Two-sided matching in the loan market ☆

Jiawei Chen [a,*], Kejun Song [b]

[a] University of California, Irvine, United States
[b] Gordon College, United States

ABSTRACT

This paper investigates the matching between banks and firms in the loan market. We estimate a many-to-one two-sided matching model using the Fox (2010) matching maximum score estimator. Using data on the U.S. loan market from 2000 to 2003, we find evidence of positive assortative matching of sizes. Moreover, we show that banks and firms prefer partners that are geographically closer, giving support to the importance of physical proximity for information gathering and expertise sharing. We also show that banks and firms prefer partners with whom they had prior loans, indicating that prior loan relationship plays an important role in the selection of current partners.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

Bank loans play a unique role in corporate financing. They are important not only for small businesses, which often lack access to public debt markets, but also for large corporations, which depend on them as a reliable source of liquidity helping to insulate them from market shocks (James and Smith, 2000; Saidenberg and Strahan, 1999). Furthermore, bank lending is an important conduit for monetary policy and is closely linked to investment and macroeconomic activity (Kashyap and Stein, 1994).

Given the importance of the loan market in the economy, it is critical for policymakers and researchers to understand the workings of the loan market. In particular, knowledge of how banks and firms choose each other (the matching between banks and firms) is important for effective policy making in the loan market and more generally for economic development.

For example, if we find that in the loan market small (large) firms generally match with small (large) banks (positive assortative matching of sizes), and assuming that this pattern cannot be easily changed in the short run, then a policymaker with the objective to increase the availability of credit to small businesses should focus on improving the economic environment for small banks and strengthening their incentive to lend. Restrictions on bank mergers and acquisitions can also

help, as consolidation in the banking sector leads to the dominance by large banks who do not lend much to small businesses.

Similarly, an assessment of the role of physical proximity in banks and firms' selection of partners will prove useful for state and local officials who aim at spurring industrial investment in the local economy, and for federal regulators who contemplate restrictions on interstate banking activities.

While the loan market has attracted a considerable amount of scholarly attention, most of the existing research has been on issues such as the determination of loan interest rates and the analysis and prediction of loan performance, and little has been done to formally investigate the matching between banks and firms. This paper fills the gap by estimating a two-sided matching model of banks and firms in the loan market. In the model, banks choose firms, firms choose banks, and they all face a tradeoff between the match quality and the transfer, as explained below.

First, from the perspective of each agent (bank or firm), matching with different partners generates different match values, giving rise to agent-specific rankings of potential partners. For example, according to Cole et al. (2004), large banks rely on standardized quantitative criteria to assess loan applications (a "cookie-cutter" approach), while small banks favor qualitative criteria based on loan officers' interactions with loan applicants (a "character" approach). At the same time, large firms are more likely to have well-documented track records and financial information, while small firms tend to focus more on relationship-building. Therefore, a large firm is a better match for a large bank's cookie-cutter approach, whereas a small firm is a better match for a small bank's character approach. In addition, large firms tend to need large loans, and small firms tend to need small loans, which also

makes them attracted to large banks and small banks, respectively. Similarly, agents' preferences for other attributes in their partners, such as physical proximity and prior loan relationship, also give rise to agent-specific rankings of potential partners.

Second, when a bank and a firm enter into a loan ("form a match"), the price and non-price characteristics of the loan determine the transfer of utility between the partners. The loan interest rate is clearly an important factor in the transfer, but other characteristics of the loan matter, too. For example, everything else equal, a borrower gives up more utility to the lender if collateral is pledged (secured loan); similarly, other characteristics such as maturity (i.e., the length of the loan, which, together with the interest rate and the discount rate, determines the net present value of all interest payments), loan size and upfront fees also influence the division of utility between the bank and the firm.[1] In this paper we model the matching between banks and firms using a transferable utility matching framework: the transfer is endogenously determined at the time of the matching, and an agent is willing to trade away match quality in order to obtain a better transfer.[2]

Estimation of the model uses the Fox (2010) matching maximum score estimator. The estimator makes use of the inequalities regarding the match values implied by the matching equilibrium. Take any two pairs that are matched in the equilibrium and swap the partners. The equilibrium condition requires that the original sum of match values be greater than or equal to the new sum of match values after the swap. The estimator maximizes the number of such inequalities that are satisfied.

Using data on the U.S. loan market from 2000 to 2003, we find evidence of positive assortative matching of sizes, that is, large banks tend to match with large firms and small banks tend to match with small firms. Moreover, we show that banks and firms prefer partners that are geographically closer, giving support to the importance of physical proximity for information gathering. We also show that banks and firms prefer partners with whom they had prior loans, indicating that prior loan relationship plays an important role in the selection of current partners. We also test a couple of other hypotheses regarding the matching and do not find support for them.

2. Model

We consider a two-sided matching model of the loan market, in which banks choose firms, firms choose banks, and the market outcome is an equilibrium matching which depends on the match values of all the bank–firm pairs (match values and the equilibrium concept are defined later).

When a bank and a firm enter into a loan ("form a match"), they also decide on the price and non-price characteristics of the loan, which determine the transfer of utility between the partners. The transfer is endogenously determined at the time of the matching, and an agent is willing to trade away match quality in order to obtain a better transfer. We therefore model the matching between banks and firms using a transferable utility framework.

2.1. Agents and quotas

Let I_t and J_t denote, respectively, the sets of banks and firms in market t, where $t = 1,2,...,T$. I_t and J_t are finite and disjoint. Below the market subscript t is dropped to simplify the notation.

In the empirical implementation of our model, markets are defined by time: a market contains the firms that borrow during a half-year and the banks that lend to them. In the data the vast majority of firms borrow only once during a half-year.[3] In such a short period of time, it is likely that a firm's financial needs can be satisfied by a single loan, whereas borrowing multiple loans would increase the administrative costs, such as the costs associated with the negotiation process. Therefore it is a reasonable approximation to model that a firm matches with only one bank in a given market.

On the other hand, a bank often lends to multiple firms during a half-year. A bank's lending activity is restricted in two ways. First, loan assessment, approval, monitoring, and review processes are relatively labor-intensive, and a bank's lending activity is restricted by the amount of resources that is available for these processes, e.g., the number of its loan officers. Consequently, the number of loans that a bank can make during a given half-year is limited.[4] Second, the total amount of loans a bank can make may be constrained by the availability of deposits, the primary source of funds for bank lending (Jayaratne and Morgan, 2000). Jayaratne and Morgan (2000) find evidence that the deposits constraint on bank lending operates only on small banks whose assets are less than \$100 million, and that larger banks are unconstrained because they have better access to capital markets. In our sample less than 1% of the banks have assets lower than \$100 million, so the lending constraint posed by inadequate deposits is less of a concern. In our study we take the limit on the total amount of loans as non-binding and take the limit on the number of loans as binding to simplify the empirical implementation and make the model tractable.

In a market, bank i can lend to q_i firms and firm j can borrow from only one bank. The model is a many-to-one two-sided matching model with endogenous transfers (Roth and Sotomayor, 1990; Shapley and Shubik, 1971). q_i is known as the *quota* of bank i in the matching literature, and every firm has a quota of one. We assume that each agent uses up its quota in equilibrium.

2.2. Matches and match values

The set of all potential loans, or *matches*, is given by $I \times J$. A *matching*, μ, is a set of matches such that $(i,j) \in \mu$ if and only if bank i and firm j are matched. We use $\mu(i)$ to denote the set of firms that borrow from bank i, and use $\mu(j)$ to denote the set of banks that lend to firm j, which is a singleton.

For a match between bank i and firm j, let $V_b(i,j)$ and $V_f(i,j)$ denote the bank's and the firm's pre-transfer payoffs, respectively. Let $u_{ij} \in \mathbb{R}$ denote the transfer from firm j to bank i, so that the bank's payoff is $V_b(i,j) + u_{ij}$ and the firm's payoff is $V_f(i,j) - u_{ij}$. The match value is $V(i,j) = V_b(i,j) + V_f(i,j)$.

A bank can match with multiple firms. Let $V_b(i,\mu(i))$ denote bank i's pre-transfer payoff from matching with the set of firms in $\mu(i)$. We assume that $V_b(i,\mu(i))$ is *additively separable* across the firms in $\mu(i)$: $V_b(i,\mu(i)) = \sum_{j \in \mu(i)} V_b(i,j)$.

2.3. Equilibrium

An *outcome* of the market consists of a matching μ and a vector of transfers u, one for each of the matches. The equilibrium concept is pairwise stability. An outcome (μ,u) is *pairwise stable* if for each pair $(i,j) \in \mu$,

$$V_b(i,j) + u_{ij} \geq V_b\left(i, j^{'}\right) + \tilde{u}_{ij'}, \tag{1}$$

where $\tilde{u}_{ij'} = V_f(i,j') - (V_f(i',j') - u_{i'j'})$, (2)

for all $i' \neq i$, $(i',j') \in \mu$. $\tilde{u}_{ij'}$ is the maximum transfer that firm j' is willing to pay to bank i if firm j' switches its lender from bank i' to bank i. Inequality (1) says that even if firm j' offers to pay bank i this maximum transfer, bank i still prefers to lend to firm j and receive u_{ij}, instead of lending to firm j' and receiving $\tilde{u}_{ij'}$.

In our model, the additive separability of $V_b(i,\mu(i))$ that we assumed is a sufficient condition for the existence of a pairwise stable outcome.[5]

In our empirical application, we don't know the transfers between banks and firms, so we follow Fox (2010) and derive the following property of a pairwise stable outcome that does not involve data on transfers.

Substituting Eq. (2) into inequality (1), we obtain

$$V_b(i,j) + u_{ij} \geq V_b(i,j') + V_f(i,j') - (V_f(i',j') - u_{i'j'}). \quad (3)$$

A symmetric inequality holds for bank i' not willing to replace firm j' with firm j:

$$V_b(i',j') + u_{i'j'} \geq V_b(i',j) + V_f(i',j) - (V_f(i,j) - u_{ij}). \quad (4)$$

Summing inequalities (3) and (4) and canceling the transfers, we obtain

$$\begin{aligned} V_b(i,j) + V_f(i,j) + V_b(i',j') + V_f(i',j') \geq V_b(i,j') + V_f(i,j') + V_b(i',j) \\ + V_f(i',j) \Rightarrow V(i,j) + V(i',j') \geq V(i,j') + V(i',j). \end{aligned} \quad (5)$$

Inequality (5) is called the *sum of match values inequality*. It says that if we take any two pairs that are matched in a pairwise stable outcome and swap the partners, the original sum of match values is greater than or equal to the new sum of match values after the swap. This characterization of the equilibrium matching is the basis for the estimation method described in the next section.

3. Estimation method

In order to form the empirical analog of inequality (5), we first parameterize the match value function as

$$V(i,j) \equiv v(X_{ij}) + \eta_i + \delta_j + \epsilon_{ij}.$$

$v(X_{ij})$ is the deterministic part of the match value that depends on X_{ij}, which consists of pair characteristics (including interactions of bank and firm characteristics). This part of the match value drives the equilibrium matching patterns and will be identified through estimation. For simplicity, we restrict attention to match value functions that are linear in X_{ij}: $v(X_{ij}) = X'_{ij}\beta$, where β is a finite dimensional parameter vector.

The terms η_i and δ_j are bank and firm fixed effects, respectively, which are observed by the agents but unknown to the econometrician. These agent-specific fixed effects enter into both sides of inequality (5) and cancel out, so they are unidentified. Theoretically, bank and firm fixed effects are valued equally by all potential partners and differencing them out leaves the matching unaffected.

The match-specific error terms $_{ij}$ are random realizations of the match values, which are also observed by the agents but unknown to the econometrician. These match-specific error terms are added to the specification

of the match value function in order to allow the empirical model to give full support to the data.

Our estimation uses the Fox (2010) matching maximum score estimator. The estimator assumes that the following rank order property holds for the stochastic structure of the model. Consider two matchings μ_1 and μ_2 that differ only in two matches: (i,j), $(i',j') \in \mu_1$ and (i,j'), $(i',j) \in \mu_2$. The rank order property says that if the sum of the deterministic match values is greater for (i,j) and (i',j') than for (i,j') and (i',j), then μ_1 is more likely to be observed than μ_2. This rank order property is a natural extension of the sum of match values inequality (5) to an econometric model, in which matchings that are more likely to be pairwise stable are more likely to occur.

Estimation uses the empirical analog of the sum of match values inequalities (5). Specifically, the maximum score estimator $\hat{\beta}$ maximizes the following maximum score objective function:

$$Q(\beta) = \sum_{t=1}^{T} \sum_{(i,j),(i',j') \in \mu_t} 1\left[X'_{ij}\beta + X'_{i'j'}\beta \geq X'_{ij'}\beta + X'_{i'j}\beta\right]. \quad (6)$$

The estimator $\hat{\beta}$ is the vector of parameter values that maximizes the number of inequalities satisfied. The objective function is a step function, and we solve the maximization problem using the differential evolution (DE) algorithm (Price et al., 2005) in Matlab. Given a linear functional form $X'_{ij}\beta$, the estimator $\hat{\beta}$ tells us the importance of each pair characteristic relative to that of a particular pair characteristic whose coefficient is normalized (to $+1$ or -1).

The maximum score estimator has a very complex limiting distribution from which to draw inference, so we use the subsampling procedure (Politis et al., 1999) to obtain confidence intervals. Delgado et al. (2001) show that subsampling gives consistent inference for the maximum score estimator.

4. Data

We obtain the data from three sources. Information on loans comes from the DealScan database produced by the Loan Pricing Corporation. To obtain information on bank characteristics, we match the banks in DealScan to those in the Reports of Condition and Income (known as the Call Reports) from the Federal Reserve Board. To obtain information on firm characteristics, we match the firms in DealScan to those in the Compustat database, a product of Standard & Poor's.

4.1. Sample

The DealScan database provides extensive coverage of bank lending to businesses in the U.S. The majority of the information comes from commitment letters and credit agreements in Securities and Exchange Commission filings, but data from large loan syndicators and the Loan Pricing Corporation's own staff of reporters are also collected. From DealScan, we obtain the identities of the borrower and the lender for each loan.

Our sample consists of 2429 loans between U.S. banks and U.S. firms from 2000 to 2003. We define markets by time and divide all loans into eight markets, each containing the lenders and the borrowers in a same half-year: January to June or July to December.[6] We use data on banks' and firms' characteristics from the quarter that precedes the market.

4.2. Variables

The matching of banks and firms (μ) is given by the names of the matched agents recorded in the loan data.

---

[6] Changing the market definition from a half-year to a year leaves our main findings unaffected, as described in Subsection 5.5 below.

We consider the following agent-specific characteristics: bank's size (total assets), bank's salaries–expenses ratio (salaries and benefits/total operating expenses), firm's size (total assets), and firm's PP&E–assets ratio (property, plant, and equipment/total assets). In the empirical model we include the interaction terms of these bank and firm characteristics in the match value function.

We also include two pair characteristics in the estimation. First, we calculate the distance between each pair using data on the ZIP codes of the banks and the firms. This pair characteristic indicates the physical proximity between the pair. Second, we construct a Prior Loan Dummy, which equals 1 if the pair had a loan in the preceding three years.[7] This dummy variable indicates whether the pair has previously matched.[8]

The hypotheses regarding how banks and firms' interaction terms and pair characteristics may affect the matching are discussed in the next section.

Bank assets and firm assets are deflated using the GDP (Chained) Price Index reported in the Historical Tables in the Budget of the United States Government for Fiscal Year 2005, with the year 2000 being the base year. Table 1 lists the definitions of the variables, and Table 2 presents summary statistics.

5. Findings

In this section, we test several hypotheses regarding the matching of banks and firms, and discuss the implications of our findings.

5.1. Positive assortative matching of sizes

The hypothesis regarding the effect of bank and firm sizes on the matching is that large banks tend to lend to large firms and small banks tend to lend to small firms. This pattern is called the positive assortative matching of sizes. Prior literature has found some evidence of this pattern; see, for example, Hubbard et al. (2002), Cole et al. (2004), and Berger et al. (2005).

The main reasons for positive assortative matching in sizes in the loan market are as follows. First, large and small banks typically use different criteria in selecting borrowers, and their different criteria tend to be favorable for large and small firms, respectively. According to Cole et al. (2004), large banks rely on standardized quantitative criteria to assess loan applications (a "cookie-cutter" approach), while small banks favor qualitative criteria based on loan officers' interactions with loan applicants (a "character" approach). At the same time, large firms are more likely to have well-documented track records and financial information, while small firms tend to focus more on relationship-building. Therefore, a large firm is a better match for a large bank's cookie-cutter approach, whereas a small firm is a better match for a small bank's character approach. In addition, large firms tend to need large loans, and small firms tend to need small loans, which also make them attracted to large banks and small banks, respectively.

To test for positive assortative matching of sizes in the loan market, two OLS regressions using the matched pairs are run: the bank's size on the firm's characteristics and the firm's size on the bank's characteristics. The results are reported in Tables 3 and 4. We find that the bank's size and the firm's size are strongly positively correlated. The coefficients on partners' sizes are both positive and have t statistics of more than 20, indicating that there is a clear pattern of positive assortative matching of sizes.

Fig. 1 provides some visual evidence. We first divide each side of the market (banks or firms) into 5 size groups according to the agents'

Table 1
Variable definitions.

Variable	Definition
Bank's size	Total assets (in billion dollars)
Bank's salaries–expenses ratio	Salaries and benefits/total operating expenses
Firm's size	Total assets (in billion dollars)
Firm's PP&E–assets ratio	Property, plant, and equipment/total assets
Matched dummy	Dummy equals 1 if the bank–firm pair are matched in the current market
Distance	Distance (in miles) between the pair, calculated using data on ZIP codes
Prior loan dummy	Dummy equals 1 if the pair had a loan in the preceding three years

sizes, with each group consisting of 20% of the agents on one side. Fig. 1 depicts the proportion of loans in our sample for each combination of bank and firm size groups. For example, the height of the bar at (2,3) represents the proportion of loans between banks in the second bank size group (with assets between the 20th and the 40th percentiles) and firms in the third firm size group (with assets between the 40th and the 60th percentiles). The figure shows a clear pattern: the tallest bars are mostly around the main diagonal (from (1,1) to (5,5)), whereas the bars far off the main diagonal (e.g., (1,5) and (5,1)) are short. This figure illustrates that most of the loans are between banks and firms that have similar size positions on their respective sides.

5.2. Other hypotheses

Here we consider several additional hypotheses regarding the matching between banks and firms.

5.2.1. Bank's size and firm's tangible assets

We hypothesize that large and small banks' different criteria in assessing borrowers, as described in the previous subsection, also interact with firms' different proportions of tangible assets and affect the matching. Specifically, a firm that has relatively less tangible assets may be a better match for a small bank which uses the character approach rather than the cookie-cutter approach, and vice versa. The reason is the following.

Tangible assets are assets that have a physical substance, such as real estate, buildings, and factories. Their values are easy to assess. In contrast, intangible assets lack physical substance and are usually hard to evaluate. They include patents, trademarks, copyrights, goodwill, etc.

With a cookie-cutter approach, a bank relies on readily available and verifiable information about the borrowers to make a decision. Therefore a firm whose assets are primarily intangible (such as an advertising agency) will be disadvantaged, as there isn't readily available and verifiable information about the value of its intangible assets.

In contrast, with a character approach, a bank evaluates loan applications using information obtained through interactions with borrowers and investigations conducted by loan officers. Therefore a firm

---

[7] In alternative specifications, we also let the prior loan dummy equal 1 if the pair had a loan in the preceding two years and four years, respectively. In order to construct the prior loan dummy, we collected data on loans in the four years that precede our sample period.

[8] A potential issue with the prior loan dummy is that whether the firm had a prior loan from the bank in question is properly a lagged endogenous variable, because it is the outcome of matches in previous periods.

Table 2
Summary statistics.

Variable	# of obs.	Mean	Std. deviation	Min	Max
Bank's size	357	64.88	128.64	0.009	662.00
Bank's salaries–expenses ratio	357	0.27	0.11	0.025	0.71
Firm's size	2429	7.53	36.93	0.001	759.25
Firm's PP&E–assets ratio	2429	0.31	0.24	0.000	1.00
Among potential pairs					
Distance	108,402	1207.51	855.04	0	5089.10
Prior loan dummy	108,402	0.01	0.11	0	1

Table 3
OLS: bank size on firm characteristics.

	Coefficient	Std. error
Constant	4.76	0.07***
Natural log of firm assets	0.49	0.02***
Firm's PP&E–assets ratio	0.31	0.18*

The dependent variable is the natural log of the bank's total assets. *** and * indicate significance at the 1% and 10% levels, respectively.

Table 4
OLS: firm size on bank characteristics.

	Coefficient	Std. error
Constant	−2.10	0.16***
Natural log of bank assets	0.38	0.02***
Bank's salaries–expenses ratio	0.73	0.42*

The dependent variable is the natural log of the firm's total assets. *** and * indicate significance at the 1% and 10% levels, respectively.

Fig. 1. Proportions of loans in different combinations of size groups.

whose assets are primarily intangible will find this approach a better match as the value of its intangible assets is more likely to be recognized.

Consequently, we hypothesize that borrowers with relatively more tangible assets tend to match with large banks, whereas borrowers with relatively more intangible assets tend to match with small banks.

In the estimation, the variable that we use to proxy the relative importance of a firm's tangible assets is its PP&E–assets ratio (property, plant, and equipment/total assets), as PP&E is the main component of a firm's tangible assets.

5.2.2. Bank's monitoring and firm's size

We hypothesize that a bank that is more focused on monitoring (devoting a larger portion of its effort to monitoring loans) may be a better match for a large firm, for the following two reasons.

First, according to Coleman et al. (2006), large firms tend to need large loans, which are typically subject to extensive monitoring by expert credit risk committees. Banks may also be required by regulators and/or board committees to monitor large loans in order to comply with credit exposure requirements. Moreover, large loans are likely to be associated with a greater volume of credit-relevant information which, even if it is publicly disclosed, needs to be monitored and evaluated by loan officers. Therefore, everything else equal, a bank that is more focused on monitoring is more suited to make such loans.

Second, the problem of informational asymmetries between the borrower and the lender may be more severe for small firms than for large firms, as large firms tend to have well-documented records whereas small firms tend to have a higher degree of information opacity. Therefore, due to possible distorted incentives arising from informational asymmetries, a small firm may shy away from banks that are more focused on monitoring.[9]

Coleman et al. (2006) argue that loan monitoring is a labor-intensive task. Although banks have invested heavily in information technology to automate part of the lending process, loans and particularly large loans still require significant labor-intensive monitoring. As a result, the effort in bank monitoring is directly related to the quantity and quality of bank staff, and hence salary expenses. We therefore follow Coleman et al. (2006) and use a bank's salaries–expenses ratio (salaries and benefits/total operating expenses) to proxy the degree to which the bank is focused on monitoring.
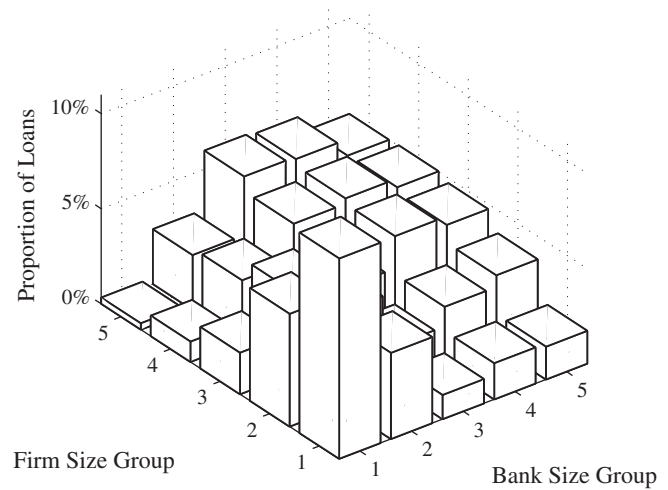
5.2.3. Physical proximity and prior loan relationship

We also include two pair characteristics in our estimation.

First, we expect the physical proximity between the pair to play an important role in matching. Physical proximity of the firm to the bank should allow the bank to have greater familiarity with the firm's business and facilitate the firm in getting expert advice from the bank. It should also reduce the transaction costs (particularly transportation costs) associated with the loan application and evaluation processes. Therefore in the estimation we include the variable distance, calculated using data on the banks' and the firms' ZIP codes. We expect a negative coefficient for this variable (implying a positive effect of physical proximity on the match value).

Second, we expect the existence of a prior loan relationship to affect the matching. Everything else equal, it is more desirable for a borrower to stay with the previous lender rather than switching to a different lender, due to the existence of switching costs. According to Kim et al. (2003), a switch between suppliers in the market for loans entails the explicit costs of closing an account with one bank and opening it elsewhere, as well as the implicit costs associated with the foregone value of the previously established customer–bank relationship. Therefore in the estimation we include a prior loan dummy, which equals 1 if the pair had a loan in the preceding three years. We expect a positive coefficient for this dummy variable.[10]

5.3. Estimation results

In the estimation we include the following interaction terms/pair characteristics as explanatory variables for the two-sided matching: Bank's Size×Firm's Size, Bank's Size×Firm's PP&E–Assets Ratio, Bank's Salaries–Expenses Ratio×Firm's Size, Distance, and Prior Loan Dummy. The results are reported in Table 5.

Based on the evidence that we described above for the positive assortative matching of sizes in the loan market, we use Bank's Size×Firm's Size as the normalized variable. We estimate first fixing the coefficient of Bank's Size×Firm's Size at +1 and then fixing it at −1. We then take the set of estimates that gives the larger objective function value as the final set of estimates. In the estimation, the point estimate for Bank's Size×Firm's Size was always +1 (consistent with positive, rather

---

[9] We thank a referee for a useful comment about this issue.

[10] We thank a referee for suggesting this issue to us.

than negative, assortative matching of sizes). The estimate of this nor-malized variable is superconsistent, and we do not report a confidence interval for it.

We run the differential evolution (DE) algorithm (Price et al., 2005) in Matlab to solve for the maximizer of the maximum score objective function (6). For point estimates, we run the optimization algorithm 30 times with different initial population members and select the set of estimates that gives the largest objective function value. For confi-dence intervals, we follow the subsampling procedure described in Politis et al. (1999) and use subsample size equal to 1/8 of the total sam-ple size.

The coefficient estimate for distance is negative and significant (the 95% confidence interval does not contain 0), indicating that banks and firms prefer to match with partners that are geographically close, consistent with our hypothesis. An implication of this finding is that state and local officials who aim at spurring industrial invest-ment in the local economy should consider providing assistance to strengthen local banks, because these are the banks who are more likely to extend credit to local businesses whereas banks located else-where won't be as helpful.

The coefficient estimate for prior loan dummy is positive and sig-nificant, indicating that banks and firms prefer to match with partners with whom they have prior loan relationship. This finding suggests that switching costs may play a significant role in the bank loan market. To the extent that these switching costs, mostly borne by the borrowers, result in the borrowers being partially "locked-in" to their previous lenders and hence restrict the competition among banks, a reduction in switching costs can help promote competition and improve welfare. Therefore, regulators should consider public policies such as the imple-mentation of switching facilities (objective and up-to-date comparison sites, switching services, etc.) to reduce switching costs in the loan market.

The very large magnitude of the estimated coefficient (1.34E + 05) for the prior loan dummy not only reflects the importance of this var-iable in affecting the matching, but also reflects the fact that the prior loan dummy is a 0–1 dummy variable, whereas observations on the other interaction terms and pair characteristics are on average much larger. For example, as reported in Table 2, observations on the variable distance (in miles) has mean 1207.51 and standard deviation 855.04, whereas observations on the prior loan dummy has mean 0.01 and stan-dard deviation 0.11.

The two other interaction terms included in the estimation, Bank's Size × Firm's PP&E–Assets Ratio and Bank's Salaries–Expenses Ratio × Firm's Size, turn out to be insignificant, indicating that the hy-potheses behind these two variables are not borne out in our data. A possible reason is that our sample contains mostly medium to large businesses that are publicly traded. That is because DealScan's main source of information on loans is the Securities and Exchange Commis-sion filings. Such filings primarily pertain to public companies, which are typically at least medium-sized. Given the relatively large sizes of the firms in our sample and the fact that they are publicly traded, the issue of information opacity is then not an important concern. Conse-quently, the distinction between tangible assets and intangible assets when it comes to valuation becomes less pronounced, and hence the interaction between the bank's size and the firm's tangible assets is not a significant factor in the matching. Furthermore, for these medium to large businesses that are publicly traded, the problem of information-al asymmetries between the borrower and the lender is not severe, and hence the interaction between the bank's monitoring and the firm's size is not a significant factor in the matching, either.

In addition to reporting the point estimates and the confidence intervals, Table 5 also reports that there are 288,091 inequalities used in the estimation. Of these inequalities, 84.3% are satisfied at the reported point estimates. The percentage of satisfied inequal-ities gives a measure of statistical fit, and in our estimation the fit is good.

Table 5

Maximum score estimates of match value function: baseline.

	Point estimate	95% Confidence interval
Bank's size × firm's size	+1	Superconsistent
Bank's size × firm's PP&E–assets ratio	−0.02	(−0.89, 2.64)
Bank's salaries–expenses ratio × firm's size	−84.98	(−243.52, 559.29)
Distance	−1.71	(−2.34, −0.34)
Prior loan dummy	1.34E + 05	(1.00E + 05, 1.82E + 05)
Number of inequalities	288,091	
% of inequalities satisfied	84.3%	

We run the differential evolution (DE) algorithm (Price et al., 2005) in Matlab to solve for the maximizer of the maximum score objective function. We estimate first fixing the coefficient of "Bank's Size × Firm's Size" at +1 and then fixing it at −1. We then take the set of estimates that gives the larger objective function value as the final set of estimates. The estimate of the parameter that can take only one of two values is superconsistent, and we do not report a confidence interval.
For point estimates, we run the optimization algorithm 30 times with different initial population members and select the set of estimates that gives the largest objective function value. For confidence intervals, we follow the subsampling procedure described in Politis et al. (1999) and use subsample size equal to 1/8 of the total sample size.

5.4. Interpreting the parameter estimates

In this subsection we provide interpretation of the parameter esti-mates reported above, by examining the relative importance of the co-variates implied by the estimates and by conducting counterfactuals.[11]

5.4.1. Relative importance of covariates in match value

Here we assess which covariate is more important to the match value, given the parameter estimates and the variation in the covari-ates in the sample.

For each of the covariates that are significant or superconsistent in the match value function (Bank's Size × Firm's Size, Distance, and Prior Loan Dummy), we calculate how the match value would change if the value of the covariate is increased by one standard deviation of the covariate. Panel (1) of Table 6 reports the results based on the point estimates. We see that if we increase Bank's Size × Firm's Size by one standard deviation, the match value between the pair would increase by 5401 (= point estimate × standard deviation of the covar-iate). In comparison, if we increase Distance by one standard devia-tion, the match value would decrease by 1461, which is only 27% of 5401, indicating that given the variation in the covariates, Bank's Size × Firm's Size plays a larger role in determining the match value than Distance. As for Prior Loan Dummy, a one-standard-deviation increase in this covariate increases the match value by 1.53E + 04, which is about 3 times larger than 5401, suggesting that this covariate has the strongest impact on the match value.

In addition to looking at the point estimates, we also consider the confidence intervals. Panels (2) and (3) of Table 6 report the results based on the lower bounds and the upper bounds of the confidence intervals, respectively. Consistent with the finding from Panel (1), these results show that among the three covariates, Prior Loan Dummy has the largest effect on the match value, followed by Bank's Size × Firm's Size, while Distance has the smallest effect.

5.4.2. Counterfactuals

Here we conduct two counterfactuals to better understand the im-plications of the parameter estimates.

The first counterfactual considers the scenario in which Distance does not matter at all in the match value, and examines how much more as-sortative matching on bank and firm sizes there is. The procedure is as follows.

---

[11] We thank a referee for suggesting useful ways to interpret the parameter estimates.

Table 6
Relative importance of covariates in match value.

(1) Using point estimate

	Point estimate	Standard deviation of covariate	Point estimate × standard deviation
Bank's size × firm's size[a]	1	5401.14	5401.14
Distance	−1.71	855.04	−1461.11
Prior loan dummy	1.34E + 05	0.11	1.53E + 04

(2) Using lower bound of confidence interval

	Lower bound of confidence interval	Standard deviation of covariate	Lower bound × standard deviation
Bank's size × firm's size[a]	1	5401.14	5401.14
Distance	−2.34	855.04	−1997.25
Prior loan dummy	1.00E + 05	0.11	1.14E + 04

(3) Using upper bound of confidence interval

	Upper bound of confidence interval	Standard deviation of covariate	Upper bound × standard deviation
Bank's size × firm's size[a]	1	5401.14	5401.14
Distance	−0.34	855.04	−293.56
Prior loan dummy	1.82E + 05	0.11	2.06E + 04

[a] The normalized covariate, bank's size × firm's size, is superconsistent.

To "turn off" the covariate Distance, we set its coefficient to zero. For all other covariates, we use the point estimates from the maximum score estimation (Table 5). We plug these coefficients in the match value function to compute the match values for all the potential bank–firm pairs. We then use a top–down algorithm, described in Appendix A, to construct the counterfactual stable matching given these match values.

For both the actual matching (where Distance matters) and the counterfactual matching (where Distance doesn't matter), we compute the size percentiles correlation, that is, the correlation coefficient between the bank's size percentile (the percentile of its size among all banks) and the firm's size percentile (the percentile of its size among all firms) in matched pairs. We find that the size percentiles correlation is 0.44 in the actual matching, and increases by 46% to 0.65 in the counterfactual matching. This result indicates that using the size percentiles correlation as the measure, turning off Distance strengthens the assortative matching on bank and firm sizes by almost half.

In the second counterfactual, we turn off the covariate Bank's Size × Firm's Size and examine how much closer (in terms of distance) the bank and the firm in matched pairs are. We find that the median distance between the bank and the firm in matched pairs is 936 miles in the actual matching, and decreases by 37% to 589 miles in the counterfactual matching. This result shows that using the median distance as the measure, turning off the assortative matching on bank and firm sizes shortens the distance between the bank and the firm in matched pairs by more than one third.

5.5. Alternative specifications

To assess the robustness of our findings, we estimate the model with the following alternative specifications.

First, choosing the 6 month time frame for the definition of a market is somewhat arbitrary, so as a robustness check, we estimate a specification in which a market contains the firms that borrow during one year (instead of 6 months) and the banks that lend to them. The results are reported in Table 7. These results show that our findings are robust to this change in market definition: for the coefficients that are estimated to be statistically significant (distance and prior loan dummy), changing the market definition causes only minor changes in the point estimates and the confidence intervals.

Second, we test whether different definitions for the prior loan dummy affect our findings. In two alternative specifications, we let the prior loan dummy equal 1 if the pair had a loan in the preceding two years and four years, respectively, instead of three years. The results are reported in Tables 8 and 9, respectively. We again find that our findings are robust, as changing the definition of the prior loan dummy causes only minor changes in the point estimates and the confidence intervals for the coefficients that are estimated to be statistically significant (distance and prior loan dummy).

6. Conclusion

This paper investigates the two-sided matching between banks and firms in the loan market using a U.S. loan sample from 2000 to 2003. Estimation uses the Fox (2010) matching maximum score estimator. We find evidence of positive assortative matching of sizes: large banks tend to match with large firms, and small banks tend to match with small firms. Moreover, we show that banks and firms prefer partners that are geographically closer, giving support to the importance of physical proximity for information gathering and expertise sharing. We also show that banks and firms prefer partners with whom they

Table 7
Maximum score estimates of match value function: alternative market definition[a].

	Point estimate	95% Confidence interval
Bank's size × firm's size	+1	Superconsistent
Bank's size × firm's PP&E–assets ratio	0.46	(−0.22, 1.11)
Bank's salaries–expenses ratio × firm's size	54.16	(−314.85, 582.98)
Distance	−1.58	(−2.11, −0.25)
Prior loan dummy	2.05E + 05	(1.79E + 05, 2.79E + 05)
Number of inequalities	656,750	
% of inequalities satisfied	79.5%	

[a] Each market contains the firms that borrow during one year and the banks that lend to them.

Table 8
Maximum score estimates of match value function: alternative definition 1 of prior loan dummy[a].

	Point estimate	95% Confidence interval
Bank's size × firm's size	+1	Superconsistent
Bank's size × firm's PP&E–assets ratio	−0.15	(−1.52, 1.94)
Bank's salaries-expenses ratio × firm's size	−78.98	(−258.16, 359.10)
Distance	−2.08	(−2.85, −0.87)
Prior loan dummy	1.40E + 05	(1.46E + 05, 1.88E + 05)
Number of inequalities	288,091	
% of inequalities satisfied	83.2%	

[a] Prior loan dummy equals 1 if the pair had a loan in the preceding two years.

Table 9
Maximum score estimates of match value function: alternative definition 2 of prior loan dummy[a].

	Point estimate	95% Confidence interval
Bank's size × firm's size	+1	Superconsistent
Bank's size × firm's PP&E–assets ratio	0.03	(−0.59, 2.40)
Bank's salaries–expenses ratio × firm's size	−124.35	(−254.20, 322.98)
Distance	−1.87	(−2.56, −0.77)
Prior loan dummy	1.28E + 05	(1.19E + 05, 1.73E + 05)
Number of inequalities	288,091	
% of inequalities satisfied	84.9%	

[a] Prior loan dummy equals 1 if the pair had a loan in the preceding four years.

had prior loans, indicating that prior loan relationship plays an important role in the selection of current partners.

We have the potential to learn a lot about the workings of various markets and their impact on the economy and the society by investigating the matching between the market participants. For example, estimation of the two-sided matching between colleges and students can shed light on how colleges and students choose each other, and how colleges' characteristics and students' attributes may interact to affect students' academic achievements, job outcomes, and professional contributions after college. Other examples include the matchings between sports teams and players (in NBA, for instance), corporations and CEOs, firms and underwriters, and so on. Because estimation of the two-sided matching model enables us to identify the interaction terms and pair characteristics that contribute to the match values, it can point the way for agents who try to obtain better matches, such as colleges that want to expand their enrollment of a certain category of students, and sports teams who try to attract a certain type of players, etc. Future research in this direction will likely prove fruitful.

Appendix A. Constructing stable matching using match values

For each market t we construct V_t, an $I_t \times J_t$ match value matrix, whose (i, j)th element is the match value between bank i and firm j (here I_t and J_t denote the number of banks and the number of firms in market t, respectively). Assume that all the match values in V_t are distinct. Based on this match value matrix, we obtain the stable matching using the following top–down algorithm. In step 1, use the original V_t as V^1 (where the superscript 1 indicates step 1) and locate the largest element in V^1. Let i^1 and j^1 denote the bank and the firm corresponding

to this element, respectively. Record i^1 and j^1 as a matched pair in the matching. Next subtract bank i^1's quota (that is, the number of firms that bank i^1 lends to; see Subsection 2.1 Agents and Quotas) by 1. Then obtain a new match value matrix V^2 by: (1) deleting firm j^1's column in V^1, and (2) deleting bank i^1's row in V^1 if and only if its quota is reduced to 0 in this step. Now go to step 2, which uses V^2 as the match value matrix and carries out the same procedure as in step 1 to obtain i^2, j^2, and V^3. Record i^2 and j^2 as a matched pair in the matching, and proceed to step 3, which uses V^3 as the match value matrix. Repeat until all the banks and firms are matched. It is straightforward to show that this procedure produces the unique stable matching given the match values in V_t.

References

Berger, A.N., Miller, N.H., Peterson, M.A., Rajan, R.G., Stein, J.C., 2005. Does function follow organizational form? Evidence from the lending practices of large and small banks. Journal of Financial Economics 76 (2), 237–269.

Chen, J., 2013. Estimation of the loan spread equation with endogenous bank–firm matching. Working paper. University of California, Irvine.

Cole, R.A., Goldberg, L.G., White, L.J., 2004. Cookie cutter vs. character: the micro structure of small business lending by large and small banks. Journal Of Financial And Quantitative Analysis 39, 227–251.

Coleman, A.D.F., Esho, N., Sharpe, I.G., 2006. Does bank monitoring influence loan contract terms? Journal of Financial Services Research 30, 177–198.

Delgado, M.A., Rodríguez-Poo, J.M., Wolf, M., 2001. Subsampling inference in cube root asymptotics with an application to Manski's maximum score estimator. Economics Letters 73, 241–250.

Fox, J., 2010. Estimating matching games with transfers. Working paper. University of Michigan.

Hubbard, R.G., Kuttner, K.N., Palia, D.N., 2002. Are there bank effects in borrowers' costs of funds? Evidence from a matched sample of borrowers and banks. Journal of Business 75, 559–581.

James, C., Smith, D.C., 2000. Are banks still special? New evidence on their role in the corporate capital-raising process. Journal of Applied Corporate Finance 13 (1), 52–63.

Jayaratne, J., Morgan, D.P., 2000. Capital market frictions and deposit constraints at banks. Journal of Money, Credit, and Banking 32 (1), 74–92.

Kashyap, A.K., Stein, J.C., 1994. Monetary policy and bank lending. In: Mankiw, N.G. (Ed.), Monetary Policy. University of Chicago Press, pp. 221–256.

Kim, M., Kliger, D., Vale, B., 2003. Estimating switching costs: the case of banking. Journal of Financial Intermediation 12, 25–56.

Politis, D.N., Romano, J.P., Wolf, M., 1999. Subsampling. Springer.

Price, K., Storn, R.M., Lampinen, J.A., 2005. Differential Evolution: A Practical Approach to Global Optimization. Springer.

Roth, A., Sotomayor, M., 1990. Two-sided matching: a study in game-theoretic modeling and analysis. Econometric Society Monograph Series.Cambridge University Press.

Saidenberg, M.R., Strahan, P.E., 1999. Are banks still important for financing large businesses? Current Issues in Economics and Finance, vol. 5:12. Federal Reserve Bank of New York, pp. 1–6.

Shapley, L.S., Shubik, M., 1971. The assignment game I: The core. International Journal of Game Theory 1, 111–130.