Abstract

We develop a theory of loan sales based on bank competition and entrepreneur directed search. We show how the interplay of the two can reduce interest rate on loans financed through on-balance-sheet activities and how this can motivate loan sales as a strategy of financing through off-balance-sheet activities. The results shed some light on the shift over the years preceding the recent financial crisis in the practice of U.S. and European banks, from the traditional ‘originate to hold’ model of credit provision, towards the ‘originate to distribute’ approach for credit extension, that is, the emergence of a ‘shadow banking system’.
Bank Competition, Directed Search, and Loan Sales*

Kevin X.D. Huang
Department of Economics
Vanderbilt University
(kevin.huang@vanderbilt.edu)

Zhe Li
School of Economics
Shanghai University of Finance and Economics
(li.zhe@mail.shufe.edu.cn)

Jianfei Sun
Antai College of Economics and Management
Shanghai Jiao Tong University
(sunjianfei@sjtu.edu.cn)

November 19, 2017

Abstract

We develop a theory of loan sales based on bank competition and entrepreneur directed search. We show how the interplay of the two can reduce interest rate on loans financed through on-balance-sheet activities and how this can motivate loan sales as a strategy of financing through off-balance-sheet activities. The results shed some light on the shift over the years preceding the recent financial crisis in the practice of U.S. and European banks, from the traditional ‘originate to hold’ model of credit provision, towards the ‘originate to distribute’ approach for credit extension, that is, the emergence of a ‘shadow banking system’.

JEL codes: E44

Keywords: Bank competition; Directed search; Loan sales; Shadow banking

*For helpful comments and suggestions, we are grateful to Gabriele Camera, Scott Davis, Miquel Faig, Gregory Huffman, Boyan Javonovic, Charles Kahn, Nobuhiro Kiyotaki, Christoffer Koch, Albert Marcet, Jia Pan, Hao Shi, Shouyong Shi, Cheng Wang, Yi Wen, Danyang Xie, Min Zhang, Xiaodong Zhu, audiences at China Meeting of the Econometric Society, CES North America Conference, Midwest Macroeconomics Meetings, North American Summer Meeting of the Econometric Society, Shanghai Macroeconomics Workshop, Tsinghua Workshop in Macroeconomics, World Congress of the International Economic Association, and seminar participants at Bellaterra Macro Seminar UAB/IAE, Federal Reserve Bank of Dallas, Fudan University, Peking University, Shanghai International Finance Summer School, and Shanghai Jiaotong University. Financial support from the Grey Fund at Vanderbilt University, and from the Shanghai Pujiang Program are gratefully acknowledged.
1 Introduction

The years preceding the recent financial crisis had observed a shift in the practice of U.S. banks from the traditional ‘originate to hold’ model of credit provision, where banks used deposits to fund loans that were kept on their balance sheets until maturity, towards an ‘originate to distribute’ approach for credit extension, under which banks sold their originated loans either in whole or in part rather than funding them with deposit liabilities.\footnote{The secondary market (where loans are sold after origination) for direct sales of individual loans grew from a mere $8 billion in 1991, to $154.8 billion in 2004, $176 billion in 2005, $238.6 billion in 2006, and further to $342 billion in 2007. The syndicated loan market (where loans are sold at origination) rose from $339 billion in 1988 to $2.2 trillion in 2007. The growth of the market for securitization of pooled loans had also been spectacular in the years leading up to the recent financial crisis. These facts are documented by Lucas et al. (2006), Drucker and Puri (2009), Ahn (2010), and Bord and Santos (2012), among others. See, also, Duffie (2007) and Loutskina and Strahan (2009) for related facts.}

A similar shift had also occurred in the European banking system (e.g., ECB 2008).

Following the recent financial crisis there has been a surge of interest in understanding the incentives of banks using the OTD model. This paper contributes to this endeavor by demonstrating one motivation of loan sales as arising from bank competition. The fact is that those years preceding the recent crisis had also witnessed increased bank competition, as technological advancement, deregulation, and globalization weakened geographic and product boundaries, which encouraged interstate\footnote{See, for example, the Senior Loan Officer Opinion Survey on Bank Lending Practices (1997-2006), conducted quarterly by the Board of Governors of the U.S. Federal Reserve System. Almost all surveyed domestic and foreign respondents (from investment and commercial banks, as well as other financial intermediaries) cited more aggressive competition from other banks or non-bank institutions as the most important factor affecting their business practices. See, also, Boot and Schmeits (2006), Ahn (2010), Ahn and Breton (2010), and Hakenes and Schnabel (2010).} cross-border banking and an ‘all-finance’ practice that opened various sources of interbank competition as well as competition from non-bank institutions and capital market, and the increased competitive pressure reflected a permanent shift in the financial system.\footnote{See, for example, DeYoung (2007) for a survey of the related literature.} Our theory suggests that the observed surge in loan sales over those years could in part represent an equilibrium response to the increase in bank competition. This is consistent with empirical evidence based on micro-level data that banks facing more intense competition are more likely to sell loans.\footnote{See, for example, Pennacchi (1988), Duffee and Zhou (2001), and Calomiris and Mason (2004).}

This paper complements the existing literature of loan sales that appeals to bank capital requirement regulation as a motivation for off-balance-sheet activities.\footnote{See, for example, Pennacchi (1988), Duffee and Zhou (2001), and Calomiris and Mason (2004).} In fact, as we show both analytically and numerically using our model, the effect on loan sales from bank
competition can be orders of magnitude greater than that from capital requirement alone. Therefore, the theory developed in this paper may help explain why enormous growth in loan sales had already occurred even before the U.S. adopted the Basel II Accord in 2005.

The mechanism embodied in our model is supported by empirical evidence as well. A bank in the model channels funds from households to entrepreneurs, who rely on bank loans to finance risky projects, by screening and monitoring the projects. The set of projects is given, so lending opportunities would be exogenous absent bank competition. Competition endogenizes loan originating opportunities faced by individual banks. Such endogenization is operationalized through directed search by entrepreneurs on the credit market following the modeling approach in Peters (1984) and Burdett et al. (2001) methodologically. The interplay of bank competition and entrepreneur directed search reduces interest rate on loans financed through on-balance-sheet activities. On the other hand, the banks' screening processes associated with their originating opportunities generate proprietary information about their applicants, which has long been established as a key factor behind loan sales. This induces the banks to use the OTD model in order to benefit from their loan originating advantages arising from such proprietary information. Therefore, in our model with bank competition, banks use loan sales as a financing strategy and sell loans for the purpose of increasing their credit supply. This is in line with much of the empirical evidence on the motivation of bank loan sales. Empirical evidence also lends direct support to the results that competition can result in a shortage of bank deposits for funding risky investments that require careful screening and monitoring, and that reduced interest rates and growth in loan sales can arise in tandem due to bank competition.

---

5This generalizes earlier studies of loan sales that treat loan originating opportunities or credit supply as exogenous. See, for example, Gordon and Pennacchi (1995), Parlour and Plantin (2008), Ahn (2010), and Ahn and Breton (2010).

6See, also, Arnold (2000), Shi (2009), Camera and Selcuk (2009), and Wright et al. (2017).


8Direct evidence based on U.S. bank holding company data from 2001 to 2007 is provided by Sarkisyan et al. (2009), who find that banks use loan sales mainly as a financing strategy. This conforms to the finding by Drucker and Puri (2009) based on four different data sources that banks sell loans mainly for the purpose of increasing their credit supply. Consistent evidence can be found in the earlier study by Cebenoyan and Strahan (2004), which shows that the chief benefit of loan selling is greater bank credit availability. Corroborating evidence can also be found in Faulkender and Petersen (2006) and Sufi (2009).


10See, for example, Guner (2006). Indeed, those years that experienced increased bank competition and growth in loan sales also saw lowered interest rates relative to previous years, with loose monetary policy.
The presentation is organized as follows. Section 2 lays out the basic environment. Section 3 analyzes a case of no bank competition. Section 4 studies the model with bank competition and entrepreneur directed search. Section 5 concludes.

2 Basic Environment

The economy takes place at infinitely many dates, \( t = 0, 1, 2, \ldots \), and is populated by overlapping generations of two-period lived households, with a unit measure of each generation, and of a large number of entrepreneurs, along with a smaller number of banks. Young households are endowed with 1 unit of time that is supplied inelastically to labor market, but old households do not work.

There are two goods in each period, a final good, which also serves as a numeraire, and an intermediate good. The final good can be directly consumed by households, stored through a storage technology, or invested in projects operated by entrepreneurs to get transformed into an intermediate good in the next period. The intermediate good and labor can then be combined to produce the final good according to \( y_t = F(m_t, l_t; z) \), where \( y_t \) is date-\( t \) output of the final good, \( z \) represents the aggregate level of technology, \( l_t \) is date-\( t \) labor input, and \( m_t \) is date-\( t \) input of the intermediate good, which is transformed from date-(\( t - 1 \)) final good, except for the initial period \( t = 0 \) in which an initial amount of intermediate good \( m_0 \) is owned by the initial old households. The production function \( F \) is of constant returns to scale with respect to the two inputs, and is strictly increasing, strictly quasi-concave, and twice continuously differentiable in both of these two variables. With the inelastic labor supply, the production function can be expressed in an intensive form as \( f(m_t; z) \equiv F(m_t, 1; z) \). There is perfect competition in this production sector, so the price of intermediate good and the wage rate of labor service are determined by their respective marginal products so that \( q_t = f'(m_t; z) \) and \( w_t = f(m_t; z) - q_t m_t \). Each initial old household simply consumes \( c^o_0 = q_0 m_0 = f'(m_0; z)m_0 \).

Each young household of generation \( t \) supplies its 1 unit of time endowment as labor inelastically and allocates its wage income \( w_t \) to finance its consumption when young, \( c^y_t \), and its consumption when old, \( c^o_{t+1} \), to maximize \( u(c^y_t) + \beta E_t u(c^o_{t+1}), \) where \( \beta \in (0, 1) \) is a discount factor, and \( E_t \) denotes the conditional expectations operator. The period utility and global imbalances as potentially other contributing factors.
function $u$ is strictly increasing, strictly concave, twice continuously differentiable, and satisfies the property that $\lim_{c \to 0^+} u(c) \to +\infty$.

At each date $t$ a set of $N$ entrepreneurs emerges with the new generation of households. These entrepreneurs have no internal funds but are endowed with projects that may transfer date-$t$ final good into date-$(t + 1)$ intermediate good. Only the entrepreneurs know how to operate their projects. The projects are indivisible with random returns: Each project takes 1 unit of the final good as input, and produces 1 or 0 unit of the intermediate good with probability $\theta$ and $1 - \theta$, respectively. The success probability $\theta$ is a random draw from a cumulative distribution function $G(\theta)$ on a support $[0,1]$.

We assume that it is costly forbidden for a household to screen the type of a project (in terms of its success probability) ex ante or to monitor the outcome of the project (in terms of its realized return) ex post, whereas it is costless for a bank to do so. This gives rise to a role of banks in intermediating borrowing and lending between entrepreneurs and households.\footnote{As in Diamond (1984), what is essential here is that the screening and monitoring costs are smaller for a bank than for a household.}

A bank exploits its traditional intermediation role by taking deposits from young households and making loans to entrepreneurs. This ‘originate to hold’ activity is subject to a capital requirement regulation in that at least $k > 0$ fraction of the bank’s loans must be financed by its equity and thus at most $1 - k$ fraction of the loans can be financed by households’ deposits. Both households and banks are interest-rate takers when making their deposit supply and demand decisions. We assume a full deposit insurance so there is no risk premium attached to the interest rate on deposits $r$, and that a bank has access to an outside market where capital can be raised at a cost $r + \rho$, where $\rho > 0$ is exogenously given. This captures in a parsimonious manner the notion that a capital requirement regulation is costly for banks. Since debt is cheaper than equity, a bank will not raise more capital than needed for meeting the capital requirement regulation. This pins down the bank’s capital structure: It will choose a debt-equity ratio of $(1 - k)/k$. As a result, if we denote the marginal cost of deposit by $\eta$, then $\eta = (1 - k) r + k (r + \rho) = r + k \rho$.

At the heart of the model is an expansion of the bank’s intermediation business to including the ‘originate to distribute’ approach for credit extension, under which the bank can sell some of its originated loans to young households instead of funding them with its
deposit liabilities. The amount of loan sales is subject to a ‘skin-in-the-game’ constraint that requires the bank to keep on its balance sheet a minimum fraction of loans that it has originated. This helps address potential adverse selection and moral hazard problems associated with the off-balance-sheet activities.

A young household can save by making deposit with banks \((D)\), purchasing loans sold by banks \((S)\), or using a storage technology that can take \(s\) units of final good in one period to yield \(y^s = z_s s^s\) units of final good in the next period, for some \(\varsigma \in (0, 1)\) and \(z_s > 0\). It is worth noting that it is the availability of this storage technology to households that makes banks behave competitively on the markets for deposits and loan sales, so that the rate of interest on \(D\) \((r)\) and the rate of return on \(S\) \((r^a\) with probability \(\bar{\theta}\) and 0 with probability \(1 - \bar{\theta}\) where \(\bar{\theta}\) is to be defined later) are both taken as given not only by households but also by banks in making their decisions. Given that households have non-satiated preferences, the expected lifetime utility of a young household is

\[
u(w - s - D - S) + \beta \left[ \theta u(z_s s^s + rD + r^aS + \pi) + (1 - \bar{\theta}) u(z_s s^s + rD + \pi) \right], \tag{1}\]

where \(\pi\) denotes the sum of profits from banks \((\pi_b)\) and entrepreneurs \((\pi_e)\). Note that the household’s consumption when young and when old must be strictly positive given the property of the utility function that \(\lim_{c \to 0^+} u(c) \to +\infty\). Since the marginal rate of return to the storage technology will approach positive infinity as \(s \to 0^+\), it must be the case that \(s\) is strictly positive as well. Since in any given period entrepreneurs must rely on bank loans to finance their projects for generating intermediate inputs in the subsequent period, which must be combined with labor services for final good production, an active banking sector in period \(t - 1\) is necessary to ensure an active production sector in period \(t\), which is essential in order for a young household of generation \(t\) to carry on, for all \(t \geq 1\). Note that a positive level of household deposit is required for the banking sector to be active, since loan sales would not be possible without any household deposit to finance loans to be sold in the first place. Therefore, it must be the case that \(D > 0\) as well.

Given \(r\), \(r^a\), and \(\bar{\theta}\), the household’s choices of \(s > 0\) and \(D > 0\) satisfy

\[
u'(w - s - D - S) = \beta \varsigma z_s s^{s-1} \left[ \theta u'(z_s s^s + rD + r^aS + \pi) + (1 - \bar{\theta}) u'(z_s s^s + rD + \pi) \right]; \tag{2}\]

\[
u'(w - s - D - S) = \beta r \left[ \theta u'(z_s s^s + rD + r^aS + \pi) + (1 - \bar{\theta}) u'(z_s s^s + rD + \pi) \right]. \tag{3}\]

It follows from the above two optimality conditions that \(r = \varsigma z_s s^{s-1}\).
Let $\varphi_a \geq 0$ denote the multiplier associated with the constraint $S \geq 0$. The household’s choice of $S$ satisfies
\[
u'(w - s - D - S) = \beta \theta^a u'(z_s s^c + rD + r^aS + \pi) + \varphi_a, \quad \text{and} \quad \varphi_a S = 0. \tag{4}
\]

Note that (3) and (4) imply
\[
\bar{\theta}^a u'(z_s s^c + rD + r^aS + \pi) + \varphi_a = r \left[ \bar{\theta} u'(z_s s^c + rD + r^aS + \pi) + (1 - \bar{\theta}) u'(z_s s^c + rD + \pi) \right]. \tag{5}
\]

Given that the period utility function $u$ is strictly increasing and strictly concave, $\varphi_a \geq 0$, $S \geq 0$, and $\varphi_a S = 0$, it is straightforward to show that: If $\bar{\theta}^a < r$ then $\varphi_a > 0$ and $S = 0$ while if $\bar{\theta}^a = r$ then $\varphi_a = 0$ and $S = 0$. Together these imply that, if $\bar{\theta}^a \leq r$ then the household’s demand for loan sales is zero.

We now show that, if $\bar{\theta}^a > r$ then $S > 0$. Suppose, on the contrary, $S = 0$ so (5) is
\[
\bar{\theta}^a u'(z_s s^c + rD + \pi) + \varphi_a = ru'(z_s s^c + rD + \pi).
\]

But, given that $\bar{\theta}^a > r$ and the marginal utility is strictly positive, the above equation can hold only if $\varphi_a < 0$, which contradicts the condition that $\varphi_a \geq 0$.

Hence, the household’s demand for loan sales is positive if and only if $\bar{\theta}^a > r$, where $\bar{\theta}^a - r$ represents a premium that compensates the household for the risk on sold loans.

In this case, we must have $\varphi_a = 0$ so (5) reduces to
\[
\left( \frac{r^a}{r} - 1 \right) = \left( 1 - \frac{1}{\bar{\theta}} \right) \frac{u'(z_s s^c + rD + \pi)}{u'(z_s s^c + rD + r^aS + \pi)}. \tag{6}
\]

Since the period utility function $u$ is twice continuously differentiable in addition to being strictly increasing and strictly concave, in light of (6), for given $\bar{\theta}$, $r$, $s$, $D$, and $\pi$, the household’s demand for loan sales $S$ is a continuously increasing function of $r^a$, for $\bar{\theta}^a > r$, and it approaches 0 as $\bar{\theta}^a$ approaches $r$ from above.

**Remark 1** The households’ demand for loan sales is zero for $\bar{\theta}^a \leq r$, but is positive and continuously increasing with $\bar{\theta}^a > r$ where $\bar{\theta}^a - r$ represents a premium that compensates the households for the risk on sold loans.

We turn now to the problems of banks and entrepreneurs. To demonstrate the role of bank competition for lending opportunities in motivating loan sales, we consider first
a benchmark situation that abstracts from bank competition. This is a case in which a monopoly bank has market power over entrepreneurs on the loan market, although it behaves competitively on the markets for deposits and loan sales due to household accessibility to the storage technology. In what follows, as in describing the household problem above, whenever the timing is clear we shall omit the subscript $t$ to help simplify presentation.

3 The Case of No Bank Competition

We examine the case of a monopoly bank in this section. In each period, the bank makes decisions on its capital and deposit, loan contract, and loan sales sequentially. A period can be divided into three stages accordingly. In the first stage, the bank chooses the amount of capital $K$ at cost $r + \rho$ and raises deposit $D$ from young households under interest rate $r$, subject to the capital requirement constraint $k(K + D) \leq K$, which we recall is binding.

In the second stage, the bank chooses a loan contract concerning lending rate $\gamma$ and the cutoff point for the quality of a project $\theta_1$: With perfect screening, only a project with an ex ante success probability greater than or equal to $\theta_1$ will get funded in this stage, and its ex post return to the bank will be $\gamma$ if the project indeed succeeds but zero if the project actually fails. In the third stage, the bank may raise additional funds by selling to the young households $S$ amount of the loans originated in the second stage subject to a ‘skin-in-the-game’ constraint that requires the bank to keep on its balance sheet at least $\lambda \in (0, 1)$ fraction of the pool of loans. These loans sold by the bank to the households have a risky return: With probability $\bar{\theta} = \int_{\theta_1}^{1} \theta dG(\theta)/[1 - G(\theta_1)]$, which is the average success probability for the pool of loans originated in the second stage, they will realize a rate of return $r^a$, but with probability $1 - \bar{\theta}$ they will return nothing. The expected rate of return on the sold loans is then $\bar{\theta}r^a$. Denote by $\theta_2$ and $\gamma^a$ respectively cutoff point for the quality of and lending rate on additional projects that may get financed in this stage when the possibility of loan sales is considered.

The bank’s expected profit conditional on its choices of capital and deposit is

$$\pi_b = \max_{\gamma, \theta_1, \gamma^a, \theta_2, S} \left[ N \int_{\theta_1}^{1} \theta dG(\theta) \gamma - \eta(K + D) + N \int_{\theta_1}^{\theta_2} \theta dG(\theta) \gamma^a - \bar{\theta}r^a S \right],$$

where the choice of $\theta_1$ in the second stage must be compatible with the on-balance-sheet funds chosen in the first stage so that $N[1 - G(\theta_1)] \leq K + D$, and where the amount of
loans originated in the second stage but sold by the bank to the households in the third stage that is used to finance \( N [G(\theta_1) - G(\theta_2)] \) amount of additional projects must respect the ‘skin-in-the-game’ constraint,

\[
S \leq (1 - \lambda) N [1 - G(\theta_1)].
\]  

(8)

With its limited liability prescribed by a loan contract, any entrepreneur would be willing to take out a loan to invest in its project if the lending rate is not greater than the expected price of intermediate good. In the meantime, the market power of the monopoly bank over entrepreneurs on the loan market would allow it to extract the entire expected surplus while leaving the entrepreneurs to earn zero profit from the invested projects. These together imply that

\[
\gamma = \gamma^a = q.
\]  

(9)

Condition (9) says that the returns to the bank from all successful projects are equal regardless of whether they are financed by on-balance-sheet or off-balance-sheet funds. In consequence, the bank may engage in loan sales if and only if \( \bar{\theta} r^a \) is not greater than \( \eta \). If \( \bar{\theta} r^a < \eta \), then off-balance-sheet funds are cheaper than on-balance-sheet funds, so the bank would raise deposit and capital in the first stage to finance loans originated in the second stage only to the extent by which these loans can be sold to the households in the third stage subject to a binding ‘skin-in-the-game’ constraint. If \( \bar{\theta} r^a = \eta \), then the two sources of funding are equally costly, so the bank’s supply of loan sales can be at any level between zero and \( (1 - \lambda) N [1 - G(\theta_1)] \).

Remark 2  The bank’s supply of loan sales is zero for \( \bar{\theta} r^a > \eta \), can be at any level between zero and \( (1 - \lambda) N [1 - G(\theta_1)] \) for \( \bar{\theta} r^a = \eta \), but equals \( (1 - \lambda) N [1 - G(\theta_1)] \) for \( \bar{\theta} r^a < \eta \).

Combining Remarks 1 and 2 we conclude that, in this case of no bank competition, if there exists an equilibrium with loan sales then the expected rate of return \( \bar{\theta} r^a \) from the sold loans must fall into the half-open, half-closed interval \( (r, \eta] \), which is nonempty provided there is a capital requirement regulation, or, \( k > 0 \). The length of the interval, \( \eta - r = k \rho \), which is governed by the degree of capital requirement, bounds the maximal risk premium that the bank is willing to compensate the households for sold loans relative to deposits. On the other side, the magnitude of risk premium needed to induce a large or even just moderate demand for loan sales from reasonably risk averse households may not
be any smaller than $k\rho$ for an empirically relevant degree of capital requirement. Hence, absent bank competition, it is conceivable that loan sales will not be of any significant scale and the ‘skin-in-the-game’ constraint may not be binding.

To get a more quantitative feel we invoke a CRRA utility function $u(c) = c^{1-\sigma} / (1-\sigma)$, a Cobb-Douglas production function $y = zm^\alpha$, and a uniform distribution function $G(\theta) = \theta$. To assign values to the model’s parameters, we set $\alpha = 0.4$, $z = 100$, $\beta = 0.68$, $N = 400$, $\rho = 0.3$, $z_s = 5$, $\varsigma = 0.5$, $\lambda = 0.2$, $\sigma = 2$, and consider a range of $k$ between 0.01 and 0.08.

Figure 1 displays the ratio of loan sales to total bank credits for the current case with no bank competition under the above parametrization. As the figure shows, loan sales indeed constitute an insignificant fraction of total bank credits. Although the share increases with $k$, the pace is moderate. As $k$ rises from 0.01, to 0.04, and then to 0.08, the share increases from 0.0062, to 0.0247, and then to 0.0495. Therefore, the present model with no bank competition would predict that, even for a capital requirement regulation as high as 8%, less than 5% of loans originated will be sold and more than 95% will be kept on the bank’s balance sheet. In the real world, by contrast, loan sales as a fraction of loans and leases in bank credits was more than 40% in 2007, the year leading up to the recent financial crisis.

4 Bank Competition as a Motivation for Loan Sales

We examine now our model of bank competition where there are $B(> 1)$ banks operating in the economy. Competition endogenizes loan originating opportunities faced by individual banks, whereas the banks’ screening processes associated with such opportunities generate proprietary information about their applicants. While the competition drives the loan rate on high quality projects financed through on-balance-sheet activities below the expected price of intermediate good, the proprietary information motivates the banks to engage in loan sales as a finance strategy in order to exploit their loan originating advantages.

As we show below, an equilibrium in this context features a binding ‘skin-in-the-game’ constraint and loan sales at a maximal scale possible.

4.1 Bank competition and directed search by entrepreneurs

One way to operationalize the above mechanism is to cast our analysis of bank competition for lending opportunities using the framework of directed search. For this purpose, and
in a similar spirit of the directed search literature, it is descriptively convenient to locate the \( B \) banks on separated islands and assume that an entrepreneur can travel to only one island so can apply to only one bank at the given point in time. Banks post loan contracts to compete for entrepreneurs to travel to their islands in the second stage of the game. Other features of the model are similar to those of the model with no bank competition, including nationwide competitive markets for both deposits and loan sales.

More specifically, in each period the banks make decisions on their capitals and deposits, loan contracts, and loan sales in three sequential stages. In the first stage, the banks choose their capitals at cost \( r + \rho \) and raise deposits from the young households taking interest rate \( r \) as given, subject to the capital requirement constraint, which is binding for all banks. In the second stage, the banks post their loan contracts. An individual bank \( i \) posts its lending rate \( \gamma_i \) and its cutoff rule \( \theta_{1,i}(n_i) \) for the quality of a project: With perfect screening, only a project with a success probability greater than or equal to \( \theta_{1,i}(n_i) \) will get funded by bank \( i \) which will get paid the rate \( \gamma_i \) if the project indeed succeeds but nothing if the project actually fails. Similar to Peters (1984), we make the selection criterion \( \theta_{1,i} \) but not the loan rate \( \gamma_i \) contingent on the number of entrepreneurs applying to bank \( i \) \( (n_i) \). This helps capture in a parsimonious manner an essential tension and tradeoff in the directed search under capacity constraint:

**Remark 3** A lower posted loan rate \( \gamma_i \) would attract a greater number of entrepreneurs \( n_i \), which would yield a higher selection criterion \( \theta_{1,i} \) and thus a lower probability for an applying entrepreneur to get financed in the second stage by bank \( i \) given the bank’s on-balance-sheet funds chosen in the first stage.

In the third stage, the banks may raise additional funds by selling to the young households loans originated in the second stage. These loans sold by the banks to the households have a risky return: They will realize a rate of return \( r^a \) with probability \( \overline{\theta} \) but nothing with probability \( 1 - \overline{\theta} \), so the expected rate of return on the sold loans is \( \overline{\theta}r^a \), where \( \overline{\theta} \) is the average success probability for the pool of loans originated in the second stage. Denote by \( \theta_{2,i} \) and \( \gamma_i^a \) the selection criterion and loan rate on additional projects that will get financed by bank \( i \) in the third stage when the possibility of loan sales is considered.

Banks’ capital structures and contractual terms are public information. Given the other banks’ and the entrepreneurs’ strategies, an individual bank \( i \)’s expected profit conditional
on its choices of capital and deposit is

\[ \pi_{b,i} = \max_{\gamma_i, \theta_{1,i}, \gamma_a, \theta_{2,i}, S_i} \left[ n_i \int_{\theta_{1,i}}^{\gamma_i} \theta dG(\theta) \gamma_i - \eta (K_i + D_i) + n_i \int_{\theta_{2,i}}^{\theta_{1,i}} \theta dG(\theta) \gamma_a - \bar{\theta} r a S_i \right], \]

where the choice of \( \theta_{1,i} \) in the second stage must be compatible with the on-balance-sheet funds chosen in the first stage so that \( n_i [1 - G(\theta_{1,i})] \leq K_i + D_i \), and where the amount of loans originated in the second stage but sold by the bank to the households in the third stage that is used to finance \( n_i [G(\theta_{1,i}) - G(\theta_{2,i})] \) amount of additional projects must respect the ‘skin-in-the-game’ constraint,

\[ S_i \leq (1 - \lambda) n_i [1 - G(\theta_{1,i})]. \] (10)

Entrepreneurs conduct directed search in choosing the islands to which they will travel. After an entrepreneur arrives at its chosen island \( i \), the success probability of its project \( \theta \) is drawn from the distribution \( G \) which is discovered by bank \( i \) upon completing a screening process. If \( \theta \) turns out to be greater than \( \theta_{1,i} \), the project is financed with loan rate \( \gamma_i \), just as prescribed by the posted contract.\(^{12}\) Otherwise, they enter into the next stage in which loan sales are considered by the bank, and in which the project will get a second chance to be financed, with loan rate \( \gamma_a \), if \( \theta \) turns out to be greater than \( \theta_{2,i} \). Given that the entrepreneur is already locked to the island at this stage of the game, the bank can now extract the entire expected surplus from the funded project, implying \( \gamma_a = q \).

Given banks’ capital structures and contractual terms, and distribution of entrepreneurs traveling to different islands, \( \{n_i\}_{i=1}^B \), the expected profit of an entrepreneur that chooses to go to island \( i \) is then given by

\[ \pi_{e,i} = \int_{\theta_{1,i}(n_i)}^{\gamma_i} \theta dG(\theta) (q - \gamma_i), \] (11)

where the entrepreneur’s optimal strategy must solve \( \max_i \{\pi_{e,i}\} \). Hence, in deciding on which island to attend in the end, the entrepreneur needs to balance the tradeoff between the ex ante probability of getting funded and the ex post profit margin, as highlighted in Remark 3. As all entrepreneurs conduct directed search under this kind of tradeoff, the expected profits from traveling to different islands must be equalized in equilibrium so that any additional entrepreneur would be indifferent about which island to attend.

We turn now to analyzing a symmetric equilibrium of this type.

\(^{12}\)We assume that banks can commit to their posted loan rates, so we do not consider the possibility of bargaining after the banks meet with entrepreneurs as in Camera and Selcuk (2009).
4.2 Symmetric equilibrium

We restrict attention to a stationary symmetric strong Nash equilibrium where allocations, prices and selection rules are identical across islands and entrepreneurs choose identical mixed strategies on which islands to attend, and where neither a bank nor an entrepreneur has any incentive to deviate from the outcome.

Definition 1 A symmetric strong Nash equilibrium with bank competition and directed search by entrepreneurs consists of prices \( \{\gamma_{i,t}, \gamma^a_{i,t}\}_{i=1}^B, r_t, r^a_t, q_t, w_t\}_{t=0}^\infty \), selection rules \( \{\{\theta_{1,i,t}, \theta_{2,i,t}\}_{i=1}^B\}_{t=0}^\infty \), and allocations \( \{\{K_{i,t}, D_{i,t}, S_{i,t}, n_{i,t}\}_{i=1}^B, m_{t}, y_{t}, y^a_{t}, s_{t}, c^a_{t}, c^o_{t}\}_{t=0}^\infty \) given initial endowment of intermediate good \( m_0 \) and policy parameters \( k \) and \( \lambda \), such that:

1. taking the wage rates, interest rates, and banks’ strategies as given, each household’s allocations solve its utility maximization problem whereas each initial old household simply consumes the return to investment of its intermediate good endowment;

2. taking the interest rates, prices, and other banks’ and all entrepreneurs’ strategies as given, each bank’s choices of capital and deposit, loan rates and selection rules, and loan sales solve its profit maximization problem;

3. taking the prices and other entrepreneurs’ and all banks’ strategies as given, an entrepreneur chooses a mixed strategy on which islands to attend to maximize its profit;

4. markets clear, in particular, for final goods (i) \( c^o_{t} + c^o_{t} + s_{t} + D_{t} + S_{t} = f(m_t; z) \) and \( c^o_{t} + c^a_{t} + s_{t} + D_{t} + S_{t} = f(m_t; z) + z_{t}s_{t-1} \) for \( t \geq 1 \), and for intermediate goods (ii) \( m_{t+1} = \sum_{i=1}^B n_{i,t} \int_{\theta_{2,i,t}}^{1} \theta dG(\theta) \) for \( t \geq 0 \);

5. allocations, prices and selection rules are identical across banks and entrepreneurs choose identical mixed strategies on which islands to attend;

6. neither a bank nor an entrepreneur has any incentive to deviate from the outcome.

A stationary symmetric strong Nash equilibrium with bank competition and directed search by entrepreneurs analyzed here belongs to a class of equilibrium under capacity constrained Bertrand competition well studied in the literature, and it is well known that there generally exists such a unique equilibrium (e.g., Peters 1984 and Burdett et al. 2001). Given that we focus on a stationary symmetric equilibrium we can omit the time or\ and the individual indexes whenever the situation is clear. It turns out that such an equilibrium in our current context features \( \gamma < q \) and a binding ‘skin-in-the-game’ constraint so loan
sales at a maximal scale possible. This is in contrast to the case of no bank competition studied in the previous section that features a generally non-binding ‘skin-in-the-game’ constraint and an insignificant scale of loan sales.

Remark 4 A stationary symmetric strong Nash equilibrium with bank competition and entrepreneur directed search features $\gamma < q$ and a binding ‘skin-in-the-game’ constraint.

Recall in the case of no bank competition studied in Section 3 it must hold that $\gamma = q$. It is intuitive why in the stationary symmetric equilibrium analyzed in this section bank competition and directed search by entrepreneurs must drive $\gamma$ below $q$ (although it has to stay above 0). Suppose, on the contrary, $\gamma = q$ (and denote by $\theta_1$ the corresponding selection criterion), so entrepreneur expected profit from applying to all banks is zero. But, then, an individual bank would have an incentive to deviate from this outcome: by lowering its posted loan rate, say $\gamma^d$, to infinitesimally below $q$ (and denoting by $\theta^d_1$ the corresponding selection criterion), entrepreneur expected profit from applying to its loans will become positive, so it will attract all entrepreneurs. This is to say that the number of entrepreneurs traveling to its island will jump discretely, from $N/B$ to $N$. Thus, even keeping its capital deposit level and scale of loan sales as before, the deviating bank could strictly increase its expected profit by lowering only negligibly its posted loan rate to measurably increase the average quality of its funded projects.

To put the above discussion into some quantitative perspective note that, since $\gamma^d < q$, it must hold that $\theta^d_1 > \theta_1$ while the two are related by $G(\theta^d_1) - G(\theta_1) = (B - 1)[1 - G(\theta^d_1)]$. It follows that \[ \int_{\theta_1}^{\theta^d_1} \theta dG(\theta) < (B - 1) \int_{\theta^d_1}^{\theta_1} \theta dG(\theta). \] Let

\[ \alpha \equiv \frac{\int_{\theta_1}^{\theta^d_1} \theta dG(\theta)}{(B - 1) \int_{\theta^d_1}^{\theta_1} \theta dG(\theta)}. \]

Then $\alpha \in (0, 1)$. With some algebra, we can show that the deviating strategy $\{\gamma^d, \theta^d_1\}$, along with its unchanged capital\deposit level and scale of loan sales, and a corresponding upward shift in the selection criterion in the third stage, from $\theta_2$ to $\theta^d_2$, where the latter is determined by $G(\theta_1) - G(\theta_2) = B[1 - G(\theta_1)] - G(\theta^d_2)]$, and where $\gamma^a$ stays equal to $q$, will strictly increase the expected profit of the deviating bank as along as $\gamma^d$ satisfies

\[ \left( \alpha + \frac{1 - \alpha}{B} \right) q < \gamma^d < q. \]
We thus have established the first half of the statement in Remark 4. To validate the second half of the statement in the remark suppose, on the contrary, the ‘skin-in-the-game’ constraint is non-binding. Then it must be the case that \( \theta_2 q = \bar{\theta} r^a \), that is, banks break even on the marginal-quality project financed in the third stage so all profitable projects are funded, and that the expected revenue from the marginal-quality project financed in the second stage is not lower than the cost of funding from either source. A bank’s profit maximization problem conditional on its choice of capital and deposit is then

\[
\pi_b = \max_{\gamma, \theta_1} \frac{K + D}{1 - G(\theta_1)} \left[ \gamma \int_{\theta_1}^{1} \theta dG(\theta) + \int_{\bar{\theta} r^a}^{\theta_1} (q\theta - \bar{\theta} r^a) dG(\theta) \right] - \eta(K + D),
\]

where we have also substituted in the tradeoff relation between the number of entrepreneurs applying to the bank and its selection criterion pertaining to its posted loan rate under the lending capacity associated with its on-balance-sheet funds \( K + D \) (e.g., Remark 3). Let \( \psi(\gamma) \) be the first order derivative of the objective function in (14) with respect to \( \gamma \).

With some algebra we obtain

\[
\psi(\gamma) = -\frac{(K + D)\pi_e \left[ \gamma \int_{\theta_1}^{1} \theta dG(\theta) + \int_{\bar{\theta} r^a}^{\theta_1} (q\theta - \bar{\theta} r^a) dG(\theta) \right]}{\theta_1 [1 - G(\theta_1)]^2 (q - \gamma)^2} < 0,
\]

where we have respected the entrepreneur expected-profit distribution prescribed by (11) and the corresponding iso-profit relation

\[
\frac{d\theta_1}{d\gamma} = -\frac{1}{\theta_1 g(\theta_1)} \frac{\theta dG(\theta)}{(q - \gamma)}.
\]

The inequality in (15) suggests nonexistence of a stationary point \( \gamma \in (0, q) \) so as to validate the second half of the statement in Remark 4.

The above reasoning should remind us of a general principle in corporate finance when only on-balance-sheet activities are considered – “never fund a negative net present value (NPV) project” – which is also true in the previous section with no bank competition. Since the ‘skin-in-the-game’ constraint is binding in the equilibrium analyzed in this section with bank competition and entrepreneur directed search, banks may in fact originate negative NPV projects in the second stage as a means to relaxing the ‘skin-in-the-game’ constraint in order to take advantage of their originating advantages.

One way to make this point more transparent is to differentiate two nuanced cases: (i) the ‘skin-in-the-game’ constraint is just binding versus (ii) the ‘skin-in-the-game’ constraint
is strictly binding. Based on the above characterization of a bank’s optimal choices in the second and third stages as a function of its choice in the first stage, and applying the envelope theorem, we can show that the optimal strategy on the bank’s on-balance-sheet activities in the first stage would lead to

$$\gamma \theta_1 - \eta = q(\theta_1 - \theta_2) - (1 - \lambda)(q\theta_2 - \bar{\theta}_r^a),$$  \hspace{1cm} (17)

where we have respected the relations $K/k = N/\{B[1 - G(\theta_1)]\}$ linking its decisions in the first and second stages, and $G(\theta_1) - G(\theta_2) = (1 - \lambda)[1 - G(\theta_1)]$ linking its decisions in the second and third stages with a binding ‘skin-in-the-game’ constraint as well as the implied monotonic relation

$$\frac{d\theta_2}{d\theta_1} = \frac{(2 - \lambda)g(\theta_1)}{g(\theta_2)}.$$  \hspace{1cm} (18)

If the ‘skin-in-the-game’ constraint is just binding, then $q\theta_2 - \bar{\theta}_r^a = 0$, and (17) makes it clear that $\gamma \theta_1 - \eta = q(\theta_1 - \theta_2) \geq 0$. This is to say that bank expected profit from the marginal-quality project financed by on-balance-sheet funds must be nonnegative. Thus the conventional wisdom concerning bank on-balance-sheet activities generalizes to our current context where banks engage in both on-balance-sheet and off-balance-sheet activities with a weakly binding ‘skin-in-the-game’ constraint.

If, however, the ‘skin-in-the-game’ constraint is strictly binding, then $q\theta_2 - \bar{\theta}_r^a > 0$, so not all profitable projects via off-balance-sheet activities can get funded due to the strongly binding ‘skin-in-the-game’ constraint, and it then becomes possible that $\gamma \theta_1 < \eta$. In such an equilibrium banks invest in some negative NPV projects in the second stage as a means to circumvent the ‘skin-in-the-game’ constraint in the third stage so they can make more profits from additional loan sales. The possibility of investing in ‘non-profitable’ projects concerning on-balance-sheet activities thus highlights banks’ motives for reselling loans in order to benefit from their originating advantages arising from entrepreneur directed search associated with bank competition.\textsuperscript{13}

Despite the nuances, both cases feature the maximal scale of loan sales as permitted by the binding ‘skin-in-the-game’ constraint. The take-home message from our analysis in this section then is that bank competition can be an important motivation for loan sales.

\textsuperscript{13}Such possibility may be high if bank competition for loan originating opportunities drives $\gamma$ much below $q$. Also, such possibility is greater, the laxer is the regulation on off-balance-sheet activities, that is, the smaller is $\lambda$, so a larger fraction of loans originated can be sold, as is highlighted by the manner in which the term $(1 - \lambda)$ enters into (17).
To put this into a more quantitative perspective, Figure 2 plots the fold increase in the ratio of loan sales to total bank credits, from the case of no bank competition, to the case of bank competition with $B = 2$ but identical remaining parametrization as in Section 3. As the figure shows, such increase is dramatic, but more so for small $k$ than for large $k$ – it is about 8 folds for $k = 0.08$, 17 folds for $k = 0.04$, but 70 folds for $k = 0.01$. As a result, loan sales constitute a significant fraction of total bank credits for all levels of $k$, all above 40%, which is comparable to what was observed in the years preceding the recent financial crisis. Thus, although even a large degree of capital requirement by itself can generate only a small scale of loan sales, bank competition in the face of an even small degree of capital requirement can generate a large scale of loan sales.

This basic conclusion holds broadly in our model, that is, it is bank competition not capital requirement that accounts for the majority of loan sales, as our theory suggests, and as above numerical experiments illustrate. We have undertaken many more numerical exercises, though we do not discuss all of the results here in order to conserve space. In all of these additional experiments, this basic conclusion holds quite generally. This is typically the case when we vary parameter values within their empirically plausible specifications. In the numerical experiments reported above, we have focused on varying $k$, as it is the key parameter influencing banks’ supply of loan sales. In other experiments not reported here, we have varied $\sigma$, the key parameter influencing households’ demand for loan sales, within its empirical plausible range $[1, 10]$, and found that the basic conclusion continues to hold. Other parameters are even less influential over the results. In general, such variations in parameter values have some quantitative impacts on the results – sometimes very modestly, and other times to a greater degree – but in no case they could alter the qualitative nature of this basic conclusion from the paper.

5 Conclusion

We have developed a general equilibrium theory of loan sales based on bank competition and entrepreneur directed search. We have demonstrated how bank competition for loan originating opportunities and entrepreneur directed search on the credit market may interact to reduce interest rate on loans financed through on-balance-sheet activities and how this can motivate loan sales as a strategy of financing through off-balance-sheet activities.
The present study complements the literature of loan sales that appeals to bank capital requirement regulation as a motivation for off-balance-sheet activities. We have shown that, in our model, the effect on loan sales from bank competition can be orders of magnitude greater than that from capital requirement alone. It is a general conclusion in our model that it is bank competition rather than capital requirement that accounts for the majority of loan sales. As such, our theory may help understand why enormous growth in loan sales had already occurred during the years before 2005 when the U.S. adopted the Basel II Accord, as those years also witnessed increased bank competition due to technological advancement, financial deregulation, and globalization. Our theory suggests that the observed surge in loan sales over those years could in part represent an equilibrium response to the increase in bank competition. At a more general level, our theory sheds some light on the emergence of a ‘shadow banking system’ during those years.

While a positive study in nature, this is not a data matching paper. Instead, our goal here is to provide an analytical framework that can be used to highlight a rudimentary role of bank competition in motivating off-balance-sheet activities. A good understanding of banks’ incentives in using the OTD model is important for addressing concerns following the recent financial crisis and ensuing recession about the implications of the OTD model for the safety and soundness of the financial system and the economy, and for assessing the broad implications of the recently developed shadow banking system. Our model already embeds key policy regulations for both on-balance-sheet and off-balance-sheet activities (i.e., the capital requirement and skin-in-the-game constraints), which are treated as given for our positive analysis here, and we have already provided some preliminary discussions about their influences on the results (e.g., Footnote 13). The framework presented in this paper if properly adapted/enriched may also prove useful in addressing related normative questions and policy issues. We intend to leave this task to future work.

\[14\] Such an endeavor is reflected by the U.S. President’s Working Group on Financial Markets in its effort to identify the sources of the financial turmoil during the onset of the crisis, and discussed in length by Ben Bernanke in his speeches at the World Affairs Council of Greater Richmond’s Virginia Global Ambassador Award Luncheon on April 10, 2008, and at the Conference Co-sponsored by the Center for Economic Policy Studies and the Bendheim Center for Finance at Princeton University on September 24, 2010. In the European Union, the Economic and Financial Affairs Council mandated the European Central Bank, in cooperation with the Banking Supervision Committee, to assess “...how the so-called ‘originate and distribute’ model ... has impacted on the incentive structures of credit markets, in a context characterized by a shift from the more traditional retail to interbank borrowing.” This is also a mission set out in the 2008 ECB Eurosystem Report, and in the plan for regulatory reform proposed in 2009 by the U.S. Committee on Capital Markets Regulation.
References


Figure 1. Share of loan sales in total bank credits (vertical axis) as a function of the degree of capital requirement $k$ (horizontal axis) in the case of no bank competition
Figure 2. Fold increase in loan sale share due to bank competition (vertical axis) as a function of the degree of capital requirement $k$ (horizontal axis)