true valuation for the pool of loan-approved borrowers and hence does not face any winner's curse problem when submitting a bid on this pool.

I now proceed to discuss the equilibrium credit volume which can be calculated from the results of proposition 4:

**Corollary 4.1.** *In equilibrium, the volume of credit is the following:*

$$m^E_\lambda(q,\gamma) = \begin{cases} 1 & \textit{if } 0 \le \lambda \le \frac{2(\bar{p}-q)}{\varepsilon} \textit{ and } \bar{p} - \frac{\gamma}{R-r} < q < \bar{p} \\ 0 & \textit{if } 0 \le \lambda \le \frac{2(q-\bar{p})}{\varepsilon} \\ \frac{1}{2} - \frac{q-\bar{p}}{\lambda\varepsilon} & \textit{if } \frac{2|q-\bar{p}|}{\varepsilon} < \lambda \le \frac{2(q-\bar{p})}{\varepsilon} + \frac{4\gamma}{\varepsilon(R-r)} \\ 1 - \frac{2\left(q-\bar{p}+\frac{\gamma}{R-r}\right)}{\lambda\varepsilon} & \textit{if } \frac{2(q-\bar{p})}{\varepsilon} + \frac{4\gamma}{\varepsilon(R-r)} < \lambda < 1 \textit{ and } q > \bar{p} - \frac{\gamma}{R-r} \\ 1 & \textit{for all } \lambda \in [0,1) \textit{ if } q \le \bar{p} - \frac{\gamma}{R-r} \end{cases} \tag{1.15}$$

Whilst this expression looks rather complex at first glance, it is actually quite easy to understand. The first three lines represent the absence of competition: they are practically the same expressions that we had found for the planner's benchmark and the monopolistic case. The only difference is that the domain where the monopolistic case applies has shrunken, and there is a new domain of $(q,\lambda)$ values where threat of entry is relevant. The new area is represented by the ultimate two lines: line four describes the case in which the bank cross-subsidizes some negative NPV loans in order to deter entry. Line five finally stands for the case in which the bank cross-subsidizes *all* negative NPV projects and additionally reduces repayments in order to attain a noncontestable portfolio.

To illustrate in which situations banks are exposed to competition, it is useful to make a similar graphical representation of the $(q,\lambda)$ plane as in the planner's solution. Figure (1.2) plots the five different domains of the above function, numbered by the corresponding line number. Areas 1 and 2 are the same as in

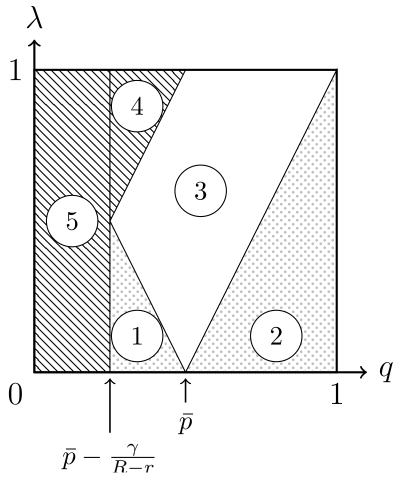


Figure 1.2: Competitive and monopolistic domains of lending.

Figure 1.1. We can see that threat of entry (as indicated by the dark shaded areas 4 and 5) affects banks that have very precise information and / or operate under very favorable conditions such that the cut-off $q$ is very low. As one can see from the expressions in lines $4$ and $5$ of equation (1.15), these areas grow as competition becomes stronger: in the limiting case $\gamma \to 0$, they cover almost the entire $(q, \lambda)$ plane with the exception of area 2.

How do the comparative statics of credit allocations under threat of entry differ from those in the monopolistic regime? The following proposition holds the answer:

**Proposition 5.** *Let $q$ be such that in contestable market equilibrium some but not all projects receive credit, i.e. $\bar{p} - \frac{\gamma}{R-r} < q < \bar{p} + \frac{\lambda\varepsilon}{2}$, and let competition be strong enough to potentially impact lenders within this regime, i.e. $\frac{\gamma}{R-r} < \frac{\lambda\varepsilon}{2}$. Then,*

- *the lower the chosen level $\lambda$ of information acquisition, the more sensitive is the volume of lending $m_E(q, \gamma)$ to changes in both monetary policy rate $\rho$ and project values $(R, r)$.*
- *holding constant the level of $\lambda$, the sensitivity $\frac{\partial m_E}{\partial q}$ of credit volume to changes in monetary policy rate and project values is twice as high under threat of competition than if the domestic bank can act as a monopolistic lender.*
- *the average default rate $\mathcal{D}_\lambda = 1 - \mathcal{A}_\lambda$ in a lender's portfolio increases as monetary policy rate falls and expected liquidation values rise. The sensitivity $\frac{\partial \mathcal{D}_\lambda}{\partial q}$ to such changes is independent of $\lambda$, but is twice as high under threat of competition than under monopolistic lending.*

*Proof.* see appendix. □

In other words, the presence of competition makes credit more volatile. An improvement in economic conditions will lead to both faster credit growth and faster deterioration of portfolio quality.

### 1.4.3 Information Choice under Competition

As the final step in solving the competition game backwards, I now find the equilibrium choices of screening precision that maximize total profits in a competitive environment. The gross profit function $\Pi_\lambda^E(q,\gamma)$ that is needed for this purpose follows immediately from the results in the last section: for all those $(q,\lambda)$ combinations for which there is no threat of entry, gross profits are exactly the same as under monopoly. In the remaining parameter ranges, the threat of competition limits equilibrium gross profits to $\gamma \cdot m_E$.

The equilibrium screening intensity $\lambda_E^*$ then solves the problem

$$\begin{aligned} \max_{\lambda} \quad & \Pi_\lambda^E - c(\lambda) \\ \text{s.t.} \quad & \lambda \geq 0 \end{aligned} \tag{1.16}$$

Even without a specific cost function, one can infer the impact of competition on screening directly from an inspection of marginal returns:

**Proposition 6.** *The equilibrium screening intensity* $\lambda_E^*$

- *is equal to the constrained efficient screening intensity* $\lambda_{SB}^*$ *whenever the monopolistic portfolio at this screening level does not attract competition, i.e. for all* $q > \bar{p} - \frac{\gamma}{R-r}$ *for which* $\lambda_{SB}^*(q) \leq \frac{2(q-\bar{p})}{\varepsilon} + \frac{4\gamma}{\varepsilon(R-r)}$.
- *falls below the constrained efficient level when competition is sufficiently strong to pose a threat of entry, i.e.* $q > \bar{p} - \frac{\gamma}{R-r}$ *and* $\lambda_{SB}^*(q) > \frac{2(q-\bar{p})}{\varepsilon} + \frac{4\gamma}{\varepsilon(R-r)}$.
- *is zero if competitive threats result in the financing of all projects regardless of their type, i.e. for all* $q \leq \bar{p} - \frac{\gamma}{R-r}$.

*Proof.* see appendix. □

This underinvestment in screening is my second key result. The intuition is clear: Whenever threat of entry prompts banks to "poison" their loan portfolios with some negative net present value loans, there is less incentive to acquire costly screening precision ex-ante. After all, banks disregard some of the information that is obtained by screening when they intentionally make bad loans. This is most apparent in the extreme case in which $q$ is so low that banks finance all projects regardless of their type: under such conditions, screening is entirely useless.

I illustrate the result graphically in Figure 1.3 with a plot of $\lambda_E^*$ over $q$ for two different levels of competition. The cost function is the same as in the planner's case, $c = c_0 \frac{\lambda}{1-\lambda}$. In the graph on the left-hand side, competition is weak and generates little impact. Borrower poaching is only profitable under the best possible economic conditions, i.e. when the cut-off $q$ is exceptionally low. Only in such exceptionally good economic states, the equilibrium choice of screening precision (shown as the solid line) will fall below the constrained-optimal level (dashed line). In all other times, the screening and lending choices of the bank and the social planner coincide. As the graph on the right-hand side of Figure 1.3 shows, the advent of more competition changes the situation dramatically: screening incentives are eroded over a wide range of economic states $q$, and for nearly all except the worst states of the economy, banks will choose a screening precision far below the constrained efficient level.

### 1.4.4 Excessive Lending

What are the consequences of inefficiently low screening for the total amount of lending? To answer this question, I evaluate the credit mass $m_\lambda^E(q, \gamma)$ at the equilibrium level $\lambda_E^*(q)$ of screening precision. Let me move the discussion immediately to the most interesting case:

Imagine that average project characteristics are good enough as to render the average project of the pool credit-worthy, i.e. $\bar{p} - \frac{\gamma}{R-r} < q < \bar{p}$. Then, as we have seen in corollary 1.1, the second-best amount of investment already exceeds the first-best benchmark of perfect information. Moreover, banks that hold *less* precise information than the planner will lend even *more* because their signal is too imprecise to actually push many bad projects below the zero expected surplus cut-off $q$. To make things worse, this is not even where lending under competition[10] actually ends! The equilibrium financing cut-off lies below $q$ because (as shown in proposition 4) banks add additional loans of negative net present value to their portfolio in order to protect themselves against poaching of their clients. The overall amount of credit in this situation is therefore excessive relative to both second-best and first-best levels and indeed deserves to be called an "inefficient lending boom":

$$m_{\lambda_E^*}^E(q, \gamma) > m_{\lambda_E^*}^{SB}(q) > m_{\lambda_{SB}^*}^{SB}(q) > m_{\lambda=1}^{FB}(q) \tag{1.17}$$

Figure 1.4 illustrates the equilibrium credit mass as a function of the economic conditions $q$ using the same equilibrium screening levels that were already shown

[10] i.e. for $\gamma < \frac{R-r}{2}\left(\bar{p} - q + \frac{\varepsilon \lambda_{SB}^*(q)}{2}\right)$

in Figure 1.3. We can see in the left-hand side picture that for a moderate amount of competition, credit becomes excessive only when economic conditions are particularly strong, i.e. $q$ is very low. The right-hand side graph displays the behavior of credit as a function of $q$ when there is more banking competition. We can observe that credit is excessive over a much wider range of economic conditions $q$. The lending boom is most pronounced for very low values of $q$, which are attained whenever the monetary policy rate $\rho$ is low and the liquidation value $r$ of failed projects is high.

But the most important factor that determines the extent of an inefficient boom is the amount of banking competition. In fact, it is possible to show that independently of the chosen screening cost function, the size of the lending boom is always greater when there is more competition:

**Proposition 7.** *Let the average project be credit-worthy, i.e.* $q \in (\bar{p} - \frac{\gamma}{R-r}, \bar{p})$, *and let competition be strong enough to pose threat of entry if monopolistic was applied, i.e.* $\gamma < \frac{R-r}{2}\left(\bar{p} - q + \frac{\varepsilon \lambda^*_{SB}(q)}{2}\right)$. *Then, the size* $m^E_{\lambda^*_E}$ *of the inefficient lending boom is increasing in the level of bank competition:*

$$\frac{\mathrm{d}}{\mathrm{d}\gamma} m^E_{\lambda^*_E}(q, \gamma) < 0. \tag{1.18}$$

*Proof.* see appendix. □

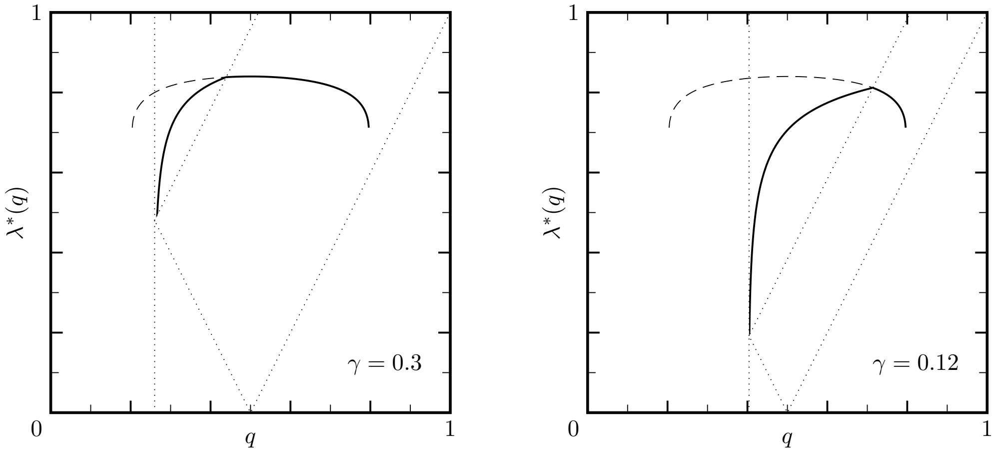


Figure 1.3: Screening precision $\lambda^*_E$ as a function of $q$ for two different levels of competition. The constrained efficient choice is indicated as a dashed line.

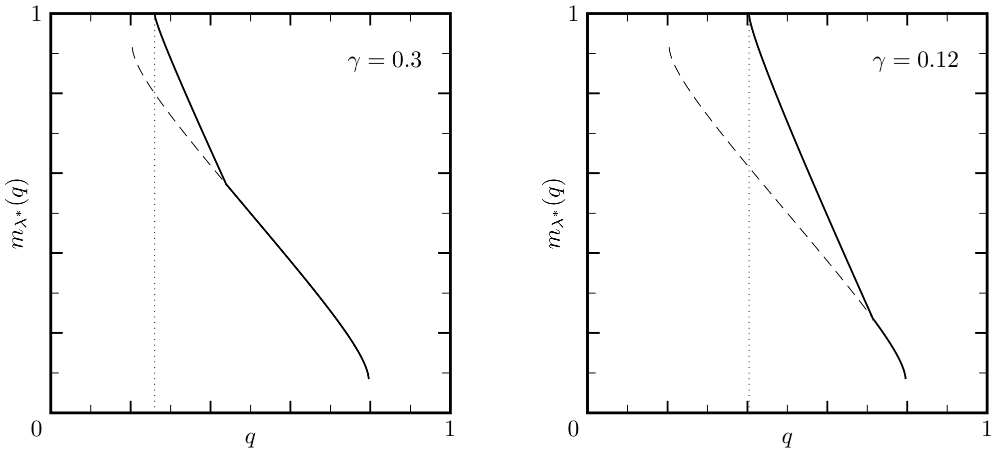


Figure 1.4: Credit mass $m^E_{\lambda^*}$ as a function of economic state $q$ for two different levels of competition. The dashed line marks the constrained efficient amount.

## 1.5 Discussion

With all central results in place, it is now time to discuss these findings in more detail and to analyze to which extent the key predictions of the model are robust to changes to my initial assumptions.

### 1.5.1 The nature of bank competition in the model

The first point that deserves to be given more detailed consideration is the nature of bank competition in my model, which is captured by the "extra monitoring cost" parameter $\gamma$. I think of $\gamma$ as a parameter that has a lot of economic meaning because it captures several distinct effects:

First, $\gamma$ can indeed reflect the additional cost of monitoring a borrower who is more difficult to observe for the bank. For example, this could be the case if the bank has no further service relationship with this customer (e.g. checking services) that could generate relevant information for monitoring the borrower's actions.[11] A reduction in $\gamma$ could thus reflect improvements in the ability of lenders to access information about remote customers, for example by exploiting advances in information technology (Hauswald and Marquez, 2003).

In an alternative interpretation of the model, $\gamma$ can be considered a switching cost: in fact, one can think of my model as the second stage of a two-period game where in the first period, banks build customer relationships by providing banking services. Banks might initially compete for extending a loan to the customer, or for handling the customer's main checking and payroll account. The exact nature of this first-period service competition remains irrelevant as long as it gives the chosen "inside" bank the exclusive opportunity to collect additional private information about the customer that can be used in the second period. In period two, the "inside" bank then competes with outside lenders for winning the customer as a borrower. In this interpretation of the model, it is more natural to think of $\gamma$ as a switching cost (due to the loss of the original bank relationship, or the cost of initiating a new business relationship with another lender) that any customer incurs who chooses to obtain his second period loan from a different lender. The existence of such switching cost are empirically well documented (e.g. Barone et al., 2011). An increase in competition in the model would be equivalent to a reduction in switching cost, which could for example occur due to changes in the regulatory environment.

Another important dimension of competition between banks is comparative advantage due to specialization. One could think of an extension of the model

[11] Mester et al. (2007) provide evidence that transaction account data help banks monitor borrowers.

in which one of the two banks (which I will refer to as the *money center bank* or $M$) has a comparative advantage relative to the other (which I will call the *local bank* or $L$) because of its specialization as a large transnational lender which gives it cheaper access to wholesale funding, i.e. $\rho_M < \rho_L$. On the other hand, the *local bank* might benefit from its geographic specialization in the form of lower marginal cost of screening, i.e. $c'_L(\lambda) < c'_M(\lambda)$. What would the corresponding equilibrium allocations under competition be? Without having to solve the model all over again, it is possible to see the answer directly by careful inspection of the symmetric model. Since I will discuss differentials in screening cost at a later point, let me for now analyze a world in which both banks have identical cost of screening but the money center bank has access to cheaper wholesale funding, $\rho_M = \rho_L - \Delta\rho$. This cost advantage enables the money center bank to compete more aggressively for customers of the local lender: it could earn a strictly positive profit by poaching any loan portfolio on which the local lender earns more than $\gamma - \Delta\rho$ per loan. The local bank will anticipate this and respond to the threat of poaching by taking on more excessive credit and reducing its screening precision $\lambda^*_L$. In fact, it is easy to see that the local bank will act exactly *as if* in a symmetric model the value of $\gamma$ had decreased to $\gamma - \Delta\rho$. The money center bank, on the other hand, will act in the opposite way: first of all, the lower refinancing cost translate into a lower required probability of success for a project to break even: $q_M = q_L - \frac{\Delta\rho}{R-r}$. Thus, the bank is able to extend its lending to more risky borrowers. To the extent that the resulting portfolio remains noncontestable, this expansion of credit will also be efficient. But more interestingly, the new loan portfolio is also more easily kept noncontestable since the local lender incurs higher cost on every loan-approved customer that it poaches from the money center bank: the local lender can only engage in profitable poaching of customers if the money center bank's profits per loan exceeds $\gamma + \Delta\rho$. The comparative advantage of lower refinancing cost therefore translates into lower competitive pressure: in equilibrium, the money center bank will behave exactly as if it was playing the symmetric game where $q = q_L - \frac{\Delta\rho}{R-r}$ and the "extra monitoring cost" parameter $\gamma$ was increased to $\gamma + \Delta\rho$.

Finally, we may ask what happens in the limit of perfect competition, which should be somehow related to the model's equilibrium in the limiting case $\gamma \to 0$. Unfortunately, in the limit $\gamma \to 0$ the fact that each bank earns zero profit also implies that there will be zero information collection. Lending would be all-or-nothing, depending only on how $q$ stands relative to $\bar{p}$. More meaningful conclusions about the perfect competition limit can only be attained if we keep $\lambda$ exogenously fixed to some positive value (e.g. $\lambda = 1$) and constrain the analysis to the second step (i.e., the portfolio choice) of the game. In this case, my model reproduces the well-known result of de Meza and Webb (1987): in the limit of perfect competition, all profits from lending to positive NPV borrowers are dissipated into additional negative NPV loans.

### 1.5.2 Improving Screening Precision: A Word of Caution

Given that the model predicts socially suboptimal levels of screening in equilibrium, we may want to think about what would happen if a regulator who is unaware of the actual origin of the problem would attempt to increase banks' screening levels by means of some intervention (for example, via a subsidy to screening, or by some minimum due diligence requirements).

By looking at the mechanics of the model, it is easy to see that such regulatory action could yield an unpleasant surprise: assume that in the competitive regime (where the monopolistic portfolio would attract outside competition and thus screening in equilibrium is inefficiently low) the screening level of banks was forced to some higher level $\lambda_{RG}$ by means of regulatory intervention such that $\lambda^*_{SB}(q) \geq \lambda_{RG} > \lambda^*_E(q,\gamma)$. The elevated screening precision will enable banks to make more profitable portfolio choices than if they had retained the less precise level of screening of the competitive equilibrium. But for this very same reason banks will also need to take more aggressive countermeasures against outside lenders who could poach their customers. They do so by adding further negative NPV clients to their borrower pool: an inspection of the fourth line in equation (1.15) reveals immediately that whenever banks operate in the competitive regime, more information precision will translate into larger credit mass. Thus, as long as regulatory intervention fails to reduce the informational asymmetries between lenders and instead only restores information acquisition at the *individual* bank, the outcome will be just a further expansion in credit which will be associated with further welfare losses.

### 1.5.3 Specialization and Industry Structure

The very same logic that higher information precision creates further excess credit also applies in the situation that we discussed in the previous section where specialization gave one lender a cost advantage in accessing funds *(money center bank)* whereas the other lender had a cost advantage in screening *(local bank)*. It is easy to see that the local bank will be at a severe disadvantage not only because of its higher cost in tapping external funds, but also exactly because of its ability to obtain private information at lower cost: since in equilibrium the local lender obtains more precise private information, its loan portfolio will be even more attractive for outsiders. In order to protect itself from competitors, the local lender therefore needs to take on more negative NPV projects than a lender with higher information acquisition cost would do. This severely limits its ability to benefit from its specialization.

If we compare the fate of the local lender to the fate of the money center bank it becomes clear that in a world in which the efficient use of information is constrained, the business model of the money center bank is the more promising one. Hence, as barriers to competition are reduced we would expect to witness a change in the structure of the banking industry: local lenders will either choose to transform themselves (via mergers) into big money center banks or exit the market.

### 1.5.4 Robustness and Generalizations

How robust are my results to changes in the assumptions of the model? Indeed, there are a few key assumptions that are indispensable:

The most important assumption in my model is the existence of a differential in private information between the *inside* lender and other *outside* banks: without this difference in information, there would be little scope for informational free-riding. The second fundamental assumption in my model is the order in which the loan negotiation process takes place: it is of fundamental importance that the uninformed lender moves last. In a game in which there is only one round of negotiations, this is not a very strong assumption because for $q > \bar{p}$ the uninformed lender would anyways never approve any loan without further information. But we can see the importance of this assumption if we were to add a second round of negotiation and would allow the informed bank to submit the last bid. Then, the superior knowledge of the informed bank would enable it to poach the best positive-NPV borrowers back from the outside lender, which would leave the outside lender with a loss-bringing portfolio of negative-NPV loans. In anticipation of this problem, the outside lender would thus refrain from bidding in first place, and the monopolistic outcome would be attained. This problem would not occur if the outside lender could submit its bid *after* the informed lender: since the second bid of the informed lender is final, it must contain information about the profitability of the selected customers that can be used by the outside lender. Thus, I would expect that the second period equilibrium in a two round game in which the uninformed bank bids after the informed bank will exhibit most of the features of the single round model that I have analyzed here.

Another important question regarding the robustness of my findings is to which extent the main result of the model – i.e. the optimality of cross-subsidized lending under competition – depends on the chosen screening technology. The good news is that in fact there is very little dependence on the chosen screening technology, as I shall show: the driving force behind most results is the fact that credit-worthiness testing creates dispersion in posterior expectations – the more precise the test, the more of it (see Ganuza and Penalva, 2010).

Let me define a completely general credit-worthiness test as follows: for a given level of test precision $\lambda$, the test comprises of a family of cumulative distribution functions $\{S_\lambda(s|p_i)\}_{p_i\in[0,1]}$ which indicate for every possible realization of a project's true type $p_i$ the probability of obtaining a signal less or equal to $s$. Together with the prior distribution $B(p)$, this induces a joint distribution from which one can calculate by Bayes' Law the posterior distribution of $p$ conditional on observing a signal realization $\sigma_{i,\lambda} = s_i$. From this posterior distribution, one can then finally calculate the expected probability conditional on the observed signal, $E[p_i|s_i]$. Just as before, I denote the cumulative density function of the random variable $E[p_i|\sigma_i]$ under test precision $\lambda$ as $\mathcal{E}_\lambda(p)$.

For any test specification that satisfies the following single-crossing assumption regarding $\mathcal{E}_\lambda(p)$, the model yields similar qualitative results:

**Assumption 1.** *Let $\lambda \in [0,1)$ and $\mu \in [0,1)$ with $\lambda > \mu$. Then, a creditworthiness test with higher precision $\lambda$ disperses posterior expectations further away from the prior $\bar{p}$ than a test with lower precision $\mu$ in the sense that*

$$\begin{aligned} \mathcal{E}_\lambda(p) &> \mathcal{E}_\mu(p) \text{ for all } p \in (0,\bar{p}), \text{ and} \\ \mathcal{E}_\lambda(p) &< \mathcal{E}_\mu(p) \text{ for all } p \in (\bar{p},1). \end{aligned} \tag{1.19}$$

The results of the second-best benchmark as in proposition 1 then follow trivially from $m_\lambda^{SB}(q) = 1 - \mathcal{E}_\lambda(q)$: for exogenous $\lambda$, the constrained optimal amount of lending is higher than under perfect information if $q < \bar{p}$, and is lower than under perfect information if $q > \bar{p}$.

With a minor twist in order to accommodate discrete posterior distributions I can also solve for the optimal portfolio under competition: remember that we defined the average success probability of the set of all projects that lie above some cut-off $z$ as $\mathcal{A}_\lambda(z) = \int_z^1 p \, \mathrm{d}\mathcal{E}_\lambda(p) / \int_z^1 \mathrm{d}\mathcal{E}_\lambda(p)$. This definition can cause problems when the posterior distribution is discrete since $\mathcal{A}_\lambda(z)$ would become a discontinuous function. By a change of variables, this problem is easily circumvented: I simply rewrite the average success probability as a function of credit mass[12] $m$ rather than cut-off probability $z$ via the transformation $m = \mathcal{E}_\lambda(p)$ and define it as

$$\hat{\mathcal{A}}_\lambda(m) = \frac{1}{m} \int_{1-m}^{1} \mathcal{E}_\lambda^{-1}(\mu) \, \mathrm{d}\mu \tag{1.20}$$

which makes it a continuous and decreasing function in $m$. Lemma 3 and corollary 3.1 continue to hold for this more general case, and one obtains an equilibrium portfolio under competition that bears much resemblance in its structure to what we have found before:

[12] It is understood that $m$ stands for a mass of $m$ projects selected in the order from best towards worse quality.

- If $\hat{\mathcal{A}}_\lambda(m_\lambda^{SB}) \leq q + \frac{\gamma}{R-r}$, the earnings per loan do not exceed $\gamma$ and the monopolistic portfolio is noncontestable. In this case, the allocation will reproduce the second-best, and a state-contingent repayment terms for loans are $(R, r)$.

- Under threat of entry $\hat{\mathcal{A}}_\lambda(m_\lambda^{SB}) > q + \frac{\gamma}{R-r}$, not every project will be financed as long as the full portfolio remains noncontestable, i.e. $q > \bar{p} - \frac{\gamma}{R-r}$. The financed mass of projects in this regime will be

$$m_\lambda^E(q) = \hat{\mathcal{A}}_\lambda^{-1}\left(q + \frac{\gamma}{R-r}\right) \tag{1.21}$$

  where $\hat{\mathcal{A}}_\lambda^{-1}(a) \equiv \sup\{m \in [0,1] : \hat{\mathcal{A}}_\lambda(m) = a\}$. Due to financing of negative-NPV projects, credit mass is excessive in this range, $m_\lambda^E(q) > m_\lambda^{SB}(q)$. Equilibrium repayment terms remain at $(R, r)$, and gross profits amount to $\gamma \cdot m_\lambda^E(q)$.

- Finally, for $q \leq \bar{p} - \frac{\gamma}{R-r}$, every project is financed at reduced repayment terms such that total gross profit equals $\gamma$.

In other words, the equilibrium with cross-subsidized loans and excessive lending is also attained under a generalized screening technology.

### 1.5.5 A Dynamic Model of Credit Cycles

I now extend the model to a simple dynamic setting. This extension to serves the purpose of capturing an additional aspect regarding information choice that was missing in the static model: typically, a banks' information choice exhibits some sort of rigidity. The baseline model does not capture any of this; $\lambda$ is simply a function of the underlying microeconomic state $q$ which is precisely known at the time when the bank makes its information choice. In reality, however, banks face large fixed cost for changing their internal organizational structure, which makes a frequent and perfect adjustment of the bank's screening precision to the ex-post realized macroeconomic state prohibitively costly. It is therefore more in line with anecdotal evidence to assume that banks make their organizational choice based on their ex-ante expectations regarding the medium-term macroeconomic outlook and are unable to quickly adjust it when the ex-post realized state turns out to require substantially different levels of screening precision.

I will show now that rigidities in screening can give rise to credit cycles. For this purpose I solve the following slightly modified version of the baseline model:

1. initially, the state of $q$ is unknown, but it is known that it will be drawn from a cumulative distribution function $\mathcal{U}(q)$.

2. the rest of the setup follows exactly the baseline model: two islands, two banks; without knowing the actual realization of $q$, every bank chooses its screening precision $\bar{\lambda}$ and pays a cost of $c(\bar{\lambda})$ that becomes sunk immediately

3. after $\lambda$ has been chosen, the actual level of the cut-off $q$ is realized

4. banks choose their domestic portfolios, observe each other's domestic lending choices and attempt to poach customers as in the baseline model

Since the last steps of the game have remained untouched, the model's findings on portfolio choice for a given $\bar{\lambda}$ are unaffected by the modification; the new setup merely changes equation (1.16) which determines the optimal level $\bar{\lambda}^*_E$ of screening under competition. The new optimization problem then reads

$$\begin{aligned} \max_{\bar{\lambda}} \quad & \int \Pi^E_{\bar{\lambda}}(q,\gamma) - c(\bar{\lambda})\,\mathrm{d}\mathcal{U}(q) \\ \text{s.t.} \quad & \bar{\lambda} \geq 0 \end{aligned} \tag{1.22}$$

The bank takes all possible realizations of $q$ weighted by their probability of occurrence into account when optimizing its expected net profits. This creates two different effects: on the one hand, bank competition reduces in a now already familiar way the benefits of bank screening for a possibly large range of low-$q$ states of nature. As long as the distribution $\mathcal{U}(q)$ assigns nonzero probability weight to those states of nature, the resulting ex-ante profit-maximizing level of screening $\bar{\lambda}^*_E$ will fall short of the amount that would be attained without competition. On the other hand, a new effect arises from the impossibility to condition information choice on the actual realization of $q$: The bank may ex-post encounter both situations in which its actual screening abilities *exceed* the level that would have been chosen if $q$ had been known ex-ante, and situations in which the contrary is true and screening precision ends up being *too low* for the ex-post realization of $q$. We shall see in a bit that this has important consequences.

But first, I want to reflect for a moment about the relation between this generalized model and the baseline model that I have discussed before. What are the circumstances for which the one is more appropriate than the other? Here is how I think of it: the baseline model does well in describing financial intermediation in "normal" economic periods in which neither asset prices nor refinancing cost exhibit dramatic and totally unforeseen changes. The prediction error on $q$ is then small enough to be negligible. The model with aggregate uncertainty is a generalization of that model and nests it as the special case in which the prior belief

$\mathcal{U}(q)$ assigns an ex-ante probability of 1 to the ex-post realized state $q = v$. The additional value of the generalized model lies in its ability to also handle "abnormal times", i.e. rare but not entirely unanticipated macroeconomic phases of high volatility and uncertainty.

Let me use the following somewhat stylized example to illustrate the effects of such macroeconomic shocks: Consider that we play a sequence $t = \{1, 2, \dots\}$ of one-shot games of the general model with aggregate uncertainty whereby at any given point $t$ in time, the cut-off parameter $q_{t+1}$ in the next period is determined as follows: with relatively high probability $\phi$ it remains at its previous level $q_t$ whereas with probability $1-\phi$ a major change of economic conditions occurs that reassigns $q_{t+1}$ to a random value anywhere within the possible range, $q_{t+1} \sim U(\bar{p} - \frac{\varepsilon}{2}, \bar{p} + \frac{\varepsilon}{2})$. This means that in the limit $\phi \to 1$ we return to playing a repeated one-shot game of the old model without aggregate uncertainty, whereas for $\phi \to 0$ the aggregate uncertainty regarding $q_{t+1}$ is maximal in every period. The profit maximization problem (1.22) becomes

$$\max_{\bar{\lambda}_t} \left[ \phi\, \Pi^E_{\bar{\lambda}_t}(q_t, \gamma) + (1-\phi) \int_{\bar{p}-\frac{\varepsilon}{2}}^{\bar{p}+\frac{\varepsilon}{2}} \Pi^E_{\bar{\lambda}_t}(q_{t+1}, \gamma)\, \varepsilon^{-1} \mathrm{d}q_{t+1} \right] - c(\bar{\lambda}_t) \quad (1.23)$$
$$\text{s.t.} \quad \bar{\lambda}_t \geq 0$$

As Figure 1.5 shows, if the probability $\phi$ of retaining the last period state of $q$ is not too low, the equilibrium information choice looks quite similar to the baseline model without aggregate uncertainty. Specifically, if there is a lending boom today (i.e. $q_t < \bar{p}$) and competition is sufficiently strong, tomorrow's ex-ante profit-maximizing screening precision $\bar{\lambda}^*_{E,t+1}$ is very low. This is because the probability of a regime change that could move $q_{t+1}$ to some new value is rather low, and even if it were to occur, only some realizations of $q_{t+1}$ will actually benefit from screening precision due to the presence of competition.

But ex-post, there can be a very different story: let us think what happens if a macroeconomic regime change actually occurs and pushes the cut-off $q_{t+1}$ to some relatively high value, i.e. $q_{t+1} > \bar{p} + \frac{\bar{\lambda}^*_{E,t+1}\,\varepsilon}{2} - \frac{2\gamma}{R-r} > \bar{p}$. Asset values are poor and refinancing cost are high enough as to give no incentive to any competitor to enter the market as an outside lender. In other words, the ex-ante threat of competition turns out to be irrelevant ex-post. One of the preemptive measures taken by the bank ex-ante to optimally respond to this threat remains in place, however, and impedes its lending activities: the bank's chosen screening precision $\bar{\lambda}^*_{E,t+1}$ is inefficiently low due to the ex-ante expectation of competition. Since the average project is not credit-worthy, we already know from corollary 4.1 that the imprecise information will prevent the bank from lending to as many borrowers as in the second-best benchmark.

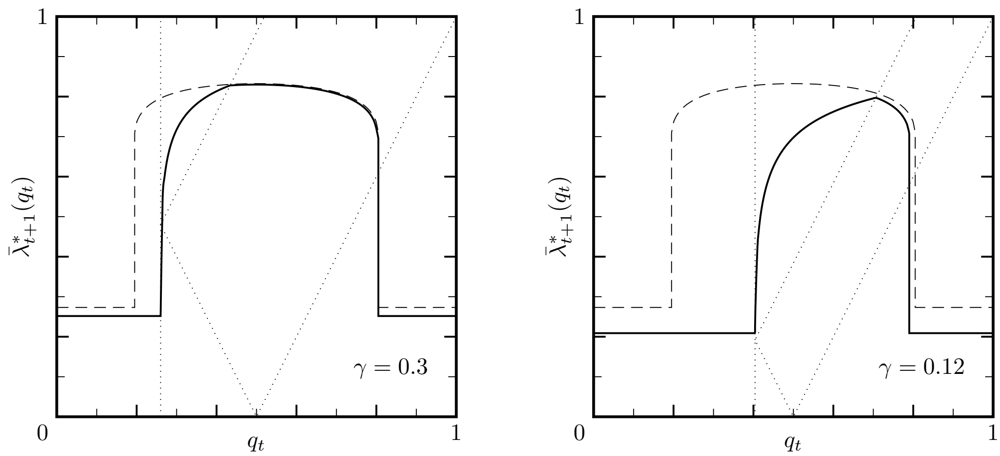


Figure 1.5: Information choice $\bar{\lambda}^*_{E,t+1}$ as a function of $q_t$ for two different levels of competition in the presence of aggregate uncertainty ($\phi = 0.8$).

As in Figure 1.3, competition is stronger in the right figure than in the left one. The solid line displays the equilibrium information choice $\bar{\lambda}^*_{E,t+1}$ as a function of the current cut-off $q_t$. The second-best benchmark is shown with dashed lines. Note that in anticipation of an aggregate shock (with probability $1 - \phi = 0.2$) some information is also collected in states $q_t$ in which information previously had no value.

In fact, due to the low precision of the screening signal, only borrowers of excellent quality will generate a signal that is sufficiently positive to secure their financing. The lack of precise information thus manifests empirically as a *flight to quality* combined with a sharp contraction in lending. This shows that one single cause – the lack of screening precision due to informational spillovers – has the potential to consistently explain two key pathologies that trouble most modern economies these days: the rise of inefficient credit booms, and the severe underprovision of credit once such economic boom goes bust.

### 1.5.6 Predictions and Empirical Findings

With the central results and robustness checks in place, it is now time to take a step back and check how the model's properties and predictions relate to the existing empirical literature.

For a given level of competition, the model predicts the existence of procyclical lending which is driven by variations in bank screening. Although bank screening effort is not directly observable, there is empirical evidence in support of a lending cycle: Berger and Udell (2004) find that credit growth following a favorable change in economic conditions is more pronounced when a substantial amount of time has passed since the last wave of defaulted loans. They explain this observation with what they call the "institutional memory hypothesis": in good times, so their argument, loan officers become increasingly inefficient in recognizing bad credit risks until they experience a bust that allows them to "re-learn" their abilities. Whilst they argue that this deterioration of screening standards is an unintended consequence of problems of the organization like high staff turnover, my model offers the perspective that a rational bank may actually find it *optimal* to let its screening skills deteriorate in good times when there is a threat of outside competition. If banks' adjustments of screening precision $\lambda$ are not instantaneous, for instance due to adjustment cost, this deterioration will be gradual and therefore observationally equivalent to the findings of Berger and Udell.

Equally in line with the screening cycle of my model are the findings of Ioannidou et al. (2009) and Jiménez et al. (2009): these two studies employ micro-level loan data for Bolivia and Spain, respectively, to investigate the impact of exogenous changes in monetary policy rates on the quality of new and existing loans. Consistent with the model's prediction that for lower values of $\rho$ (and thus $q$) banks screen their loan applications less precisely and extend their lending towards inefficient projects, they observe that loans that were made in times of low monetary policy rates are substantially more prone to default.

The question whether adverse selection problems that are induced by incumbent lenders matter for the entry decisions of banks has been addressed empirically by Gobbi and Lotti (2004). They show in an analysis of the Italian retail loan market that higher potential earning opportunities as measured by spreads between loan and deposit rates do not attract competition from outside banks whereas earning opportunities in financial services that do not require private information do attract entry. Likewise, Bofondi and Gobbi (2006) observe that entrants in local Italian loan markets suffer from substantially higher default rates than incumbent banks, even more so if they do not have a local branch office. These findings corroborate the validity of my model's central assumption that entry in local credit markets is deterred by the inferior quality of outsiders' information relative to the incumbent lender.

But does available empirical evidence also support the model's message with respect to the effects of competition on asset selection and lending procyclicality? Before I answer these questions, it is important to hold on and think for a moment how the degree of competition should be measured empirically. This is far from trivial, and the model itself with its contestable market structure provides a perfect example of where the problem lies. It is not necessary for a competitor to actually enter a market in order to have an effect on allocations: the mere threat of entry can already be enough. Studies that use bank concentration ratio or the number of active lenders in a market as a proxy for the degree of competition may thus measure something entirely different from what we were talking about so far: Claessens and Laeven (2004) warn that bank concentration can be a poor indicator of competition and suggest the use of a more sophisticated measure which quantifies the degree of *contestability* in a loan market. Given appropriate data, such measure can be constructed from the response of revenues earned to changes in input prices. Clearly, it would be desirable to use a measure of contestability instead of bank concentration in any empirical analysis related to the questions raised by my model, and results obtained with bank number or concentration may have to be treated with care.

With this caveat said, I now turn to the relevant findings regarding the effects of competition on lending. According to Micco and Panizza (2005), competition increases the procyclicality of lending: they analyze a panel of ninety-three countries for the 1990–2002 period and find a strong negative relationship between bank concentration and credit volatility. Although there are various alternative explanations for this increase in volatility, [13] the finding is clearly compatible with the view taken in my model.

The most imminent evidence that competition spurs inefficient lending and contributes to credit growth via insufficient screening comes from a recent paper of Dell'Ariccia et al. (2008) on loan rejections in the U.S. market for mortgages during the build-up of the U.S. subprime housing bubble: In an analysis of over 50 million loan-level observations they find that controlling for differences in economic conditions mortgages denial rates dropped faster in areas with a larger number of competitors. Even more important, they observe that the erosion of lending standards was most pronounced in loan markets in which banks were facing the entry of large outside lenders, which is a strong indication that excessive lending may have occurred in response to threat of entry.

### 1.5.7 Policy Implications

The unfortunate tendency of banks to lend too widely during booms has been often attributed to irrationality or overconfidence. By offering a rational perspective on these known facts, my model can serve as a framework for the discussion of macroprudential policies that are related to the procyclicality of lending.

#### Curtailing Lending Booms

Which advice does my model give to a policy maker who aims at curtailing an excessive lending boom and improving upon the distorted outcomes?

The lending booms in my model arise from the incumbent bank's defense against the potential entry of competition which deteriorates its portfolio and

[13] For example, it seems plausible that long-term borrower relationships are established more easily in concentrated markets which can also result in smoother funding (Petersen and Rajan, 1995).

screening choices. Thus, unless the mechanisms that give rise to informational spillovers can be somehow eliminated, naive policies aiming at improving borrower screening, for example by subsidizing screening effort or imposing minimum screening requirements, will inevitably generate even more excessive lending (as we have seen in section 1.5.2) and be welfare detrimental. Returning to a monopolistic setting by creating barriers to entry (i.e., raising $\gamma$) may not be a good idea either, albeit for reasons which remain outside the model: as already mentioned, monopoly usually fails to attain Pareto optimality because monopolistic pricing fails to extract all surplus and renders projects inefficiently small. The model's prediction that return to monopoly raises surplus should therefore be taken with a grain of salt.

Hence, the primary focus of the policy maker should lie on actions that increase the cut-off parameter $q$. The most traditional way to attain more restrictive lending practices goes by raising the refinancing cost $\rho$: central banks usually try to limit the size of a lending boom by taking a more restrictive monetary policy stance.[14] As we can see from the definition of the cut-off $q \equiv (\rho - r)/(R - r)$, this measure is especially effective when the risky part $R - r$ of project payoffs is small because even minor changes in $\rho$ then have major effect on $q$. Since the relevant trade-offs for the conduct of such policy are not modeled, the discussion of its welfare implications must remain incomplete: it is only possible to state that any policy that increases $q$ without generating additional transfers reduces the surplus generated by the banking sector. In practice, the heterogeneity of credit markets may limit the ability of monetary policy to address lending booms effectively: in the model, we have analyzed a situation in which there was only one single credit market for which the average quality of projects was exogenously given as $\bar{p}$, so were applicant heterogeneity $\varepsilon$ and project payoffs $R$, $r$. In reality, there are hundreds of distinct credit markets, e.g. car loans, mortgages, or loans to specific industries, which are all very different along each of these dimensions. Since monetary policy affects all of them, it may be impossible to fight a lending boom in some of these markets without creating a harmful contraction of lending in others.

An alternative policy measure that can be better targeted to certain credit markets is to selectively reduce the strength of the safety net that is offered to the lender by the option to repossess the project upon default of the borrower. This becomes especially relevant if project liquidation values $r$ are relatively high. Banks may then become "lazy" in their screening activities because even a project run by a bad borrower will not expose them to the risk of substantial loss (Manove et al., 2001). It is easy to see how this problem arises within the model: a high value of

[14]see e.g. the criticism of Taylor (2007) of the lax monetary policy of the FED during the rise of the U.S. housing bubble.

$r$ will not only result in a lower cut-off $q$ and give rise to lower lending standards, but the risky part $R - r$ of the project will also be reduced, which by proposition 2 diminishes screening effort even independently from the change in $q$.

What can be done in order to improve the situation? Imagine that there was some tax $\tau < r$ on repossessing the project of a borrower in default. Then, the lender could only extract a payoff of $\hat{r} \equiv r - \tau$ in case the project fails, and would choose a loan portfolio as if the cut-off $q$ had the value $\hat{q} \equiv (\rho - \hat{r})/(R - \hat{r})$ which is clearly higher than $q$. In other words, lending standards would tighten and there would be less credit. Moreover, as expected project payoffs fall, banks would find themselves less exposed to competition, which would lead them to make less cross-subsidized loss-bringing loans. Finally, all these changes together with the elevated risky component of payoffs would increase banks' incentives to screen. If these measures are temporarily put in place during an episode of a massive lending boom (i.e. with $q$ close to $\bar{p} + \frac{\gamma}{R-r}$), the reduction in inefficient lending plus the gains from tax proceeds can be so large that the new tax increases total welfare.

**Dynamic Provisioning**

Another particularly interesting case for which my model provides a new line of argumentation is the idea of *dynamic provisioning*. This macroprudential policy tool is being used in Spain since 2000 and has received revived interest in the light of the recent financial turmoil. The idea is relatively simple: if the regulator requires banks to accumulate a certain amount of *general* loan loss provisions during boom times, those provisions can be used during the downturn to compensate for the sharply increasing losses in the loan portfolio. But this policy tool has repeatedly come under heavy criticism because making loss provisions even before losses materialize is presumably not compatible with the principle of "true and fair value" accounting that is dictated by IFRS book keeping rules.

An implicit assumption which underlies this criticism is that every loan on the bank's balance sheet should be considered as an asset of positive net present value at the time at which it is originated. If this was indeed the case, then the accumulation of forward-looking provisions could be rightfully criticized as some type of cookie-jar accounting which disguises the true value of the firm: loss provisions would be made for completely healthy loans without even knowing whether losses will actually be incurred. Dynamic provisioning would attain protection against risk in downturns only at the cost of sacrificing the ability of balance sheet information to provide a clear, precise and fully adequate picture about the state of the firm at a given moment in time.

If one adopts the view of bank competition and asset selection that I have developed in this paper, there is no reason for such criticism. If banks indeed find it optimal to mix projects with negative net present value into their portfolio in order to keep competitors out of their markets, the very same principles of true and fair value reporting command that provisions against those foreseeable losses must be made at the relevant moment, i.e. upon loan origination. This is exactly what a well-implemented dynamic provisioning policy can achieve: it signals that losses from negative NPV loans will certainly bind some of the future profits from other loans in the loan portfolio. There is no trade-off between dynamic provisioning and the ability of books to reflect the truthful state of the firm at a given instant: such distortions would rather arise if traditional booking methods were followed because investors would be mislead into thinking that the bank holds a broad selection of equally profitable loans.

## 1.6 Conclusions

In this chapter, I have constructed a model of bank lending and bank competition that offers an explanation for the provision of poorly screened excessive bank credit when banks face fierce competition. This occurs whenever uninformed lenders can free-ride on the information generated by other lenders by observing the loan offers of rival banks and poaching their loan-approved customers. Incumbent lenders optimally react to such threat of entry by generating an adverse selection problem for the entrant by expanding their lending activity towards non-creditworthy customers and by simultaneously lowering their screening effort. This results in excessive credit and insufficient screening. The model also offers a theoretical explanation for the mounting empirical evidence that low monetary policy rates may prompt banks to lend to riskier and possibly inefficient projects.

Clearly, there is substantial room for further extension. For the purpose of clarity of exposition, I have made assumptions that cast the model in a very simple form and allow me to highlight the welfare-diminishing effects of competition in the presence of informational spillovers. But this is clearly not the full story, and when employing the model to answer policy questions, the welfare-improving effects of competition must be put back into place: project size should be endogenous rather than fixed. This could prevent the monopolist from extracting the full project surplus: monopolistic pricing would render projects inefficiently small, and the monopolistic case would no longer be constrained efficient. In such an (arguably more realistic) world, competition would offer important welfare gains but could also generate detrimental credit booms. The question to which extent these welfare gains from competition are offset by the welfare-diminishing impact on lending standards that I have analyzed in this chapter, and which level of competition should optimally prevail, is an interesting avenue for future research.

## 1.A Appendix

### 1.A.1 Derivation of Posterior Beliefs after Screening

From the definition of the test, it is immediate that the cumulative density function of $\sigma_i$ conditional on the project's actual quality $p_i$ being $p$ is given by a weighted combination of the cumulative of the random distribution and a step function,

$$P(\sigma_i \leq s | p_i = p) = (1-\lambda)\left(\frac{1}{2} + \frac{s-\bar{p}}{\varepsilon}\right) + \lambda\mathcal{H}(s-p) \tag{1.24}$$

where $\mathcal{H}(\cdot)$ is the Heaviside step function that evaluates to zero if the argument is negative, and to one otherwise. The corresponding probability density function[15] then reads

$$P(\sigma_i = s | p_i = p) = (1-\lambda)\frac{1}{\varepsilon} + \lambda\delta(s-p) \tag{1.25}$$

where $\delta(x)$ denotes the Dirac delta distribution that satisfies

$$\int_{-\infty}^{\infty} \delta(x) = 1 \text{ and } \delta(x) = \begin{cases} 0 & \text{for } x \neq 0 \\ +\infty & \text{for } x = 0 \end{cases}$$

By Bayes' Law,

$$P(p_i = p | \sigma_i = s) = \frac{P(\sigma_i = s|p_i = p)P(p_i = p)}{P(\sigma_i = s)} = (1-\lambda)\varepsilon^{-1} + \lambda\delta(s-p)$$

and the conditional expectation becomes

$$\begin{aligned} E[p_i|\sigma_i = s] &= \int p\left[(1-\lambda)\varepsilon^{-1} + \lambda\delta(s-p)\right]\mathrm{d}p \\ &= \int (1-\lambda)\varepsilon^{-1}p\,\mathrm{d}p + \lambda s \\ &= (1-\lambda)E[p] + \lambda s \end{aligned} \tag{1.26}$$

### 1.A.2 Proofs

**Proof of Proposition 1:**

The result for $m_\lambda^{SB}$ follows trivially from $m_\lambda^{SB} = 1 - \mathcal{E}_\lambda(q)$. To verify the indicated average success probability $\mathcal{A}_\lambda^{SB}$, observe that the statement is true for $m_\lambda^{SB} \in$

[15] strictly, this is not a function, but a distribution in the Schwartz sense

$\{0, 1\}$. Thus, we can assume $0 < m_\lambda^{SB} < 1$. Then,

$$\begin{aligned}
\mathcal{A}_\lambda^{SB} &= \frac{1}{m_\lambda^{SB}} \int_q^1 p \, \mathrm{d}\mathcal{E}_\lambda(p) = \frac{1}{m_\lambda^{SB}} \int_q^{\bar{p}+\frac{\lambda\varepsilon}{2}} (\lambda\varepsilon)^{-1} p \, \mathrm{d}p \\
&= \frac{1}{m_\lambda^{SB}\lambda\varepsilon} \left[\frac{p^2}{2}\right]_q^{\bar{p}+\frac{\lambda\varepsilon}{2}} = \frac{1}{m_\lambda^{SB}\lambda\varepsilon} \left[\frac{(\bar{p}+\frac{\lambda\varepsilon}{2})^2}{2} - \frac{q^2}{2}\right] \\
&= \frac{\lambda\varepsilon}{2m_\lambda^{SB}} \left(\frac{\bar{p}-q}{\lambda\varepsilon} + \frac{1}{2}\right) \left(\bar{p}+q+\frac{\lambda\varepsilon}{2}\right) = \frac{\bar{p}+q+\frac{\lambda\varepsilon}{2}}{2} \\
&= \bar{p} + \frac{\lambda\varepsilon}{2} \left(\frac{q-\bar{p}}{\lambda\varepsilon} + \frac{1}{2}\right) = \bar{p} + \frac{\lambda\varepsilon}{2}(1 - m_\lambda^{SB})
\end{aligned}$$

□

**Proof of Proposition 2:**

Marginal gross returns from information are

$$\frac{\partial}{\partial\lambda}(R-r)\frac{(2\bar{p}-2q+\lambda\varepsilon)^2}{8\lambda\varepsilon} = (R-r)\left[\frac{\varepsilon}{8} - \frac{(\bar{p}-q)^2}{2\lambda^2\varepsilon}\right]$$

which is always non-negative and attains a maximal value of $(R-r)\varepsilon/8$ when $q = \bar{p}$. Note that for $R \to r$ this goes to zero which proves that there will be no information acquisition in the limit of disappearing risk. □

**Proof of Lemma 3:**

I show the lemma in three steps:

1. In equilibrium, the payoff from choosing any noncontestable symmetric portfolio $\mathcal{Q}$ is equal to its surplus $\Pi[\mathcal{Q}]$.

   This is obvious since, because of symmetry, no offer of the entrant can attract a subgroup of borrowers that is better than the average of $Q$. From noncontestability it is clear that the entrant can not profitably enter the market even when the incumbent extracts all surplus.

2. If $\mathcal{P}$ is a (not necessarily symmetric) noncontestable portfolio, and $\mathcal{Q}$ is a symmetric noncontestable portfolio with $\Pi[\mathcal{Q}] \geq \Pi[\mathcal{P}]$, then by choosing the symmetric portfolio $\mathcal{Q}$ the incumbent obtains at least the same payoff in the contestable market game's equilibrium than by choosing $\mathcal{P}$.

By the previous step, choosing $\mathcal{Q}$ yields a game payoff of $\Pi[\mathcal{Q}]$ for the incumbent. Thus, if the converse were true, there would exist some portfolio $\mathcal{P}$ which, if chosen, would yield for the incumbent in equilibrium a payoff strictly higher than $\Pi[\mathcal{P}]$. Since splitting a portfolio's profits between the two players whilst keeping repayments constant is already a zero sum game, the entrant would necessarily have to bear negative profits; this holds even more when the necessarily lower repayment rates demanded by the entrant and his extra monitoring cost of $\gamma$ per loan are considered. The entrant would do better by staying out of the market which contradicts that this is an equilibrium strategy.

3. If $\mathcal{P}$ is a (not necessarily symmetric) noncontestable portfolio, then there always exists a *symmetric* noncontestable portfolio $\mathcal{Q}$ with weakly higher surplus $\Pi[\mathcal{Q}] \geq \Pi[\mathcal{P}]$.

   Define $D_{\max} = \sup_{i \in \mathcal{P}} \{D_i\}$ and $\pi[\mathcal{P}] \equiv \frac{\Pi[\mathcal{P}]}{|\mathcal{P}|}$, and denote the portfolio arising from financing every project $i \in \mathcal{P}$ at repayment rates $(D_{\max}, r)$ as $\hat{\mathcal{P}}$, and let $\hat{F}$ be the portfolio in which the full measure of *all* projects financed at repayment rates $(D_{\max}, r)$.
   If $\Pi[\hat{F}] > \gamma$, portfolio $\hat{F}$ can be made noncontestable by lowering repayment rates symmetrically to yield profits exactly equal to $\gamma$. The resulting profits are weakly higher than $\Pi[\mathcal{P}]$ since $\gamma \cdot 1 \geq \gamma|\mathcal{P}|$. The case in which $\pi[\hat{\mathcal{P}}] \leq \gamma$ is trivial: $\mathcal{Q} = \hat{\mathcal{P}}$. Thus it remains to show the statement for $\pi[\hat{\mathcal{P}}] > \gamma$ and $\pi[\hat{F}] \leq \gamma$.
   Denote the measure of loans *not* in $\mathcal{P}$ as $N$, index them arbitrarily with $j$ (where $0 \leq j \leq |N|$), and denote as $N_k$ the portfolio subset of loans in $N$ with index no higher than $k$, financed at $(D_{\max}, r)$. Then, the function $k \mapsto \pi[\hat{\mathcal{P}} \cup N_k]$ is continuous and satisfies by construction $\pi[\hat{\mathcal{P}} \cup N_0] > \gamma$, and $\pi[\hat{\mathcal{P}} \cup N_{|N|}] \leq \gamma$. By the intermediate value theorem there exists some $l$ with $0 < l \leq |N|$ such that $\pi[\hat{\mathcal{P}} \cup N_l] = \gamma$. Since $|\hat{\mathcal{P}} \cup N_l| \geq |\mathcal{P}|$, the symmetric noncontestable portfolio $\mathcal{Q} \equiv \hat{\mathcal{P}} \cup N_l$ must yield weakly higher surplus than $\mathcal{P}$.

Thus, let $\mathcal{P}$ be an arbitrary noncontestable portfolio. Then by the last step we can construct a noncontestable symmetric $\mathcal{Q}$ with the same or higher surplus, and due to the first two steps we know that the game payoff attained by playing $\mathcal{Q}$ is at least the same as the payoff from playing $\mathcal{P}$, q.e.d.

**Proof of Proposition 4:**

To prove the statement, I solve the maximization problem in corollary 3.1 for which I have already shown that every solution is an equilibrium portfolio choice in the game with competition. There are three cases:

1. If the solution to the unconstrained profit maximization problem is noncontestable, it is obvious that the monopolistic allocation maximizes eq (1.14). All projects with expected success probability above $q = \frac{\rho - r}{R - r}$ are financed at terms $(R, r)$. This occurs whenever the success probability of the monopolistic portfolio satisfies

$$\begin{aligned} \mathcal{A}_\lambda^{SB} &\leq q + \frac{\gamma}{R - r} \\ \frac{\bar{p} + q}{2} + \frac{\lambda \varepsilon}{4} &\leq q + \frac{\gamma}{R - r} \\ \Rightarrow \lambda &\leq \frac{2(q - \bar{p})}{\varepsilon} + \frac{4\gamma}{\varepsilon(R - r)} \end{aligned} \tag{1.27}$$

2. If the noncontestability constraint in equation (1.14) binds (i.e. $\lambda > \frac{2(q-\bar{p})}{\varepsilon} + \frac{4\gamma}{\varepsilon(R-r)}$), it must hold with equality:

$$\begin{aligned} \max_{\mathcal{Q}=(\mathcal{S},\,(D,d))} \quad & \Pi[\mathcal{Q}] \\ \text{s.t.} \quad & \Pi[\mathcal{Q}] = \gamma |\mathcal{Q}| \end{aligned} \tag{1.28}$$

Thus, any equilibrium has to maximize the size $|\mathcal{Q}|$ of the portfolio whilst maintaining an average profit per loan of exactly $\gamma$. Let me assume additionally that $q$ is such that the portfolio $\mathcal{F}$ that comprises of all projects financed at $(R, r)$ is noncontestable (i.e. $q \geq \bar{p} - \frac{\gamma}{R-r}$).

To solve this problem, I first characterize what the largest possible portfolio $\mathcal{Q}$ is that satisfies $\Pi[\mathcal{Q}] = \gamma|\mathcal{Q}|$ when repayment terms are exogenously fixed at some $(D, d)$, where $D \leq R$ and $d \leq r$. To yield exactly $\gamma$ per loan, the average probability of success in the portfolio must be equal to

$$\mathcal{A}_\lambda[\mathcal{Q}] = \frac{\int\limits_{\mathcal{S}\in\mathcal{Q}} p \,\mathrm{d}\mathcal{E}_\lambda(p)}{\int\limits_{\mathcal{S}\in\mathcal{Q}} \mathrm{d}\mathcal{E}_\lambda(p)} = \frac{\rho - d + \gamma}{D - d} \tag{1.29}$$

This condition is met by many symmetric portfolios that all have repayment terms $(D, d)$. But the size of the portfolio can only be maximal if the portfolio is chosen in a pecking order that starts with the best projects and proceeds in the order of expected success probabilities towards worse ones: to see this, imagine that the portfolio selection would not follow a cut-off structure, and that some nonzero mass $m$ of projects with average success probability $\pi_m$ was denied financing whereas a subset of projects of equal size but worse average success probability $\pi_n < \pi_m$ was financed. Then, by exchanging in the portfolio the worse projects projects for the better ones, one can increase the nominator of eqn (1.29) by $m \cdot (\pi_m - \pi_n)$ whereas the denominator stays the same. This "excess" in success probability can be used to enlarge the portfolio size by financing an additional mass of previously unfinanced projects that have worse quality than the portfolio average, whereby the mass is chosen such as to restore the average success probability to its original value. This shows that a portfolio which does not respect the pecking order of "best loans come first" will not be maximal in size.

Thus, there will be a unique cut-off $\hat{q}$ that will be determined by meeting the average success probability condition:

$$\begin{aligned} \frac{2\bar{p} + 2\hat{q} + \lambda\varepsilon}{4} &= \frac{\rho - d + \gamma}{D - d} \\ \Rightarrow \hat{q} &= \frac{2(\rho - d + \gamma)}{D - d} - \bar{p} - \frac{\lambda\varepsilon}{2} \end{aligned} \tag{1.30}$$

This expression makes it obvious that the cut-off is decreasing in $D$, so portfolio size is increasing in $D$; thus when $D$ is also chosen, we must have $D = R$. Finally, the derivative of (1.30) with respect to $d$ yields $\frac{2(\rho - D + \gamma)}{(D-d)^2}$ which for $D = R$ (as it is the case here) is negative if $\gamma + \rho < R \Rightarrow \frac{\gamma}{R-r} + \frac{\rho - r}{R-r} < 1$, i.e. $q + \frac{\gamma}{R-r} < 1$. Note that the left side of the inequality is the average success probability from eq. (1.29). Thus the inequality holds, and the size of the portfolio is increasing in $d$, which implies $d = r$. We then have a cut-off $\hat{q}$ of

$$\begin{aligned} \hat{q} &= \frac{2(\rho - r + \gamma)}{R - r} - \bar{p} - \frac{\lambda\varepsilon}{2} \\ &= 2\left(q + \frac{\gamma}{R - r}\right) - \bar{p} - \frac{\lambda\varepsilon}{2} \end{aligned} \tag{1.31}$$

3. If $\mathcal{F}$ is contestable, i.e. $q \leq \bar{p} - \frac{\gamma}{R-r}$, any noncontestable portfolio that maximizes size must have a size of $1$ and yield gross profits of $\gamma$. For a symmetric portfolio with $D = \frac{\gamma + \rho - r(1-\bar{p})}{\bar{p}}$ and $d = r$, gross profits are exactly $\gamma$:

$$\begin{aligned} \Pi &= \bar{p}D + (1 - \bar{p})r - \rho \\ &= \gamma + \rho - r(1 - \bar{p}) + (1 - \bar{p})r - \rho \\ &= \gamma \end{aligned} \tag{1.32}$$

**Proof of Proposition 5:**

The first two statements follow directly from taking the derivative of eq. (1.15) with respect to $q$ for the monopolistic and the competitive regime,

$$\frac{\partial m(q,\gamma)}{\partial q} = \begin{cases} \frac{-1}{\lambda\varepsilon} & \text{if } \bar{p} + \frac{\lambda\varepsilon}{2} - \frac{2\gamma}{R-r} < q < \bar{p} + \frac{\lambda\varepsilon}{2} \\ \frac{-2}{\lambda\varepsilon} & \text{if } \bar{p} - \frac{\gamma}{R-r} < q \leq \bar{p} + \frac{\lambda\varepsilon}{2} - \frac{2\gamma}{R-r} \end{cases} \tag{1.33}$$

Combining this result with $\frac{\partial \mathcal{D}_\lambda}{\partial m} = -\frac{\partial \mathcal{A}_\lambda}{\partial m} = \frac{\lambda\varepsilon}{2}$ yields

$$\frac{\partial \mathcal{D}_\lambda}{\partial q} = \frac{\partial \mathcal{D}_\lambda}{\partial m} \cdot \frac{\partial m}{\partial q} = \begin{cases} -\frac{1}{2} & \text{if } \bar{p} + \frac{\lambda\varepsilon}{2} - \frac{2\gamma}{R-r} < q < \bar{p} + \frac{\lambda\varepsilon}{2} \\ -1 & \text{if } \bar{p} - \frac{\gamma}{R-r} < q \leq \bar{p} + \frac{\lambda\varepsilon}{2} - \frac{2\gamma}{R-r} \end{cases} \tag{1.34}$$

which proves the third statement.

**Proof of Proposition 6:**

To prove the statement, let's look at the gross profit function in the presence of competition: for $q > \bar{p} - \frac{\gamma}{R-r}$, it reads

$$\Pi_\lambda(q,\gamma) = \begin{cases} \max\{0, (R-r)(\bar{p}-q)\} & \text{if } 0 \leq \lambda \leq \frac{2|q-\bar{p}|}{\varepsilon} \\ (R-r)\frac{(2\bar{p}-2q+\lambda\varepsilon)^2}{8\lambda\varepsilon} & \text{if } \frac{2|q-\bar{p}|}{\varepsilon} < \lambda \leq \frac{2(q-\bar{p})}{\varepsilon} + \frac{4\gamma}{\varepsilon(R-r)} \\ \gamma\left[1 - \frac{2\left(q-\bar{p}+\frac{\gamma}{R-r}\right)}{\lambda\varepsilon}\right] & \text{if } \frac{2(q-\bar{p})}{\varepsilon} + \frac{4\gamma}{\varepsilon(R-r)} < \lambda < 1 \end{cases} \tag{1.35}$$

whilst for $q \leq \bar{p} - \frac{\gamma}{R-r}$ gross profits are independent of $\lambda$,

$$\Pi_\lambda(q,\gamma) = \gamma \text{ for all } \lambda \in [0,1) \tag{1.36}$$

Despite its piecewise structure, the function is continuous everywhere. Specifically, note that the transition between monopolistic and competitive regimes occurs at $\lambda = \frac{2(q-\bar{p})}{\varepsilon} + \frac{4\gamma}{\varepsilon(R-r)}$, which substituted into the competitive profit function yields the same profits as when approaching the threshold from the monopolistic side: $\Pi = \frac{\gamma^2}{(q-\bar{p})(R-r)+2\gamma}$

Taking derivative with respect to $\lambda$, one obtains the marginal returns to information for $q > \bar{p} - \frac{\gamma}{R-r}$ as

$$\frac{\partial \Pi_\lambda}{\partial \lambda}(q,\gamma) = \begin{cases} 0 & \text{if } 0 \leq \lambda \leq \frac{2|q-\bar{p}|}{\varepsilon} \\ (R-r)\left(\frac{\varepsilon}{8} - \frac{(\bar{p}-q)^2}{2\lambda^2\varepsilon}\right) & \text{if } \frac{2|q-\bar{p}|}{\varepsilon} < \lambda \leq \frac{2(q-\bar{p})}{\varepsilon} + \frac{4\gamma}{\varepsilon(R-r)} \\ \frac{2\gamma\left(q-\bar{p}+\frac{\gamma}{R-r}\right)}{\lambda^2\varepsilon} & \text{if } \frac{2(q-\bar{p})}{\varepsilon} + \frac{4\gamma}{\varepsilon(R-r)} < \lambda < 1 \end{cases} \tag{1.37}$$

whereas there are zero marginal returns to information for $q \leq \bar{p} - \frac{\gamma}{R-r}$. Hence we already know that $\lambda_E^* = 0$ for this region, which proves the third statement of the proposition. The first statement of the proposition is equally obvious because profits and marginal returns to information for this case are the same as in the planner's benchmark. It thus only remains to show that marginal returns under threat of competition are lower than those attained under monopoly. Denote the difference in marginal returns to information as

$$\begin{aligned}\Delta(\lambda) &\equiv (R-r)\left(\frac{\varepsilon}{8}-\frac{(\bar{p}-q)^2}{2\lambda^2\varepsilon}\right)-\frac{2\gamma\left(q-\bar{p}+\frac{\gamma}{R-r}\right)}{\lambda^2\varepsilon} \\ &= \frac{(R-r)\varepsilon}{8}-\frac{((q-\bar{p})(R-r)+2\gamma)^2}{2(R-r)\lambda^2\varepsilon} \qquad (1.38)\end{aligned}$$

This term is non-negative if

$$\begin{aligned}&\frac{1}{8}(R-r)\varepsilon\cdot 2(R-r)\lambda^2\varepsilon-((q-\bar{p})(R-r)+2\gamma)^2\geq 0 \\ &\Leftrightarrow \left((R-r)\frac{\lambda\varepsilon}{2}\right)^2-((q-\bar{p})(R-r)+2\gamma)^2\geq 0 \qquad (1.39)\end{aligned}$$

Substituting into this expression the threshold between monopolistic and competitive regimes, $\lambda=\frac{2(q-\bar{p})}{\varepsilon}+\frac{4\gamma}{\varepsilon(R-r)}$, we see that the left-hand side of this inequality becomes zero. For any larger value of $\lambda$ it will obviously remain positive. Thus, marginal returns to information are continuous at the threshold, and for higher $\lambda$ marginal returns to information under competition are strictly less than under monopoly. q.e.d.

**Proof of Proposition 7:**

Under the conditions stated in the proposition, a reduction of competition has two effects on credit: As one can see by taking the total derivative wrt. $\gamma$,

$$\frac{\mathrm{d}}{\mathrm{d}\gamma}\left(m^E_{\lambda^*_E(\gamma)}(q,\gamma)\right) = \frac{\partial m^E_\lambda}{\partial\gamma}+\left.\frac{\partial m^E_\lambda}{\partial\lambda}\right|_{\lambda=\lambda^*_E}\cdot\frac{\partial\lambda^*_E}{\partial\gamma} \qquad (1.40)$$

there is a *direct* effect and an *indirect* one via $\lambda$. The two effects have opposite sign: for constant information level, less competition results in less cross-subsidization and thus less credit. However, less competition also raises information precision which again prompts more cross-subsidization and more credit. I show here that the first effect prevails:

First, note that

$$\frac{\partial m^E_\lambda}{\partial\lambda}=\frac{\partial}{\partial\lambda}\left(1-\frac{2(q-\bar{p}+\frac{\gamma}{R-r})}{\lambda\varepsilon}\right)=\frac{1-m^E_\lambda}{\lambda} \qquad (1.41)$$

and, as a consequence of $\Pi_\lambda^E = \gamma m_\lambda^E$,

$$\frac{\partial \Pi_\lambda^E}{\partial \lambda} = \gamma \frac{1 - m_\lambda^E}{\lambda} \tag{1.42}$$

From here I take derivatives and obtain

$$\begin{aligned} \frac{\partial^2 \Pi_\lambda^E}{\partial \lambda^2} &= -\frac{2\gamma}{\lambda^2}(1 - m_\lambda^E) \text{ and} & (1.43)\\ \frac{\partial^2 \Pi_\lambda^E}{\partial \lambda \partial \gamma} &= \frac{1 - m_\lambda^E}{\lambda} + \frac{2\gamma}{\lambda^2 \varepsilon (R - r)} & (1.44) \end{aligned}$$

I then apply the implicit function theorem on the first-order condition that fixes $\lambda_E^*$ in order to put bounds on the sensitivity of $\lambda_E^*$ with respect to $\gamma$:

$$\frac{\mathrm{d}}{\mathrm{d}\,\gamma} \lambda_E^*(q, \gamma) = -\frac{\frac{\partial^2 \Pi_\lambda^E}{\partial \lambda \partial \gamma}}{\frac{\partial^2 \Pi_\lambda^E}{\partial \lambda^2} - c''(\lambda)} \leq \frac{\frac{\partial^2 \Pi_\lambda^E}{\partial \lambda \partial \gamma}}{-\frac{\partial^2 \Pi_\lambda^E}{\partial \lambda^2}} = \frac{\frac{1 - m_\lambda^E}{\lambda} + \frac{2\gamma}{\lambda^2 \varepsilon (R - r)}}{\frac{2\gamma}{\lambda^2}(1 - m_\lambda^E)} \tag{1.45}$$

The inequality results from the fact that I assumed a convex cost function, so $c'' \geq 0$. Simplifying further, I obtain

$$\frac{\mathrm{d}}{\mathrm{d}\,\gamma} \lambda_E^*(q, \gamma) \leq \frac{\lambda}{2\gamma} + \frac{1}{\varepsilon (R - r)(1 - m_\lambda^E)} \tag{1.46}$$

and

$$\frac{\partial m_\lambda^E}{\partial \lambda} \frac{\mathrm{d}\lambda}{\mathrm{d}\gamma} \leq \frac{1 - m_\lambda^E}{2\gamma} + \frac{1}{\varepsilon \lambda (R - r)} \tag{1.47}$$

The sum of direct and indirect effect then yields

$$\frac{\mathrm{d}}{\mathrm{d}\gamma} \left( m_{\lambda_E^*(\gamma)}^E (q, \gamma) \right) \;\; \leq \;\; \frac{1 - m_\lambda^E}{2\gamma} - \frac{1}{\varepsilon \lambda (R - r)} \tag{1.48}$$

Substituting the definition of $m_\lambda^E$, I have

$$\frac{\mathrm{d}}{\mathrm{d}\gamma} \left( m_{\lambda_E^*(\gamma)}^E (q, \gamma) \right) \;\; \leq \;\; \frac{q - \bar{p}}{\gamma \lambda \varepsilon} + \frac{1}{\lambda \varepsilon (R - r)} - \frac{1}{\varepsilon \lambda (R - r)} = \frac{q - \bar{p}}{\gamma \lambda \varepsilon} \tag{1.49}$$

Since I assumed that the average project is credit-worthy, $q < \bar{p}$ holds and the equation is negative. q.e.d.

### 1.A.3 Closed-Form Solutions

For the specific cost function $c(\lambda) = c_0 \frac{\lambda}{1-\lambda}$ the model has closed-form solutions for $\lambda^*(q)$:
In the monopolistic case, the first-order condition has the form

$$\frac{c_0}{(\lambda-1)^2} = b - \frac{a}{\lambda^2}$$

with $a \equiv \frac{(R-r)(\bar{p}-q)^2}{2\varepsilon}$ and $b \equiv \frac{(R-r)\varepsilon}{8}$. Solving for $\lambda$ leads to a quartic equation for which the solution is known in standard tables. Two roots can be discarded because they describe local minima, so only two solutions $\lambda_1^*$ and $\lambda_2^*$ remain. The correct solution is identified as the only one that lies within the $[0, 1]$ interval (depending on $q$, either the one or the other branch applies). Define

$$\begin{aligned}
\gamma &\equiv a - b - c_0 \\
\phi &\equiv 108a^2b - 108ab^2 - 108ab\gamma - 2\gamma^3 \\
\beta &\equiv \frac{\sqrt[3]{2}\gamma^2}{3a\left(\phi + \sqrt{\phi^2 - 4\gamma^6}\right)^{1/3}} + \frac{\left(\phi + \sqrt{\phi^2 - 4\gamma^6}\right)^{1/3}}{3\sqrt[3]{2}a}
\end{aligned}$$

The two solutions are then given by

$$\lambda_1^* = \frac{1}{2}\sqrt{2 + \beta - \frac{4\gamma}{3a} - \frac{2(a-2b-\gamma)}{a\sqrt{1-\beta-\frac{2\gamma}{3a}}}} - \frac{1}{2}\sqrt{1-\beta-\frac{2\gamma}{3a}} + \frac{1}{2} \quad (1.50)$$

$$\lambda_2^* = \frac{1}{2}\sqrt{2 + \beta - \frac{4\gamma}{3a} + \frac{2(a-2b-\gamma)}{a\sqrt{1-\beta-\frac{2\gamma}{3a}}}} + \frac{1}{2}\sqrt{1-\beta-\frac{2\gamma}{3a}} + \frac{1}{2} \quad (1.51)$$

In the competitive case, the first-order condition is quadratic in $\lambda$ and can be solved to yield as the unique non-negative solution

$$\lambda_E^*(q,\gamma) = \frac{1}{1+\sqrt{\frac{c_0}{\delta}}} \text{ where } \delta = 2\gamma\varepsilon^{-1}(q - \bar{p} + \tfrac{\gamma}{R-r}) \quad (1.52)$$

# Chapter 2

# GEEKS BEARING GIFTS: COMPETITION AND INFORMATION SHARING IN KNOWLEDGE-BASED SERVICE INDUSTRY

## 2.1 Introduction

The endorsement of open source software development principles by big software firms such as Sun Microsystems and IBM has changed the face of software industry dramatically during the past ten years. As late as in the late nineties, it was generally understood that the source code of virtually any commercial software project was an intellectual asset which had to be guarded carefully against unauthorized dissipation because it contained core elements of business-relevant knowledge of the firm. This thought was challenged in Eric Raymond's influential book "The Cathedral and the Bazaar" (Raymond, 1999): impressed by the tremendous success of the volunteer-driven open development models around the Linux operating system, he advocated the general adoption of a more open development model in software industry.

Today, just a decade later, profit-maximizing firms are actively developing and distributing intellectual assets of billions of dollars of worth under open source licenses. "Open source" implies that the entire product's source code is freely available for download under a license that permits unlimited use, copying, modification and redistribution of the code and any derived work, thus making the product a public good. Examples of commercially developed open source products include the Java language, the OpenOffice software suite, and the JBoss application server. The commercial momentum behind open source software nowadays even extends to projects that used to be driven by volunteer contributions: According to a study by Kroah-Hartman et al. (2008), over 70 percent of the more recent contributions to the Linux kernel were made by programmers who were paid for their work.

To the economist, the most immediate question is which incentives lie behind this free provision of a public good. Incentives for participation in open source software development have so far been discussed primarily in the context of individual non-professional contributors, for instance by Lerner and Tirole (2002). The literature currently only offers a handful of explanations why profit-maximizing firms might want to participate in open source software development at as high intensity as empirically documented. Existing arguments include the sale of complementarities (Lerner and Tirole, 2005), positive network externality to a substitute closed source product (Mustonen, 2005), the higher profitability of an open source platform in two-sided markets where applications are proprietary (Economides and Katsamakas, 2006), and the option to generate revenues through dual-licensing (Comino and Manenti, 2007). However, neither of these approaches offers a compelling explanation for the numerous cases in which the commercial open source producer does not control any complementary market to generate revenue, and in which the deployed software is not a platform or a complement to a commercial offer of the firm.

In this paper, we embark on a different path and take Eric Raymond's provocative assessment that *"...software is largely a service industry operating under the persistent but unfounded delusion that it is a manufacturing industry"* (Raymond, 1999, p. 120) as an invitation to characterize the general properties of competition in a knowledge-based service industry: We aim to see whether commercial open source software development can be understood more naturally from the perspective of a competitive service sector.

Indeed, knowledge-based service firms operate very differently from manufacturing industry: rather than delivering a standardized commodity product at a fixed price, they provide unique and customer-specific solutions in a process of close interaction with their customers. The customers are usually firms themselves: a typical example of what we have in mind is a scenario in which an IT systems integrator like IBM provides IT solutions to a large enterprise like GM. The generated service value is often of highly intangible nature, which makes it hard to measure and virtually impossible to contract upon. The crucial role of customer participation is another distinct dimension of a service: because the provision of any service involves a transformation of the customer himself, the degree of customer participation directly affects the value of the service outcome. By learning from customer participation, the service firm also increases its stock of customer-specific knowledge which enables it to deliver higher service value to the same client in the future.

In this paper, we show that the aforementioned distinct dimensions of knowledge-based service industry can yield forms of competition that are vastly distinct from those known from manufacturing industry. In a typical manufacturing scenario, it is highly desirable for a producer to outpace all competitors in terms of production-relevant knowledge because the surplus that goes to the producer keeps increasing with the knowledge stock. As we will show, the opposite holds in a knowledge-based service industry. The continuous need to elicit customer participation in the value generation process, and the impossibility to agree on this participation by contractual means (because of measurement problems) makes it favorable for the service firm with the highest knowledge stock to share a portion of its production-relevant knowledge with competitors. By deliberately strengthening competitors, the firm can mitigate customer concerns about being held up in the future and can generate not only higher social welfare but also higher individual profits due to stronger customer participation. To the extent to which software source code embodies parts of the internal knowledge of an IT firm, we are therefore able to explain the commercial provision of open source software by leading IT service firms as a form of optimal knowledge sharing in knowledge-based service industry.

The perspective of our paper is related to, but different in several aspects, from the view taken in Shepard (1987) and Farrell and Gallini (1988) who show that a monopolist may have incentive to create competition in order to be able to commit to a low price or a high quality. First, their models are essentially static in the sense that they describe transactions in commodity goods whose value to the buyer is constant over time. This stands in sharp contrast to the service-centered perspective of our model where customers need to decide strategically which service provider they choose and how much effort they exert in the joint co-creation of value because their future payoffs are directly affected by these choices. Second, unlike Shepard (1987), we assume an environment in which knowledge is tacit and not transferable to rival firms through licensing, for instance due to limited enforceability of intellectual property rights. As a consequence, in our model, the incumbent cannot use a fixed fee to extract the benefit that accrues to a rival firm from an increase in the latter's stock of knowledge. This situation is quite coherent with what happens when an incumbent makes a software available as open source.

Before we proceed to present our service-centered model, we briefly mention contributions from the literature that we find relevant and related to our work, but that did not fit into the discussion above. A very early work by Kuan (2001) has emphasized the aspect of consumer integration into production through open source projects; von Hippel and von Krogh (2003) have also made important points with respect to this user innovation aspect. Casadesus-Masanell and Ghemawat (2006) employ a "demand-side learning" parameter in their dynamic model of competition between an open source and a closed source product; this idea comes close to the concept of knowledge accumulation from user collaboration that we have in mind. However, unlike in their model, we will assume collaboration as costly to the user in order to model key strategic dimensions of a knowledge-based service exchange: our model captures B2B (Business to Business) transactions while their model captures B2C (Business to Consumers) transactions where consumers are not strategic.

The rest of this chapter is structured as follows: in section 2.2, we present our general model of knowledge-based service industry. We then dissect the general model into two illustrative special cases: section 2.3 analyzes a reduced model with only ex-ante investment, whereas section 2.4 takes ex-ante investment as exogenous and focuses on customer ex-post collaboration. Section 2.5 develops two important extensions to the model, and section 2.6 concludes.

## 2.2 A Model of Knowledge-Based Service Industry

### 2.2.1 Service Transactions and Service Value

In order to model service transactions, we first need to clarify along which dimensions we consider a knowledge-based service transaction to be different from a commodity good purchase. We adopt the view of Zeithaml, Parasuraman, and Berry (1985) who propose the following four defining characteristics of a service as opposed to a commodity transaction:

1. Intangibility
2. Heterogeneity (inability to standardize)
3. Inseparability (of production and consumption)
4. Perishability (inability to inventoried).

These properties stand in sharp contrast to typical Industrial Organization models of commodity industry (in particular, models of B2C transactions) which assume that a producer first produces a standardized, tangible and storable commodity which is subsequently sold in a separate step to mass consumers. Intangibility and perishabilty will make contracting on service value difficult.[1] The notion of heterogeneity fits well the B2B service transactions that we have in

[1] In fact, we will assume that the value created by a service firm is not contractible.

mind. But the in our opinion most relevant and most unique aspect of a service transaction is the inseparability of consumption and production. It implies that service value is co-created by the consumer and the service firm, which leads us to adopt the following definition of a service:

**Definition.** *(Fitzsimmons and Fitzsimmons, 2001): A time-perishable, intangible experience performed for a customer acting in the role of co-producer.*

The lack of standardization and the inseparability of consumption and production make it obvious that a production function (which captures produced quantities of a standardized commodity good) is of little use for quantifying production in service industry. It seems more reasonable to capture service production by a function $\mathbb{V}(\cdot)$ that describes the *value* that is co-created by the consumer and the service firm from a given quantity of input factors.

We postulate that in the context of knowledge-based service industry, the created service value is a function of three input factors:

$$\begin{aligned} \mathbb{V}: \quad & \mathbb{R}^3 && \to \mathbb{R} \\ & (K, s, x) && \mapsto \mathbb{V}(K, s, x) \end{aligned} \tag{2.1}$$

The key input factor supplied by the service firm is its relevant *customer-specific knowledge stock* $K$. Customer-specific knowledge enables the service firm to understand the customer's problem in more detail and to deliver a better solution which increases service value. We postpone a more detailed discussion of knowledge stocks to the next subsection. The two other factors capture the customer's effort in collaborating towards a successful service outcome. We distinguish between *ex-ante investment* $s$ before the service contract is signed, and *ex-post collaboration* $x$ after the contract has been signed:

By *ex-ante investment* $s$, we mean the size of project-related irreversible investment that the customer makes before actually soliciting an offer from service firms. Basically, what we have in mind is that the customer, for instance a big firm, decides at the beginning of each period her overall business plan and the strategic role of the IT service in this plan. This in turn determines the size of the IT service that she will ask service providers to deliver, and the primary budget of the project. This decision on the level of ex-ante investment has some irrevocable element since revising it requires revising the entire business plan.

By *ex-post collaboration level* $x$, we mean the intensity and quality of tasks and practices that are performed by the customer at her own cost, but that positively affect the value of the service. Examples of such tasks in an IT service transaction environment are: generating a robust customer-side requirement analysis in the run-up for project specification; providing the service firm with documentation on internal processes and business organization; ensuring the disposability of customer staff for coordination with the service firm (for activities such as specification refinement, testing and the preparation of migration); supporting maintenance issues that can only be resolved jointly (e.g., bug regression); and promoting high levels of user skills and IT awareness among end-users and management.

As the above discussion illustrates, the dimensions of customer investment and customer collaboration comprise of many aspects that are difficult to measure objectively and thus impossible to contract upon. We therefore assume that $s$ and $x$ are observable but *not contractible* as in the incomplete contracting literature (Grossman and Hart, 1986; Hart and Moore, 1988), and as a consequence, the service value is not contractible either. Furthermore, we assume that the client cannot commit in advance to future choices of $s$ and $x$ and so the client and a firm cannot sign a long-term contract.

In order to focus on the client's choice of ex-ante investment and collaboration level, we deliberately abstract from the moral hazard problem on the part of service providers, which arises because $\mathbb{V}(K, s, x)$ is not contractible. More precisely, some papers on reputation such as Kreps (1990), Shapiro (1983), Choi (1998), Tadelis (2002) and Bar-Isaac (2007) study an agent's incentive to work (or shirk) when the quality of the service (or product) that he produces is not contractible and the client should pay the price before the service is produced (or before he consumes the product). Our $\mathbb{V}(K, s, x)$ can be interpreted as the value generated either when reputational concern allows the service provider to overcome his incentive problem or when he cannot overcome the problem. In other words, we focus on the client's choice of collaboration for given degree of incentive problem on the part of the service providers.

Both ex-ante investment and ex-post collaboration are costly for the customer. We model the customer's cost of ex-ante investment $s$ and ex-post collaboration $x$ by introducing cost functions $C_s(s)$ and $C_x(x)$, where we assume that these functions are convex and twice differentiable and satisfy $C_s(0) = C_x(0) = 0$, $\lim_{s\to 0} C_s'(s) = \lim_{x\to 0} C_x'(x) = 0$, and $C_s'' > 0$, $C_x'' > 0$. With respect to the cost incurred by the service provider, we postulate that the cost to deliver the service is constant with respect to $K$, $s$ and $x$. For the sake of simplified notation, we will set it equal to zero.

Finally, we assume that the service value function $\mathbb{V}(K, s, x)$ has the usual analytical properties of a neoclassical production function: it is differentiable twice, has positive but diminishing returns to $K$, $s$ and $x$, and all cross-derivatives are positive,

$$\frac{\partial^2}{\partial s \partial K}\mathbb{V}(K, s, x) > 0, \quad \frac{\partial^2}{\partial x \partial K}\mathbb{V}(K, s, x) > 0, \quad \frac{\partial^2}{\partial s \partial x}\mathbb{V}(K, s, x) > 0.$$

In other words, higher specific knowledge $K$ allows for higher returns to ex-ante investment and customer collaboration, and vice versa. We also assume that no value can be produced without knowledge: $\lim_{K \to 0} \mathbb{V}(K, s, x) = 0$.

### 2.2.2 Knowledge Stocks and Knowledge Accumulation

The final ingredient to our model of knowledge-based service industry is a notion of learning effects from repeated service transactions: whenever a knowledge-based service firm delivers a service, its interaction with the client allows it to privately observe information that increases the service firm's knowledge stock about this particular customer. A higher knowledge stock, in turn, increases the service value that the firm can generate in future service transactions with the same customer. For example, in the context of B2B IT service transactions a service firm may be able to learn about ways of integrating its solutions with the existing IT architecture of its client, about the structure of its client's business processes, or about unexploited opportunities to improve the client's manufacturing process with the help of information technology. Such knowledge is valuable because it allows the IT firm to deliver more tailor-made solutions in the future.

Clearly, the amount of learning depends on how much information the customer is willing to disclose, which is closely related to the customer's chosen level of participation before and during the service transaction. For our model of knowledge-based service industry we thus assume that the specific knowledge stock of a service firm that performs a service transaction for a customer evolves according to

$$K_{t+1} = K_t + k(K_t, s_t, x_t) \tag{2.2}$$

where $k(K, s, x)$ is a learning function that is increasing in all three parameters, concave and satisfies $\lim_{K \to 0} k(K, s, x) = 0$ for all $s, x$. We postulate that a higher knowledge stock of the firm increases the marginal learning from both customer collaboration and from project size:

$$\frac{\partial^2 k}{\partial K \partial x} > 0, \; \frac{\partial^2 k}{\partial K \partial s} > 0 \text{ and } \frac{\partial^2 k}{\partial s \partial x} > 0.$$

Finally, it is important to emphasize that for most knowledge-based services, the knowledge stock $K_t$ is of tacit nature and can't be transferred to rival firms by licensing. We will adopt this assumption throughout our model. Often the reason for this restriction lies in the limited availability or enforceability of intellectual property rights: for example, IT service providers who have detailed knowledge of possible software integration paths in a particular client's IT infrastructure can not claim any kind of intellectual property rights on such knowledge. As a consequence, a service firm cannot use a fixed fee to extract the benefit that accrues to a rival firm from an increase in the latter's stock of knowledge. Instead, a firm can only decide whether to keep its knowledge private or to give it away for free. This situation is quite coherent with what happens when an incumbent makes a software available under an open source license.

### 2.2.3 General Model of Knowledge-Based Service Transactions

We now have everything in place to write down a dynamic model of knowledge-based service industry: There are $N \geq 2$ service firms that are competing to provide a knowledge-based service to one single customer. The firms are endowed with initial knowledge stocks $K_1^1, K_1^2, \ldots, K_1^N$. Without loss in generality, we shall assume that firms are arranged such that $K_1^1 \geq K_1^2 \geq \cdots \geq K_1^N$. Time is discrete, $t = \{1, 2, \ldots\}$. Every period of the model comprises of a game in which the players play the following steps:

1. the customer chooses the size of the project by making an ex-ante investment of $s_t$ at a cost of $C_s(s_t)$; these cost become sunk immediately

2. the customer receives bids $p_i(s_t)$ from all firms $i \in \{1, 2, \ldots, N\}$ who offer to execute the project; the customer chooses the firm that leaves her the highest surplus, taking into account how the chosen service firm's learning will increase or reduce her surplus in the following periods. In case the customer is indifferent between firms, we assume that she chooses the incumbent (i.e., the firm with the highest knowledge stock).

3. the customer provides on-project collaboration effort $x_t^i$ and incurs the corresponding cost $C_x(x_t^i)$

4. at the end of the period, the service value $\mathbb{V}(K_t^i, s_t, x_t^i)$ materializes, firms' knowledge stocks update according to the corresponding equations of motion, and the customer obtains a total period surplus of

$$S_t(K_t^i, s_t, x_t^i) = \mathbb{V}(K_t^i, s_t, x_t^i) - C_s(s_t) - C_x(x_t^i) - p_i(s_t) \tag{2.3}$$

Firms maximize the net present value of overall profits, whereas the customer maximizes the net present value of overall customer surplus; both the customer and the firm discount the future by the same discount factor $\delta < 1$.

We believe that this model captures in a quite general manner key aspects of knowledge-based services, i.e. the co-creation of value between customers and service firms, and the existence of a dynamic learning channel. Although this general model with infinite horizon could be solved with recursive techniques, the intuition for its properties is conveyed better by restricting the model to two special cases: in what follows, we will solve two two-period versions of the general model where we focus on only one of the two customer choice variables at a time: in the first model, we will solve for the customer's optimal choice of ex-ante investment $s_t$ under different knowledge sharing policies, whereas our second model takes investment size $s_t$ as exogenously fixed and completes the picture by analyzing the customer's optimal choice of ex-post collaboration $x_t$.

## 2.3 Knowledge Sharing and Ex-ante Investment

In this section, we will focus on understanding the impact of various knowledge sharing policies on customer ex-ante investment and service firm profits. We thus choose to completely abstract from the aspect of ex-post customer collaboration. This can be attained easily by assuming that the learning function $k(K, s, x)$ does not depend explicitly on the collaboration effort $x$,

$$k(K, s, x) = k_s(K, s).$$

With this simplification, the client's choice of ex-post collaboration $x$ in one period will leave the service offers of future periods unaffected, and in equilibrium the client thus simply chooses the optimal static collaboration level $\bar{x}^i = \bar{x}^i(K^i, s)$ with firm $i$ according to the first order condition $\frac{\partial}{\partial x}\mathbb{V}(K^i, s, \bar{x}^i) = C'_x(\bar{x}^i)$. We can therefore transform the problem and conveniently hide all terms related to ex-post collaboration $x^i$ in a new definition of the service value function[2]:

$$\mathcal{V}(K^i, s) \equiv \mathbb{V}(K^i, s, \bar{x}^i(K^i, s)) - C_x(\bar{x}^i(K^i, s)) \tag{2.4}$$

In what follows, we compare for a given level of the incumbent's initial knowledge stock $K_1^I$ the equilibrium ex-ante investments, customer surplus and profits of three games $\{\Gamma_C, \Gamma_O, \Gamma_R\}$. Each game corresponds to the incumbent's adoption of one of the following three knowledge sharing policies: *closed source* (C), *open source* (O) and *reciprocal open source* (R). We precise on the definition and the economic meaning of these policies:

**Closed Source (C):** If the incumbent selects closed source, he retains exclusivity over his customer-specific knowledge. This makes him the only firm that is able to offer positive service value. We capture this by assuming that the initial knowledge stock of all entrants is zero, $K_1^E = 0$. Moreover, under closed source only the firm $i$ that is chosen to provide the service can realize positive learning effects from the co-creation of service value with

[2] Note that, by the nature of the transformation, all our assumptions regarding the properties of $\mathbb{V}$ and its derivatives carry over to $\mathcal{V}$.

the customer:

$$K_{t+1}^j = \begin{cases} K_t^j + k_s(K_t^j, s_t) & \text{for the chosen firm } j = i \\ K_t^j & \text{for all other firms } j \neq i \end{cases} \tag{2.5}$$

In other words, the choice of a closed source license allows the incumbent to keep all his static knowledge stock and all dynamic knowledge gains private.

**Open Source (O):** If the incumbent adopts open source, he gives up his monopoly position over the customer: by releasing (some of) the customer-specific software which it uses to deliver the service under an open source license, the incumbent makes a portion of his knowledge stock freely available to entrants. This creates competition because the new entrants can use this knowledge to bid on providing a service to the same customer. In the model, we capture this by assuming that the release of the software as open source means that the incumbent shares a positive fraction[3] of his initial knowledge stock with entrants, so $0 < K_1^E < K_1^I$.

The mere adoption of an open source license does not imply that dynamic knowledge gains are also shared: for example, if code is made available under a BSD license, there is no requirement for products that modify or improve the original code to be released again as open source. Similarly, firms that resort to dual-licensing (see Comino and Manenti, 2007) are not committed to sharing any dynamic knowledge gains because they typically reserve the right to ship their own improvements under proprietary licensing terms. Thus we assume that knowledge stocks evolve under the same law of motion as under closed source and follow equation (2.5).

**Reciprocal Open Source (R):** As a third knowledge-sharing policy, we consider the case in which the incumbent publishes his software exclusively under a reciprocal open source license such as the GNU General Public License. Such reciprocal open source licenses state that licensees must share all improvements to the software under the license terms of the original code (and thus make them open source). Many commercial software firms that have released software under such license terms have even taken additional measures to further underline their commitment to keeping their development

[3]We assume that the incumbent still retains an advantage over the entrants because knowledge sharing via open source is not perfect. For example, some relevant parts of the code might be rather obfuscated and may be well-understood only by the incumbent's developers, or the incumbent's developers hold additional private knowledge that is not reflected in the source code. In an extension of our analysis, we also discuss what determines the fraction of knowledge stock that the incumbent would optimally want to share.

process open for everyone, for example by transferring the oversight of the development process to a non-profit foundation.

We interpret such choices on the side of the incumbent as a commitment to sharing not only static knowledge stocks but also all future dynamic knowledge gains. Moreover, due to the reciprocal terms of the license any entrant will be equally obliged to share all dynamic increases of his knowledge stock. We will model this situation by postulating that all dynamic changes to the knowledge stock of the chosen firm $i$ will equally accrue to all its competitors,

$$K_{t+1}^j = K_t^j + k_s(K_t^i, s_t) \quad \text{for all firms } j \in \{1, \ldots, N\} \tag{2.6}$$

Note that for the sake of simplicity we have assumed that the rival firms' gain is equal to the full amount of the dynamic knowledge gain of the service firm that was actually chosen for the service contract.

We will now solve the three games $\Gamma_C, \Gamma_O, \Gamma_R$ which have the following structure: First, the initial knowledge stocks of entrants is set according to the static knowledge sharing policy of the corresponding game, i.e.

$$K_1^2 = K_1^3 = \cdots = K_1^N = \begin{cases} 0 & \text{for } \Gamma_C \\ K_1^E & \text{for } \Gamma_O, \Gamma_R \end{cases}$$

Next, the firms compete for two periods $t = \{1, 2\}$ to deliver a knowledge-based service to the customer. The structure and timing of each period follows our general service model[4]:

1. the customer chooses ex-ante collaboration $s_t$ and pays $C_s(s_t)$

2. all service firms quote their prices $p^i(s_t)$

3. the customer chooses the firm that offers the best overall surplus, taking into account how her choice affects future periods

4. a service value of $\mathcal{V}(K_t^i, s_t)$ is produced and knowledge stocks are updated according to the dynamic knowledge sharing policy of the game, i.e. equation (2.5) for $\Gamma_C, \Gamma_O$ and equation (2.6) for $\Gamma_R$.

[4]in its reduced form with static ex-post collaboration effort

### 2.3.1 Benchmark: Planner's Solution

Before we establish the equilibria of the service transaction games under closed source and open source, we first pin down the socially optimal choices as a benchmark: which levels of ex-ante investment should a social planner choose whose exclusive concern is to maximize social welfare? Clearly, the highest social welfare is attained if the planner allocates the service transaction to the most knowledgeable service provider; hence, the planner solves the problem

$$\max_{s_1,s_2} \mathcal{V}(K_1^I, s_1) - C_s(s_1) + \delta \left[\mathcal{V}(K_1^I + k_s(K_1^I, s_1), s_2) - C_s(s_2)\right] \tag{2.7}$$

We thus find that the socially optimal investments $s_1^*$ and $s_2^*$ are uniquely pinned down by the first-order conditions

$$\left.\frac{\partial \mathcal{V}}{\partial s}\right|_{(K_1^I + k_s(K_1^I, s_1), s_2^*)} = C_s'(s_2^*) \text{ and} \tag{2.8}$$

$$\left.\frac{\partial \mathcal{V}}{\partial s}\right|_{(K_1^I, s_1^*)} + \delta \left.\frac{\partial \mathcal{V}}{\partial K}\right|_{(K_1^I + k_s(K_1^I, s_1^*), s_2^*)} \cdot \left.\frac{\partial k_s}{\partial s}\right|_{(K_1^I, s_1^*)} = C_s'(s_1^*) \tag{2.9}$$

The meaning of these equations is quite intuitive: whilst in the second period the planner simply chooses the statically optimal amount of ex-ante investment, in the first period he additionally takes dynamic gains from knowledge accumulation into account and thus selects an investment level higher than the statically optimal one.

### 2.3.2 Game $\Gamma_C$: Closed Source

We can now proceed to solve the game $\Gamma_C$ of closed source service provision. The incumbent firm will enjoy monopoly power over the customer since all entrants' knowledge stocks (and thus ability to produce value) are zero. The incumbent will win both periods of the game, and entrants will never acquire any knowledge. But monopoly power will enable the incumbent to extract all surplus: in the second period, the incumbent will charge a price of

$$p_2^C(s_2^C) = \mathcal{V}(K_1^I + k_s(K_1^I, s_1^C), s_2^C)$$

and leave the customer with a second period surplus of $-C_s(s_2^C)$. Hence, the customer will optimally abstain from making any irreversible ex-ante investment before writing a contract with the incumbent service firm: $s_2^C = 0$. Going backwards, the same logic also applies to the first period, and we have

**Proposition 1.** *Under closed source, there is zero ex-ante investment in both periods: $s_1^C = s_2^C = 0$. The total surplus that is generated amounts to $\mathcal{V}(K_1^I, 0) + \delta \cdot \mathcal{V}(K_1^I + k(K_1^I, 0), 0)$. All surplus accrues to the incumbent.*

We see that the combination of monopolistic market power with lack of commitment not to hold up the customer for her irreversible ex-ante investments eradicates all incentives to make such investment in first place. Thus, relative to the planner's solution, social welfare will be lost: this effect is most pronounced for highly customized services whose value $\mathcal{V}(K, s)$ increases steeply with $s$, and it is least relevant for transactions like commodity purchases that do not derive much additional value from irreversible customer ex-ante investment.

### 2.3.3 Game $\Gamma_O$: Open Source (without commitment)

Under open source, the incumbent shares some of his initial knowledge stock with all entrants: all entrants obtain free access to the source code of the software which endows them with a positive initial knowledge stock $K_1^E < K_1^I$ and enables them to enter the market. In this section, we solve the model assuming that the incumbent chooses a non-reciprocal open source licensing model (such as BSD license) and that neither the incumbent nor any of the entrants can commit to sharing their dynamic knowledge gains with competitors.

Again, we solve the game by backward induction. For a given ex-ante investment $s_1$ in the first period, knowledge stocks in the second depend on which firm was awarded the service project in the first period. We therefore distinguish two cases:

#### Case 1: Second period if the incumbent was awarded the first period

Suppose that the incumbent has run the project in the previous period. Then learning effects increase the knowledge stock of the incumbent to $K_1^I + k_s(K_1^I, s_1)$ whereas all entrants' knowledge stocks remain at $K_1^E$. Since any of the entrants can at most offer the entire service value that would be generated, the incumbent can charge a price of up to

$$p_{2,I}^I(s_2) = \mathcal{V}(K_1^I + k_s(K_1^I, s_1), s_2) - \mathcal{V}(K_1^E, s_2)$$

without losing the customer to an entrant. We will use the subscript $2, I$ to represent the second-period payoffs when when the incumbent was awarded the project in period one. For a given level of second-period investment $s_2$, the customer surplus and each firm's profit in period two reads

$$\begin{aligned}
S_{2,I} &= \mathcal{V}(K_1^E, s_2) - C_s(s_2) \\
\Pi_{2,I}^I &= \mathcal{V}(K_1^I + k_s(K_1^I, s_1), s_2) - \mathcal{V}(K_1^E, s_2) \\
\Pi_{2,I}^E &= 0 \text{ (for all entrant firms)}
\end{aligned}$$

Let us denote the equilibrium level of ex-ante investment under open source in the second period when the incumbent was awarded the first period as $s^O_{2,I}$. It maximizes customer surplus and is thus determined by the first-order condition

$$C'_s(s^O_{2,I}) = \left.\frac{\partial V}{\partial s}\right|_{(K^E_1, s^O_{2,I})} \tag{2.10}$$

We note that the second period ex-ante investment $s^O_{2,I}$ is independent of the previous period collaboration level $s_1$. Moreover, it is easy to see that $s^O_{2,I}$ is strictly greater than zero (which was the level $s^C_2$ attained under closed source) but falls short of the socially optimal level $s^*_2$.

**Case 2: Second period if an entrant was awarded the first period**

If the customer chooses an entrant in the first period as her service provider, the incumbent's knowledge stock in period two will remain at its initial level $K^I_1$ whereas the knowledge stock of the selected entrant will be augmented by learning effects from the first period service transaction: $K^E_2 = K^E_1 + k_s(K^E_1, s_1)$. The outcome of the second period competition now depends to a considerable extent on whether the entrant's dynamic learning effect is sufficiently large to surpass the incumbent in terms of knowledge stock, i.e. whether

$$K^E_1 + k_s(K^E_1, s_1) > K^I_1. \tag{2.11}$$

We will refer to this situation as *substantial learning*. If learning is substantial, the entrant firm will win the service contract in the second period as well and it will make positive profits; otherwise, the incumbent wins the second period contract and extracts the part of customer surplus that is not protected by the customer's outside option of choosing the entrant. Let us denote as $H$ the firm $\in \{E, I\}$ with the higher second period knowledge stock, and as $L$ the firm with the second-highest second period knowledge stock, both conditional on the first period being run by the entrant:

$$\begin{aligned} K^H_2 &\equiv \max\{K^I_1, K^E_1 + k_s(K^E_1, s_1)\} \\ K^L_2 &\equiv \min\{K^I_1, K^E_1 + k_s(K^E_1, s_1)\} \end{aligned}$$

Then, by the same arguments as in the previous case, the most knowledgeable firm $H$ (which can now be either the incumbent or the entrant) will win the contract in period two, will charge a price of

$$p_{2,E}(s_2) = \mathcal{V}(K^H_2, s_2) - \mathcal{V}(K^L_2, s_2)$$

and leave the customer with a surplus of

$$S_{2,E} = \mathcal{V}(K_2^L, s_2) - C_s(s_2)$$

such that the customer's optimal second period choice of ex-ante investment $s_{2,E}^O$ will be determined by the first-order equation

$$\left.\frac{\partial \mathcal{V}}{\partial s}\right|_{(K_2^L, s_{2,E}^O)} = C_s'(s_{2,E}^O)$$

We observe that, since $K_2^L > K_1^E$ and $\frac{\partial^2 \mathcal{V}}{\partial K \partial s} > 0$, the customer's marginal gains from ex-ante investment are higher if an entrant rather than the incumbent was chosen for the first period. Thus, awarding the project to the entrant in the first period will induce strictly higher second period ex-ante investment than if the first period project is awarded to the incumbent. Nevertheless, because $K_2^L < K_1^I + k_s(K_1^I, s_1)$ the amount of ex-ante investment in both cases falls short of the socially optimal amount that we have established in the planner's benchmark. In addition, we show in appendix 2.A.1 that the higher amount of ex-ante investment under the entrant also generates a higher surplus for the customer in period two than if the customer had chosen the more knowledgeable incumbent in period one.

Summarizing the results of the game's second period, we have:

**Observation** In period two, ex-ante investment is lower than the social optimum. The underinvestment is most severe when the first period project was awarded to the incumbent, and less severe when when the entrant was chosen in the first period. There is an intrinsic conflict between choosing the right service firm in period one and choosing an efficient level of ex-ante investment in period two.

Before moving on to the first period, it will be useful to clarify already at this stage how the equilibrium ex-ante investment level in period two depends on the choice of $s_1$. Note that the customer's incentive to invest ex-ante in the second period depends only on the level of $K_1^E$ (if the incumbent was awarded the first period) or the level of $K_2^L$ (if an entrant was awarded the first period). Only the latter could possibly be a function of $s_1$ because $K_2^L$ increases in $s_1$ if learning is non-substantial. Therefore, we note that the equilibrium choice of $s_2$ does not depend on $s_1$ if the entrant's learning is substantial or if the incumbent wins the first period. Otherwise, the equilibrium level of $s_2$ will be increasing in $s_1$.

### First Period

In period one, if the incumbent wants to win the contract, he can always win it. However, it may be optimal not to win the contract. We first conduct the analysis under the assumption that the incumbent wins the contract and then study when it is optimal for him to actually do so.

Solving backwards, let us for now take $s_1$ as given and analyze the prices $p_1^i(s_1)$ that the competing service firms will charge in equilibrium as they attempt to win the business of the customer in the first period. If learning is not substantial ($K_1^E + k(K_1^E, s_1) < K_1^I$), the best offer an entrant can make is a price of $p_1^E = 0$. If learning is substantial, an entrant can pledge the second period profit that he can realize if he is chosen as the service provider. Thus, entrants will offer to run the project in the first period for a price of

$$p_1^E(s_1) = -\left[\delta \cdot \max\{0, \mathcal{V}(K_1^E + k_s(K_1^E, s_1), s_{2,E}^O) - \mathcal{V}(K_1^I, s_{2,E}^O)\}\right]$$

In order to win the customer in period one, the incumbent has not only to match this offer but he also needs to compensate the customer for the partial loss of her second period surplus that she will incur if she chooses the incumbent in period one:

$$p_1^I(s_1) = \mathcal{V}(K_1^I, s_1) - \mathcal{V}(K_1^E, x_1) - p_1^E - \delta\,(S_{2,E} - S_{2,I})$$

As a consequence, the overall surplus that accrues to the customer under open source when he chooses the incumbent is equal to

$$S_{total}^O \;=\; \mathcal{V}(K_1^I, s_1) - p_1^I(s_1) - C_s(s_1) + \delta \cdot S_{2,I}$$

which after substituting expressions and further simplification[5] becomes

$$S_{total}^O = \mathcal{V}(K_1^E, s_1) - C_s(s_1) + \delta\left[\mathcal{V}(K_1^E + k_s(K_1^E, s_1), s_{2,E}^O) - C_s(s_{2,E}^O)\right] \tag{2.12}$$

Basically, the customer's total surplus is given by the payoff that she can achieve when she has free access to the entrant's knowledge. A minor qualification is that $s_{2,E}^O(s_1)$ maximizes $\mathcal{V}(K_1^E + k_s(K_1^E, s_1), s_2) - C_s(s_2)$ only when the innovation is not substantial.

Finally we can solve the first step of the game: the customer chooses the level of ex-ante investment $s_1$ in order to maximize her overall surplus $S_{total}^O$. The optimal level $s_1^O$ of ex-ante investment in period one is therefore characterized by the first-order condition[6]

$$\left.\frac{\partial \mathcal{V}}{\partial s}\right|_{(K_1^E, s_1^O)} + \delta\left[\left.\frac{\partial \mathcal{V}}{\partial K}\right|_{(K_1^E + k_s(K_1^E, s_1^O), s_{2,E}^O)} \cdot \left.\frac{\partial k_s}{\partial s}\right|_{(K_1^E + k_s(K_1^E, s_1^O), s_{2,E}^O)}\right] = C_s'(s_1^O) \tag{2.13}$$

[5]In particular, we exploit the fact that the entrant's second period profits and the second period customer surplus under the entrant must sum to $\mathcal{V}(K_1^E + k_s(K_1^E, s_1), s_{2,E}^O) - C_s(s_{2,E}^O)$.

[6]Note that from the envelope theorem, we can neglect the indirect effect through the change $s_{2,E}^O$ when learning is not substantial: furthermore, when learning is substantial, $s_{2,E}^O$ does not depend on $s_1$.

This equation has a deeper economic interpretation: It shows that ex-ante investment in the first period generates two sources of benefits to the customer. First, it generates the immediate benefit of increasing the value of the period one service and the customer appropriates a part of it. Second, it increases the value that an entrant can produce in period two from accumulating customer-specific knowledge if the same firm was also chosen for the first period. Since this raises the value of the outside option of the customer, the customer benefits from it. In other words, the customer can sell the right to learn about her service needs.

We can now compare the equilibrium value of period one investment size $s_1^O$ under open source to first-best benchmark $s_1^*$ which yields the following insight:

**Proposition 2.** *In the game $\Gamma_O$ of ex-ante investment under open source without commitment, the customer's ex-ante investment in period one is strictly higher than under closed source, i.e. $s_1^O > s_1^C$. The position of $s_1^O$ relative to the socially optimal level $s_1^*$ is undetermined: depending on the functional forms of $\mathcal{V}(\cdot,\cdot)$, of $k_s(\cdot,\cdot)$ and the magnitude of the discount factor $\delta$, $s_1^O$ can be inefficiently low, efficient or excessive. For sufficiently small discount factor $\delta$, $s_1^O$ will always be inefficiently low.*

*Proof.* see appendix. □

It is interesting to note that open source can induce socially excessive ex-ante investment in the first period. This occurs because the marginal impact of an increase of knowledge stock on productivity is higher for the entrant than for the incumbent due to diminishing returns to knowledge.

Finally, we calculate the profits for the incumbent: the incumbent's payoff is the total value it produces minus the total value that entrants who win the customer for two periods would generate, so

$$\begin{aligned}\Pi^{I,O}_{total} &= \mathcal{V}(K_1^I, s_1^O) - \mathcal{V}(K_1^E, s_1^O) + \delta\left[C_s(s_{2,E}^O) - C_s(s_{2,I}^O)\right] \\ &\quad + \delta\left[\mathcal{V}(K_1^I + k_s(K_1^I, s_1^O), s_{2,I}^O) - \mathcal{V}(K_1^E + k_s(K_1^E, s_1), s_{2,E}^O)\right]\end{aligned}$$

The term related to cost, $C_s(s_{2,E}) - C_s(s_{2,I})$, represents the customer's lower investment cost in period two in case the incumbent is selected in period 1. Note that the incumbent does not bear the customer's period one cost of ex-ante investment since it is already sunk.

### When does the incumbent want to win the first period?

Up to now we have assumed that the incumbent wants to win the period one contract. However, if the initial gap in terms of knowledge stock is large and learning is non-substantial, the incumbent may be better off letting an entrant win the period one contract in order to boost the customer's period two ex-ante investment. Since there is free entry under open source, an entrant's overall profit is always zero whereas the customer gets all the benefit from having free access to the entrant's technology. Thus, the customer's payoff for a given $s_1$ is the same as in the previous subsection and the equilibrium choice of $s_1$ remains unchanged. We can therefore directly compare the incumbent's payoff conditional on winning the period one contract and his payoff conditional on losing it: the incumbent will choose to lose the first period contract whenever his payoff conditional on the entrant serving the first period contract,

$$\Pi^{I,O}_{onlyperiod2} = \delta \left\{ \mathcal{V}(K_1^I, s_{2,E}^O) - \mathcal{V}(K_1^E + k_s(K_1^E, s_1^O), s_{2,E}^O) \right\},$$

exceeds the overall profits $\Pi^{I,O}_{total}$ from serving the customer for both periods.

In order to better understand the conditions under which it remains optimal for the incumbent to win the first period contract, let us consider just for a moment a static one-period version of the game with fixed knowledge stocks $(K^I, K^E, \dots)$: in a static model, the customer's optimal investment choice $\bar{s}$ will be pinned down by the first-order condition

$$\left.\frac{\partial \mathcal{V}}{\partial s}\right|_{(K^E, \bar{s})} = C_s'(\bar{s})$$

which implies that $\bar{s}$ is an increasing function of the entrants' knowledge stock: $\bar{s} = \bar{s}(K^E)$. The incumbent's profit function will be

$$\Pi^I(K^E) = \mathcal{V}(K^I, \bar{s}(K^E)) - \mathcal{V}(K^E, \bar{s}(K^E)).$$

Clearly, if $K^E$ is sufficiently close to $K^I$, profits will be decreasing in $K^E$: as the knowledge gap between the incumbent and entrants becomes smaller, the incumbent's profits fall towards zero. However, if entrants have substantially smaller knowledge stocks than the incumbent (i.e., $K^E \ll K^I$), profits can be increasing in $K^E$ because higher entrant knowledge will induce higher customer investment $\bar{s}(K^E)$. Assume that $\Pi^I(K^E)$ is strictly concave and reaches its maximum for some value $K^{E*} \in \left[0, K^I\right)$. Going back to our dynamic model, $K^E \geq K^{E*}$ will then be a sufficient condition under which the incumbent will never yield the first period to an entrant: to see this, note that the incumbent's second period profit in the dynamic model depends on $K^E$ exactly like the profit function in a static game, and thus the incumbent's second period (and overall) profits would decrease if the entrant were to gain further knowledge.

The fact that the incumbent may find it optimal not to win the first period also suggests that knowledge-based service firms may use market segmentation as a mechanism to induce customer investment. Although we deliberately focused on the case of one buyer, it is easy to extend the result to a setting with two buyers where the knowledge accumulated from one customer is partially applicable to another and vice versa. In this situation, even though the incumbent's production technology has constant return to scale and hence he can win both customers if he wants, the incumbent may deliberately leave a customer to a competitor in order to boost customer investment.

### 2.3.4 Game $\Gamma_R$: Reciprocal Open Source

In this section, we finally explore the equilibrium of the game $\Gamma_R$ in which the incumbent can commit, for example by using a reciprocal open source license, to incorporate all dynamic knowledge gains from the provision of the service in his open source software. We again solve backwards, starting with the second period. Since all firms are committed to sharing all their knowledge gains, the incumbent will always have a larger knowledge stock than any entrant, irrespectively of which firm was chosen in period one. Hence, the incumbent always wins both periods. For the sake of brevity, we restrict our discussion to this particular case. The full analysis of all subgames of game $\Gamma_R$ (as needed to establish the first period price charged by the incumbent) can be found in the appendix.

If the incumbent was selected in the first period, his second period knowledge stock will be $K_2^I = K_1^I + k_s(K_1^I, s_1)$ whereas all entrants will have knowledge stocks of $K_2^E = K_1^E + k_s(K_1^I, s_1)$. Thus, the incumbent firm will always win the second period and charge a price of

$$p_{2,I}^R(s_2) = \mathcal{V}(K_1^I + k_s(K_1^I, s_1), s_2) - \mathcal{V}(K_1^E + k_s(K_1^I, s_1), s_2)$$

which leaves the customer with a surplus of

$$S_{2,I}^R = \mathcal{V}(K_1^E + k_s(K_1^I, s_1), s_2) - C_s(s_2).$$

The corresponding first-order condition for the customer's second period's choice of ex-ante investment $s_{2,I}^R$ therefore is

$$\left.\frac{\partial \mathcal{V}}{\partial s}\right|_{(K_1^E + k_s(K_1^I, s_1), s_{2,I}^R)} = C_s'(s_{2,I}^R)$$

which immediately unveils the following: under reciprocal open source, ex-ante investment $s_{2,I}^R$ in period two is greater than ex-ante investment $s_{2,I}^O$ under open source without commitment, and it is always increasing in the first period investment level $s_1$. The reason for this is that dynamic knowledge gains are now shared among all firms which raises the customer's outside options in period two. Thus, much of the additional second period surplus from dynamic knowledge accumulation accrues to the customer. This is to be contrasted with the equilibrium of the game $\Gamma_O$ of open source without commitment, where all the additional second period surplus from dynamic knowledge accumulation is extracted by the service firm.

Going back to the first period, the incumbent will charge a price of

$$p_1(s_1) = \mathcal{V}(K_1^I, s_1) - \mathcal{V}(K_1^E, s_1) + \delta \left[S_{2,I}^R - S_{2,E}^R\right].$$

We see that the incumbent can charge a premium over the immediate value offered by the entrant because the second period customer surplus is higher if the incumbent obtains the first period due to superior learning effects.

The customer will choose a level of first period ex-ante investment $s_1^R$ that maximizes her overall surplus,

$$S_{total}^R(s_1) = \mathcal{V}(K_1^E, s_1) - C_s(s_1) + \delta \left[\mathcal{V}(K_1^E + k_s(K_1^E, s_1), s_{2,E}^R) - C_s(s_{2,E}^R)\right],$$

so her optimal choice $s_1^R$ will satisfy the following first-order condition:

$$\left.\frac{\partial \mathcal{V}}{\partial s}\right|_{(K_1^E, s_1^R)} + \delta \left.\frac{\partial \mathcal{V}}{\partial K}\right|_{(K_1^E + k_s(K_1^E, s_1^R), s_{2,I}^R)} \cdot \left.\frac{\partial k_s}{\partial s}\right|_{(K_1^E, s_1^R)} = C_s'(s_1^R)$$

Remarkably, this equation is almost identical to our result for the first-order condition under open source without commitment: the only difference is that the term in $\frac{\partial \mathcal{V}}{\partial K}$ is now evaluated at a strictly higher second period investment level $s_{2,I}^R > s_{2,I}^O$. Comparing this finding with our previous results yields immediately the following insight:

**Proposition 3.** *In the game $\Gamma_R$ of ex-ante investment under reciprocal open source, the incumbent always wins both periods. The first period ex-ante investment $s_1^R$ is always higher than the corresponding first period ex-ante investment $s_1^O$ under open source without commitment; its position relative to the benchmark level $s_1^*$ is undetermined, i.e. $s_1^R$ can be inefficiently low, efficient or excessive. Second period investment $s_{2,I}^R$ under reciprocal open source is strictly higher than under open source without commitment but it falls short of the efficient level $s_2^*$.*

*Proof.* It follows immediately from comparing the relevant first-order conditions. □

Finally, we can calculate the profits of the incumbent under reciprocal open source. We find:

$$\begin{aligned}
\Pi_{total}^{I,R} = \; & \mathcal{V}(K_1^I, s_1^R) - \mathcal{V}(K_1^E, s_1^R) \\
& + \delta \left[\mathcal{V}(K_1^I + k_s(K_1^I, s_1^R), s_{2,I}^R) - V(K_1^E + k_s(K_1^E, s_1^R), s_{2,E}^R)\right] \\
& + \delta \left[C_s(s_{2,E}^R) - C_s(s_{2,I}^R)\right]
\end{aligned}$$

Comparing this expression with the corresponding profit function $\Pi^{I,O}_{total}$ under open source without commitment we can see that the profit functions for the incumbent are exactly the same under open source (O) and reciprocal open source (R); only investment levels are different. Whilst this may seem surprising at first glance, it is actually intuitive to understand: whilst there is no sharing of knowledge gains with rivals under open source (O), the incumbent does compensate the customer in period one for the potential knowledge gains that rivals could obtain from running the first period. These inter-period transfers result in overall payoffs that are not very different from the outcome under reciprocal open source (R).

### 2.3.5 Comparison: Determinants of Open v.s. Closed Source

We can now attempt to understand under which conditions the incumbent will optimally choose an open source license (O) or (R) or a closed source license (C) to deliver his knowledge-based service. Hence, we need to compare social welfare and profits of the games $\Gamma_C$, $\Gamma_O$ and $\Gamma_R$.

Before we solve this problem, let us first understand which license would be optimal to choose if the incumbent were to sell a commodity good (rather than a service). For a typical commodity good, there are no knowledge gains, so we assume $k_s(K, s) = 0$. Moreover, the value of a commodity transaction can depend positively on the knowledge $K$ of the good producer but there typically is no benefit from any irreversible ex-ante investment on the customer side; therefore, we assume that $\mathcal{V}(K, s)$ only depends on $K$, i.e. $\mathcal{V}(K, s) = \bar{\mathcal{V}}(K)$. Comparing between the three available licensing options, we observe that overall surplus is the same for each of them, but only closed source (C) allows the incumbent to extract all of it. Hence, we observe:

**Observation** In the model of ex-ante investment, an incumbent will never choose open source (O) or reciprocal open source (R) for a commodity good transaction: in terms of profits, these choices are strictly dominated by closed source (C).

For the provision of knowledge-based services, however, closed source is likely to be suboptimal. Before we discuss this scenario in detail we shall, for the sake of clarity, we introduce the following notation: we denote as $SW(K, s_1, s_2)$ the social welfare that is produced over the course two periods if a service firm with an initial knowledge stock of $K$ runs the project in both periods and the customer invests amounts of $s_1$ and $s_2$, respectively:

$$SW(K, s_1, s_2) \equiv \mathcal{V}(K, s_1) - C_s(s_1) + \delta \left[\mathcal{V}(K + k(K, s_1), s_2) - C_s(s_2)\right]$$

This function has the following important property:

**Lemma 4.** *Social welfare $SW(K, s_1, s_2)$ is strictly concave in $s_1$, and it is also strictly concave in $s_2$.*

*Proof.* Concavity in both variables follows directly from the fact that $\frac{\mathrm{d}^2 C_s}{\mathrm{d}s^2} > 0$ and that $\frac{\partial^2 \mathcal{V}}{\partial K^2}$, $\frac{\partial^2 \mathcal{V}}{\partial s^2}$ and $\frac{\partial^2 k_s}{\partial s^2}$ are all negative. $\square$

We can now rank the social welfare that is generated under each of the three games of ex-ante investment:

**Proposition 5.** *If open source (O) does not lead to excessive investment in period one (i.e., $s_1^O \leq s_1^*$), social welfare under open source (O) is strictly higher than under closed source (C). Even if open source leads to excessive investment in period one, social welfare under open source is higher than under closed source as long as $SW(K_1^I, s_1^O, 0) > SW(K_1^I, 0, 0)$.*

*Proof.* The result is a direct consequence of the concavity of $SW(K, s_1, s_2)$ in $s_1$ and $s_2$: whenever $s_1^* > s_1^O > s_1^C = 0$ we can use the fact that $s_2^* > s_{2,I}^O > s_{2,I}^C = 0$ together with concavity to conclude that

$$\begin{aligned} SW(K_1^I, s_1^*, s_2^*) &> SW(K_1^I, s_1^O, s_2^*) > SW(K_1^I, s_1^O, s_2^O) > \dots \\ &> SW(K_1^I, s_1^O, 0) > SW(K_1^I, 0, 0) = SW(K_1^I, s_1^C, s_2^C). \end{aligned}$$

Note that the second last inequality relies only on the concavity of $SW(K, s_1, s_2)$ in $s_2$ and remains true independently of whether $s_1^O$ is excessive or not. $\square$

By the same argument, we also have

**Corollary 5.1.** *If reciprocal open source (R) does not lead to excessive investment in period one (i.e., $s_1^R \leq s_1^*$), social welfare under reciprocal open source (R) is strictly higher than under open source (O). Even if reciprocal open source leads to excessive investment in period one, $SW(K_1^I, s_1^R, s_2^O) > SW(K_1^I, s_1^O, s_2^O)$ is a sufficient condition for social welfare under reciprocal open source (R) to be higher than under open source (O). Then, if ranked ("$\succsim$") by social welfare, we have that $R \succsim O \succsim C$.*

Even though with open source licenses the incumbent must surrender a fraction of the total surplus to the customer, the gain in overall welfare (as compared to closed source) can be sufficiently large that the incumbent's profits are higher under open source (O) or reciprocal open source (R) than under closed source: to see this, note that we can rewrite the previously derived profit functions under closed source, open source and reciprocal open source as

$$\begin{aligned} \Pi_{total}^{I,C} &= SW(K_1^I, s_1^C, s_{2,I}^C), \\ \Pi_{total}^{I,O} &= SW(K_1^I, s_1^O, s_{2,I}^O) - SW(K_1^E, s_1^O, s_{2,E}^O), \text{ and} \\ \Pi_{total}^{I,R} &= SW(K_1^I, s_1^R, s_{2,I}^R) - SW(K_1^E, s_1^R, s_{2,E}^R). \end{aligned}$$

We can thus express the difference in profit between open source (O) and closed source (C) as

$$\Pi^{I,O}_{total} - \Pi^{I,C}_{total} = \left[SW(K^I_1, s^O_1, s^O_{2,I}) - SW(K^I_1, s^C_1, s^C_{2,I})\right] - SW(K^E_1, s^O_1, s^O_{2,E}).$$

The first term reflects the surplus gain under open source due to more efficient investment size and has positive sign under the conditions that we have discussed previously. The second term is negative and represents the loss in profit due to competition by entrants. The comparison between reciprocal open source (R) and closed source (C) yields a very similar result. Summarizing our results, we find:

**Observation** In the model of ex-ante investment, a knowledge-based service firm will choose open source (O) or reciprocal open source (R) rather than closed source (C) whenever the gains in social welfare from more efficient ex-ante investment size outweigh the loss in profit due to the presence of entrants.

## 2.4 Knowledge Sharing and Ex-post Collaboration

We now turn to the question how much collaboration effort the customer will provide ex-post (after signing a service contract) under each of the three different licensing regimes $\{C, O, R\}$. Whenever the service provider cannot commit to share his dynamic knowledge gains, customers need to weigh carefully the benefits versus the downsides of more collaboration (for example, in the form of information disclosure): although more collaboration increases today's service value, it may also reduce the customer benefits from tomorrow's services. This is because the increasing knowledge gap between the incumbent and the entrants enables the incumbent to extract more customer surplus in the future. We will see that a firm's commitment to freely share all dynamic knowledge gains with its rivals (as it is, for example, the case under a reciprocal open source license) can help ameliorate customer concerns about hold-up and induce stronger customer collaboration.

As we have done before, we analyze separate games $\{\tilde{\Gamma}_C, \tilde{\Gamma}_O, \tilde{\Gamma}_R\}$ for each of the three licensing options of closed (C), open (O) and reciprocal (R) open source, taking the general model that we discussed in section 2.2.3 as the foundation of our setup. Since we are interested in ex-post collaboration levels, we simplify our analysis by assuming that the size of the service project is fixed. Hence, the customer's ex-ante investment always takes some constant exogenous value $\bar{s} > 0$. This allows us to define a reduced form of the service value function

$$V(K, x^i) \equiv \mathbb{V}(K, \bar{s}, x^i) - C_s(\bar{s})$$

and to introduce a more compact notation for the learning function:

$$k_x(K, x) \equiv k(K, \bar{s}, x)$$

As before, we will assume that under closed source (C) and open source (O) the dynamic knowledge gains are not shared with competitors whereas they are fully shared under reciprocal open source (R).[7] We also make the same assumptions regarding the initial knowledge stock of the entrants as before: at the beginning of the first period of the game, knowledge stocks are set to their initial values

$$K_1^i = \begin{cases} K_1^I & \text{for the incumbent } i = 1 \\ 0 & \text{for all entrants } i \in \{2,3,\ldots,N\} \text{ in game } \tilde{\Gamma}_C \\ K_1^E & \text{for all entrants } i \in \{2,3,\ldots,N\} \text{ in games } \tilde{\Gamma}_O, \tilde{\Gamma}_R \end{cases}$$

Subsequently, we play the following game for two periods where each period comprises of the following steps:

1. Each service firm submits a bid $p_t^i$.
2. The customer selects the service firm $i \in \{1,\ldots,N\}$ that offers her the highest overall surplus, taking into account how her choice affects her surplus in future periods.
3. The customer determines her level of collaboration $x_t^i$ with this firm and incurs collaboration cost $C_x(x_t^i)$.
4. The service value materializes and knowledge stocks of all firms are adjusted according to the corresponding equation of motion.

Again, payoffs originating from the second period will be discounted with a factor of $\delta > 0$.

### 2.4.1 Benchmark: Planner's Solution

If we were to ask a social planner to choose the actions that maximize overall welfare, the planner would always choose to let the firm with the highest knowledge stock deliver the service. The two first-order conditions that pin down the socially

[7] In particular, we assume that if firm $i$ wins the first period service contract and the customer collaborates with this firm with intensity $x_1^i$, the second period knowledge stocks will be

$$K_2^j = \begin{cases} K_1^i + k_x(K_1^i, x_1^i) & \text{for firm } j = i \\ K_1^j + \lambda k_x(K_1^i, x_1^i) & \text{for all other firms } j \neq i \end{cases}$$

with $\lambda = 0$ for closed source (C) and open source (O), and $\lambda = 1$ for reciprocal open source (R).

optimal levels of collaboration in both periods, $x_1^*$ and $x_2^*$ are

$$\left.\frac{\partial V}{\partial x}\right|_{(K_1^I+k_x(K_1^I,x_1),x_2^*)} = C_x'(x_2^*) \text{ and} \quad (2.14)$$

$$\left.\frac{\partial V}{\partial x}\right|_{(K_1^I,x_1^*)} + \delta\left.\frac{\partial V}{\partial K}\right|_{(K_1^I+k_x(K_1^I,x_1^*),x_2^*)} \cdot \left.\frac{\partial k_x}{\partial x}\right|_{(K_1^I,x_1^*)} = C_x'(x_1^*) \quad (2.15)$$

The first of those two equations is of a form that we will continue to incur throughout the rest of our analysis. Thus, we simplify notation by introducing the function $\bar{x}(K)$ to refer to the *statically optimal collaboration level* that is implicitly defined by the first-order condition

$$\left.\frac{\partial V}{\partial x}\right|_{(K,\bar{x}(K))} = C_x'(\bar{x}(K)).$$

Since $\frac{\partial^2 V}{\partial x \partial K} > 0$, $\bar{x}(K)$ is strictly increasing in $K$. Using this new notation, we can write the optimality condition for $x_2^*$ as

$$x_2^* = \bar{x}(K_1^I + k_x(K_1^I, x_1))$$

Since the optimality condition for $x_1^*$ comprises of both the static term and a strictly positive dynamic term, we can also bound the socially optimal first collaboration in period one from below:

$$x_1^* > \bar{x}(K_1^I)$$

The economic interpretation of the two optimality conditions for $x_1^*$ and $x_2^*$ is straightforward: optimal collaboration in the second period is equal to the statically optimal level whereas the optimal collaboration level in the first period exceeds the statically optimal level because the planner fully accounts for the dynamic knowledge gains from collaboration.

### 2.4.2 Game $\tilde{\Gamma}_C$: Closed Source

Under closed source, the incumbent has monopoly power over the customer and thus extracts all customer surplus. However, in sharp contrast to ex-ante investment, the monopolist's surplus extraction does not completely eliminate the customer's incentive to collaborate ex post: in the second period, the customer will choose irrespectively of the already paid and sunk service price $p_2^C$ her optimal level $x_2^C$ of ex-post collaboration is simply the statically optimal level,

$$x_2^C = \bar{x}(K_1^I + k_x(K_1^I, x_1))$$

Anticipating that the customer's optimal ex-post collaboration choice will be $x_2^C$, the incumbent will set a price for the second period that leaves the customer with exactly zero surplus if she collaborates optimally:

$$p_2^C = V(K_1^I + k_x(K_1^I, x_1), x_2^C) - C_x(x_2^C)$$

The fact that the second period customer surplus is always zero irrespectively of the incumbent's knowledge stock implies that the customer's actions in the first period do not affect her second period surplus. Her first-period collaboration choice $x_1^C$ is thus also governed by the optimality condition of a static model, hence

$$x_1^C = \bar{x}(K_1^I)$$

We can see that the customer's collaboration in the first period is inefficiently low from a social point of view: this occurs because she does not participate in the benefits of the incumbent's dynamic knowledge gains. Moreover, as $x_1^C < x_1^*$, the incumbent's second period knowledge stock $K_2^I = K_1^I + k_x(K_1^I, x_1^C)$ will fall short of the socially optimal amount $K_1^I + k_x(K_1^I, x_1^*)$, which in turn implies that customer collaboration in the second period will also be inefficiently low.[8]

In summary, we find:

**Proposition 6.** *In the game $\tilde{\Gamma}_C$ of ex-post collaboration under closed source, the incumbent firm extracts all surplus. Customer collaboration is inefficiently low in both periods and does not account for dynamic knowledge gains: $x_1^C = \bar{x}(K_1^I)$ and $x_2^C = \bar{x}(K_1^I + k_x(K_1^I, x_1^C))$. Incumbent profits (and total surplus) amount to*

$$\Pi_{total}^C = V(K_1^I, x_1^C) - C_x(x_1^C) + \delta\left[V(K_1^I + k_x(K_1^I, x_1^C), x_2^C) - C_x(x_2^C)\right].$$

### 2.4.3 Game $\tilde{\Gamma}_O$: Open Source (without commitment)

As we shall see, the case in which the incumbent discloses his software as open source but does not commit to sharing any dynamic knowledge gains bears substantial similarity to the situation under closed source. In particular, the incumbent's inability (or unwillingness) to commit not to hold up his customer for knowledge that was obtained from earlier customer collaboration will leave the customer hesitant to collaborate beyond the statically optimal level in first place.

[8]Note that, from our assumptions regarding the cross-derivatives of $\mathbb{V}(K, s, x)$ follows immediately that $V(K, x)$ has the same properties, in particular $\frac{\partial^2 V}{\partial x \partial K} > 0$, which explains this result.

### If the incumbent is awarded the first period

We solve again by backward induction, starting with the second period. We will use indexes $_{2,ij}$ to denote second period quantities if firm $i$ serves in the first and firm $j$ serves in the second period. If the incumbent has run the first period project and received a customer collaboration of $x^O_{1,I}$, the incumbent's knowledge stock in period two amounts to $K^I_2 = K^I_1 + k_x(K^I_1, x^O_{1,I})$. Since the entrants' knowledge stocks remain constant, the incumbent will charge for the second period project a price of

$$p^O_{2,II} = V(K^I_1 + k_x(K^I_1, x_1), x^O_{2,II}) - C_x(x^O_{2,II}) - \left[V(K^E_1, \bar{x}(K^E_1)) - C_x(\bar{x}(K^E_1))\right].$$

The customer will choose an optimal collaboration level $x^O_{2,II}$ of

$$x^O_{2,II} = \bar{x}(K^I_1 + k_x(K^I_1, x_1))$$

and will obtain a second period surplus of

$$S^O_{2,I} = V(K^E_1, \bar{x}(K^E_1)) - C_x(\bar{x}(K^E_1)).$$

Note that the customer's period two surplus is independent of the collaboration choice $x_{1,I}$ in period one. We can thus already conclude that with the incumbent, the customer's optimal period one collaboration $x^O_{1,I}$ will only reach the statically optimal level $x^O_{1,I} = \bar{x}(K^I_1)$ and fall short of the efficient benchmark, i.e. $x^O_{1,I} < x^*_1$.

### If an entrant is awarded the first period

The problem that first period collaboration incentives are compromised due to lack of commitment on the side of the service firm would be less severe if the customer were to choose an entrant in period one: second period knowledge stocks would amount to $K^I_2 = K^I_1$ for the incumbent and $K^E_2 = K^E_1 + k_x(K^E_1, x_{1,E})$ for the entrant. Just as in the previous model of ex-ante investment, second period outcomes if the entrant has served period one depend on whether the entrant's knowledge gain from learning allows him to surpass the incumbent's knowledge stock (*"substantial learning")*. We shall again use the indexes $\{H, L\}$ to refer to the firm with the highest and second highest period two knowledge stocks, respectively: $K^H_2 \equiv \max(K^I_1, K^E_1 + k_x(K^E_1, x_{1,E}))$, and $K^L_2 \equiv \min(K^I_1, K^E_1 + k_x(K^E_1, x_{1,E}))$.

The most knowledgeable firm $H$ wins the period two contract and charges a price $p_{2,EH} = V(K^H_2, \bar{x}(K^H_2)) - C_x(\bar{x}(K^H_2)) - \left[V(K^L_2), \bar{x}(K^L_2)) - C_x(\bar{x}(K^L_2))\right]$, leaving the customer with a second period surplus of

$$S^O_{2,EH} = V(K^L_2, \bar{x}(K^L_2)) - C_x(\bar{x}(K^L_2))$$

if the customer collaborates at her optimal level $x^O_{2,EH} = \bar{x}(K^H_2)$. We can see that just like before, the second period customer surplus $S_{2,EH}$ depends on earlier collaboration $x_{1,E}$ only if learning is non-substantial. If learning is substantial, the customer nevertheless participates in the surplus gain from her first period collaboration: since the entrant will make positive profits in the second period, these profits (which do depend on the first period collaboration level $x_{1,E}$) can be pledged when the entrant attempts to win the customer in period one: hence, overall customer surplus for both periods will reflect some benefit arising from dynamic knowledge gains,

$$\begin{aligned} S^O_{E,total} &= V(K^E_1, x^O_{1,E}) - C_x(x^O_{1,E}) \\ &\quad +\delta\left[V(K^E_1 + k_x(K^E_1, x^O_{1,E}), \bar{x}(K^E_2)) - C_x(\bar{x}(K^E_2))\right] \end{aligned}$$

and the customer chooses a first period collaboration level $x^O_{1,E}$ according to the first-order condition[9]

$$\left.\frac{\partial V}{\partial x}\right|_{(K^E_1, x^O_{1,E})} + \delta \left.\frac{\partial V}{\partial K}\right|_{(K^E_1 + k_x(K^E_1, x^O_{1,E}), \bar{x}(K^E_2))} \cdot \left.\frac{\partial k_x}{\partial x}\right|_{(K^E_1, x^O_{1,E})} = C'_x(x^O_{1,E}).$$

We observe the following: not only does first period collaboration under the entrant account for dynamic knowledge gains, but collaboration might even become socially excessive: this situation can arise because marginal returns to knowledge in the second term on the left-hand side are higher if the knowledge stock is lower, which will occur whenever innovation is non-substantial. On the other hand, marginal returns to collaboration as shown in the first left-hand side term are lower (due to the lower knowledge stock of the entrant) than in the efficient benchmark. Thus, first period collaboration with the entrant may higher or lower than the efficient level, depending on which of the two effects prevails.

However, we can show that the possibility that the entrant may win the first period in equilibrium (and collaboration levels may come close to or even exceed the efficient benchmark) is of purely theoretical nature:

**Proposition 7.** *In the game $\tilde{\Gamma}_O$ of collaboration under open source (without commitment) the incumbent can always win both periods, and he always has incentive to do so. The equilibrium collaboration in both periods is the same as under closed source: $x^O_{1,I} = x^C_{1,I}$, and $x^O_{2,II} = x^C_{2,II}$. In particular, collaboration levels are inefficiently low in both periods, and $x^O_{1,I}$ does not account for any dynamic knowledge gains. Finally, social welfare is the same as in the game $\tilde{\Gamma}_C$ of closed source, but the incumbent's profits are strictly lower than under closed source.*

*Proof.* see appendix. □

[9]Note that as a consequence of the envelope theorem in this first-order condition all terms related to changes in $\bar{x}(K^E_2)$ as a function of $x^O_{1,E}$ cancel out.

### 2.4.4 Game $\tilde{\Gamma}_R$: Reciprocal Open Source

If the incumbent commits to sharing his dynamic knowledge gains by using a reciprocal open source license, customer collaboration will be substantially higher than in the previous cases – in fact, it may even exceed the efficient level in both periods.

Because no entrant can ever match the offer of the incumbent when all knowledge gains are fully shared among firms, the incumbent will always win both periods. We therefore show our results only for this case and delegate the rest to the appendix. Let's first look at the second period. If the incumbent has already served period one, knowledge stocks will be $K_2^I = K_1^I + k_x(K_1^I, x_{1,I}^R)$ for the incumbent and $K_2^E = K_1^E + k_x(K_1^I, x_{1,I}^R)$ for the entrant, which allows for a surplus extraction of

$$\begin{aligned} p_{2,II}^R &= V(K_1^I + k_x(K_1^I, x_{1,I}^R), \bar{x}(K_2^I)) - C_x(\bar{x}(K_2^I)) \\ &\quad - \left[V(K_1^E + k_x(K_1^I, x_{1,I}^R), \bar{x}(K_2^E)) - C_x(\bar{x}(K_2^E))\right] \end{aligned}$$

by the incumbent firm. The customer will thus select a collaboration level of $x_{2,II}^R = \bar{x}(K_1^I + k_x(K_1^I, x_{1,I}^R))$ in the second period, and will choose her first period collaboration level $x_{1,I}$ as to maximize her overall surplus,

$$\begin{aligned} \max_{x_{1,I}^R} \quad & V(K_1^I, x_{1,I}^R) - p_{1,I}^R - C_x(x_{1,I}^R) \\ & +\delta\left[V(K_1^E + k_x(K_1^I, x_{1,I}^R), \bar{x}(K_2^E)) - C_x(\bar{x}(K_2^E))\right]. \end{aligned}$$

Hence, her equilibrium level of collaboration $x_{1,I}^R$ with the incumbent in the first period is determined by the first-order condition

$$\left.\frac{\partial V}{\partial x}\right|_{(K_1^I, x_{1,I}^R)} + \delta \left.\frac{\partial V}{\partial K}\right|_{(K_1^E + k_x(K_1^I, x_{1,I}^R), \bar{x}(K_2^E))} \cdot \left.\frac{\partial k_x}{\partial x}\right|_{(K_1^I, x_{1,I}^R)} = C_x'(x_{1,I}^R). \quad (2.16)$$

An inspection of this first-order condition, combined with previous results, unveils the following:

**Proposition 8.** *In the game $\tilde{\Gamma}_R$ of collaboration under reciprocal open source, the incumbent always wins both periods. The equilibrium levels of customer collaboration are strictly higher in both periods than under closed source or under open source without commitment. It is undetermined how equilibrium collaboration $x_{1,I}^R$ ranks relative to the socially optimal level: it can be inefficiently low, efficient or even excessive from a social point of view. However, in the limit $K_1^E \to K_1^I$ the socially optimal levels of collaboration are attained in both periods.*

*Proof.* see appendix. □

### 2.4.5 Comparison of the three Licenses

We are now in a position to compare the three licensing options (C), (O) and (R) by adding a period zero of the game in which the incumbent selects a licensing model $l \in \{C, O, R\}$ and subsequently plays the corresponding subgame $\tilde{\Gamma}_l$.

Before we turn to the main analysis, let us first for a moment consider the incumbent's optimal licensing strategy for a highly commoditized good that is not customizable to an individual customer's need. We capture this by assuming that the delivered value is constant and independent of any collaboration effort on the customer side, i.e. $V(K, x) = \bar{V}(K)$, and that there is no learning about the customer either, i.e. $k_x(K, x) = 0$. Since generated value does not depend on collaboration and there is no learning either, all three licensing options generate the same social welfare. However, only license (C) allocates the entire surplus to the incumbent. We observe:

**Observation** In the model of ex-post collaboration, a firm that produces a non-customizable commodity good (rather than a service) will never choose any of the open-source licenses (O) or (R). The incumbent will always opt for closed source (C) since it allows for the extraction of the entire surplus.

In sharp contrast to this result, closed source (C) may no longer be optimal if the incumbent attempts to sell a knowledge-based service rather than a commodity good: let us return to our original assumption that $V(K, x)$ is strictly increasing and concave, and that there are positive learning effects $k_x(K, x)$ from customer collaboration. As we have seen, customer collaboration in both periods then remains inefficiently low under closed source (C) because the customer does not receive any of the additional surplus that is generated from dynamic knowledge gains; the situation under open source without commitment (O) is exactly the same, it only differs from closed source by the aspect that the incumbent has to surrender some surplus to the customer. Hence, in the model of ex-post collaboration, the incumbent will never choose open source without commitment (O) because it is always dominated by closed source (C).

However, under certain circumstances it can be optimal for the incumbent to choose a reciprocal open source license (R) instead of a closed source license (C): in order to see this, first note that if reciprocal open source does not induce socially excessive collaboration in period one, social welfare[10] under reciprocal open source is guaranteed to be higher than under closed source. The reason is again the same as in the case of ex-ante investment: social welfare is a concave function of customer collaboration; hence, collaboration levels that are closer to the first-best benchmark will generate a higher social welfare. We summarize our findings:

[10] We define here social welfare as the sum of service provider profits and customer surplus.

**Observation** In the model of ex-post collaboration, the incumbent will never adopt open source without commitment (O) because it is always dominated by closed source (C). However, the incumbent may find it profitable to adopt reciprocal open source (R) to boost customer collaboration. Whether or not the incumbent finds it profitable to adopt reciprocal open source depends on whether the gains in social welfare from better collaboration are sufficient to compensate for the incumbent's reduced ability to extract surplus in the presence of the newly created competition.

## 2.5 Extensions

In this section, we present two extensions to our model which explore what happens if some of the assumptions that underlie our model are relaxed.

### 2.5.1 Endogenous Choice of Shared Knowledge Stock

So far, we have treated the decision of whether to adopt open or closed source as a binary choice and have considered $K_1^E$ as exogenously given. But this falls short of recognizing that many IT services to big clients are provided by the development of an entire portfolio of software programs. The incumbent then has more fine-grained control over the amount of information he wishes to share with rival firms via open source licensing. It therefore appears more appropriate to consider $K_1^E$ an endogenous quantity.

In this section we ask how much knowledge $K_1^E \in [0, K_1^I]$ the incumbent would optimally share with the public by releasing open source software. The game $\Gamma_E$ of endogenous information choice has the following structure:

1. first, the incumbent decides on the quantity $K_1^E \in [0, K_1^I]$ of knowledge stock that is to be shared via open source licensing with all entrants
2. second, the incumbent plays the game $\Gamma_R$ of reciprocal open source[11], whereby the initial knowledge stock $K_1^E$ of entrants is set to the previously chosen value.

Remember that the incumbent's profit function can be rewritten in terms of social welfare functions, which allows us to cast the incumbent's profit maximization problem into the following form:

$$\max_{K_1^E} \Pi_{total}^{I,R}(K_1^E) = SW(K_1^I, s_1^R, s_{2,I}^R) - SW(K_1^E, s_1^R, s_{2,E}^R).$$

[11]For simplicity, we only present the case (R) of reciprocal open source, but similar results would be obtained for the case (O) of open source without commitment.

Observing that the induced ex-ante investments $s_1^R$ and $s_{2,I}^R$ in both periods approach the socially optimal values in the limit $K_1^E \to K_1^I$, we thus immediately find:

**Observation** When $K_1^E$ is chosen endogenously by the incumbent, he selects $K_1^E$ as to maximize the difference between social welfare and the outside option. The chosen entrant knowledge stock $K_1^E$ will therefore always be lower than the socially optimal level $K_1^{E*} = K_1^I$ that would induce the highest social welfare.

### 2.5.2 Forking, and Competition by Openness

Another important assumption that we have implicitly made in our model was that whenever the incumbent can not commit to share dynamic knowledge gains and chooses the license (O), all entrants will equally lack such commitment power. This assumption would always be fulfilled if the only commitment vehicle for sharing dynamic knowledge gains was the choice of a reciprocal open source license: since the right to re-license a software under more restrictive or liberal open source licensing terms lies exclusively with the copyright holder, open source essentially constrains an entrant to always deliver the software under the same open source license that was chosen by the incumbent.

However, in practice there are other ways by which an entrant can commit to share dynamic knowledge gains even if the incumbent does not impose such commitment in the software's licensing terms. Whilst an entrant can't simply apply different licensing terms to an existing open source software, nothing prevents him from creating a fork of the project (i.e., a derivative work that is based on the incumbent's original source code but which is subsequently developed along a different path than the original project) which gives him full control over the development process of the forked project. In particular, the entrant can choose to commit to a "more open" development process than the incumbent: this can, for example, be attained by making all project related communications, code repositories and decision processes public (usually under the roof of an independent non-profit organization) and by building a reputation for accepting relevant third-party contributions on a regular basis.[12]

If these measures are sufficiently credible to convince the customer that the entrant will share all future dynamic knowledge gains whereas the incumbent is unable (or unwilling) to commit to the same level of knowledge sharing, the incumbent may lose the customer to the entrant: we make a minor modification to

[12]For example, this strategy has been used with great success by the relatively young database service firm SkySQL: in the face of growing doubts about Oracle's commitment to sharing dynamic knowledge gains (especially, security patches) for its MySQL open source database product, SkySQL created a fork of Oracle MySQL called MariaDB which it keeps developing in a fully open development process under the roof of the (non-profit) MariaDB Foundation.

the game $\tilde{\Gamma}_O$ of ex-post collaboration under open source to demonstrate this effect. Let us assume that the incumbent does not share any dynamic knowledge gains whereas knowledge stocks under the entrant follow the same equation of motion as in the game $\tilde{\Gamma}_R$ of reciprocal open source. If the customer selects the entrant in period one, period two knowledge stocks will be

$$\begin{aligned} K_2^I &= K_1^I + k_x(K_1^E, x_{1,E}) \\ K_2^E &= K_1^E + k_x(K_1^E, x_{1,E}) \end{aligned}$$

which implies that the customer will choose the incumbent in period two and obtain a second period surplus of

$$S_{2,EI}^{R|O} = V(K_1^E + k_x(K_1^E, x_{1,E}), \bar{x}(K_2^E)) - C_x(\bar{x}(K_2^E)).$$

In order to win the customer in period one, the incumbent would thus need to match the overall customer surplus if the entrant was chosen in the first period,

$$\begin{aligned} S_{total,EI}^{R|O} &= V(K_1^E, x_{1,E}^R) - C(x_{1,E}^R) \\ &\quad + \delta[V(K_1^E + k_x(K_1^E, x_{1,E}^R), \bar{x}(K_2^E)) - C_x(\bar{x}(K_2^E))], \end{aligned}$$

where $x_{1,E}^R$ is the amount of period one customer collaboration with the entrant from game $\tilde{\Gamma}_R$. However, due to his lack of commitment to share dynamic knowledge gains, the incumbent only produces (if chosen for both periods) an overall social welfare of

$$\begin{aligned} SW_{total,II}^{R|O} &= V(K_1^I, \bar{x}(K_1^I)) - C_x(\bar{x}(K_1^I)) \\ &\quad + \delta[V(K_1^I + k_x(K_1^I, \bar{x}(K_1^I)), \bar{x}_2) - C_x(\bar{x}_2)]. \end{aligned}$$

where $\bar{x}_2 \equiv \bar{x}(K_1^I + k_x(K_1^I, \bar{x}(K_1^I)))$ denotes the statically optimal collaboration level in period two.

For sufficiently high entrant knowledge stock $K_1^E$, the total social welfare that the incumbent can produce will be less than the total surplus that the customer obtains if she chooses the entrant in period one: to see this, observe that as $K_1^E \to K_1^I$, the first period collaboration level with the entrant $x_{1,E}^R$ approaches the socially optimal one. Second period customer surplus $S_{2,EI}^{R|O}$ then converges towards the second period social welfare in the first-best benchmark. Thus, eventually the overall customer surplus $S_{total,EI}^{R|O}$ must be strictly higher than the social welfare $SW_{I,total}^{R|O}$ that can be produced by the incumbent whose period one collaboration is inefficiently low due to lack of commitment.

Therefore, we conclude that for knowledge-based service firms, the commitment to share dynamic knowledge gains (for example, by forking a software project and making it more "open") can constitute a powerful weapon for market entry, and the incumbent firm may end up having to make the same commitment in order not to lose the market to the entrant.

## 2.6 Concluding remarks

We have presented a simple dynamic model of a knowledge-based service industry that focuses on customer participation in the service transaction. We studied two related models: In the model of ex-ante investment size, we found that knowledge sharing through open source and market sharing is a strategy that the dominant firm can employ to boost customer investment. In the model of ex-post customer collaboration, we found that open source will only boost customer collaboration and increase welfare if the service firm can commit to share dynamic knowledge gains, for example by adopting a reciprocal open source license.

Moreover, we observe that an entrant's commitment to share dynamic knowledge gains by opening up the development process of an open source product can constitute an aggressive entry strategy. The dominant firm may then embrace reciprocal open source for two reasons: either to boost the collaboration level or in order not to lose the competition.

Although our discussion has been concentrated on IT service industry, the line of argumentation presented in our model is far more general and not limited to software industry in any way. Specifically, we think that our theory can, by its focus on knowledge accumulation in a repeated service relationship, also be used to analyze the competitive dynamics between top management consulting firms such as McKinsey and Boston Consulting Group.

We deliberately focused on the case of a single buyer. In the future, it would be interesting to extend our analysis to the case of multiple buyers when knowledge is transferable across different buyers. This would raise coordination issues among the buyers. More generally, our paper is a first step toward to the IO of a knowledge-based service industry and more studies are needed to characterize the characteristics of this industry with respect to the traditional manufacturing industry.

## 2.A Appendix

### 2.A.1 Comparing Second Period Customer Surplus

To show that in the game $\Gamma_O$ the customer has always higher period two surplus if se chooses the entrant in period one, we first prove the following lemma:

**Lemma 9.** *Let $W(K) \equiv \mathcal{V}(K, s_K) - C_s(s_K)$ be the function of social welfare in a static one-period game where investment $s_K(K)$ is determined by the static first-order condition $C_s'(s_K) = \frac{\partial V}{\partial K}\big|_{(K,s_K)}$. Then, $W(K)$ is monotonically increasing in $K$, i.e. $\frac{\mathrm{d}}{\mathrm{d}K}W(K) > 0$.*

*Proof.* $\frac{\mathrm{d}}{\mathrm{d}K}W(K) = \frac{\partial V}{\partial K}\big|_{(K,s_K)} + \left[\frac{\partial V}{\partial s}\big|_{(K,s_K)} - C_s'(s_K)\right] \cdot \frac{\mathrm{d}s_K}{\mathrm{d}K} = \frac{\partial V}{\partial s}\big|_{(K,s_K)}$. □

This simple result, which is a direct consequence of the envelope theorem, enables us to now compare the customer's second period surplus if she has chosen the entrant in the first period to the surplus that accrues to her if she chooses the incumbent:

$$\begin{aligned}\Delta S &= S_{2,E} - S_{2,I} \\ &= \mathcal{V}(K_2^L, s_{2,E}^O) - C_s(s_{2,E}^O) - \left[\mathcal{V}(K_1^E, s_{2,I}^O) - C_s(s_{2,I}^O)\right] \\ &= W(K_2^L) - W(K_1^E)\end{aligned}$$

Remembering the definition of $K_2^L$, we have that $K_1^L > K_1^E$, and thus the difference in customer surplus is positive.

### 2.A.2 Proof of proposition 2:

The proposition follows directly from inspecting the first-order equation (2.13) and comparing it to the efficient benchmark: First, note that the induced investment $s_1^O$ is strictly positive and thus clearly larger than $s_1^C$. Compared to the first order equation (2.9) which determines the efficient benchmark level $s_1^*$, we see that the first term in equation (2.13), $\frac{\partial \mathcal{V}}{\partial s}\big|_{(K_1^E, s_1^O)}$, is smaller than its counterpart in the efficient benchmark because $K_1^E < K_1^I$, and $\frac{\partial^2 \mathcal{V}}{\partial K \partial s} > 0$. Hence, for sufficiently small $\delta$ investment will be inefficiently low. The second term (i.e., the expression within the square bracket) is ambiguous: on the one hand, we had assumed diminishing marginal returns to knowledge which means that marginal returns to knowledge stock are higher for the entrant than the for the incumbent, and thus $\frac{\partial \mathcal{V}}{\partial K}\big|_{(K_1^E + k_s(K_1^E, s_1^O), s_{2,E}^O)}$ can be larger with open source than under the benchmark. On the other hand, knowledge stock $K_1^E + k_s(K_1^E, s_1^O)$ under the entrant is lower than the efficient level, and so is $s_{2,E}^O$. This diminishes both of the two factors within the square bracket relative to the efficient benchmark. Depending on which effect prevails, $s_1^O$ will be inefficiently low, efficient or even excessive.

### 2.A.3 Full results for the game $\Gamma_R$:

In this section, we provide all results for the game of ex-ante investment under reciprocal open source that were not shown in the main text.

**Second period if an entrant has served the first period:**

We solve backwards and start our analysis by assuming that an entrant has served the first period when the customer's ex-ante investment was $s_1$. The resulting knowledge stocks in period two are

$$K_2^I = K_1^I + k_s(K_1^E, s_1)$$
$$K_2^E = K_1^E + k_s(K_1^E, s_1)$$

Clearly, the incumbent has a higher second period knowledge stock than the entrant and will thus win period two. The entrant can at most offer a customer surplus of $\mathcal{V}(K_2^E, s_{2,E}^R) - C_s(s_{2,E}^R)$ which means that the incumbent charges a price of

$$p_{2,E}^R = \mathcal{V}(K_1^I + k_s(K_1^E, s_1), s_{2,E}^R) - \mathcal{V}(K_1^E + k_s(K_1^E, s_1), s_{2,E}^R)$$

and leaves the customer with a period two surplus of

$$S_{2,E}^R = V(K_1^E + k_s(K_1^E, s_1), s_{2,E}^R) - C_s(s_{2,E}^R).$$

Hence, if an entrant has served the first period the customer will optimally choose an ex-ante investment $s_{2,E}^R$ in period two that satisfies the first-order condition

$$\left.\frac{\partial V}{\partial s}\right|_{(K_1^E + k_s(K_1^E, s_1), s_{2,E}^R)} = C_s'(s_{2,E}^R).$$

**Second period if the incumbent has served the first period:**

If, on the other hand, the incumbent has served the first period (which will always happen in equilibrium), second period knowledge stocks are higher than if the client had chosen the entrant in period one:

$$K_2^I = K_1^I + k_s(K_1^I, s_1)$$
$$K_2^E = K_1^E + k_s(K_1^I, s_1)$$

Thus, the incumbent charges a price of

$$p_{2,I}^R = \mathcal{V}(K_1^I + k_s(K_1^I, s_1), s_{2,I}^R) - \mathcal{V}(K_1^E + k_s(K_1^I, s_1), s_{2,I}^R)$$

and leaves the customer with a period two surplus of

$$S_{2,I}^R = \mathcal{V}(K_1^E + k_s(K_1^I, s_1), s_{2,I}^R) - C_s(s_{2,I}^R)$$

which is strictly greater than the surplus the incumbent could have offered in period two if the entrant had run the first period project. As a consequence, there will be higher ex-ante investment in period two: the customer will choose an amount $s_{2,I}^R$ that is determined by the first-order condition

$$\left.\frac{\partial \mathcal{V}}{\partial s}\right|_{(K_1^E + k_s(K_1^I, s_1), s_{2,I}^R)} = C_s'(s_{2,I}^R).$$

**First Period:**

In period one, the incumbent can charge a price that extracts the difference to the outside option in the same period plus the additional surplus which the customer receives in the second period if she selects the incumbent in period one: If the customer chooses the incumbent in period one, her overall surplus is

$$S_{total,I}^R = \mathcal{V}(K_1^I, s_1) - p_{1,I} - C_s(s_1) + \delta S_{2,I}^R$$

whereas an entrant could offer at most to provide the first round service for free, resulting in a customer surplus of

$$S_{total,E}^R = \mathcal{V}(K_1^E, s_1) - C_s(s_1) + \delta S_{2,E}^R.$$

Thus, the incumbent can charge a first period price of

$$p_{1,I}^R = \mathcal{V}(K_1^I, s_1) - \mathcal{V}(K_1^E, s_1) + \delta \left[S_{2,I}^R - S_{2,E}^R\right]$$

and leave the customer with an overall surplus of

$$S_{total,I}^R = \mathcal{V}(K_1^E, s_1) - C_s(s_1) + \delta \left[V(K_1^E + k_s(K_1^E, s_1), s_{2,E}^R) - C_s(s_{2,E}^R)\right].$$

which implies that optimal first period ex-ante investment $s_1^R$ will satisfy the first-order condition

$$\left.\frac{\partial \mathcal{V}}{\partial s}\right|_{(K_1^E, s_1^R)} + \delta \left.\frac{\partial \mathcal{V}}{\partial K}\right|_{(K_1^E + k_s(K_1^E, s_1^R), s_{2,I}^R)} \cdot \left.\frac{\partial k_s}{\partial s}\right|_{(K_1^E, s_1^R)} = C_s'(s_1^R).$$

The incumbent obtains overall profits of

$$\begin{aligned}
\Pi_{total}^{I,R} &= \mathcal{V}(K_1^I, s_1^R) - \mathcal{V}(K_1^E, s_1^R) \\
&\quad + \delta \left[\mathcal{V}(K_1^I + k_s(K_1^I, s_1^R), s_{2,I}^R) - V(K_1^E + k_s(K_1^E, s_1^R), s_{2,E}^R)\right] \\
&\quad + \delta \left[C_s(s_{2,E}^R) - C_s(s_{2,I}^R)\right].
\end{aligned}$$

### 2.A.4 Proof of proposition 7:

We will only show that the incumbent will never choose to yield the first period to an entrant because the fact that first period collaboration with the incumbent $x^O_{1,I}$ remains at the inefficiently low static level was already shown in the main text. For the case in which learning is substantial, it is immediate that the incumbent will never yield the first period. We thus restrict ourselves to analyzing the case in which learning is non-substantial.

If the incumbent yields the first period to an entrant, the incumbent's overall profits are

$$\begin{aligned} \Pi^O_{total,p2} &= \delta \left[ V(K^I_1, \bar{x}(K^I_1)) - C_x(\bar{x}(K^I_1)) \right] \\ &\quad -\delta \left[ V(K^E_1 + k_x(K^E_1, x^O_{1,E}), \bar{x}(K^E_2)) - C_x(\bar{x}(K^E_2)) \right] \end{aligned}$$

If, on the other hand, the incumbent serves the first period, his profits are

$$\begin{aligned} \Pi^O_{total,p1\,p1} &= p^I_1 + \delta \left[ V(K^I_1 + k_x(K^I_1, x_1), x^O_{2,II}) - C_x(x^O_{2,II}) \right] \\ &\quad -\delta \left[ V(K^E_1, \bar{x}(K^E_1)) - C_x(\bar{x}(K^E_1)) \right] \end{aligned}$$

so the incumbent will be indifferent between winning or losing the first period if he can win the first period at a price of

$$\begin{aligned} p^I_{1,indiff} &= \delta \left[ V(K^I_1, \bar{x}(K^I_1)) - C_x(\bar{x}(K^I_1)) \right] \\ &\quad -\delta \left[ V(K^E_1 + k_x(K^E_1, x^O_{1,E}), \bar{x}(K^E_2)) - C_x(\bar{x}(K^E_2)) \right] \\ &\quad -\delta \left[ V(K^I_1 + k_x(K^I_1, x_1), x^O_{2,II}) - C_x(x^O_{2,II}) \right] \\ &\quad +\delta \left[ V(K^E_1, \bar{x}(K^E_1)) - C_x(\bar{x}(K^E_1)) \right] \end{aligned}$$

Note that the sum of the first and the third line gives a negative overall contribution to $p^I_{1,indiff}$. Hence, we know that a price $\tilde{p}^I_1$ that results from elimination of these two lines is strictly higher and the incumbent thus prefers to win the first period at this price,

$$\begin{aligned} \tilde{p}^I_1 &= -\delta \left[ V(K^E_1 + k_x(K^E_1, x^O_{1,E}), \bar{x}(K^E_2)) - C_x(\bar{x}(K^E_2)) \right] \\ &\quad +\delta \left[ V(K^E_1, \bar{x}(K^E_1)) - C_x(\bar{x}(K^E_1)) \right] \end{aligned}$$

At this price, the customer would enjoy a total surplus of

$$\begin{aligned} S^I_{tot} &= V(K^I_1, x^O_{1,I}) - C_x(x^O_{1,I}) \\ &\quad +\delta \left[ V(K^E_1 + k_x(K^E_1, x^O_{1,E}), \bar{x}(K^E_2)) - C_x(\bar{x}(K^E_2)) \right] \end{aligned}$$

if she chooses the incumbent. Since the entrant never wins the second period, he can offer at most a first period price of zero, which would leave a customer surplus

$$\begin{aligned} S^E_{tot} &= V(K^E_1, x^O_{1,E}) - C_x(x^O_{1,E}) \\ &\quad +\delta \left[ V(K^E_1 + k_x(K^E_1, x^O_{1,E}), \bar{x}(K^E_2)) - C_x(\bar{x}(K^E_2)) \right]. \end{aligned}$$

Since $V(K_1^I, x_{1,I}^O) - C_x(x_{1,I}^O)$ is always greater than $V(K_1^E, x_{1,E}^O) - C_x(x_{1,E}^O)$, it follows that the customer will never choose the entrant at this price $\tilde{p}_1^I$. Since this price is already higher than the lowest price that satisfies the incumbent's participation constraint, it is clear that the incumbent will always choose to win the first period. □

### 2.A.5 Proof of proposition 8:

Observe that, with the exception of the second term on the left-hand side, the first-order condition (2.16) is exactly the same as the first-order condition for the socially optimal level of collaboration, eq. (2.15). The second term deviates from the first-best condition only by the partial derivative $\frac{\partial V}{\partial K}$ which is evaluated for different values of $K$ and $x$: the argument in $K$ is lower than in the benchmark (it is $K = K_1^E + k(K_1^I, x_1)$ instead of $K = K_1^I + k_x(K_1^I, x_1)$), which, keeping $x$ constant, would imply that the term involving $\frac{\partial V}{\partial K}$ should be greater than its benchmark counterpart since we had assumed that $\frac{\partial V}{\partial K} < 0$. However, $x$ is not the same as in the benchmark either ($\frac{\partial V}{\partial K}$ is evaluated $x = \bar{x}(K_2^E)$ instead of the larger $x = \bar{x}(K_2^I)$). This effect thus goes in the opposite direction and, absent any change in $K$, would diminish $\frac{\partial V}{\partial K}$ relative to the benchmark. Depending on which of the two effects prevails, the first period collaboration will be excessive, efficient or inefficiently low.

Finally, if period one collaboration is excessive, so is the collaboration level in period two: the excessive collaboration fosters more learning than under the first-best, i.e. $k_x(K_1^I, x_{1,I}^R) > k_x(K_1^I, x_1^*)$. Hence, the incumbent's second period knowledge stock is greater than in the planner's benchmark. Since we had assumed that $\frac{\partial^2 V}{\partial K \partial x} > 0$, the higher knowledge stock then induces a higher (and thus socially excessive) level of collaboration in the second period. The same argument can also be applied to see that inefficiently low collaboration in period one will trigger inefficiently low collaboration in period two as well. □

### 2.A.6 Game $\tilde{\Gamma}_R$: complete derivation

In this section we present the complete discussion of game $\tilde{\Gamma}_R$. In order to know the price which the incumbent will charge in period one, we first need to find the customer surplus that the entrant can offer in both periods:

**If the entrant was awarded period one:**

Assume the entrant was awarded the first period, and the customer has collaborated with the entrant with intensity $x_1$. Then, the knowledge stocks at the beginning of period two are

$$
\begin{aligned}
K_{2,E}^{I} &= K_1^I + k_x(K_1^E, x_{1,E}^R), \\
K_{2,E}^{E} &= K_1^E + k_x(K_1^E, x_{1,E}^R).
\end{aligned}
$$

We immediately see that the incumbent will have a higher knowledge stock and win the second period, producing a service value of $V(K_{2,E}^I, \bar{x}(K_{2,E}^I))$ and charging a price of

$$
\begin{aligned}
p_{2,EI}^R &= V(K_{2,E}^I, \bar{x}(K_{2,E}^I)) - C_x(\bar{x}(K_{2,E}^I)) \\
&\quad - \left[V(K_{2,E}^E, \bar{x}(K_{2,E}^E)) - C_x(\bar{x}(K_{2,E}^E))\right]
\end{aligned}
$$

which leaves the customer with a second period surplus of

$$S_{2,EI}^R = V(K_{2,E}^E, \bar{x}(K_{2,E}^E)) - C_x(\bar{x}(K_{2,E}^E)).$$

Then, the total customer surplus in both periods if the entrant is chosen for the first period is

$$
\begin{aligned}
S_{total,1E}^R &= V(K_1^E, x_{1,E}^R) - C_x(x_{1,E}^R) - p_{1,E}^R \\
&\quad +\delta\left[V(K_1^E + k_x(K_1^E, x_{1,E}^R), \bar{x}(K_{2,E}^E)) - C_x(\bar{x}(K_{2,E}^E))\right]
\end{aligned}
$$

which implies that the first-order condition that fixes equilibrium customer collaboration $x_{1,E}^R$ in period one must be

$$\left.\frac{\partial V}{\partial x}\right|_{(K_1^E, x_{1,E}^R)} + \delta \left.\frac{\partial V}{\partial K}\right|_{(K_1^E + k_x(K_1^E, x_{1,E}^R), \bar{x}(K_{2,E}^E))} \cdot \left.\frac{\partial k_x}{\partial x}\right|_{(K_1^E, x_{1,E}^R)} = C_x'(x_{1,E}^R).$$

In other words, just like under open source without commitment, customer collaboration with the entrant can be excessive or inefficiently low. In any case, since the entrant can not win period two, the best offer he can make is a price of zero, so customer surplus if she goes with the entrant in period one is equal to

$$
\begin{aligned}
S_{total,E}^R &= V(K_1^E, x_{1,E}^R) - C_x(x_{1,E}^R) \\
&\quad +\delta\left[V(K_1^E + k_x(K_1^E, x_{1,E}^R), \bar{x}(K_{2,E}^E)) - C_x(\bar{x}(K_{2,E}^E))\right].
\end{aligned}
$$

In particular, note that the second period surplus is lower if the entrant has run the first period than if the incumbent has done so; this arises because the entrant realizes weaker learning effects due to his lower knowledge stock. The incumbent will fully extract this additional second period surplus that he generates in case he wins the first period by charging a proportionately higher price in period one:

**If the incumbent was awarded the first period:**

As we have discussed in the main text, if the customer chooses the incumbent in period one, knowledge stocks evolve as

$$\begin{aligned} K_{2,I}^I &= K_1^I + k_x(K_1^I, x_{1,E}^R), \\ K_{2,I}^E &= K_1^E + k_x(K_1^I, x_{1,E}^R). \end{aligned}$$

and the customer will obtain a second period surplus of

$$S_{2,II}^R = V(K_1^E + k_x(K_1^I, x_{1,I}^R), \bar{x}(K_{2,I}^E)) - C_x(\bar{x}(K_{2,I}^E)),$$

and her overall surplus as a function of the price $p_{1,I}^R$ charged by the incumbent in period one will be given by

$$\begin{aligned} S_{total,I}^R &= V(K_1^I, x_{1,I}^R) - C_x(x_{1,I}^R) - p_{1,I}^R \\ &\quad + \delta \left[ V(K_1^E + k_x(K_1^I, x_{1,I}^R), \bar{x}(K_{2,I}^E)) - C_x(\bar{x}(K_{2,I}^E)) \right]. \end{aligned}$$

The incumbent will charge a price that gives the customer exactly the same surplus as if she chooses the entrant in period one so that $S_{total,I}^R = S_{total,E}^R$. Solving for $p_{1,I}^R$, we obtain

$$p_{1,I}^R = V(K_1^I, x_{1,I}^R) - C_x(x_{1,I}^R) - \left[ V(K_1^E, x_{1,E}^R) - C_x(x_{1,E}^R) \right] + \delta \left[ S_{2,II}^R - S_{2,EI}^R \right].$$

As mentioned before, one can see that the higher second period surplus under the incumbent is fully extracted via the first period price. The customer is left with an overall surplus of

$$\begin{aligned} S_{total,I}^R &= V(K_1^E, x_{1,E}^R) - C_x(x_{1,E}^R) \\ &\quad + \delta \left[ V(K_1^E + k_x(K_1^E, x_{1,E}^R), \bar{x}(K_{2,E}^E)) - C_x(\bar{x}(K_{2,E}^E)) \right] \end{aligned}$$

whereas the incumbent realizes a total profit of

$$\begin{aligned} \Pi_{total,I}^R &= V(K_1^I, x_{1,I}^R) - C_x(x_{1,I}^R) - \left[ V(K_1^E, x_{1,E}^R) - C_x(x_{1,E}^R) \right] \\ &\quad + \delta \left[ V(K_1^I + k_x(K_1^I, x_{1,I}^R)), \bar{x}(K_{2,I}^I)) - C_x(\bar{x}(K_{2,I}^I)) \right] \\ &\quad - \delta \left[ V(K_1^E + k_x(K_1^E, x_{1,E}^R), \bar{x}(K_{2,E}^E)) - C_x(\bar{x}(K_{2,E}^E)) \right] \end{aligned}$$

# Bibliography


Abdul Abiad, Enrica Detragiache, and Thierry Tressel. A new database of financial reforms. *IMF Staff Papers*, 57(2):281–302, June 2010.

Heski Bar-Isaac. Something to prove: Reputation in team. *Rand Journal of Economics*, 38(2):495–511, 2007.

Guglielmo Barone, Roberto Felici, and Marcello Pagnini. Switching cost in local credit markets. *International Journal of Industrial Organization*, 29:694–707, 2011.

Allen N. Berger and Gregory F. Udell. The institutional memory hypothesis and the procyclicality of bank lending behavior. *Journal of Financial Intermediation*, 13(4):458 – 495, 2004.

Marcello Bofondi and Giorgio Gobbi. Informational barriers to entry into credit markets. *Review of Finance*, 10(1):39–67, 2006.

Claudio Borio and Philip Lowe. Asset prices, financial and monetary stability: exploring the nexus. BIS Working Papers 114, Bank for International Settlements, July 2002.

Thorsten Broecker. Credit-worthiness tests and interbank competition. *Econometrica*, 58(2):429–452, 1990.

Ramon Casadesus-Masanell and Pankaj Ghemawat. Dynamic Mixed Duopoly: A Model Motivated by Linux vs. Windows. *Management Science*, 52(7):1072–1084, 2006.

Jay Pil Choi. Brand extension as informational leverage. *The Review of Economic Studies*, 65(4):655–669, 1998.

S. Claessens and L. Laeven. What drives bank competition? some international evidence. *Journal of Money, Credit and Banking*, 36(3):563–583, 2004.

Stefano Comino and Fabio M. Manenti. Dual licensing in open source software markets. Department of Economics Working Papers 0718, Department of Economics, University of Trento, Italia, 2007.

David de Meza and David Webb. Too much investment: A problem of asymmetric information. *The Quarterly Journal of Economics*, 102(2):281–292, 1987.

Giovanni Dell'Ariccia and Robert Marquez. Lending booms and lending standards. *The Journal of Finance*, 61(5):2511–2546, 2006.

Giovanni Dell'Ariccia, Ezra Friedman, and Robert Marquez. Adverse selection as a barrier to entry in the banking industry. *RAND Journal of Economics*, 30 (3):515–534, 1999.

Giovanni Dell'Ariccia, Deniz Igan, and Luc Laeven. Credit booms and lending standards: Evidence from the subprime mortgage market. IMF Working Paper WP/08/106, International Monetary Fund, 2008.

Alexis Direr. Multiple equilibria in markets with screening. *Journal of Money, Credit and Banking*, 49(4):791–798, 2008.

Nicholas Economides and Evangelos Katsamakas. Two-Sided Competition of Proprietary vs. Open Source Technology Platforms and the Implications for the Software Industry. *Management Science*, 52(7):1057–1071, 2006.

Barry Eichengreen and Kris Mitchener. The great depression as a credit boom gone wrong. BIS Working Papers 137, Bank for International Settlements, 2003.

Joseph Farrell and Nancy T. Gallini. Second-sourcing as a commitment: Monopoly incentives to attract competition. *Quarterly Journal of Economics*, 103(4):673–694, 1988.

J.A. Fitzsimmons and M.J. Fitzsimmons. *Service Management: Operations, Strategy and Information Technology*. McGraw Hill, New York, NY, 2001.

Juan-José Ganuza and José S. Penalva. Signal orderings based on dispersion and the supply of private information in auctions. *Econometrica*, 78(3):1007–1030, 2010.

Giorgio Gobbi and Francesca Lotti. Entry decisions and adverse selection: An empirical analysis of local credit markets. *Journal of Financial Services Research*, 26(3):225–244, 2004.

Sanford J Grossman and Oliver D Hart. The costs and benefits of ownership: A theory of vertical and lateral integration. *Journal of Political Economy*, 94(4): 691–719, 1986.

Oliver D Hart and John Moore. Incomplete contracts and renegotiation. *Econometrica*, 56(4):755–85, July 1988.

Robert Hauswald and Robert Marquez. Information technology and financial services competition. *Review of Financial Studies*, 16(3):921–948, July 2003.

Robert Hauswald and Robert Marquez. Competition and strategic information acquisition in credit markets. *The Review of Financial Studies*, 19(3):967–1000, 2006.

Michael Hume and Andrew Sentance. The global credit boom: Challenges for macroeconomics and policy. *Journal of International Money and Finance*, 28 (8):1426–1461, 2009.

Vasso Ioannidou, Steven Ongena, and José-Luis Peydró. Monetary policy, risk-taking and pricing: Evidence from a quasi-natural experiment. Discussion Paper 04S, European Banking Center, 2009.

Gabriel Jiménez, Steven Ongena, José-Luis Peydró, and Jesús Saurina Salas. Hazardous times for monetary policy: What do twenty-three million bank loans say about the effects of monetary policy on credit risk-taking? Working Paper Series 0833, Banco de Espana, 2009.

David Kreps. Corporate culture and economic theory. In *Perspectives on Positive Political Economy*. Cambridge University Press, 1990.

Greg Kroah-Hartman, Jonathan Corbet, and Amada McPherson. Linux kernel development: How fast it is going, who is doing it, what they are doing, and who is sponsoring it. Technical report, The Linux Foundation, April 2008. URL `http://www.linux-foundation.org/publications/linuxkerneldevelopment.php`.

Jennifer W. Kuan. Open Source Software as Consumer Integration Into Production. *SSRN eLibrary*, 2001.

Josh Lerner and Jean Tirole. Some simple economics of open source. *Journal of Industrial Economics*, 50(2):197–234, 2002.

Josh Lerner and Jean Tirole. The economics of technology sharing: Open source and beyond. *Journal of Economic Perspectives*, 19:99–120, 2005.

Guido Lorenzoni. Inefficient credit booms. *Review of Economic Studies*, 75(3): 809–833, 2008.

Michael Manove, A. Jorge Padilla, and Marco Pagano. Collateral versus project screening: a model of lazy banks. *RAND Journal of Economics*, 32(4):726–744, 2001.

Robert Marquez. Competition, adverse selection, and information dispersion in the banking industry. *Review of Financial Studies*, 15(3):901–926, 2002.

Enrique G. Mendoza and Marco E. Terrones. An anatomy of credit booms: Evidence from macro aggregates and micro data. Working Papers 14049, National Bureau of Economic Research, 2008.

Enrique G. Mendoza and Marco E. Terrones. An anatomy of credit booms and their demise. Working Paper Series 18379, National Bureau of Economic Research, 2012.

Loretta Mester, Leonard Nakamura, and Micheline Renault. Transaction accounts and loan monitoring. *Review of Financial Studies*, 20(3):477–493, 2007.

Alejandro Micco and Ugo Panizza. Bank concentration and credit volatility. Working Papers Central Bank of Chile 342, Central Bank of Chile, December 2005.

Mikko Mustonen. When does a firm support substitute open source programming? *Journal of Economics & Management Strategy*, 14(1):121–139, 2005.

Yoshiaki Ogura. Learning from a rival bank and lending boom. *Journal of Financial Intermediation*, 15(4):535–555, 2006.

Mitchell A. Petersen and Raghuram G. Rajan. The effect of credit market competition on lending relationships. *The Quarterly Journal of Economics*, 110(2): 407–443, 1995.

Raghuram G. Rajan. Why bank credit policies fluctuate: A theory and some evidence. *The Quarterly Journal of Economics*, 109(2):399–441, 1994.

Eric Raymond. *The Cathedral and the Bazaar*. O'Reilly Media, 1st edition, 1999. ISBN-13: 978-1565927247.

Michael H. Riordan. Competition and bank performance: A theoretical perspective. In C. Mayer and X. Vives, editors, *Capital Markets and Financial Intermediation*. Cambridge University Press, 1993.

Martin Ruckes. Bank competition and credit standards. *The Review of Financial Studies*, 17(4):1073–1102, 2004.

Moritz Schularick and Alan Taylor. Credit booms gone bust: Monetary policy, leverage cycles and financial crises, 1870-2008. Working Paper Series 15512, National Bureau of Economic Research, 2009.

Carl Shapiro. Premiums for high quality products as returns to reputations. *The Quarterly Journal of Economics*, 98(4):659–680, 1983.

Andrea Shepard. Licensing to enhance demand for new technologies. *RAND Journal of Economics*, 18(3):360–368, 1987.

Steven Tadelis. The market for reputations as an incentive mechanism. *Journal of Political Economy*, 110(4):854–882, 2002.

John Taylor. Housing and monetary policy. Working Paper Series 14631, National Bureau of Economic Research, 2007.

Aaron Tornell and Frank Westermann. Boom-bust cycles in middle income countries: Facts and explanation. *IMF Staff Papers*, 49(Special i):111–155, 2002.

Eric von Hippel and Georg von Krogh. Open source software and the "private-collective" innovation model: Issues for organization science. *Organization Science*, 14(2):209–223, 2003.

Ernst-Ludwig von Thadden. Asymmetric information, bank lending and implicit contracts: the winner's curse. *Finance Research Letters*, 1(1):11–23, 2004.

V.A. Zeithaml, A. Parasuraman, and L. L. Berry. Problems and strategies in services marketing. *Journal of Marketing*, 49:33–36, 1985.