Estimation of the Loan Spread Equation with Endogenous Bank-Firm Matching

Jiawei Chen*

March 2013

Abstract

This paper estimates the loan spread equation taking into account the endogenous matching between banks and firms in the loan market. To overcome the endogeneity problem, I supplement the loan spread equation with a two-sided matching model and estimate them jointly. Bayesian inference is feasible using a Gibbs sampling algorithm that performs Markov chain Monte Carlo simulations. I find that medium-sized banks and firms tend to be the most attractive partners, and that liquidity is also a consideration in choosing partners. Furthermore, banks with higher monitoring ability charge higher spreads, and firms that are more leveraged or less liquid are charged higher spreads.

Keywords: Loan Spread Equation, Two-Sided Matching, Bayesian Inference, Gibbs Sampling

---

*Department of Economics, 3151 Social Science Plaza, University of California, Irvine, CA 92697-5100. E-mail: jiaweic@uci.edu. I thank Jan Brueckner, Linda Cohen, Joseph Harrington, Ivan Jeliazkov, Ali Khan, Robert Moffitt, Dale Poirier, Matt Shum, Tiemen Woutersen, and seminar participants at Brown, Iowa, Johns Hopkins, St. Louis Fed, UC Irvine, USC, Williams College, and the 13th Advances in Econometrics Conference for their helpful comments.

1 Introduction

Bank loans are an important source of credit to firms. For instance, in 2010 the U.S. marketplace for bank loans is roughly $1.5 trillion in size, representing 17% of the overall corporate debt issuance and making bank loans a significant component of the capital markets (Stoeckle, 2011). Moreover, bank loans constitute the critical "lending" channel of monetary policy transmission and have substantial impact on investment and aggregate economic activity (Kashyap and Stein, 1994).

Not surprisingly, empirical researchers have long been interested in the pricing of bank loans. In particular, loan spreads (markups of loan interest rates over a benchmark rate) are regressed on the characteristics of banks, firms, and loans to examine the relationship between collateral and risk in financial contracting (Berger and Udell, 1990), and to provide evidence of the bank lending channel of monetary transmission (Hubbard, Kuttner, and Palia, 2002). However, the non-randomness of the bank-firm pairs in the loan samples is typically ignored.

In this paper, we argue that banks and firms prefer to match with partners that have higher quality, so banks choose firms, firms choose banks, and the matching outcome is endogenously determined. We show that, as a result, the regressors in the loan spread equation are correlated with the error term. In order to overcome this endogeneity problem, we develop a two-sided matching model to supplement the loan spread equation. We find that medium-sized banks and firms tend to be the most attractive partners, and that liquidity is also a consideration in choosing partners. Furthermore, banks with higher monitoring ability charge higher spreads, and firms that are more leveraged or less liquid are charged higher spreads.

Both firms and banks have strong economic incentives to select their partners. When a bank lends to a firm, the bank not only supplies credit to the firm but also provides monitoring, expert advice, and endorsement based on reputation (e.g. Diamond, 1984 and 1991). Empirical evidence suggests that those "by-products" are important for firms. For instance, Billet, Flannery and Garfinkel (1995) and Johnson (1997) show that banks' monitoring ability and reputation have significant positive effects on borrowers' performance in the stock market.

The size of a bank—the amount of its total assets—also plays an important role in firms' choices. First, a larger bank is likely to have better diversified assets and a lower risk, making it more attractive to firms. Second, the small size of a bank may place a constraint on its lending, which is undesirable for a borrowing firm, since its subsequent loan requests could be denied and it might have to find a new lender and pay a switching cost. Third, large banks usually have more organizational layers and face more severe information distortion problems than small banks, so they are generally less effective in processing and communicating borrower information, making them less able to provide valuable client-specific monitoring and expert advice. Fourth, Brickley, Linck and Smith (2003) observe that employees in small to medium-sized banks own higher percentages of their banks' stocks than employees in large banks. As a result the loan officers in small to medium-sized banks have stronger incentives and will devote more effort to collecting and processing borrower information, which helps the banks better serve their clients. Thus the size of a bank has multiple effects on its quality perceived by firms and those effects operate in opposite directions. Which bank size is most attractive is determined by the net effect.

Banks' characteristics affect how much benefit borrowing firms will receive, so firms prefer banks that are better in those characteristics, e.g., banks with higher monitoring ability, better reputation, suitable size, and so on. Banks are ranked by firms according to a composite quality index that combines those characteristics.

Now consider banks' choices. In making their lending decisions, loan officers in a bank screen the applicants (firms) and provide loans only to those who are considered creditworthy. Firms with lower leverage ratios (total debt/total assets) or higher current ratios (current assets/current liabilities) are usually considered less risky and more creditworthy. Larger firms also have an advantage here, because they generally have higher repaying ability and better diversified assets, and are more likely to have well-documented track records and lower degrees of information opacity.

However, the large size of a firm also has negative effects on its attractiveness. Because larger firms have stronger financial needs, the loan made to a larger firm usually has a larger amount and accounts for a higher percentage of the bank's assets, thus reducing the bank's diversification. Since banks prefer well diversified portfolios, the large size of a borrowing firm may be considered unattractive. In addition, lending to a large firm means that the bank's control over the firm's investment decisions will be relatively small, which is undesirable.[1] Therefore, the size of a firm also has multiple effects on its quality perceived by banks, and which firm size is most attractive depends on the relative magnitudes of those effects. Firms are ranked by banks according to a composite quality index that combines firms' characteristics, such as their risk and their sizes.

The above analysis shows that there is endogenous two-sided matching in the loan market: banks choose firms, firms choose banks, and they all prefer partners that have higher quality. Thus we need to address the endogeneity issue when we estimate the loan spread equation, as discussed below.

In our model banks' and firms' quality are multidimensional, but to illustrate the consequence of the endogenous matching, we assume for a moment that a bank's quality is solely determined by its liquidity risk, and that a firm's quality is solely determined by its information opacity. Further assume that banks' liquidity risk, firms' information opacity, and non-price loan characteristics such as maturity and loan size are determinants of loan spreads. The spread equation is:

$$r_{ij} = \alpha_0 + \kappa L_i + \lambda I_j + N'_{ij}\alpha_3 + \nu_{ij},\ \nu_{ij} \sim N(0, \sigma^2_\nu), \tag{1}$$

where r_{ij} is the loan spread if bank i lends to firm j, L_i is bank i's liquidity risk, I_j is firm j's information opacity, and N_{ij} is the non-price loan characteristics.

Liquidity risk and information opacity are not perfectly observed, and the bank's ratio of cash to total assets and the firm's ratio of property, plant, and equipment (PP&E) to total assets are used as their proxies, respectively. Assume

$$L_i = \rho C_i + \eta_i,\ \eta_i \sim N(0, \sigma^2_\eta),\ \text{and}$$

$$I_j = \sigma P_j + \delta_j,\ \delta_j \sim N(0, \sigma^2_\delta),$$

where C_i is bank i's ratio of cash to total assets, and P_j is firm j's ratio of PP&E to total assets. Now equation (1) becomes

$$\begin{aligned} r_{ij} &= \alpha_0 + \kappa(\rho C_i + \eta_i) + \lambda(\sigma P_j + \delta_j) + N'_{ij}\alpha_3 + \nu_{ij} \\ &= \alpha_0 + \kappa\rho C_i + \lambda\sigma P_j + N'_{ij}\alpha_3 + \kappa\eta_i + \lambda\delta_j + \nu_{ij}. \end{aligned} \tag{2}$$

[1] See Rajan (1992) for a discussion on banks' control over borrowers' investment decisions.

Note that the error term contains η_i and δ_j, the unobserved quality. Because of the endogenous matching, the unobserved quality of a bank or a firm affects its matching outcome, and therefore correlates with its partner's characteristics. As a result, the regressors in the spread equation are correlated with the error term, giving rise to an endogeneity problem. Such endogeneity problem introduced by the use of proxies in the matching setting is pointed out by Ackerberg and Botticini (2002) in their analysis of contract choices.

The current study takes a full information approach to address the endogeneity problem. We develop a many-to-one two-sided matching model in the loan market that supplements the spread equation to permit non-random matching of banks and firms.[2] In the matching model, the set of participants are different in different markets, which provides exogenous variation in agents' matching outcome and solves the endogeneity problem for the same intuitive reason as the traditional instrumental variable method. In addition to addressing the endogeneity problem in estimation of the loan spread equation, the matching model also allows us to investigate the factors that determine banks' and firms' quality, enabling us to better understand how banks and firms choose each other in the loan market.

Our matching model is a special case of the College Admissions Model, for which an equilibrium matching always exists (Gale and Shapley, 1962; Roth and Sotomayor, 1990). Two-sided matching models are applied to markets in which agents are divided into two sides and each participant chooses a partner or partners from the other side. Examples include the labor market, the marriage market, the education market, and so on. There are a few studies on two-sided matching in financial markets. Fernando, Gatchev, and Spindt (2005) study the matching between firms and their underwriters, and Sorensen (2007) studies the matching between venture capitalists and the companies in which they invest. Park (2008) introduces the use of matching models into the merger setting and studies the incentives of acquirers and targets in the mutual fund industry.

We obtain Bayesian inference using a Gibbs sampling algorithm (Geman and Geman, 1984; Gelfand and Smith, 1990; Geweke, 1999) with data augmentation (Tanner and Wong,

[2] A similar full information approach is used by Ackerberg and Botticini (2002) to overcome the endogeneity problem, in which they include an ordered probit matching equation to supplement the contract choice equation.

1987; Albert and Chib, 1993). The method iteratively simulates each block of the parameters and the latent variables conditional on all the others to recover the joint posterior distribution. It transforms an integration problem into a simulation problem and overcomes the difficulty of integrating a highly nonlinear function over thousands of dimensions, most of which correspond to the latent variables. As a result, computational burden is substantially reduced. Related MCMC algorithms have been applied to the estimation of the optimal job search model (Lancaster, 1997) and the selection model of hospital admissions (Geweke, Gowrisankaran, and Town, 2003), among others. Sorensen's (2007) paper on venture capital is the first study that uses the method to estimate a two-sided matching model.

Our empirical analysis uses a sample of $1,369$ U.S. loan facilities between 455 banks and $1,369$ firms from 1996 to 2003. We find that positive assortative matching of sizes is prevalent in the loan market, that is, large banks tend to match with large firms, and vice versa. We show that for agents on both sides of the market there are similar relationships between quality and size, which lead to similar size rankings for both sides and explain the positive assortative matching of sizes. Medium-sized banks and firms are the most preferred partners, and liquidity is also a consideration in choosing partners. Furthermore, banks with higher monitoring ability charge higher spreads, and firms that are more leveraged or less liquid are charged higher spreads.

The remainder of the paper is organized as follows: Section 2 provides the specification of the model, Section 3 presents the empirical method, Section 4 describes the data, Section 5 presents and interprets the empirical results, and Section 6 concludes.

2 Model

The first component of our model is a spread equation, in which the loan spread is a function of the bank's characteristics, the firm's characteristics, and the non-price characteristics of the loan. A two-sided matching model in the loan market supplements the spread equation to permit non-random matching of banks and firms.

2.1 Spread Equation

We are interested in estimating the following spread equation:

$$r_{ij} = \alpha_0 + B_i'\alpha_1 + F_j'\alpha_2 + N_{ij}'\alpha_3 + \epsilon_{ij} \equiv U_{ij}'\alpha + \epsilon_{ij},\ \ \epsilon_{ij} \sim N(0, \sigma_\epsilon^2), \tag{3}$$

where r_{ij} is the loan spread if bank i lends to firm j, B_i is a vector of bank i's characteristics, F_j is a vector of firm j's characteristics, and N_{ij} is the non-price loan characteristics.

Prior studies, such as Hubbard, Kuttner and Palia (2002) and Coleman, Esho and Sharpe (2006), suggest that the bank's monitoring ability and risk, as well as the firm's risk and information opacity are important determinants of the loan spread. Those characteristics are not perfectly observed, so we follow the literature and use proxies for them in the spread equation. Because estimation of our model is numerically intensive, we choose a parsimonious specification to keep estimation feasible, focusing on a set of key variables.

Bank's Monitoring Ability. According to the hold-up theory in Rajan (1992) and Diamond and Rajan (2000), a bank that has superior monitoring ability can use its skills to extract higher rents. Moreover, Leland and Pyle (1977), Diamond (1984, 1991) and Allen (1990) show that banks' monitoring plays an important role in firms' operation and provides value to them. Therefore, we expect a bank that has higher monitoring ability to charge a higher spread.

A bank's salaries-expenses ratio, defined as the ratio of salaries and benefits to total operating expenses, is a proxy for its monitoring ability. Coleman, Esho and Sharpe (2006) show that monitoring activities are relatively labor-intensive, and that salaries can reflect the staff's ability and performance in these activities.

Bank's Risk. A bank's risk comes from two sources, inadequate capital and low liquidity. Both of them affect the bank's cost of funds and may have an impact on the spread (see, for example, Hubbard, Kuttner and Palia, 2002).

A bank's capital-assets ratio is a proxy for its capital adequacy, and its ratio of cash to total assets is a proxy for its liquidity risk. The size of a bank (its total assets) is also a proxy for its risk, since a larger bank is likely to have better diversified assets and lower risk.

Firm's Risk. Proxies for a firm's risk include the leverage ratio (total debt/total assets), the current ratio (current assets/current liabilities), and the size of the firm.

Risk is positively related to the leverage ratio, so a firm that has a higher leverage ratio is charged a higher spread, all else being equal. On the other hand, a firm with a higher current ratio is more liquid and less risky, so it is typically charged a lower spread. Due to the diversification effects of increasing firm size, firm risk is negatively associated with firm assets, and a larger firm can usually get a loan with a lower spread.

Firm's Information Opacity. In general smaller firms pose larger information asymmetries because they typically lack well-documented track records, so the size of a firm is also a proxy for information opacity.

Another proxy for a firm's information opacity is the ratio of property, plant, and equipment (PP&E) to total assets, which indicates the relative significance of tangible assets in the firm. A firm with relatively more tangible assets poses smaller information asymmetries. Consequently it can borrow at a lower spread, all else being equal.

Non-Price Loan Characteristics. Non-price loan characteristics are included on the right-hand side of the spread equation as control variables. They are maturity (in months), natural log of the loan facility size, purpose dummies such as "acquisition" and "recapitalization", type dummies such as "a revolver credit line with duration shorter than one year", and a secured dummy. The definitions of these variables are presented in Section 4. In this paper we follow the approach in the literature (for example, Berger and Udell, 1990; Hubbard, Kuttner and Palia, 2002; John, Lynch, and Puri, 2003) and take the non-price loan characteristics as exogenous.

2.2 Two-Sided Matching Model

A two-sided matching model is developed to supplement the spread equation and address the endogeneity problem resulting from the non-random matching between banks and firms:

$$
\begin{aligned}
r_{ij} &= \alpha_0 + B_i'\alpha_1 + F_j'\alpha_2 + N_{ij}'\alpha_3 + \epsilon_{ij} \equiv U_{ij}'\alpha + \epsilon_{ij},\ \epsilon_{ij} \sim N(0, \sigma_\epsilon^2), && (4)\\
Q_i^b &= B_i'\beta + \eta_i,\ \eta_i \sim N(0, \sigma_\eta^2), && (5)\\
Q_j^f &= F_j'\gamma + \delta_j,\ \delta_j \sim N(0, \sigma_\delta^2), && (6)\\
m_{ij} &= \mathrm{I}(\text{bank } i \text{ lends to firm } j), && (7)
\end{aligned}
$$

where Q_i^b is the quality index of bank i, Q_j^f is the quality index of firm j, and I(.) is the indicator function. r_{ij}, N_{ij} are observed iff the match indicator $m_{ij} = 1$. η_i and δ_j are allowed to be correlated with ϵ_{ij}.

In the two-sided matching model, whether m_{ij} equals one or zero is determined by both banks' choices and firms' choices, and the outcome corresponds to the unique equilibrium matching (defined later), which depends on the Q_i^b's and the Q_j^f's.

We assume that in the matching process, each agent has complete information on the quality indexes of the agents on the other side of the market. We consider this assumption a reasonable approximation of the portion of the loan market that we analyze, since we focus on the upper portion of the market, which consists of large banks and large, publicly traded firms. As a result, a two-sided matching model is more suitable here than a job-search model.

Miller and Bavaria (2003) and Yago and McCarthy (2004) document that during the 1990's, "market-flex language" became common in the loan market, which lets the pricing of a loan be determined after the loan agreement is made. This practice is consistent with the fact that banks commonly employ loan pricing formulas and strive to keep the weights assigned to different risk factors constant for a given period of time across all borrowers, leaving not much room for negotiation on loan spreads (Bhattacharya, 1997, Pages 688-689). Because of these institutional features, in this paper we model the loan spreads as determined by the characteristics of banks, firms, and loans, and do not incorporate endogenous transfers between partners (transferable utility) into our matching model.[3]

[3] Chen and Song (2013) consider a transferable utility matching model in the loan market, in which the

Agents, Quotas and Matches. 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. The market subscript t is sometimes dropped to simplify the notation.

In the empirical implementation of our model, a market is specified to contain all the firms that borrow during a half-year and all the banks that lend to them. In the data the vast majority of firms borrow only once during a half-year, whereas a bank often lends to multiple firms. We therefore model the loan market using a many-to-one two-sided matching model, also known as the College Admissions Model (Gale and Shapley, 1962; Roth and Sotomayor, 1990). In market t, bank i lends to q_{it} firms and firm j borrows from only one bank. q_{it} 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.

The set of all potential loans, or *matches*, is given by $M_t = I_t \times J_t$. A *matching*, μ_t, is a set of matches such that $(i, j) \in \mu_t$ if and only if bank i and firm j are matched in market t.

Let $\mu_t(i)$ denote the set of firms that borrow from bank i in market t, and let $\mu_t(j)$ denote the set of banks that lend to firm j in market t, which is a singleton. We then have

$$m_{ij} = 1 \iff (i, j) \in \mu_t \iff j \in \mu_t(i) \iff i \in \mu_t(j) \iff \{i\} = \mu_t(j).$$

Equilibrium Matching. The matching of banks and firms is determined by the equilibrium outcome of a two-sided matching process. The payoff firm j receives if it borrows from bank i is Q_i^b, and the payoff bank i receives if it lends to the firms in the set $\mu_t(i)$ is $\sum_{j \in \mu_t(i)} Q_j^f$. Consequently, each bank prefers firm j to firm j' iff $Q_j^f > Q_{j'}^f$, and each firm prefers bank i to bank i' iff $Q_i^b > Q_{i'}^b$. The quality indexes are assumed to be distinct so there are no "ties".

A matching is an equilibrium if it is stable, that is, if there is no blocking coalition of agents. A coalition of agents is blocking if they prefer to deviate from the current matching

(price and non-price) characteristics of loan are endogenously determined at the time of the matching, and an agent is willing to trade away match quality in order to obtain better terms in the loan. They investigate the two-sided matching in the loan market but not the loan spread equation—the latter assumes that loan spreads are determined by the characteristics of banks, firms, and loans, and therefore is not compatible with the transferable utility matching framework.

and form new matches among them. Formally, μ_t is an equilibrium matching in market t iff there does not exist $\tilde{I} \subset I_t$, $\tilde{J} \subset J_t$ and $\tilde{\mu}_t \neq \mu_t$ such that $\tilde{\mu}_t(i) \subset \tilde{J} \cup \mu_t(i)$ and $\sum_{j \in \tilde{\mu}_t(i)} Q_j^f > \sum_{j \in \mu_t(i)} Q_j^f$ for all $i \in \tilde{I}$, and $\tilde{\mu}_t(j) \in \tilde{I}$ and $Q^b_{\tilde{\mu}_t(j)} > Q^b_{\mu_t(j)}$ for all $j \in \tilde{J}$.

The above stability concept is group stability. A related stability concept is pair-wise stability. A matching is pair-wise stable if there is no blocking pair. In the College Admissions Model, Roth and Sotomayor (1990) prove that pair-wise stability is equivalent to group stability and that an equilibrium always exists. Appendix A shows that there exists a unique equilibrium matching in our model, which is a special case of the College Admissions Model.

The unique equilibrium matching is characterized by a set of inequalities, based on the fact that there is no blocking bank-firm pair. For each bank, stability requires that its worst borrower be better than any other firm whose lender is worse than this bank. Similarly, for each firm, stability requires that its lender be better than any other bank whose worst borrower is worse than this firm. Appendix B derives the lower and upper bounds on the agents' quality indexes, $\underline{Q}_i^b$, $\underline{Q}_j^f$, $\overline{Q}_i^b$, and $\overline{Q}_j^f$, such that:

$$\mu_t = \mu_t^e \iff Q_i^b \in (\underline{Q}_i^b, \overline{Q}_i^b), \forall i \in I_t \text{ and } Q_j^f \in (\underline{Q}_j^f, \overline{Q}_j^f), \forall j \in J_t, \tag{8}$$

where μ_t^e denotes the unique equilibrium matching in market t. This characterization of the equilibrium matching is used in the estimation method in the next section.

3 Estimation

Two-sided matching in the loan market presents numerical challenges when it comes to estimation. Maximum likelihood estimation requires integrating a highly nonlinear function over thousands of dimensions, most of which correspond to the latent quality indexes. Instead we use a Gibbs sampling algorithm that performs Markov chain Monte Carlo (MCMC) simulations to obtain Bayesian inference, and augment the observed data with simulated values of the latent data on quality indexes so that the augmented data are straightforward to analyze. The method iteratively simulates each block of the parameters and the latent variables conditional on all the others to recover the joint posterior distribution. It transforms a high-dimensional integration problem into a simulation problem and substantially reduces the computational burden.

In our model, variation in the set of participants across different markets provides exogenous variation in agents' matching outcome, which solves the endogeneity problem for the same intuitive reason as the traditional instrumental variable method.

3.1 Error Terms and Prior Distributions

Estimation of the quality index equations is subject to the usual identification constraints in discrete choice models, so σ_η and σ_δ are set to one to fix the scales, and the constant and market characteristics are excluded to fix the levels.

To allow for correlation among the error terms, we assume $\epsilon_{ij} = \kappa\eta_i + \lambda\delta_j + \nu_{ij}$, $\nu_{ij} \sim N(0, \sigma_\nu^2)$, with

$$\begin{pmatrix} \epsilon_{ij} \\ \eta_i \\ \delta_j \end{pmatrix} \sim N\left(\mathbf{0}, \begin{bmatrix} \kappa^2 + \lambda^2 + \sigma_\nu^2 & \kappa & \lambda \\ \kappa & 1 & 0 \\ \lambda & 0 & 1 \end{bmatrix}\right).$$

The signs in the two-sided matching model are identified by requiring λ to be non-positive, as theory predicts that firms with higher unobserved quality (lower unobserved risk or lower degrees of unobserved information opacity) are charged lower loan spreads, everything else being equal.[4]

The prior distributions are multivariate normal for α, β, γ, normal for κ, and truncated normal for λ (truncated on the right at 0). The means of these prior distributions are zeros, and the variance-covariance matrices are $10I$, where I is an identity matrix. The prior distribution of $1/\sigma_\nu^2$ is gamma, $1/\sigma_\nu^2 \sim G(2, 1)$. The above are diffuse priors that include reasonable parameter values well within their supports. We try larger variances and other changes in the priors and the estimates are left almost unchanged. For any parameter, the variance of the prior distribution is at least 233 times the variance of the

[4] For the model to be identified, the sign of one of the parameters in β, γ, κ and λ must be specified, because both (β, γ, κ, λ) and ($-\beta$, $-\gamma$, $-\kappa$, $-\lambda$) would be admissible given the same observed (endogenous and exogenous) variables. Here we follow theory prediction and assume λ to be non-positive. If the truth is $\lambda > 0$, we would obtain a set of estimates of the quality index equations in which all the good factors have negative coefficients and all the bad factors have positive coefficients. In a sense, it is theory prediction, not the empirical method, that tells which of the two exactly opposite scenarios is reasonable.

posterior distribution, showing that the information contained in the Bayesian inference is substantial.

3.2 Conditional Posterior Distributions

In the model, the exogenous variables are B_i, F_j, and N_{ij}, which are abbreviated as X. The observed endogenous variables are r_{ij} (the loan spread) and m_{ij} (the match indicator). The unobserved quality indexes are Q_i^b and Q_j^f. The parameters are α, β, γ, κ, λ, and $1/\sigma_\nu^2$, which are abbreviated as θ. In market t, let X_t, r_t, μ_t and Q_t^* represent the above variables, where μ_t embodies all the m_{ij}'s and Q_t^* denotes all the quality indexes.

The joint density of the endogenous variables and the quality indexes conditional on the exogenous variables and the parameters is as follows:

$$
\begin{aligned}
p(r_t, \mu_t, Q_t^* \mid X_t, \theta) = \mathrm{I}\left(Q_i^b \in (\underline{Q}_i^b, \overline{Q}_i^b),\ \forall i \in I_t \text{ and } Q_j^f \in (\underline{Q}_j^f, \overline{Q}_j^f),\ \forall j \in J_t\right) \\
\times \prod\nolimits_{(i,j)\in\mu_t} \phi\left(r_{ij} - U_{ij}'\alpha - \kappa(Q_i^b - B_i'\beta) - \lambda(Q_j^f - F_j'\gamma); 0, \sigma_\nu^2\right) \\
\times \prod\nolimits_{i\in I_t} \phi(Q_i^b - B_i'\beta; 0, 1) \times \prod\nolimits_{j\in J_t} \phi(Q_j^f - F_j'\gamma; 0, 1),
\end{aligned}
\tag{9}
$$

where $\mathrm{I}(.)$ is the indicator function and $\phi(.;\mu,\sigma^2)$ is the $N(\mu,\sigma^2)$ pdf. To obtain the likelihood function for market t $L_t(\theta) = p(r_t, \mu_t \mid X_t, \theta)$, we need to integrate $p(r_t, \mu_t, Q_t^* \mid X_t, \theta)$ over all possible values of the quality indexes. Due to endogenous matching in the market, the bounds on each agent's quality index depend on other agents' quality indexes, so the integral can not be factored into a product of lower-dimensional integrals. The Gibbs sampling algorithm with data augmentation transforms this high-dimensional integration problem into a simulation problem and makes estimation feasible.

To keep our study tractable, we model the markets as independent, so the product of $p(r_t, \mu_t, Q_t^* \mid X_t, \theta)$ for $t = 1, 2, ..., T$ gives the joint density $p(r, \mu, Q^* \mid X, \theta)$ for all the markets. From Bayes' rule, the density of the posterior distribution of Q^* and θ conditional on the data is

$$
\begin{aligned}
p(Q^*, \theta \quad | \quad & X, r, \mu) = p(\theta) \times p(r, \mu, Q^* \mid X, \theta)/p(r, \mu \mid X) \\
\propto \quad & p(\theta) \times p(r, \mu, Q^* \mid X, \theta),
\end{aligned}
\tag{10}
$$

where $p(\theta)$ is the prior densities of the parameters. The conditional posterior distributions are described in Appendix C. They are truncated normal for Q_i^b, Q_j^f, and λ, multivariate normal for α, β, and γ, normal for κ, and gamma for $1/\sigma_\nu^2$.

3.3 Simulation

In the algorithm, the parameters and the quality indexes are partitioned into blocks. Each of the parameter vectors (α, β, γ, κ, λ, and $1/\sigma_\nu^2$) and the quality indexes is a block. In market t the number of quality indexes is equal to the number of agents, $|I_t| + |J_t|$, so altogether we have $\sum_{t=1}^{T}(|I_t|+|J_t|)+6$ blocks. In each iteration of the algorithm, each block is simulated conditional on all the others according to the conditional posterior distributions, and the sequence of draws converge in distribution to the joint distribution.[5]

Estimation results reported in Section 5 are based on $20,000$ draws from which the initial $2,000$ are discarded to allow for burn-in. Visual inspection of the draws shows that convergence to the stationary posterior distribution occurs within the burn-in period. To formally examine whether the posterior simulator is correct and whether convergence has been achieved, three sets of tests are conducted. First, joint distribution tests of the posterior simulator (Geweke, 2004) using $1,224$ test functions (the 48 first moments and $1,176$ second moments of the 48 parameters in the model) yield one rejection in tests of size 0.05 and none in tests of size 0.01, showing that the posterior simulator is correct. Second, the Raftery-Lewis test (Raftery and Lewis, 1992) using all the draws shows that a small amount of burn-in (6 draws) and a total of $8,700$ draws are needed for the estimated 95% highest posterior density intervals to have actual posterior probabilities between 0.94 and 0.96 with probability 0.95, indicating that satisfactory accuracy can be achieved using the draws we have. Finally, based on draws $2,001 \sim 3,800$ (the first 10% after burn-in) and draws $11,001 \sim 20,000$ (the last 50% after burn-in), Geweke's convergence diagnostic (Geweke, 1992) is less than 1.96 in absolute value for all parameters, showing that convergence of the MCMC algorithm has been achieved.

[5] The sufficient condition for convergence set forth in Roberts and Smith (1994) is satisfied.

4 Data

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

4.1 Sample

The DealScan database contains detailed information on lending to large businesses in the U.S. dating back to 1988. For each loan facility, DealScan reports the identities of the borrower and the lender, the pricing information (spread and fees), and the information on non-price loan characteristics, such as maturity, secured status, purpose of the loan, and type of the loan.

We focus on loan facilities between U.S. banks and U.S. firms from 1996 to 2003, and divide them into sixteen 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. Our sample consists of $1,369$ loan facilities between 455 banks and $1,369$ firms.

4.2 Variables

Information on loan spreads comes from the All-In Spread Drawn (AIS) reported in the DealScan database. The AIS is expressed as a markup over the London Interbank Offering Rate (LIBOR). It equals the sum of the coupon spread, the annual fee, and any one-time fee divided by the loan maturity. In DealScan, the AIS is given in basis points (1 basis point = 0.01%). We divide the AIS by 100 to obtain r_{ij}, expressed in percentage points.

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

[6] Changing the market definition from a half-year to a year leaves our main findings unaffected.

The right-hand side of the spread equation includes a constant, year dummies, and three groups of exogenous variables. The first group includes the following bank characteristics: salaries-expenses ratio (salaries and benefits/total operating expenses), capital-assets ratio (total equity capital/total assets), ratio of cash to total assets (cash/total assets), and four size dummies. Each size dummy corresponds to one fifth of the banks with the cutoffs being approximately \$5 billion, \$13 billion, \$32 billion, and \$76 billion in assets. The size dummy for the smallest one fifth is dropped. The size dummies enable us to detect nonlinear relationships between sizes and loan spreads.

The second group includes the following firm characteristics: leverage ratio (total debt/total assets), current ratio (current assets/current liabilities), ratio of property, plant, and equipment (PP&E) to total assets (PP&E/total assets), and four size dummies. Each size dummy corresponds to one fifth of the firms with the cutoffs being approximately \$65 million, \$200 million, \$500 million, and \$1,500 million in assets. The size dummy for the smallest one fifth is dropped.

The third group includes the following non-price loan characteristics: maturity (in months), natural log of facility size, purpose dummies, type dummies, and a secured dummy. The loan purposes reported in DealScan are combined into five categories: acquisition (acquisition lines and takeover), general (corporate purposes and working capital), miscellaneous (capital expenditure, equipment purchase, IPO related finance, mortgage warehouse, project finance, purchase hardware, real estate, securities purchase, spinoff, stock buyback, telecom build-out, and trade finance), recapitalization (debt repayment/debt consolidation/refinancing and recapitalization), and other. The purpose dummy for "other" is dropped. There are three categories of loan types: revolver/line < 1 year (a revolving credit line whose duration is less than one year), revolver/line ≥ 1 year, and other. The type dummy for "other" is dropped. A secured dummy is also included, which equals one if the borrower is required to pledge collateral for the loan, and equals zero otherwise.

The right-hand side variables in the quality index equations are bank characteristics and firm characteristics, respectively. Bank assets, firm assets, and facility size 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 provides the definitions and sources of the variables, and Table 2 presents summary statistics.

5 Findings

In this section, we first present evidence that positive assortative matching of sizes is prevalent in the loan market, that is, large banks tend to match with large firms, and vice versa. We then show that for agents on both sides of the market there are similar relationships between quality and size: after controlling for other factors, the medium-sized agents are regarded as having the highest quality, followed by the largest agents, and the smallest agents are at the bottom of the list. Consequently there are similar size rankings on both sides, which explain the positive assortative matching of sizes. Liquidity is also a consideration in choosing partners. Finally, the effects of bank characteristics, firm characteristics, and non-price loan characteristics on loan spreads are examined.

5.1 Positive Assortative Matching of Sizes

It is recognized in the literature that large banks tend to lend to large firms and vice versa. See, for example, Hubbard, Kuttner and Palia (2002) and Berger et al. (2005). To verify this positive assortative matching of sizes, two OLS regressions using the matched pairs are run: the bank's size (natural log of total assets) on the firm's characteristics and the firm's size (natural log of total assets) on the bank's characteristics. The results (not reported) show that in matched pairs the bank's size and the firm's size are strongly positively correlated. The coefficients on partner's size are both positive and have t statistics around 20, indicating that there is indeed positive assortative matching of sizes.

5.2 Quality Indexes

Table 3 reports the posterior means and standard deviations of the coefficients in the quality index equations (5) and (6).

Sizes of the Agents. All the size dummies have positive coefficients and for most of them the 90% highest posterior density intervals (HPDIs) do not include zero, indicating that on both sides of the market, the group of the smallest agents—the omitted group—is considered the worst in terms of quality.[7] On the lenders' side, the smallest banks suffer from severe lending constraints and low reputation associated with their small sizes. On the borrowers' side, the smallest firms are considered the least creditworthy because they have low repaying ability and less diversified assets, and lack well-documented track records.

A closer look at the coefficients reveals that on both sides of the market, it is the medium-sized agents who have the highest quality. Banks with assets between the 40th and the 80th percentiles (group 3 and group 4) and firms with assets between the 40th and the 60th percentiles (group 3) are the most attractive. The largest agents are less attractive than the medium-sized ones, but are better than the smallest ones.

As the size of a bank increases, it has lower risk and greater lending capacity, making it more attractive. On the other hand, larger banks typically have more severe information distortion problems, and their loan officers have weaker incentives in collecting and processing borrower information. For the group of the largest banks, these negative effects outweigh the banks' advantages over the medium-sized banks in terms of risk and lending capacity.

Similarly, as the size of a firm increases, its repaying ability grows, its assets are more diversified, and it poses smaller information asymmetries. However, the group of the largest firms are less attractive than the medium-sized firms because lending to the largest firms means that the bank's assets will be less diversified and that its control over the firms' investment decisions will be weaker, and these disadvantages of the largest firms outweigh their advantages over the medium-sized firms.

Note that the negative effect of a firm's large size on its quality is likely understated, since in our model the limit on the number of loans a bank can make is binding and the limit on the total amount of loans is non-binding. If we take into account that sometimes the binding limit is on the total amount of loans, then lending to a large firm should be less attractive: the size of the loan will typically be large, which means that the bank may have

[7] HPDIs are used for model comparison in an *ad hoc* fashion. With two models $M_1 : \beta_j = 0$ and $M_2 : \beta_j \neq 0$, a finding that the HPDI under M_2 does not include zero is evidence against M_1. On the other hand, using posterior odds ratios for model comparison that involves equality restrictions typically requires the elicitation of informative priors. See Koop (2003), pp 38-45.

to sacrifice more than one lending opportunity elsewhere in order to lend to this large firm, impairing the bank's assets diversification.

The size rankings for both sides of the loan market are similar. From the highest quality to the lowest quality, the size ranking is 4-3-2-5-1 for the banks and 3-4-2-5-1 for the firms, where the numbers represent the size groups. All else being equal, the medium-sized agents have higher quality than the largest ones, which in turn have higher quality than the smallest ones. That explains the positive assortative matching of sizes. Medium-sized banks lend to medium-sized firms because both groups are the top candidates on their respective sides. Among the remaining agents, who face restricted choice sets, the largest banks and the largest firms are the top candidates, so they are matched. Finally, the smallest banks and the smallest firms have the lowest quality, and they have no choice but to match with each other.

Other Factors. On the banks' side, the coefficient on the ratio of cash to total assets is positive with a 95% HPDI that does not include zero, reflecting the negative impact of banks' liquidity risk on their quality. The coefficients on the salaries-expenses ratio and the capital-assets ratio are both positive, consistent with the hypothesis that banks with higher monitoring ability and/or higher capital adequacy are more attractive. Zero is included in the 90% HPDIs for these two coefficients, though, suggesting that in our sample the influence of these two ratios on the banks' quality is weak.

On the firms' side, the current ratio has a positive coefficient whose 95% HPDI does not include zero, supporting the view that a firm's quality is negatively related to its risk, especially the liquidity risk, for which the current ratio is a proxy. The other two variables both have the expected signs. The coefficient on the leverage ratio is negative, indicating that firms with higher leverage ratios are less attractive because they are riskier. The coefficient on the ratio of PP&E to total assets has a positive sign, suggesting that firms with relatively more tangible assets have higher quality because they pose smaller information asymmetries. The fact that the 90% HPDIs for these two coefficients include zero indicates that in our sample these two ratios are not important concerns of the banks when they rank the borrowers.

Effects on Matching Preference. To assess the importance of a variable in the matching process, we calculate an agent's probability advantage in being preferred to another agent due to a difference in that variable, everything else being equal. In order to obtain assessment that is independent of our choice of the ratio variables' units, we examine certain percentiles of these variables in the data. Table 4 reports the 10th, 30th, 50th, 70th, and 90th percentiles of each of the ratio variables.

Table 5 uses such percentiles and reports the variables' effects on matching preferences. Specifically, for each variable, the agents (banks or firms) are divided into five percentile groups: 0th-20th, 20th-40th, 40th-60th, 60th-80th, and 80th-100th. For the size variables, these groups simply correspond to the size groups in Table 3. For the ratio variables, the median of each group is used to represent that group. Each cell in Table 5 then reports one of the last four group's probability advantage in being preferred to the default group (the lowest 20 percent), everything else being equal.

The probability advantage that bank i has relative to bank i' is

$$\begin{aligned} & \Pr(B_i'\beta + \eta_i > B_{i'}'\beta + \eta_{i'}) - \Pr(B_{i'}'\beta + \eta_{i'} > B_i'\beta + \eta_i) \\ = \; & 2 \times \Pr(B_i'\beta + \eta_i > B_{i'}'\beta + \eta_{i'}) - 1 \\ = \; & 2 \times \Pr(\eta_{i'} - \eta_i < B_i'\beta - B_{i'}'\beta) - 1 \\ = \; & 2 \times \Phi\left(\frac{B_i'\beta - B_{i'}'\beta}{\sqrt{2}}\right) - 1, \end{aligned}$$

where $\Phi(.)$ is the standard normal cdf.[8] The probability advantage that firm j has relative to firm j' is obtained analogously. Consider a borrower's choice between two banks. If the two banks have no difference in their observed characteristics, then the choice is completely determined by the unobserved quality, and the probability of each bank being preferred to the other is 50%. Now suppose one bank is in the smallest size group and the other is in the middle size group, then the probability that the middle-sized bank is preferred to the small bank is 64.47%, representing a probability advantage that equals $64.47\% - 35.53\% = 28.94\%$ and giving the middle-sized bank a nearly 2-to-1 advantage if both hope to match with the same borrower. Table 5 shows that most of the size dummies have positive and noticeable effects, indicating that size plays an important role in agents' quality. In particular, medium-

[8] Point estimate of the probability advantage is obtained as $2 \times E_{\beta|X,r,\mu}\Phi\left(\frac{B_i'\beta - B_{i'}'\beta}{\sqrt{2}}\right) - 1$.

sized agents are the most preferred partners on both sides of the market. Furthermore, the two proxies for liquidity (the ratio of cash to total assets for banks and the current ratio for firms) have the largest effects among the ratio variables, suggesting that liquidity is also a consideration in choosing partners. For example, a bank with a ratio of cash to total assets in the highest 20 percent enjoys a probability advantage of 13.80% relative to a bank with a ratio in the lowest 20 percent. These effects, however, are not so substantial as those of the size dummies, and we therefore conclude that size appears to be the most important factor in the matching.

5.3 Loan Spread Determinants

The covariance between the error terms in the loan spread equation and the bank quality index equation, κ, is found to have a 95% HPDI that does not include zero (Table 3). That is evidence that the matching process is correlated with the loan spread determination and can not be ignored. To see that the m_{ij}'s (that is, the match indicators) are correlated with the ϵ_{ij}'s, rewrite the spread equation as follows, noting that each firm borrows once in a market:

$$\begin{aligned} r_j &= \alpha_0 + m_j' B\alpha_1 + F_j'\alpha_2 + N_j'\alpha_3 + \epsilon_j & (11) \\ &= \alpha_0 + m_j' B\alpha_1 + F_j'\alpha_2 + N_j'\alpha_3 + \kappa m_j'\eta + \lambda\delta_j + \nu_j,\ \nu_j \sim N(0, \sigma_\nu^2), & (12) \end{aligned}$$

where r_j is the spread that firm j pays, $m_j = (m_{1j}, m_{2j}, ..., m_{Ij})'$ is the vector of match indicators, $B = (B_1, B_2, ...B_I)'$ is the matrix of bank characteristics, N_j is the non-price loan characteristics of the loan firm j borrows, and $\eta = (\eta_1, \eta_2, ..., \eta_I)'$ is the vector of error terms in the bank quality index equation. If m_j were in fact independent of ϵ_j—as it would be if firms were randomly matched to banks—then m_j would be exogenous in equation (11). However, the evidence above against $\kappa = 0$ implies that m_j is correlated with ϵ_j. Thus the regressors are correlated with the error term, giving rise to an endogeneity problem. In this paper, the inclusion of a two-sided matching model to supplement the spread equation overcomes this endogeneity problem. Estimates of the spread equation are presented in Table 6.

Bank Characteristics. The coefficients on the bank size dummies are all negative and most of them have 90% HPDIs that do not include zero, supporting the view that larger banks are likely to have better diversified assets and hence lower risk, so that they charge lower loan spreads. As expected, these coefficients exhibit a downward trend. Compared to banks with assets below the 20th percentile, banks with assets between the 20th and the 60th percentiles charge loan spreads that are around 15 basis points lower, whereas banks with assets above the 60th percentile charge loan spreads that are nearly 30 basis points lower.

The salaries-expenses ratio has a positive coefficient whose 99% HPDI does not include zero, showing that banks with superior monitoring ability indeed charge higher loan spreads. The coefficients on the capital-assets ratio and the ratio of cash to total assets have 90% HPDIs that include zero, suggesting that in our sample, banks' capital adequacy risk and liquidity risk do not have significant impact on loan spreads.

To ensure that our assessment of the ratio variables' impact on loan spread is independent of our choice of these variables' units, we again use the percentiles reported in Table 4 to calculate the differences in the spreads charged by banks in different percentile groups. The results are reported in Table 7. For example, the table shows that a bank with a salaries-expenses ratio in the highest 20 percent charges a premium of 39 basis points relative to a bank with a ratio in the lowest 20 percent. In our sample, the average loan spread is 1.89%, or 189 basis points, so a 39-basis-point premium represents an increase of more than 20%, clear evidence that banks charge a premium for superior monitoring ability.

Firm Characteristics. In Table 6, all the firm size dummies have negative coefficients whose 95% HPDIs do not include zero, consistent with the hypothesis that larger firms are charged lower loan spreads because they are less risky and have lower degrees of information opacity. The coefficients on the firm size dummies also exhibit a downward trend. For example, compared to firms with assets below the 20th percentile, firms with assets between the 20th and the 40th percentiles are charged loan spreads that are 22 basis points lower, whereas firms with assets above the 80th percentile are charged loan spreads that are 46 basis points lower.

Two firm ratios have 99% HPDIs that do not include zero: the leverage ratio (positive) and the current ratio (negative). A higher leverage ratio or a lower current ratio (indicating a lower liquidity) represents a higher borrower risk, so the signs of the coefficients confirm that firms with higher risk are charged higher loan spreads. The coefficient on the ratio of PP&E to total assets has a 90% HPDI that includes zero, suggesting that the ratio does not significantly affect borrowers' costs of funds.

Table 7 shows the characteristics' impact on loan spread. For example, we see that a firm with a leverage ratio in the highest 20 percent is charged a premium of 48 basis points relative to a firm with a ratio in the lowest 20 percent. On the other hand, a firm with a current ratio in the highest 20 percent is given a discount of 13 basis points relative to a firm with a ratio in the lowest 20 percent. These numbers attest to the substantial impact of firms' leverage and liquidity on loan spreads charged.

Non-Price Loan Characteristics. In Table 6, three non-price loan characteristics have coefficients whose 90% HPDIs do not include zero: the natural log of facility size, the revolver/line>=1 year dummy, and the secured dummy.

The negative coefficient on the natural log of facility size is likely due to economies of scale in bank lending. The processes of loan approval, monitoring, and review are relatively labor-intensive, and the labor costs in these processes increase less than proportionally when the size of the loan increases. As a result, a larger loan has a lower average labor costs and is therefore charged a lower loan spread.

The dummy for revolving credit lines whose durations are greater than or equal to one year has a negative coefficient. Since that type of loans are by far the most common, accounting for 67% of all loans, the negative coefficient may reflect that other types of loans are non-standard or even custom-made, and are charged higher loan spreads to compensate for the banks' extra administrative costs resulting from the loans' non-standard nature.

Related to the interpretation of the secured dummy, there are two major theories on the relationship between collateral and risk. In the first theory (referred to as the "sorting-by-observed-risk paradigm" by Berger and Udell, 1990), observably risky borrowers are *required* to pledge collateral, while observably safe borrowers are not. In the second theory (referred to as the "sorting-by-private-information paradigm" by Berger and Udell), because of informational asymmetry, borrower risk is unobservable to the banks and certain borrowers *choose* to pledge collateral to signal their quality and/or to lower borrowing costs. While it is likely that both theories apply to some cases, Berger and Udell (1990) find that empirically the sorting-by-observed-risk paradigm clearly dominates the sorting-by-private-information paradigm (also see the references and anecdotal evidence cited in their paper, for example on Page 23). Furthermore, given the nature of the loans included in our sample—large bank loans made to large and publicly traded companies, informational asymmetry is not a first-order concern. Therefore, the sorting-by-observed-risk paradigm is most relevant here and we thus take the secured status of a loan as a requirement of the bank due to the borrower being judged as risky, rather than a choice of the borrower (empirically testing between the two theories is beyond the scope of this paper).

The sorting-by-observed-risk interpretation is also consistent with the common definition of an unsecured loan. According to Fitch (2006), an unsecured loan is also called a character loan or a good faith loan, and is granted by the lender on the strength of the borrower's creditworthiness, rather than a pledge of assets as collateral. The coefficient on the secured dummy reported in Table 6 shows that a sizeable 88-basis-point premium is charged if the borrower is below the threshold for an unsecured loan, consistent with the finding in prior empirical studies (for example, Berger and Udell, 1990; John, Lynch, and Puri, 2003; Focarelli, Pozzolo, and Casolaro, 2008) that the value of the recourse against collateral does not fully offset the higher risk of secured borrowers.

6 Conclusion

We have the potential to learn a lot about financial markets and the effects of monetary policy by investigating the pricing of bank loans. For example, empirical evidence on determinants of loan spreads can provide insights into risk premiums in financial contracting and transmission mechanisms of monetary policy. This paper shows that there is endogenous matching in the bank loan market, and that we face an endogeneity problem in estimation of the loan spread equation when some characteristics of banks or firms are not perfectly observed and proxies are used. To control for the endogenous matching, we develop a two-sided matching model to supplement the loan spread equation. Because estimation of the model is numerically intensive, we choose a parsimonious specification to keep estimation feasible, focusing on a set of key variables. We obtain Bayesian inference using a Gibbs sampling algorithm with data augmentation, which transforms a high-dimensional integration problem into a simulation problem and overcomes the computational difficulty.

Using a sample of $1,369$ U.S. loan facilities between 455 banks and $1,369$ firms from 1996 to 2003, we find evidence of positive assortative matching of sizes in the market, that is, large banks tend to match with large firms, and vice versa. We then show that for agents on both sides of the market there are similar relationships between quality and size, which lead to similar size rankings for both sides and explain the positive assortative matching of sizes. Medium-sized banks and firms are the most preferred partners. Liquidity is also a consideration in choosing partners. Furthermore, banks with higher monitoring ability charge higher spreads, whereas larger banks charge lower spreads. On the other side of the market, firms that are more leveraged or less liquid are charged higher spreads, whereas larger firms are charged lower spreads.

The two-sided matching model not only addresses the endogeneity problem in estimation of the loan spread equation, but also provides a way to assess agents' quality and to understand how agents choose each other. The latter is an important issue in various two-sided markets. For instance, in an empirical study of academic achievements or job outcomes of college students, a two-sided matching model can be used to estimate the colleges' quality and the students' ability. Other examples include the matchings between teams and athletes (in NBA, for instance), corporations and CEOs, firms and underwriters, and so on. Furthermore, the two-sided matching model enables us to identify the factors that contribute to agents' quality, which can point the way for agents who try to improve their standing, such as colleges that want to attract better students. This suggests that understanding the quality indexes can play an important role in such markets. We view those issues as interesting avenues for future research.

Appendix A. Uniqueness of Equilibrium Matching

The existence of an equilibrium matching in the College Admissions Model is proved in Roth and Sotomayor (1990). Below we show that there exists a unique equilibrium matching in our model, which is a special case of the College Admissions Model.

Re-index the banks and the firms according to the preference orderings, so that $i \succ_j i'$, $\forall i > i'$, $\forall j$, and $j \succ_i j'$, $\forall j > j'$, $\forall i$, where $i \succ_j i'$ denotes that firm j prefers bank i to bank i' and $j \succ_i j'$ denotes that bank i prefers firm j to firm j'. Let q_{it} be the quota of bank i. The following J-step algorithm produces the unique equilibrium matching, in which there is perfect sorting. In step 1, firm J matches with bank I. In step 2, firm $J-1$ matches with bank I if $q_{It} \geq 2$, otherwise it matches with bank $I-1$. In step 3, firm $J-2$ matches with bank I if $q_{It} \geq 3$, otherwise it matches with bank $I-1$ if $q_{It} + q_{I-1,t} \geq 3$, otherwise it matches with bank $I-2$. And so on.

First, μ is an equilibrium matching. Suppose not, then there exists at least one blocking pair (i', j') such that $i' > \mu(j')$ and $j' > \min\{j : j \in \mu(i')\}$. That is a contradiction, since by construction if $i' > \mu(j')$ then $j'' > j'$, $\forall j'' \in \mu(i')$, so $j' > \min\{j : j \in \mu(i')\}$ can not be true.

Second, the equilibrium matching is unique. Suppose not, then there exists $\tilde{\mu} \neq \mu$ such that $\tilde{\mu}$ is also an equilibrium matching. There is at least one match that is in μ but not in $\tilde{\mu}$. Now consider the first step in the algorithm that forms a match that is not in $\tilde{\mu}$. Call that match (i', j'). It follows that $\min\{j : j \in \tilde{\mu}(i')\} < j'$ and that $\tilde{\mu}(j') < i'$, since all the matches formed in the earlier steps are in both μ and $\tilde{\mu}$. Therefore (i', j') is a blocking pair for $\tilde{\mu}$, a contradiction.

Appendix B. Inequalities Characterizing the Unique Equilibrium Matching

The unique equilibrium matching is characterized by a set of inequalities, based on the fact that there is no blocking bank-firm pair. For each bank, stability requires that its worst current borrower be better than any other firm whose current lender is worse than this bank. Similarly, for each firm, stability requires that its current lender be better than any other bank whose worst current borrower is worse than this firm.

Consider a matching in market t, μ_t. Suppose bank i and firm j are not matched in μ_t. (i, j) is a blocking pair iff $Q_j^f > \min\limits_{j' \in \mu_t(i)} Q_{j'}^f$ and $Q_i^b > Q_{\mu_t(j)}^b$. So (i, j) is not a blocking pair iff $Q_j^f < \min_{j' \in \mu_t(i)} Q_{j'}^f$ or $Q_i^b < Q_{\mu_t(j)}^b$. Equivalently, (i,j) is not a blocking pair iff $Q_j^f < \overline{Q}_{ji}^f$ and $Q_i^b < \overline{Q}_{ij}^b$, where

$$\overline{Q}_{ji}^f = \begin{cases} \min_{j' \in \mu_t(i)} Q_{j'}^f & if\ Q_i^b > Q_{\mu_t(j)}^b \\ \infty & \text{otherwise,} \end{cases}$$

and

$$\overline{Q}_{ij}^b = \begin{cases} Q_{\mu_t(j)}^b & if\ Q_j^f > \min_{j' \in \mu_t(i)} Q_{j'}^f \\ \infty & \text{otherwise.} \end{cases}$$

Now suppose bank i and firm j are matched in μ_t. Bank i or firm j is part of a blocking pair iff $Q_j^f < \max_{j' \in f(i)} Q_{j'}^f$ or $Q_i^b < \max_{i' \in f(j)} Q_{i'}^b$, where $f(i)$ is the set of firms that do not currently borrow from bank i but would prefer to do so, and $f(j)$ is the set of banks that do not currently lend to firm j but would prefer to do so. These two sets contain the feasible deviations of the agents and are given by

$$\begin{aligned} f(i) &= \{j \in J_t \backslash \mu_t(i) : Q_i^b > Q_{\mu_t(j)}^b\}, \text{ and} \\ f(j) &= \{i \in I_t \backslash \mu_t(j) : Q_j^f > \min_{j' \in \mu_t(i)} Q_{j'}^f\}. \end{aligned}$$

Therefore, neither bank i nor firm j is part of a blocking pair iff $Q_j^f > \underline{Q}_{ji}^f$ and $Q_i^b > \underline{Q}_{ij}^b$, where $\underline{Q}_{ji}^f = \max_{j' \in f(i)} Q_{j'}^f$ and $\underline{Q}_{ij}^b = \max_{i' \in f(j)} Q_{i'}^b$.

Let μ_t^e denote the unique equilibrium matching in market t. The above analysis leads to the following characterization of the equilibrium matching:

$$\mu_t = \mu_t^e \iff Q_i^b \in (\underline{Q}_i^b, \overline{Q}_i^b),\ \forall i \in I_t \text{ and } Q_j^f \in (\underline{Q}_j^f, \overline{Q}_j^f),\ \forall j \in J_t,$$

where

$$\begin{aligned} \underline{Q}_i^b &= \max_{j \in \mu_t(i)} \underline{Q}_{ij}^b, \\ \overline{Q}_i^b &= \min_{j \notin \mu_t(i)} \overline{Q}_{ij}^b, \\ \underline{Q}_j^f &= \underline{Q}_{j,\mu_t(j)}^f, \text{ and} \\ \overline{Q}_j^f &= \min_{i \notin \mu_t(j)} \overline{Q}_{ji}^f. \end{aligned}$$

Appendix C. Conditional Posterior Distributions

We obtain the conditional posterior distributions by examining the kernels of the conditional posterior densities. The conditional posterior distribution of Q_i^b is $N(\hat{Q}_i^b, \hat{\sigma}^2_{Q_i^b})$ truncated to the interval $(\underline{Q}_i^b, \overline{Q}_i^b)$, where

$$\hat{Q}_i^b = B_i'\beta + \frac{\kappa \sum_{j\in\mu_t(i)}[r_{ij} - U_{ij}'\alpha - \lambda(Q_j^f - F_j'\gamma)]}{\sigma_\nu^2 + \kappa^2 q_{it}}, \text{ and}$$

$$\hat{\sigma}^2_{Q_i^b} = \frac{\sigma_\nu^2}{\sigma_\nu^2 + \kappa^2 q_{it}}.$$

The conditional posterior distribution of Q_j^f is $N(\hat{Q}_j^f, \hat{\sigma}^2_{Q_j^f})$ truncated to the interval $(\underline{Q}_j^f, \overline{Q}_j^f)$, where

$$\hat{Q}_j^f = F_j'\gamma + \frac{\lambda[r_{\mu_t(j),j} - U'_{\mu_t(j),j}\alpha - \kappa(Q^b_{\mu_t(j)} - B'_{\mu_t(j)}\beta)]}{\sigma_\nu^2 + \lambda^2}, \text{ and}$$

$$\hat{\sigma}^2_{Q_j^f} = \frac{\sigma_\nu^2}{\sigma_\nu^2 + \lambda^2}.$$

The prior distributions of α, β, γ, and κ are $N(\bar{\alpha}, \bar{\Sigma}_\alpha)$, $N(\bar{\beta}, \bar{\Sigma}_\beta)$, $N(\bar{\gamma}, \bar{\Sigma}_\gamma)$, and $N(\bar{\kappa}, \bar{\sigma}^2_\kappa)$, respectively. The prior distribution of λ is $N(\bar{\lambda}, \bar{\sigma}^2_\lambda)$ truncated on the right at 0. The prior distribution of $1/\sigma_\nu^2$ is gamma, $1/\sigma_\nu^2 \sim G(a, b)$, $a, b > 0$.

The conditional posterior distribution of α is $N(\hat{\alpha}, \hat{\Sigma}_\alpha)$, where

$$\hat{\Sigma}_\alpha = \left\{\bar{\Sigma}_\alpha^{-1} + \sum_{t=1}^{T}\sum_{(i,j)\in\mu_t}\frac{1}{\sigma_\nu^2}U_{ij}U_{ij}'\right\}^{-1}, \text{ and}$$

$$\hat{\alpha} = -\hat{\Sigma}_\alpha\left\{-\bar{\Sigma}_\alpha^{-1}\bar{\alpha} - \sum_{t=1}^{T}\sum_{(i,j)\in\mu_t}\frac{1}{\sigma_\nu^2}U_{ij}(r_{ij} - \kappa(Q_i^b - B_i'\beta) - \lambda(Q_j^f - F_j'\gamma))\right\}.$$

The conditional posterior distribution of β is $N(\hat{\beta}, \hat{\Sigma}_\beta)$, where

$$\hat{\Sigma}_\beta = \left\{\bar{\Sigma}_\beta^{-1} + \sum_{t=1}^{T}\sum_{i\in I_t}\frac{\sigma_\nu^2 + \kappa^2 q_{it}}{\sigma_\nu^2}B_iB_i'\right\}^{-1}, \text{ and}$$

$$\hat{\beta} = -\hat{\Sigma}_\beta\left\{-\bar{\Sigma}_\beta^{-1}\bar{\beta} + \sum_{t=1}^{T}\left[\sum_{(i,j)\in\mu_t}\frac{\kappa}{\sigma_\nu^2}B_i(r_{ij} - U_{ij}'\alpha - \kappa Q_i^b - \lambda(Q_j^f - F_j'\gamma)) - \sum_{i\in I_t}Q_i^bB_i\right]\right\}.$$

The conditional posterior distribution of γ is $N(\hat{\gamma}, \hat{\Sigma}_\gamma)$, where

$$\hat{\Sigma}_\gamma = \left\{\bar{\Sigma}_\gamma^{-1} + \sum_{t=1}^{T}\sum_{j\in J_t}\frac{\sigma_\nu^2 + \lambda^2}{\sigma_\nu^2}F_jF_j'\right\}^{-1}, \text{ and}$$

$$\hat{\gamma} = -\hat{\Sigma}_\gamma \left\{ -\bar{\Sigma}_\gamma^{-1}\bar{\gamma} + \sum_{t=1}^{T} \left[\sum_{(i,j)\in\mu_t} \frac{\lambda}{\sigma_\nu^2} F_j (r_{ij} - U'_{ij}\alpha - \kappa(Q_i^b - B'_i\beta) - \lambda Q_j^f) - \sum_{j\in J_t} Q_j^f F_j \right] \right\}.$$

The conditional posterior distribution of κ is $N(\hat{\kappa}, \hat{\sigma}_\kappa^2)$, where

$$\hat{\sigma}_\kappa^2 = \left\{ \frac{1}{\bar{\sigma}_\kappa^2} + \sum_{t=1}^{T} \sum_{i\in I_t} \frac{q_{it}(Q_i^b - B'_i\beta)^2}{\sigma_\nu^2} \right\}^{-1}, \text{ and}$$

$$\hat{\kappa} = -\hat{\sigma}_\kappa^2 \left\{ -\frac{\bar{\kappa}}{\bar{\sigma}_\kappa^2} - \sum_{t=1}^{T} \sum_{(i,j)\in\mu_t} \frac{(r_{ij} - U'_{ij}\alpha - \lambda(Q_j^f - F'_j\gamma))(Q_i^b - B'_i\beta)}{\sigma_\nu^2} \right\}.$$

The conditional posterior distribution of λ is $N(\hat{\lambda}, \hat{\sigma}_\lambda^2)$ truncated on the right at 0, where

$$\hat{\sigma}_\lambda^2 = \left\{ \frac{1}{\bar{\sigma}_\lambda^2} + \sum_{t=1}^{T} \sum_{j\in J_t} \frac{(Q_j^f - F'_j\gamma)^2}{\sigma_\nu^2} \right\}^{-1}, \text{ and}$$

$$\hat{\lambda} = -\hat{\sigma}_\lambda^2 \left\{ -\frac{\bar{\lambda}}{\bar{\sigma}_\lambda^2} - \sum_{t=1}^{T} \sum_{(i,j)\in\mu_t} \frac{(r_{ij} - U'_{ij}\alpha - \kappa(Q_i^b - B'_i\beta))(Q_j^f - F'_j\gamma)}{\sigma_\nu^2} \right\}.$$

Let $n = \sum_{t=1}^{T} |J_t|$ denote the total number of loans in all the markets. The conditional posterior distribution of $1/\sigma_\nu^2$ is $G(\hat{a}, \hat{b})$, where

$$\hat{a} = a + \frac{n}{2}, \text{ and}$$

$$\hat{b} = \left[\frac{1}{b} + \frac{1}{2} \sum_{t=1}^{T} \sum_{(i,j)\in\mu_t} (r_{ij} - U'_{ij}\alpha - \kappa(Q_i^b - B'_i\beta) - \lambda(Q_j^f - F'_j\gamma))^2 \right]^{-1}.$$

References

[1] Ackerberg, D. A. and M. Botticini (2002), "Endogenous Matching and the Empirical Determinants of Contract Form", *Journal of Political Economy*, Vol. 110:3, 564-91.

[2] Albert, J. and S. Chib (1993), "Bayesian Analysis of Binary and Polychotomous Response Data", *Journal of the American Statistical Association*, Vol. 88, 669-79.

[3] Allen, F. (1990), "The Market for Information and the Origin of Financial Intermediation", *Journal of Financial Intermediation*, Vol. 1, 3-30.

[4] Berger, A. N., N. H. Miller, M. A. Peterson, R. G. Rajan and J. C. Stein (2005), "Does Function Follow Organizational Form? Evidence from the Lending Practices of Large and Small Banks", *Journal of Financial Economics*, Vol. 76:2, 237-69.

[5] Berger, A. N. and G. F. Udell (1990), "Collateral, Loan Quality, and Bank Risk", *Journal of Monetary Economics,* Vol. 25, 21-42.

[6] Bhattacharya, H. (1997), *Banking Strategy, Credit Appraisal and Lending Decisions: A Risk-Return Framework*, Oxford University Press.

[7] Billet, M., M. Flannery and J. Garfinkel (1995), "The Effect of Lender Identity on a Borrowing Firm's Equity Return", *Journal of Finance*, Vol. 50:2, 699-718.

[8] Brickley, J. A., J. S. Linck and C. W. Smith (2003), "Boundaries of the Firm: Evidence from the Banking Industry", *Journal of Financial Economics*, Vol. 70, 351-83.

[9] Chen, J. and K. Song (2013), "Two-Sided Matching in the Loan Market", *International Journal of Industrial Organization*, Vol. 31, 145-52.

[10] Coleman, A. D. F., N. Esho, and I. G. Sharpe (2006), "Does Bank Monitoring Influence Loan Contract Terms?" *Journal of Financial Services Research*, Vol. 30, 177-98.

[11] Diamond, D. (1984), "Financial Intermediation and Delegated Monitoring", *Review of Economic Studies*, Vol. 51, 393-414.

[12] Diamond, D. (1991), "Monitoring and Reputation: The Choice Between Bank Loans and Directly Placed Debt", *Journal of Political Economy*, Vol. 99, 689-721.

[13] Diamond, D. and R. Rajan (2000), "A Theory of Bank Capital", *Journal of Finance*, Vol. 55:6, 2431-65.

[14] Fernando, C. S., V. A. Gatchev, and P. A. Spindt (2005), "Wanna Dance? How Firms and Underwriters Choose Each Other", *Journal of Finance*, Vol. 60:5, 2437-69.

[15] Fitch, T. (2006), *Dictionary of Banking Terms*, Barron's Educational Series, Inc.

[16] Focarelli, D., A. F. Pozzolo, and L. Casolaro (2008), "The Pricing Effect of Certification on Syndicated Loans", *Journal of Monetary Economics*, Vol. 55, 335-49.

[17] Gale, D. and L. Shapley (1962), "College Admissions and the Stability of Marriage", *American Mathematical Monthly*, Vol. 69, 9-15.

[18] Gelfand, A. and A. Smith (1990), "Sampling-Based Approaches to Calculating Marginal Densities", *Journal of the American Statistical Association*, Vol. 85, 398-409.

[19] Geman, S. and D. Geman (1984), "Stochastic Relaxation, Gibbs Distributions and the Bayesian Restoration of Images", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 6, 721-41.

[20] Geweke, J. (1992), "Evaluating the Accuracy of Sampling-Based Approaches to the Calculation of Posterior Moments" (with discussion), *Bayesian Statistics*, J. M. Bernardo, J. O. Berger, A. P. Dawid, and A. F. M. Smith (eds.), Vol. 4, 169-93, Oxford University Press.

[21] Geweke, J. (1999), "Using Simulation Methods for Bayesian Econometric Models: Inference, Development and Communication" (with discussion and rejoinder), *Econometric Reviews*, Vol. 18, 1-126.

[22] Geweke, J. (2004), "Getting It Right: Joint Distribution Tests of Posterior Simulators", *Journal of the American Statistical Association*, Vol. 99, 799-804.

[23] Geweke, J., G. Gowrisankaran, and R. Town (2003), "Bayesian Inference for Hospital Quality in a Selection Model", *Econometrica*, Vol. 71:4, 1215-39.

[24] Hubbard, R. G., K. N. Kuttner, and D. N. Palia (2002), "Are There Bank Effects in Borrowers' Costs of Funds? Evidence from a Matched Sample of Borrowers and Banks", *Journal of Business*, Vol. 75, 559-81.

[25] John, K., A. Lynch, and M. Puri (2003), "Credit Ratings, Collateral, and Loan Characteristics: Implications for Yield", *Journal of Business*, Vol. 76, 371-409.

[26] Johnson, S. (1997), "The Effects of Bank Reputation on the Value of Bank Loan Agreements", *Journal of Accounting, Auditing and Finance*, Vol. 24, 83-100.

[27] Kashyap, A. K. and J. C. Stein (1994), "Monetary Policy and Bank Lending", *Monetary Policy*, N. G. Mankiw (ed.), 221-56, University of Chicago Press.

[28] Koop, G. (2003), *Bayesian Econometrics*, John Wiley and Sons.

[29] Lancaster, T. (1997), "Exact Structural Inference in Optimal Job-Search Models", *Journal of Business and Economic Statistics*, Vol. 15:2, 165-79.

[30] Leland, H. and D. Pyle (1977), "Informational Asymmetries, Financial Structure and Financial Intermediation", *Journal of Finance*, Vol. 32:2, 371-87.

[31] Miller, S. and S. Bavaria (2003), "A Guide to the U.S. Loan Market", Standard & Poor's, October.

[32] Park, M. (2008), "An Empirical Two-Sided Matching Model of Acquisitions: Understanding Merger Incentives and Outcomes in the Mutual Fund Industry", working paper, University of Minnesota.

[33] Raftery, A. E. and S. Lewis (1992), "How Many Iterations in the Gibbs Sampler?" *Bayesian Statistics*, J. M. Bernardo, J. O. Berger, A. P. Dawid, and A. F. M. Smith (eds.), Vol. 4, 763-73, Oxford University Press.

[34] Rajan, R. (1992), "Insiders and Outsiders: The Choice Between Informed and Arms-Length Debt", *Journal of Finance*, Vol. 47, 1367-406.

[35] Roberts, G. O. and A. F. M. Smith (1994), "Simple Conditions for the Convergence of the Gibbs Sampler and Metropolis-Hastings Algorithms", *Stochastic Processes and Their Applications*, Vol. 49, 207-16.

[36] Roth, A. and M. Sotomayor (1990), *Two-Sided Matching: A Study in Game-Theoretic Modeling and Analysis*, Econometric Society Monograph Series, Cambridge University Press.

[37] Sorensen, M. (2007), "How Smart is Smart Money? A Two-Sided Matching Model of Venture Capital", *Journal of Finance*, Vol. 62, 2725-62.

[38] Stoeckle, G. (2011), "Investment Insights: Bank Loans Within the Broader Market", Invesco, March.

[39] Tanner, T. and W. Wong (1987), "The Calculation of Posterior Distributions by Data Augmentation", *Journal of the American Statistical Association*, Vol. 82, 528-49.

[40] Yago, G. and D. McCarthy (2004), “The U.S. Leveraged Loan Market: A Primer”, Milken Institute Research Report, October.

Table 1. Variable Definitions and Sources

VARIABLE	DEFINITION	SOURCE
Dependent Variable		
Loan Spread	All-In Spread Drawn above LIBOR/100; in percentage points	DealScan
Independent Variables		
Bank Characteristics		
Salaries-Expenses Ratio	Salaries and Benefits/Total Operating Expenses	Call Reports
Capital-Assets Ratio	Total Equity Capital/Total assets	Call Reports
Ratio of Cash to Total Assets	Cash/Total assets	Call Reports
Bank_Size2	Dummy = 1 if the bank has $5 billion to $13 billion assets	Call Reports
Bank_Size3	Dummy = 1 if the bank has $13 billion to $32 billion assets	Call Reports
Bank_Size4	Dummy = 1 if the bank has $32 billion to $76 billion assets	Call Reports
Bank_Size5	Dummy = 1 if the bank has more than $76 billion assets	Call Reports
Firm Characteristics		
Leverage Ratio	Total Debt/Total Assets	Compustat
Current Ratio	Current Assets/Current Liabilities	Compustat
Ratio of Property, Plant, and Equipment to Total Assets	PP&E/Total Assets	Compustat
Firm_Size2	Dummy = 1 if the firm has $65 million to $200 million assets	Compustat
Firm_Size3	Dummy = 1 if the firm has $200 million to $500 million assets	Compustat
Firm_Size4	Dummy = 1 if the firm has $500 million to $1,500 million assets	Compustat
Firm_Size5	Dummy = 1 if the firm has more than $1,500 million assets	Compustat
Non-Price Loan Characteristics		
Maturity	Loan Facility Length in Months	DealScan
Natural Log of Facility Size[2]	Log(Tranche Amount)	DealScan
Acquisition	Dummy = 1 if specific purpose is Acquisition	DealScan
General	Dummy = 1 if specific purpose is General	DealScan
Miscellaneous	Dummy = 1 if specific purpose is Miscellaneous	DealScan
Recapitalization	Dummy = 1 if specific purpose is Recapitalization	DealScan
Revolver/Line < 1 Yr.	Dummy = 1 if the loan is a revolving credit line with duration < 1 year	DealScan
Revolver/Line >= 1 Yr.	Dummy = 1 if the loan is a revolving credit line with duration >= 1 year	DealScan
Secured	Dummy = 1 if the loan is secured	DealScan

Table 2. Summary Statistics

Variable	Number of Observations	Mean	Standard Deviation	Minimum	Maximum
Loan Spread	1369	1.89	1.20	0.15	10.80
Salaries-Expenses Ratio	455	0.25	0.09	0.03	0.59
Capital-Assets Ratio	455	0.09	0.03	0.05	0.32
Ratio of Cash to Total Assets	455	0.07	0.04	0.00	0.44
Bank Assets ($ Million)	455	72311.19	124219.52	15.98	625255.55
Leverage Ratio	1369	0.26	0.23	0.00	1.95
Current Ratio	1369	2.26	2.29	0.08	31.68
Ratio of PP&E to Total Assets	1369	0.31	0.25	0.00	0.96
Firm Assets ($ Million)	1369	1807.32	6327.17	1.06	172827.99
Maturity	1369	32.91	22.97	2.00	280.00
Facility Size ($ Million)	1369	192.54	491.94	0.20	10202.00
Acquisition	1369	0.10	0.30	0	1
General	1369	0.46	0.50	0	1
Miscellaneous	1369	0.04	0.21	0	1
Recapitalization	1369	0.29	0.45	0	1
Revolver/Line < 1 Yr.	1369	0.06	0.24	0	1
Revolver/Line >= 1 Yr.	1369	0.67	0.47	0	1
Secured Status	1369	0.66	0.48	0	1

Table 3. Estimates of Quality Index Equations

	Mean	Std. Dev.
Bank Quality Index		
Salaries-Expenses Ratio	0.17	0.58
Capital-Assets Ratio	0.31	1.89
Ratio of Cash to Total Assets	2.58	1.17 **
Bank_Size2	0.41	0.15 ***
Bank_Size3	0.53	0.15 ***
Bank_Size4	0.54	0.15 ***
Bank_Size5	0.30	0.15 **
Firm Quality Index		
Leverage Ratio	-0.05	0.13
Current Ratio	0.03	0.01 **
Ratio of PP&E to Total Assets	0.02	0.12
Firm_Size2	0.11	0.08
Firm_Size3	0.29	0.09 ***
Firm_Size4	0.17	0.09 *
Firm_Size5	0.11	0.09
κ	0.17	0.09 **
λ	-0.10	0.08
$1/\sigma_v^2$	1.51	0.06 ***

1. The dependent variables are the quality indexes.
2. Posterior means and standard deviations are based on 20,000 draws from the conditional posterior distributions, discarding the first 2,000 as burn-in draws.
3. *, **, and *** indicate that zero is not contained in the 90%, 95%, and 99% highest posterior density intervals, respectively.

Table 4. 10th, 30th, 50th, 70th, and 90th Percentiles of Ratio Variables

	Percentiles				
	10th	**30th**	**50th**	**70th**	**90th**
Bank Ratio Variables					
Salaries-Expenses Ratio	0.13	0.20	0.24	0.28	0.35
Capital-Assets Ratio	0.06	0.07	0.08	0.09	0.11
Ratio of Cash to Total Assets	0.03	0.05	0.06	0.08	0.12
Firm Ratio Variables					
Leverage Ratio	0.00	0.10	0.23	0.36	0.53
Current Ratio	0.76	1.28	1.72	2.39	3.95
Ratio of PP&E to Total Assets	0.05	0.13	0.22	0.40	0.73

Table 5. Effects of Bank and Firm Characteristics on Matching Preference

	Percentile Groups			
	20th - 40th	**40th - 60th**	**60th - 80th**	**80th - 100th**
Bank Characteristics				
Salaries-Expenses Ratio	0.66%	1.03%	1.41%	2.01%
Capital-Assets Ratio	0.18%	0.33%	0.49%	0.88%
Ratio of Cash to Total Assets	2.80%	5.26%	7.96%	13.80%
Bank Size	22.53%	28.94%	29.48%	16.76%
Firm Characteristics				
Leverage Ratio	-0.26%	-0.59%	-0.94%	-1.39%
Current Ratio	0.86%	1.60%	2.71%	5.29%
Ratio of PP&E to Total Assets	0.09%	0.20%	0.40%	0.78%
Firm Size	6.41%	16.23%	9.29%	6.38%

1. Each cell reports the group's probability advantage in being preferred to the default group (the lowest 20 percent), everything else being equal. For a ratio variable, the median of each group is used to represent that group.

Table 6. Estimates of Loan Spread Equation

	Mean	Std. Dev.
Constant	1.60	0.21 ***
Salaries-Expenses Ratio	1.83	0.33 ***
Capital-Assets Ratio	1.51	1.19
Ratio of Cash to Total Assets	0.15	0.68
Bank_Size2	-0.18	0.10 *
Bank_Size3	-0.14	0.11
Bank_Size4	-0.27	0.10 ***
Bank_Size5	-0.28	0.08 ***
Leverage Ratio	0.91	0.11 ***
Current Ratio	-0.04	0.01 ***
Ratio of PP&E to Total Assets	-0.07	0.10
Firm_Size2	-0.22	0.08 ***
Firm_Size3	-0.32	0.10 ***
Firm_Size4	-0.27	0.11 **
Firm_Size5	-0.46	0.13 ***
Maturity	0.00	0.00
Natural Log of Facility Size	-0.20	0.02 ***
Acquisition	0.04	0.11
General	0.02	0.09
Miscellaneous	0.20	0.13
Recapitalization	0.06	0.09
Revolver/Line < 1 Yr.	0.16	0.11
Revolver/Line >= 1 Yr.	-0.11	0.06 *
Secured	0.88	0.06 ***

1. The dependent variable is the loan spread.
2. Posterior means and standard deviations are based on 20,000 draws from the conditional posterior distributions, discarding the first 2,000 as burn-in draws.
3. *, **, and *** indicate that zero is not contained in the 90%, 95%, and 99% highest posterior density intervals, respectively.
4. Dummies for years 1997-2003 are included on the RHS of the spread equation.

Table 7. Effects of Bank and Firm Characteristics on Loan Spread

	Percentile Groups			
	20th - 40th	40th - 60th	60th - 80th	80th - 100th
Bank Characteristics				
Salaries-Expenses Ratio	0.13	0.20	0.27	0.39
Capital-Assets Ratio	0.02	0.03	0.04	0.08
Ratio of Cash to Total Assets	0.00	0.01	0.01	0.01
Bank Size	-0.18	-0.14	-0.27	-0.28
Firm Characteristics				
Leverage Ratio	0.09	0.20	0.32	0.48
Current Ratio	-0.02	-0.04	-0.06	-0.13
Ratio of PP&E to Total Assets	-0.01	-0.01	-0.02	-0.05
Firm Size	-0.22	-0.32	-0.27	-0.46

1. Each cell reports the difference in loan spread (in percentage points) between the group and the default group (the lowest 20 percent), everything else being equal. For a ratio variable, the median of each group is used to represent that group.