How Smart is Smart Money? A Two-Sided Matching Model of Venture Capital

MORTEN SØRENSEN*

Abstract

I find that companies funded by more experienced VCs are more likely to go public. This follows both from the direct influence of more experienced VCs and from sorting in the market, which leads experienced VCs to invest in better companies. Sorting creates an endogeneity problem, but a structural model based on a two-sided matching model is able to exploit the characteristics of the other agents in the market to separately identify and estimate influence and sorting. Both effects are found to be significant, with sorting almost twice as important as influence for the difference in IPO rates.

---

* University of Chicago, GSB. I want to thank Tim Bresnahan for inspiration and encouragement throughout this project and Susan Athey, Patrick Bajari, Peter Coles, Liran Einav, Thomas Hellmann, Yael Hochberg, Steve Kaplan, Laura Lindsey, Joao Manoel de Mello, Manju Puri, Garth Saloner, Ed Vytlacil, Pai-Ling Yin, an anonymous referee, and many seminar participants for helpful discussions and comments. I gratefully acknowledge financial support from the Danish Research Academy, the Forman Fellowship, and the Kapnick Foundation, which provided a grant through the Stanford Institute for Economic Policy Research.

Venture capitalists (VCs) invest in privately held entrepreneurial companies. They are actively involved in the monitoring and management of these companies (Gompers and Lerner (1999), Gorman and Sahlman (1989), Sahlman (1990)), they typically own a substantial fraction of the companies, usually in the form of convertible securities (Casamatta (2003), Cornelli and Yosha (2003), Hellmann (1998), Kaplan and Strömberg (2003)), and they are commonly represented on the boards of directors (Lerner (1995)). Kaplan and Schoar (2005) find significant persistence in the VCs' returns and conclude that the most likely explanation is heterogeneity in the skills of these investors.

I find that companies with more experienced investors are more likely to go public. This finding is driven by two distinct effects, *influence*, which means that experienced VCs add more value and bring companies public at a higher rate, and *sorting*, which means that more experienced VCs invest in better companies. The main objective here is to develop an econometric model to distinguish these two effects and to estimate and thereby quantify their relative importance. Both effects are found to be statistically and economically significant, with sorting appearing almost twice as important as influence for explaining the observed differences in IPO rates across VC investors.

More experienced VCs can influence and add value to companies in several ways. First, experienced VCs may be better at monitoring and managing companies. Second, they may have access to larger networks, drawing on a greater number of contacts with suppliers, customers, and potential managers (Hellmann and Puri (2002), Hochberg, Ljungqvist, and Lu (2006)). Third, the reputation of an experienced VC may communicate unobserved qualities about the company to the market, increasing the market value of the company (Megginson and Weiss (1991)). In contrast, sorting arises from the matching between VCs and companies. Companies care about the identity of the investor. When faced with multiple offers, companies routinely turn down the investor with the best financial offer in favor of an investor that adds more value in other ways (Hsu (2004)). Since companies more willingly accept financing from better VCs, these VCs have more feasible investments to choose from. Anecdotal evidence suggests that VCs consider this access to a proprietary deal flow to be a distinct competitive advantage. Sorting predicts that investments by experienced VCs perform better because the companies themselves are inherently better, *not* because the

VCs increase the value of these companies. Clearly, sorting and influence are not mutually exclusive effects. The question is how to empirically separate these two effects and measure their relative importance.

Viewed in a classical regression framework, the investor's experience becomes endogenous when sorting causes experienced investors to invest in companies that are better along a number of dimensions that are unobserved in the data. Companies with better unobserved characteristics, as captured by the error term in the regression, match with more experienced investors. The error term becomes positively correlated with experience, and the estimated coefficient is biased upwards relative to the investors' actual influence. A simple example illustrating this bias is given below. This problem is particularly severe for entrepreneurial companies. Entrepreneurial companies have short operating and financial histories. Little information is systematically observed for these companies, and VCs emphasize intangible qualities such as the quality of the management team, competition, customer adoption, and general uncertainties facing the business when they evaluate these investments (Kaplan and Strömberg (2004)). Not surprisingly, the results below indicate substantial endogeneity, and the sorting on the unobserved characteristics appears stronger than the sorting on the observed characteristics.

The classical solution to the endogeneity problem is to estimate the model using instrumental variables. For each investment, the instrument must be independent of the outcome but related to the experience of the investor making the investment. Unfortunately, the economics underlying the decisions make it difficult to find such an instrument. The matching of a given company with an investor with a certain level of experience is determined by their mutual investment decisions, and when their characteristics only relate to the investment decision to the extent the characteristics affect expected returns, they are not valid instruments. For example, the distance between the investor and the company is one candidate instrument,[1] as VCs are more likely to invest in companies that are located closer (Lerner (1995)). However, because it is also easier for VCs to monitor and manage these companies, the VCs' preference for these investments may be due to the higher return from the improved monitoring and managing, in which case distance is not independent of outcome and hence it is not a valid instrument. The same argument applies to other characteristics as well.

To overcome the problem of missing instruments, I develop a new structural model that exploits the implications of sorting to separate sorting from influence. Sorting implies that in a market with more experienced investors, a given investor is pushed down in the relative ranking and is left with worse companies. Hence, an investor's investment decisions depend on the characteristics of other agents in the market. However, the outcome of the investment is independent of these other characteristics, and the other investors' characteristics present a source of exogenous variation. This exogenous variation is similar to an instrumental variable, and the structural model uses it to identify influence and sorting. Other papers have successfully exploited this kind of variation as a source of exogenous variation (i.e., Berry, Levinsohn, and Pakes (1995), Bresnahan (1987)) although the methodology employed here differs substantially from their methodologies.

Specifically, the structural model has two parts. The first part consists of the outcome equation, which specifies the outcome of each investment. With sorting and the resulting endogeneity problem, estimation of this equation alone yields inconsistent estimates. The second part controls for sorting, and it is based on a two-sided matching model called the college admissions model (see Gale and Shapley (1962) and Roth and Sotomayor (1990)). This model forms the basis for an empirical model of market sorting, and the resulting empirical matching model is a discrete choice model that generalizes existing models, such as the Probit model, by allowing for interactions among the choices made by different agents. Together, the two parts are analogous to the two stages of the two-stage estimator in the Heckman selection model (Heckman (1976, 1979)). The matching model controls for the sorting and the selection of the observed investments and eliminates bias in the estimation of the outcome equation. Estimation is numerically intensive, but Bayesian estimation using Gibbs sampling (Gelfand and Smith (1990), Geweke (1999)) is a feasible estimation method. The estimation procedure here is related to the procedure used by Geweke, Gowrisankaran, and Town (2003), but it extends their analysis by using the empirical matching model to control for endogeneity rather than the multinomial Probit model, which is unable to capture the sorting that is the source of identification in this paper.[2]

The approach outlined above has certain limitations. The matching model is a static equilibrium model, and it does not capture dynamic features of the market or issues concerned with the timing of investments (see Inderst and Müller (2004)). Second, the matching model assumes complete information, with experienced VCs having access to a superior deal flow because more companies prefer these VCs as their investors, *not* because these VCs are better at screening entrepreneurs (an informational issue).[3] Further, the computational complexity of the structural model necessitates some compromises in the specification of the estimated model. Computation time increases substantially with the number of explanatory variables, and only simple specifications are computationally tractable. I have been unable to estimate the model with fixed effects, random effects, or other more general error structures. Finally, it should be noted that the methodological approach is a structural approach. The model explicitly imposes a number of economic restrictions on the analysis, and the validity of the estimated coefficients depends on these assumptions. The benefit of this approach is that the model solves the problem of missing instruments, more direct methods are unable to do this. In addition, the explicit statement of the underlying assumptions makes the identification and interpretation of the parameters transparent.[4] Hopefully, future extensions of the analysis can address the remaining shortcomings.

The outline of the paper is as follows. Section I illustrates the bias arising in markets with sorting. Section II presents the formal matching model. Section III develops the empirical model and Section IV presents the data. Section V discusses empirical findings, and Section VI concludes. A description of the estimation procedure is in the Appendix.

I. Bias From Two-Sided Matching

Consider a market with two investors, i = A, B, and four companies, j = 1, 2, 3, 4. Each investor has a single characteristic, say experience. Experience is observed, and let the experience of investors A and B be $X_A = 1$ and $X_B = 2$, respectively. Each company also has a single characteristic, say the quality of the management team, but this is unobserved in the data. Let the qualities of the four companies' management be $X_1 = 1$, $X_2 = 2$, $X_3 = 3$, and $X_4 = 4$. Given two companies and four investors, there are eight potential matches, the outcomes of which are determined by the following outcome equation:

$$Y_{ij} = \beta_0 + X_i\beta_1 + X_j\beta_2 + \varepsilon_{ij}\,. \tag{1}$$

Here, Y_{ij} is the outcome, and X_i and X_j are the characteristics of the investor and the company. For simplicity, let ε_{ij} always equal zero, and let the true parameters be $\beta_0 = 0$, $\beta_1 = 0$, and $\beta_2 = 10$. Since β_1 is zero, the outcome does not depend on the investor's experience, and the influence is zero. Assigning a company to a more or less experienced investor leaves its outcome unchanged.

Let investor A invest in companies 1 and 2, and let investor B invest in companies 3 and 4. The bold numbers in Figure 1 illustrate the observed outcomes of these matches.

*** FIGURE 1 ABOUT HERE ***

Since companies' qualities are unobserved, one can imagine using the four observed outcomes to estimate the short outcome equation

$$Y_{ij} = \beta_0 + X_i\beta_1 + \varepsilon_{ij}\,. \tag{2}$$

The ordinary least squares (OLS) estimates of equation (2) are $\widehat{\beta}_0 = -5$ and $\widehat{\beta}_1 = 20$, and the estimated influence of 20 is biased upwards relative to the true value of zero. This bias arises from the systematic matching of companies 3 and 4, which have better unobserved characteristics, to the more experienced investor B. Note, that the model is symmetric. The reverse effect can be present at the same time, and an unobserved investor characteristic that is related to the matching would bias the estimated effect of the companies' characteristics.

To solve this problem, consider observing an additional market with four similar companies, but with an additional investor C. Let $X_C = 3$. The presence of investor C reduces the other investors' relative ranks in the market, leaving investor B with investments in companies 2 and 3. The outcomes of these investments are given by the remaining (not bold) numbers in Figure 1. Again, the estimate of influence from the short outcome equation, now with eight observed outcomes, is biased upwards. However, a direct comparison of the two markets shows that company 2 (and 4) has investors with different experiences, yet these investments have identical outcomes. It follows that the outcomes are unaffected by experience and that the influence is actually zero. Here, the change in the matching between the two markets serves the role of the exogenous variation. For the empirical application, it remains to determine which companies are comparable across different markets. This requires a matching model.

II. Two-Sided Matching Model

In venture capital an investment is a mutual decision, requiring the consent of both the investor and the company. This is captured by a two-sided matching model. This model is a static equilibrium model from cooperative game theory, and the particular model developed here is a variation of the two-sided matching model known as the college admissions model (introduced by Gale and Shapley (1962)). Roth and Sotomayor (1990) present a broad analysis of the properties of this and related models.

A. Main Assumptions

Before the formal model is presented, its main assumptions are discussed. First, the college admissions model is a one-to-many matching model. Each investor can match with several companies, but each company can only match with a single investor. In practice, it is common for companies to receive funding from multiple investors (a *syndicate*) over several investment rounds. One investor is the *lead investor*, and this investor is responsible both for the main contact with the company and for bringing other investors into the deal.[5] The analysis focuses on the initial investment by this investor. Lerner (1994) finds that in the initial investment round VCs with similar levels of experience tend to syndicate, and the lead investor's experience then approximates the experience of the syndicate overall. Lerner (1994) finds further that more experienced VCs tend to enter later investment rounds of companies with better performance. However, including these later financing rounds would complicate the analysis, since the early performance cannot be attributed to these later investors, and separating companies' performance in individual investment rounds is difficult. To the extent that later-round investors influence company performance, part of the lead investor's initial influence can be interpreted as the ability to attract these better investors in later rounds.

Second, investors are restricted in the number of investments they make. An alternative, although less tractable, assumption would be to restrict the amount of capital invested. However, the restriction on the number, rather than the amount, reflects anecdotal evidence (Quindlen (2000)) that VCs' scarce resource is time, not money, and that deals require roughly equal amounts of time. In a systematic study of VCs' investment analyses, Kaplan and Strömberg (2004) find that time commitment is a common concern for VCs when evaluating potential investments.

Finally, the model imposes restrictions on preferences. Each potential match has a valuation, and the valuation represents the expected net present value (NPV), at the time of the investment decision.[6] Investments in better performing companies have higher valuations, and a match with a VC that adds more value also has a higher valuation. The model assumes that the investor expects to receive the fraction λ of the NPV, and the company expects to receive $(1-\lambda)$, where λ is given. This assumption rules out transfers. For example, a weaker company cannot attract a more experienced investor by offering a higher fraction of shares. Theoretically, this is justified on two grounds. First, Holmstrom and Tirole (1997) argue that liquidity constrained entrepreneurs with potential moral hazard (i.e., shirking) have limited *pledgeable expected income*. If the entrepreneur initially pledges too much of the company, the incentives are muted. Second, much of the return is distributed by share allocations in later investment rounds. This ex-post bargaining involves new investors and hence it is difficult ex-ante to commit to transfers. In this situation, standard bargaining models predict a fixed sharing rule. The benefit of the the fixed sharing rule is that it ensures that the model has a unique equilibrium. The standard college admissions model without this restriction has multiple equilibria, and it is well known that the likelihood function is not well defined for games with multiple equilibria (Bresnahan and Reiss (1991)).

Empirically, there is little evidence of transfers. Hsu (2004) finds that more reputable investors initially finance companies at somewhat lower (10% to 14%) pre-money valuations, but it is unclear if this difference is undone in later rounds. Kaplan and Strömberg (2003) find that investment contracts specify future changes in ownership stakes depending on performance, making it difficult to interpret initial differences without accounting for these contingencies.

B. Agents and Matchings

The model has two disjoint sets of agents. The set of investors is I, and the set of companies is J. Each company can match with a single investor, and each investor can invest in a limited number of companies. Let investor i's quota be q_i. The set of *potential investments* (also known as *matches* in the matching literature and *deals* in the VC literature) is $M = I \times J$. A *matching* is a collection of *potential investments*, $\mu \subset M$. Investor i's portfolio is denoted $\mu(i)$, and company j's investor is denoted $\mu(j)$. Using this notation, a match between investor i and company j can be stated in three equivalent ways: as $ij \in \mu$, as $j \in \mu(i)$, or as $i = \mu(j)$.[7]

Investors have preferences over investments in portfolios of companies, and companies have preferences over matches with individual investors. To state these preferences, let each potential match have a *valuation*, and let the valuation of the match ij be denoted V_{ij}. These valuations are assumed to be distinct, and the same valuations determine the preferences of both the investors and the companies. The valuation is divided between the investor and the company according to a fixed sharing rule determined by $\lambda \in (0,1)$, where the investor receives the fraction λ, and the company receives $(1-\lambda)$. Investor i's preferences over portfolios and company j's preferences over investors are then represented by the two profit functions:

$$\Pi_i(\mu(i)) = \lambda \sum_{j \in \mu(i)} V_{ij} \tag{3}$$

$$\Pi_j = (1-\lambda) V_{\mu(j)j} \,. \tag{4}$$

According to these preferences, company *j* prefers a match with investor *i'* to a match with investor *i* when $V_{i'j} > V_{ij}$, and investor *i* prefers an investment in company *j'* to an investment in company *j* when $V_{ij'} > V_{ij}$. Note that these preferences do not depend on λ, and λ is not estimated by the empirical model.

C. Equilibrium

The equilibrium concept used in two-sided matching models is *stability*. A matching is said to be stable when no agent prefers to deviate from the matching and form a new match. The simplest deviation involves a single investor and a single company. When they both prefer to abandon their existing matches and form a new match together, they form a *blocking pair*. A matching without any such blocking pairs is *pair-wise stable*.[8] More complicated deviations involve larger groups of agents, and the corresponding equilibrium concept is *group stability*. In the college admissions model a matching is pair-wise stable if and only if it is group stable (see Roth and Sotomayor (1990) for a formal definition and proof), thus it is without loss of generality to consider only pair-wise stability. Further, in the college admissions model a stable matching always exists (Gale and Shapley (1962)). Since the model here is a special case of the college admissions model, these properties directly apply.

The valuations and the fixed sharing rule impose additional restrictions on preferences, and imply two new results, namely, the equilibrium is unique, and the equilibrium condition can be expressed as a collection of inequalities. Both of these results are important for the empirical application. Without a unique equilibrium, the statistical likelihood function is not well defined (Bresnahan and Reiss (1991)), and without a simple equilibrium condition, the empirical model is intractable. The propositions are given below, and their proofs are in Sørensen (2005).

PROPOSITION 1: *The college admissions model with preferences given by equations (3) and (4) and with distinct valuations has a unique equilibrium.*

To characterize the equilibrium, it is convenient to define $\overline{V}_{ij}$ and $\underline{V}_{ij}$ and the sets *S*(*i*) and S(*j*) as follows:

$$\overline{V}_{ij} \equiv \max[V_{\mu(j)j}, \min_{j' \in \mu(i)} V_{ij'}] \tag{5}$$

$$\underline{V}_{ij} \equiv \max[\max_{i' \in S(j)} V_{i'j}, \max_{j' \in S(i)} V_{ij'}] \tag{6}$$

$$S(i) \equiv \left\{ j \in J : V_{ij} > V_{\mu(j)j} \right\} \tag{7}$$

$$S(j) \equiv \left\{ i \in I : V_{ij} > \min_{j' \in \mu(i)} V_{ij'} \right\}. \tag{8}$$

For an unmatched pair *ij*, $\overline{V}_{ij}$ is the opportunity cost of deviating and forming a new match together. The valuation of company *j*'s current match is $V_{\mu(j)j}$, and the valuation of the worst company in investor *i*'s portfolio is $\min_{j' \in \mu(i)} V_{ij'}$. These matches are abandoned to form the new match, and the opportunity cost is the maximum of their valuations. When V_{ij} exceeds $\overline{V}_{ij}$, investor *i* and company *j* both prefer to deviate and form a new match together. When V_{ij} is below $\overline{V}_{ij}$ for all unmatched pairs, no investor-company combination forms a blocking pair, and the matching is stable.

For a matched pair *ij*, $\underline{V}_{ij}$ is the opportunity cost of remaining together. The sets *S*(*i*) and *S*(*j*) contain the feasible deviations for investor *i* and company *j*, respectively. The feasible deviations for investor *i* are the companies that prefer a match with this investor to their current match. Similarly, the feasible deviations for company *j* are the investors that are willing to abandon one of their current matches to match with this company. The opportunity cost is the valuation of the best feasible deviation. When V_{ij} exceeds $\underline{V}_{ij}$ for all matched pairs, they all prefer their current match to their best feasible deviation and the matching is stable.

The following proposition characterizes the equilibrium using both $\overline{V}_{ij}$ and $\underline{V}_{ij}$. The conditions are equivalent, but both are important for the estimation of the model since they impose different bounds on the latent valuation variables. The first condition presents bounds for the unmatched pairs, and the second for the matched pairs (see equations (A6) and (A7) in the Appendix).

PROPOSITION 2: *The matching μ is stable if and only if $V_{ij} < \overline{V}_{ij}$ for all $ij \notin \mu$. Equivalently, the matching μ is stable if and only if $V_{ij} > \underline{V}_{ij}$ for all $ij \in \mu$.*

Finally, a bit of additional notation is useful. Let the vector of all valuations in the market be $V \in R^{|M|}$. For a given matching μ, the set of valuations for which μ is the stable matching is $\Gamma_\mu \subset R^{|M|}$. The condition in Proposition 2 can then be stated concisely as

$$V \in \Gamma_\mu \Leftrightarrow \left[V_{ij} < \overline{V}_{ij} \quad \text{for all } ij \notin \mu\right] \Leftrightarrow \left[V_{ij} > \underline{V}_{ij} \text{ for all } ij \in \mu\right] \; . \tag{9}$$

III. Structural Empirical Model

The valuations represent agents' preferences. They are unobserved in the data, and in the empirical model they are latent variables. The valuation of each potential match $ij \in M$ is given by the valuation equation

$$V_{ij} = W_{ij}'\alpha + \eta_{ij} , \tag{10}$$

where $\eta_{ij} \in R^k$ is a vector of observed characteristics for investor *i* and company *j*, and $\alpha \in R^k$ contains parameters to be estimated. The error term η_{ij} contains factors that are unobserved in the data. Substituting the valuation equation into the equilibrium condition gives the equilibrium condition

$$\mu \text{ is stable } \Leftrightarrow \eta \in \Gamma_\mu - W\alpha , \tag{11}$$

where $\eta \in R^{|M|}$ are the error terms and $W \in R^{|M| \times k}$ are the observed characteristics for the entire market. The term $W\alpha \in R^{|M|}$ denotes matrix multiplication of W_{ij} with α for each potential match (so

$W\alpha = \{W_{ij}'\alpha, ij \in M\}$). Let $1[\cdot]$ be the indicator function. The likelihood function of the matching model is then

$$L(\mu;\, \alpha) = Pr(\eta \in \Gamma_{\mu} - W\alpha) = \int 1[\eta \in \Gamma_{\mu} - W\alpha]dF(\eta)\ . \tag{12}$$

When several independent matching markets are observed, the likelihood function is the product over these markets, and, at least in principle, α can be estimated directly by maximizing this function.

The empirical matching model is a discrete choice model, and its parameters are only identified up to scale and level. This is natural, since the valuations represent preferences and these are unaffected by a change in the level or scale of the valuations. This means that the constant term (and other characteristics that are constant within each market) is excluded from *W*, since the corresponding coefficient is unidentified. This normalizes the level of the parameters. The scale is normalized by setting the variance of the error term equal to one.[9]

The second part of the structural model is the outcome equation. For each $ij \in M$, let

$$Y_{ij} = X_{ij}'\beta + \varepsilon_{ij}\ , \tag{13}$$

where X_{ij} contains observed characteristics and β contains parameters to be estimated. The error term contains factors that are unobserved in the data. Some of these factors may be observed by the investor and the company at the time of the investment decision. Since only the binary outcome IPO_{ij} is observed in the data, Y_{ij} is a latent variable, and following a standard Probit specification, the outcome of the investment is a public offering (IPO_{ij} = 1) when Y_{ij} is greater than zero, that is,

$$IPO_{ij} = 1[Y_{ij} \geq 0]\ . \tag{14}$$

Like other discrete choice models, the sign and level of the coefficients are identified, but the scale is unidentified and is normalized by fixing the variance of the error term. Usually, this variance is set equal to one, but for computational reasons a slightly different normalization is used below.

A. Distribution of Error Terms and Interpretation of Parameters

The error terms are assumed to be independent of *X* and *W*, and this assumption identifies the parameters of the model. Since the outcome equation is defined for all $ij \in M$, the estimated parameters predict the outcomes of all potential matches, not just the observed ones. The estimated coefficient on the investor's experience reflects the predicted change in the outcome for a given company following an increase in the investor's experience. This change is the effect after controlling for the sorting in the market, and thus represents the investor's influence.

The coefficients in the valuation equation capture preferences over matches. When the coefficient on the investor's experience is positive, matches with more experienced investors have higher valuations and rank higher in the companies' preferences. Similarly, if the coefficient on the stage of the company is positive, investors prefer investment in late-stage companies over early-stage companies. The larger these coefficients, the stronger the preferences, and the stronger is the sorting over the observed characteristics. If the coefficients are close to zero and insignificant, then either the matching is random or it depends on characteristics not included in the valuation equation.

For tractability, the joint distribution of (ε_{ij}, η_{ij}) is assumed to be independent for different matches and to follow the bivariate normal distribution

$$\begin{pmatrix} \varepsilon_{ij} \\ \eta_{ij} \end{pmatrix} \sim N\left(0, \begin{bmatrix} 1+\delta^2 & \delta \\ \delta & 1 \end{bmatrix}\right). \tag{15}$$

With a normal prior distribution (*conjugate prior*), the posterior distributions are also normal or truncated normal.[10] However, normality is not essential for the estimation or identification of the model, and it could be relaxed. The variances of the two error terms normalize the scales of the two equations. The variance of η_{ij} is set to one, and the variance of ε_{ij} is set to $1 + \delta^2$. This normalization is convenient for the estimation and is without loss of generality (see the Appendix).

The covariance between the error terms captures unobserved factors that affect both the outcome and the valuation of a match. For example, a company may have a particularly strong management team. This is unobserved in the data, but it is partly observed by the VC before investing and it is important for the outcome. To the extent it affects the valuation of the investment and its outcome, it enters the error terms in both the outcome and valuation equations, inducing a positive correlation between these two error terms. The covariance therefore reflects factors that are unobserved in the data but affect the outcome and are taken into account in the initial valuation. This captures sorting over characteristics that are unobserved in the data.

B. Interaction, Estimation, and Identification

In the matching model investment decisions interact. When an investor invests in a company, other investors cannot invest in this company and their investment decisions interact. Interaction leads to sorting, and interaction and sorting are fundamental properties of the model that have implications for both estimation and identification.

For the estimation, interaction means that each agent's investment decision cannot be analyzed in isolation. Unlike the Probit model, for example, the likelihood function does not factor into a product over the likelihood of each agent's action (matching decision).[11] To evaluate the likelihood function, all error terms must be integrated simultaneously. The dimensionality of this integral runs into the thousands, and currently it is not possible to evaluate such high-dimensional integrals with the speed and precision required for maximum likelihood estimation (Judd (1998)). However, Bayesian estimation using Markov Chain Monte Carlo (MCMC) circumvents this integration problem. Berger (1993), Tanner (1998), and Robert and Casella (2004) provide introductions to this estimation method.

While interaction and sorting complicate the estimation, they also provide a solution to the endogeneity problem. As illustrated in the initial example, the endogeneity problem arises from sorting on unobserved company characteristics, captured by the error term in the matching model. Because of sorting and interaction, the presence of other agents (and their characteristics more generally) affects investment decisions and leads investors with differing experiences to invest in companies with similar unobserved characteristics (similar error terms) for exogenous reasons. Implicitly, this facilitates the direct comparison illustrated in Section I and identifies the parameters. The identifying assumption is thus that the presence and characteristics of the agents in each market are exogenously given and independent of the error terms of the model.

Bayesian estimation requires imposing prior distributions on the parameters. These are taken to be independent normal distributions, and the prior distributions of α, β, and δ are denoted $N\left(\overline{\alpha},\Sigma_{\alpha}\right)$, $N\left(\overline{\beta},\Sigma_{\beta}\right)$, and $N\left(\overline{\delta},\Sigma_{\delta}\right)$, respectively. Together with normal distributions over the error terms, the normal priors mean that the distributions simulated by the estimation procedure are either normal or truncated normal, which significantly improves the numerical tractability of the model. In the estimation below, the prior distributions have means of zero and variances of 10. Other studies more carefully construct prior distributions (i.e., Geweke, Gowrisankaran, and Town (2003)), but here the primary goal is simply to use as uninformative a prior as possible while maintaining a proper posterior distribution. For all estimated parameters, the prior variances are at least 110 times greater than the posterior variance. This suggests that the prior is fairly uninformative, and that the posterior distributions and the estimated parameters primarily reflect information in the data. Further increases in the prior variances leave the estimated parameters and the empirical results largely unchanged.

IV. Data Description

The Venture Xpert data provided by SDC (now owned by Thomson Financial) are used for the estimation. SDC began compiling data on VC investments in 1977 and has supplemented their data with investments dating back to the early 1960s. These data are used in several previous studies of venture capital (e.g., Kaplan and Schoar (2005), Lerner (1995)). The completeness of the data are investigated by Gompers and Lerner (1999) and Kaplan, Sensoy, and Strömberg (2002), who find that the SDC data contain most VC investments and missing investments tend to be among the less significant ones.

A. Description of Sample

The full database contains 22,747 VC investments made between 1975 and 1995. I estimate the model using a restricted sample of these observations. In particular, the sample is restricted to investments received by young companies,[12] with investments in restructurings or buy-outs of more mature companies eliminated. The sample contains only the first investment for each company, and in cases in which several investors participate in this round, the sample includes only the investment by the lead investor. This investor is identified as the VC firm that makes the largest total investment in the company.[13] The sample is then restricted to investments made from 1982 to 1995. It takes companies several years to go public after the initial investment, and restricting the sample to investments up to the end of 1995 leaves companies seven years to go public (the status of the companies is current as of March 2003). The 14-year period ensures a sufficient sample size.

Next, I divide the investments in the sample into markets. Each market contains the investments in companies located in the same geographical state and receiving investments within the same half-year (January to June or July to December). The two states with the most VC investments are California and Massachusetts, and the sample is restricted to companies located in these two states. This leaves 56 markets: two states, for fourteen years, with two markets in each year. Finally, the sample is restricted to investments by VC firms that make a total of 10 or more investments in these 56 markets. This restriction ensures that the investors in the final sample are active VCs, and excludes investments by smaller more idiosyncratic investors. The final sample consists of 75 investors and 1,666 investments in 1,666 companies.

B. Description of Variables

Descriptive statistics of the variables are presented in Table I. In the model the two endogenous variables are the outcomes and the matching. The variable *IPO*, which records investment outcomes, equals one if the investment results in a public offering, and zero otherwise. This is a coarse measure of investment outcomes, but it is frequently used in the literature (i.e., Brander, Amit, and Antweiler (2002)). Gompers and Lerner (2000) compare this measure to broader measures that also include companies that are acquired or in registration. They find that the different measures give qualitatively similar results. The second endogenous variable is the matching of the investors and the companies. The identities of the VC and the company are observed, and this matching is observed directly in the data.

Only a few coarse exogenous characteristics are observed for the sample companies. This reflects a general problem for studies of entrepreneurial companies. By nature, these companies have short operating and financial histories, and thus many important characteristics are difficult to summarize in a systematic way. The following variables are included in the analysis. First, *YEAR* captures the year of the investment. Second, *STAGE* contains the company's stage of development at the time of the investment. A company is an early-stage company when it is at the seed or start-up stage, and it is a late-stage company when it is at the expansion or later stage. This distinction roughly corresponds to whether the company has regular revenues or not.[14] In the sample, 82% of the companies are early-stage companies. Third, *INDUSTRY_GROUP* divides the companies into six major industry groups: "Communications and Media," "Computer Related," "Semiconductors and other Electronics," "Biotechnology," "Medical, Health, and Life Sciences," and "Other."

The observed investor characteristics are the amount of available capital and investor experience. In particular, *FUND_SIZE* contains the size of the VC fund making the investment, measured in millions of dollars. VC firms are organized as collections of funds, and the funds are usually prevented from investing in the same companies to avoid conflicts of interests among the participants in different funds. The size of the individual fund is thus an appropriate measure of capital available for investing in a company. The fund size is defined as the sum of the fund's investments in the data.[15] For a number of investments, only the firm and not the fund is disclosed. For these investments, the fund size is set to the average of the sizes of all the firm's funds.[16]

At the time of each investment, *TOTAL_EXPERIENCE* measures the investor's experience by counting the number of investment rounds the VC firm has participated in since 1975. All investments in the unrestricted sample are counted, including investments in companies in all states, at all stages, across both initial and later rounds. Data for investments before 1975 are less accurate, but the VC industry was much smaller then. It expanded significantly after regulatory changes in the late 1970s, and relatively few investments are lost by excluding the earlier period. Note that since experience is calculated at the time of each investment, an investor's experience increases over the sample.

Experience is interpreted as a measure of an investor's ability to influence and add value to a company. Experienced investors may have greater ability for two reasons. First, participating in more investments allows investors to learn more about managing and monitoring companies and to expand their network with managers, suppliers, and customers. In this case, experience increases ability. Alternatively, investors may have an inherent ability, and to the extent investors with higher abilities are more likely to survive, greater experience signals higher inherent ability. While this distinction is interesting, it is not pursued here. The two reasons are not exclusive, and regardless of whether ability is caused by experience or ability causes experience, experience measures ability.

Alternative experience measures used in the literature are the age of the VC firm, the cumulative total amount invested, and the number of companies (see Gompers (1996) and Hochberg, Ljungqvist, and Lu (2006)). Age is a less attractive measure of investor ability, since it does not distinguish active and inactive investors. The number of investment rounds is also preferable to the number of companies, since some investors are involved at early stages and help companies develop and mature while others mainly invest at final stages. These two types of investors may invest in the same number of companies, but the experience of the first type is likely more relevant for influencing companies. These investors participate in more investment rounds in each company, and counting rounds rather than companies provides a simple adjustment for the involvement of the investors. Similarly, the number of investment rounds is preferable to the amount invested. Later rounds tend to involve larger amounts, and the latter measure emphasizes these rounds and is less indicative of VCs' ability.

*** TABLE I ABOUT HERE ***

V. Empirical Findings

To examine the relationship between the experience of the VCs and the outcome of their investments, the 1,666 companies are divided into 10 groups according to the experience of the investor. The first group contains companies with investors with experiences between 0 and 24, the second group contains companies with investors with experiences ranging from 25 to 49, and so forth up to the tenth and final group, which contains companies with investors with an experience greater than 225. In the sample, 95% of the investments are made by investors with an experience between 0 and 225. Figure 2 illustrates the number of companies in each group. Figure 3 illustrates the fraction of companies in each group that go public, and it appears that investments made by experienced investors are more likely to go public. The group with the least experienced investors in the sample observes the lowest IPO rate, at 19.6%. The highest IPO rate, 46.2%, is achieved by investors with experiences between 150 and 175. The decrease in the IPO rate for investors with higher experience is partly explained by the greater sampling variation due to the low number of observations for these investors. This evidence suggests a positive relationship between experience and investment outcomes.

*** FIGURE 2 AND FIGURE 3 ABOUT HERE ***

The experience of each investor increases over time, and thus investments by the same investor fall into increasingly higher groups. If the IPO rates were also increasing over time, this would create a spurious relationship between the IPO rate and experience, similar to the pattern observed in Figure 3. The amount of capital available to an investor is also likely to increase with the investor's experience, and if, for some reason, the amount of capital contributes to the performance of the investments, this may also explain the observed pattern. To investigate these explanations, a Probit model is estimated. In this model, the probability that an investment results in a public offering is estimated as a function of the characteristics of the agents participating in the deal. Estimates of this model are presented in Table II and Table III.

*** TABLE II AND TABLE III ABOUT HERE ***

In all specifications, investor experience enters positively and significantly. The first two specifications in Table II allow experience to enter linearly or in logs. The last specification allows the time of the investments to enter linearly rather than as individual year controls, and for this specification the standard errors are calculated both with and without clustering at the investor level.[17] In all cases, investor experience has a positive and statistically significant effect on the outcome of the investment. Investments made by more experienced investors are thus more likely to result in public offerings.

To provide a sense of the economic magnitude of the above result, the IPO rate for an investment made by an entirely inexperienced investor is compared to the IPO rate for a more experienced investor. Using the estimates from Specification 1 in Table III, an investor without any experience has an estimated IPO rate of 21.4%, whereas the IPO rate for an investor with an experience of 225 is 38.9%. Thus, after controlling for the time and fund size, an investment by the more experienced investor is 82% more likely to succeed than the investment by the inexperienced investor, a substantial economic difference.

The other coefficients in Tables II and III are not surprising. The amount of capital available to the investor has a small but significant positive effect on the investment outcome. In the specifications that include the location of the company (California or Massachusetts), location has no effect on the outcome. More surprisingly, the year of the financing has little effect on the outcomes. In Specifications 1 and 2 in Table III, the year of the investment enters through controls for the individual years (the base year is 1982), and no apparent trend in the estimated coefficients obtains. In Specification 3, the time of the investment enters as a linear time trend, and the estimated coefficient is insignificant. There is no indication that the IPO rate increases with the development of the venture capital industry through this period, and there is only weak evidence that companies receiving their first financing at the end of the sample period are less likely to go public because of the subsequent downturn in the market. Apparently, the growth in the VC market over the sample period has followed proportional growth in the number of companies going public.

C. Specification of Structural Model

Estimation of the structural model is numerically intensive, and only parsimonious specifications are numerically tractable. First, *STATE*, which indicates the location of the company, is insignificant in all specifications of the Probit model, and thus this variable is excluded from the estimation of the structural model.[18] Next, a comparison of Specification 1 and Specification 2 in Table II suggests that letting experience enter linearly or taking the logarithm of experience is immaterial, and thus just the linear specification is estimated below. Further, there is no apparent time trend in the data, and thus the individual year dummies are replaced with a linear time variable. Finally, the structural model is derived using a simple error structure that excludes correlations between the error terms for different matches involving the same company or investor. In particular, one may suspect some correlation between investments made by the same investor. To provide one measure of the magnitude of this correlation, the Probit model in Specification 3 of Table II is estimated both with and without clustering of the error terms at the venture capital firm level. There is no indication that imposing the simpler error structure substantially affects the estimates of the model.

D. Sorting on Observed Variables

From an econometric perspective there are two distinct kinds of sorting, sorting on observed variables and sorting on unobserved variables. Sorting on observed variables is straightforward to verify. Table IV presents three regressions of investor experience on the observable characteristics of the companies. Late-stage and biotechnology companies have investors with significantly more experience than other companies. In Specification 1, an investor in a late-stage company has an experience that is 19 larger than an investor in an early-stage company. This difference is 28% of average experience, which corresponds to 69.6 past investments. On average, an investor in a biotechnology company has an experience of 12 more past investments than an investor in the “Other” category. This difference is 17% of the sample average.

Experience is increasing over time, and companies later in the data have more experienced investors. Including the average experience of the investors in the market in the regression (Specifications 2 and 3) ensures that the estimated coefficients capture sorting within each market rather than a general trend towards late-stage and biotechnology investments.

*** TABLE IV ABOUT HERE ***

The Probit estimates in Table III provide a measure of the economic magnitude of the sorting on the observed variables. Specification 1 does not include the observed characteristics of the companies, and hence the estimated coefficient reflects the total effect, including sorting on both observed and unobserved variables and investor influence. Specifications 2 and 3 include the observed characteristics of the companies in the specification. By including these characteristics, the estimates control for sorting on observable variables. As the table reports, late-stage companies and companies in the biotechnology industry are more likely to go public, and the estimated effect of experience now captures only the sorting on the unobserved variables and investor influence. The decrease in the estimated coefficient on experience when moving from Specification 1 to Specifications 2 and 3 measures the effect of the sorting on the observed variables on the outcome of the investment. This effect is small, at roughly one tenth of the total effect of experience, and it is smaller than one standard error of the estimated coefficient. As argued above, the observed company characteristics are coarse, and it is not surprising that sorting on these characteristics is less important than sorting on the unobserved characteristics.

The estimate of the valuation equation of the structural model provides additional evidence of sorting on the observed variables. Estimates of this equation are presented in the lower half of Table V. The valuations represent the agents' preferences over matches, and the estimated coefficients in the valuation equation represent the agents' preferences over the characteristics included in the equation.

*** TABLE V ABOUT HERE ***

The estimated coefficient on experience is positive and statistically significant.[19] Matches with investors with greater experience have higher valuations and are more attractive on average. Depending on the interpretation of experience, this may be because experienced investors have acquired better skills, or because experience signals higher inherent ability. Clearly, companies have additional information about investors, and any given investor may be more or less attractive than indicated by their experience. However, on average, companies prefer investors with higher experience.

In Table V, the positive and significant coefficients on the stage and biotechnology variables show that VCs prefer late-stage and biotechnology companies to other companies. Taken together, the findings that companies prefer more experienced investors and that investors prefer late-stage and biotechnology companies imply that there is sorting in the market, with more experienced investors investing in late-stage and biotechnology companies. This is consistent with the regression of experience on company characteristics in Table IV.

It is not obvious how to interpret the coefficients in the valuation equation. As in a Probit model, the coefficients are only identified up to scale, and they are normalized by the variance of the error term. In a Probit model the economic magnitudes of the coefficients are usually interpreted by calculating their marginal effect. An analogous concept is derived here for the matching model. Consider an investor facing a choice between two companies. If the companies have identical observed characteristics, the choice depends entirely on the unobserved factors, and the probability that the investor prefers one to the other is 50%. If one of these otherwise identical companies is a late-stage company and the other is an early-stage company, the coefficients in the valuation equation predict that the probability that the investor prefers the late-stage company is 54.7%. The marginal increase in the probability is 4.7%. These marginal probabilities are presented in Table V in the dP/dW column.[20] The probability that the investor prefers to invest in a biotechnology company relative to a company in the "Other" industry group is 57.9%. From the perspective of a company comparing two otherwise identical investors, one of which has one additional unit of experience, the company prefers the experienced investor with probability 50.82%. The marginal increase of 0.82% translates into a preference for an investor with an experience of 225 relative to an investor with no experience with probability 86.7%. These numbers indicate that companies have stronger preferences over investors with different degrees of experience than investors have over companies at different stages and in different industries. This is not surprising. Experience is a more direct measure of an investor's ability and is closely tied to companies' preferences. The stage and industry of a company are, at best, indirect measures of quality, and these characteristics appear to be less important for investors' preferences.

E. Sorting on Unobserved Variables

Sorting on unobserved variables is captured by the error term in the valuation equation. When experienced investors invest in companies with better unobserved characteristics, the error terms in the valuations of their matches are greater on average. The vertical axis of Figure 4 plots the average of the error terms in the valuations of the companies each investor invests in. A high value means that the investor invests in companies with better unobserved variables. To reduce the noise in the figure, only investors that make at least two investments in the market are included. On the horizontal axis is the relative experience of the investor, given as the percentile rank of the investor's experience in the market. The investor with the highest experience in the market has a relative experience of one.

*** FIGURE 4 ABOUT HERE ***

Figure 4 shows an increasing relationship between the error terms and the investors' relative experience. Investors with relatively higher experience can invest in companies with better unobserved characteristics, and this reflects the sorting on the unobserved variables. In Figure 4, the average of the error terms for investors between the 75th and 100th percentiles is 1.6, and the average for investors between 0 and the 25th percentiles is 0.8. Not only are more experienced investors more likely to invest in the more attractive late-stage and biotechnology companies (reflecting sorting on observed variables), but moving from the bottom quartile to the top quartile in terms of experience increases the unobserved valuation of the investments by roughly 0.8 (= 1.6 - 0.8). This increase is comparable to the coefficients in the valuation equation. In particular, it is almost twice the valuation gained from investing in a late-stage and biotech company, 0.1657 + 0.2808 = 0.45 (see Table V), and the magnitude of sorting on unobserved variables appears to be at least as significant as the sorting on observed variables.

F. Testing for Sorting on Unobserved Variables

The correlation between the error terms in the valuation and outcome equations provides a statistical measure of the sorting on unobserved variables. When an unobserved variable affects the outcome in a positive way, the error term in the outcome equation is positive. When this variable is recognized in the valuation, the error term in the valuation equation is also positive, and sorting on unobserved variables causes a positive correlation between these two error terms. Absent sorting on unobserved variables, the two error terms would be independent. Thus, a statistical test for sorting is to test whether the covariance between the error terms is significantly greater than zero. The covariance is normalized to being positive, and the appropriate test is a one-sided test. In Table V, the covariance is 0.22, which is statistically greater than zero at the 1% level (one-sided). The hypothesis of no sorting on unobserved variables is therefore rejected.

G. Estimating Influence

The coefficients in the outcome equation (in Table V) estimate the dependence of the outcome on the agent's characteristics after controlling for sorting in the market. The coefficient on experience is positive and significant, which provides direct evidence of investor influence. The marginal effect of experience is 0.000346. An investment by an investor with no experience in an average company leads to a public offering with probability 14.3%. Holding the company fixed but replacing the investor with an investor that has an experience of 225 increases the IPO rate to 22.5%. The influence of the more experienced investor increases the probability of going public by 57.3%, a substantial improvement.

The other estimated parameters in the outcome equation are not surprising. Fund size has a small but significant positive effect on the investment outcomes. Late-stage companies have a 4.9% larger probability of going public than early-stage companies. In the sample, biotechnology is the best performing industry, and the average biotechnology company has a 39% larger probability of going public than the average company in the "Other" industry group.

On average, experienced investors bring companies public at a higher rate, but one may suspect individual variation across investors' skill. To determine the magnitude of the individual components, a fixed investor effect is included in the specification. Unfortunately, the computational requirements of the model increases with the number of explanatory variables, and it is not possible to include separate effects for all 75 investors. Instead, fixed effects are included for the five most experienced investors in sample. This specification is Specification 3 in Table VI. The estimated coefficients provide evidence of significant differences in the performance of individual investors. Specifically, *INV1*, the most experienced investor, and *INV3*, the third most experienced investor, achieve significantly better outcomes than the other investors (the base group is the 70 remaining investors); *INV2* and *INV4* are performance neutral, and *INV5* underperforms relative to the other investors. The economic magnitudes are significant. Matching with *INV1* increases the IPO rate by 17% and matching with *INV5* decreases the IPO rate by 14% relative to the average investor with similar experience. This preliminary result suggests that there is a strong investor-specific component, and thus that the choice of a particular investor is important for companies.

*** TABLE VI ABOUT HERE ***

The finding that more experienced investors have substantial influence on investment outcomes complements previous empirical studies of VCs. Hellmann and Puri (2000, 2002) compare companies that receive VC financing to other companies and find that VC-backed companies bring products to market sooner and replace management faster. However, due to endogeneity, it is not obvious whether these differences are caused by the VCs, or whether VCs favor investments in companies with products closer to shipping and with management that is easier to replace. The results here suggest that within a narrow group of VCs (i.e., rather than comparing VC to non-VC investments) there are substantial differences in influence. These differences are likely greater when comparing VCs to other investors, since these investors differ along institutional and organizational dimensions in addition to experience.

The endogeneity problem means that studies that determine investor influence by regressing investment outcomes on characteristics are likely to overestimate influence. If estimates of the Probit model in Tables II and III were interpreted as investor influence (implicitly assuming no sorting on unobserved characteristics), the impact of experience would be overstated by as much as 90%. In the Probit analysis that controls for all observable characteristics (Table III, Specification 2) the marginal effect of experience is 0.000657. The correct marginal effect from the outcome equation is 0.000346 (Table V), and if the first estimate were interpreted as influence, it would overestimate the effect by as much as 90%.

H. Relative Magnitudes of Influence and Sorting

The total effect of experience is now decomposed into the effect of sorting and the effect of influence. In Figure 5, experience is on the horizontal axis and the IPO rate is on the vertical axis. The solid line represents the observed IPO rate as a function of investor experience, and is calculated from the Probit estimates (Specification 1 in Table II). The slope reflects the total effect of influence and sorting. The broken line represents the IPO rate and is calculated from the estimates of the outcome equation (Table V). It reflects the IPO rate for an average company when matched with investors with different levels of experience. If the matching of the investors and the companies were random, there would be no sorting, and the two lines would be on top of each other.

*** FIGURE 5 ABOUT HERE ***

In Figure 5, the distances *A* and *B* represent the impact of the sorting on the IPO rate, and distance *C* represents the impact of influence on the IPO rate. The effect of sorting differs for more and less experienced investors. For experienced investors, sorting leads to better matches and higher IPO rates than if the matching were random. For inexperienced investors, the effect is ambiguous. On the one hand, sorting leaves inexperienced investors with a smaller set of feasible investments, as the experienced investors pick the best companies. On the other hand, sorting leads inexperienced investors to make better investments from this smaller set of feasible investments. The first effect is negative and the second is positive, and relative to a market with random investments, inexperienced investors can be either better or worse off. In Figure 5, the distance *A* is positive, which suggests that the positive effect dominates. However, this finding is particularly sensitive to the limitations in the error structure, and should be interpreted with caution.

For an investor with zero experience, sorting raises the IPO rate from 14.3% to 21.5%, represented by the distance *A*. For an investor with an experience of 225, sorting raises the IPO rate from 22.5% to 38.8% relative to random matching. This increase is represented by the distance *B* in Figure 5. The increase in the IPO rate for the experienced investor is 16.3%, which is more than twice the increase of 7.2% for the inexperienced investor. The influence of the more experienced investors is represented by the distance *C*. Absent sorting, an inexperienced investor would have an IPO rate of 14.3% and an investor with an experience of 225 would have an IPO rate of 22.5% when investing in an average company. Here the company is held constant, and thus the increase of 8.2% measures the influence of the experienced investor. Comparing these numbers gives us a first estimate of the relative impact of sorting and influence. For an investor with an experience of 225, the combined effect raises the IPO rate from 14.3% to 38.8%, represented by *B* plus *C*. Influence accounts for 8.2 of the percentage points (represented by *C*) and sorting accounts for the remaining 16.3% (represented by *B*). Thus, roughly speaking, influence explains one-third and sorting explains the remaining two-thirds of the increase in the IPO rate when moving from the least to the most experienced investor.

I. Comparing Coefficients in Outcome and Valuation Equations

It is interesting to note the relationship between the coefficients in the outcome and valuation equations. Both equations are only identified up to scale, and only the relative magnitudes of the coefficients are comparable. The structural empirical model does not impose any relationship between the coefficients in these two equations, and they are estimated from different endogenous variables. The outcome equation is estimated from the observed outcomes, whereas the valuation equation is estimated from the latent valuations, which are determined by the observed matches. Nothing inherent in the model or estimation leads these coefficients to be similar.

However, the economic assumptions underlying the model suggest that the coefficients are related. When investors and companies generate preferences and valuations from a fixed sharing rule, matches with better outcomes have higher valuations. This appears consistent with the estimated coefficients. All coefficients in the two equations have the same signs, and their relative magnitudes are somewhat similar. For example, in the valuation equation, “Biotechnology” is the most attractive and “Computer Related” is the least attractive of the industries, and this ranking also holds in the outcome equation. This supports the model and the interpretation of the valuations as the NPV of the individual matches.

J. Alternative Market Definitions

In the structural model, the definition of a market is somewhat arbitrary. Table VII presents estimates of the model with alternative market definitions. Specifically, the model is estimated with four different market definitions, and the estimated coefficients are remarkably robust. In the baseline specification, a market contains investments during the same half-year in companies located in the same state (either California or Massachusetts). In the second specification, a market is extended to a whole year rather than a half-year. In the third specification, California and Massachusetts are treated as the same market, and in the final specification a market lasts a whole year and California and Massachusetts are considered the same market. The estimated coefficients on stage and industry, and their standard errors, are robust to the different market definitions. The coefficient on experience varies between 0.0011 and 0.0014, and no apparent pattern emerges in the estimated coefficients.

*** TABLE VII ABOUT HERE ***

Estimating the model with different market definitions provides a heuristic test of the identifying assumptions of the model. The critical identifying assumption is that investors and companies are exogenously assigned to different markets. This assumption would fail if investors and companies choose to participate in a market based on the unobserved characteristics of the other agents in the market. For example, if California has particularly good companies, this could attract Massachusetts investors. Essentially, this means that investors sort between the two states, as well as within each of them, and a more appropriate market definition would treat the two states as being the same market. The results in Table VII show that changing the market definition in this way leaves the estimated coefficients largely unaffected. Similarly, investors and companies may attempt to time the market and increase their activity when the market is particularly good. This corresponds to temporal sorting. Again, Table VII shows that extending the market definition temporally leaves the results largely unchanged. Overall, this informal test provides no indication that the identifying assumptions are violated or that the analysis is sensitive to the particular market definition.

VI. Summary and Conclusions

This paper develops new econometric methods to analyze the market for venture capital investments. It finds that companies with experienced investors are more likely to go public. This follows both from the influence of these investors and from sorting in the market. Influence means that experienced investors add more value than inexperienced investors. Sorting means that experienced investors invest in companies that are inherently better, and hence are associated with higher IPO rates. The results show that both sorting and influence are significant. Companies prefer experienced investors, investors prefer investments in late-stage and biotechnology companies, and these agents tend to match. Sorting also occurs based on variables that are unobserved in the data. These variables include, for example, the unobserved quality of the management team or the product. Not surprisingly, sorting on unobserved variables is at least as strong as sorting on observed variables.

Distinguishing influence and sorting is important for understanding venture capital. VCs tend to emphasize their influence and ability to work with entrepreneurs, while some entrepreneurs are less enthusiastic, some even claiming that VCs are merely “Vulture Capitalists,” offering expensive financing and little else. The results suggest that good VCs have substantial positive influence on companies. However, the analysis considers the 75 most active VCs, and these are probably more influential than the average VC. The results indicate a decline in the influence within this sample, and extrapolating suggests that some inexperienced investors offer very little influence, in which case the epithet above may be appropriate for these investors.

Of course, sorting means that entrepreneurs who are funded by these VCs are not the most promising entrepreneurs either.

The analysis leaves interesting questions unanswered. One may ask, for instance, whether VC financing is more important for certain kinds of firms. In the specification of the outcome equation, experience enters linearly and influence applies equally to all firms. In principle, more flexible specifications can measure synergies between certain investors and entrepreneurs. However, data and computational tractability are limiting factors.

It is also tempting to consider the efficiency of the VC market. In the structural model, the matching is inefficient, since it does not maximize the sum of the valuations. The source of the inefficiency is the restriction on transfers, and with unrestricted transfers the equilibrium would maximize this sum. It is possible to estimate the potential improvement in the valuations, but this is unlikely to give a good measure of inefficiencies for several reasons. First, for this analysis of efficiency the assumption of transferability is critical. The assumption imposes a market failure on the model, and the inefficiency is assumed rather than estimated. For the present analysis, it is sufficient that the matching model provides a good approximation of the sorting, even if the valuations are not exactly correct. Second, the main inefficiencies in this market may be institutional rather than artifacts of the matching. Sahlman (1990) argues that the contractual and organizational structures are primarily designed to overcome problems with agency, informational asymmetries, and illiquidity. Investigating the efficiency of the VC market is interesting but must address a broader range of concerns than just matching.

Finally, the empirical matching model is of independent interest, and it may help further our understanding of other markets with sorting and influence. Some examples are the labor market and the matching of employers and employees, the lending market and the matching of borrower and lenders, the market for corporate mergers and the matching of acquirers and targets, and the market for education and the matching of schools and students (see Roth and Sotomayor (1990)). A better understanding of these markets hinges on distinguishing sorting and influence, in their various forms.

Appendix: Estimation Procedure

The model is estimated using Bayesian estimation using MCMC simulation, which has attractive properties for estimating discrete choice models (Geweke, Keane, and Runkle (1994)). Albert and Chib (1993) and Tanner and Wong (1987) show that treating latent variables as parameters significantly simplifies simulation of the resulting augmented posterior distribution, and the MCMC procedure known as Gibbs sampling (Gelfand and Smith (1990), Geweke (1999)) can simulate this distribution. The procedure is iterative. Each iteration produces a draw from a Markov chain, and under weak regularity conditions that are satisfied here (Roberts and Smith (1994)), the simulated distribution converges to the augmented posterior distribution. The Markov chain is generated by drawing each individual dimension of the joint target distribution conditional on the draws of the other dimensions, and the simulated univariate conditional distributions are derived below.

For notation, let the markets be indexed by $m = 1, \ldots, N$. Let $V_m \equiv \{V_{ij},\, ij \in M_m\}$ and $Y_m^* \equiv \{Y_{ij}, ij \in \mu_m\}$ be the latent valuation and outcome variables in market *m*, and let *V* and Y^* contain these variables in all markets. Let Y_{-ij}^* contain all outcome variables except Y_{ij}, and define V_{-ij} similarly. Let W_{ij} denote the vector of exogenous variables in the valuation equation for the match between investor *i* and company *j*. Let $W_m \equiv \{W_{ij},\, ij \in M_m\}$ contain these variables in market *m*, and let $W \equiv \{W_m,\, m = 1, \ldots, N\}$ contain these variables in all markets. Define X_{ij}, X_m, and *X* similarly for the exogenous variables in the outcome equation. Similarly, define the outcome variables IPO_{ij}, IPO_m, and *IPO*, and the variables containing the matchings, μ_{ij}, μ_m, and μ (here, μ_{ij} is a binary variable that equals one when the match is formed, and μ_m is a subset of M_m). Let $\theta \equiv (\alpha, \beta, \delta)$ contain the parameters of the model. Finally, the densities defined by the model are denoted ϕ and the densities derived for the simulation are denoted π.

The latent outcome variables are defined for all $ij \in M_m$. It would be natural to define $Y_m \equiv \{Y_{ij}, ij \in M_m\}$, instead of just $Y_m^* \equiv \{Y_{ij}, ij \in \mu_m\}$. However, because the model imposes no restrictions on the outcome variables for the unobserved matches, these can be "integrated out." The star superscript denotes the restriction to observed matches. This integration does not change the posterior distribution or the estimated parameters, but it reduces the number of simulated variables and it causes slight changes in the distributions derived below.

The prior distribution is a normal distribution. Let the prior density be denoted $\phi_0(\theta)$. This density is given by

$$\phi_0(\theta) = C \times \exp\left(-.5(\theta-\theta_0)'\Sigma_\theta^{-1}(\theta-\theta_0)\right), \tag{A1}$$

where, here and below, *C* is a generic normalizing constant that ensures densities integrate to one. The matrix Σ_θ is the covariance matrix of the prior distribution, and in all estimated specifications it is ten times the identity matrix. Further increases in the prior variance leave estimated parameters largely unchanged. Corresponding to the parameters, the covariance matrix is decomposed into Σ_α, Σ_β, and Σ_δ.

The error term in the outcome equation can be decomposed into orthogonal terms as $\varepsilon_{ij} = \eta_{ij}\delta + \xi_{ij}$, where the joint distribution of ξ_{ij} and η_{ij} is

$$\begin{pmatrix} \xi_{ij} \\ \eta_{ij} \end{pmatrix} \sim N\left(0, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right). \tag{A2}$$

This is without loss of generality, as the joint distribution of $(\varepsilon_{ij}, \eta_{ij})$ is still given by equation (15). Let $\phi_m(V_m, Y_m^* \mid \theta, X_m, W_m)$ be the density of the latent variables defined by the valuation and outcome equations. The valuation equation implies $\eta_{ij} = V_{ij}' - W_{ij}'\alpha$ and the density of the latent variables in market *m* is

$$\begin{aligned} \phi_m(V_m, Y_m^* \mid \theta, X_m, W_m) = C \times \prod_{ij \in M_m} \exp\left(-.5(V_{ij} - W_{ij}'\alpha)^2\right) \\ \times \prod_{ij \in \mu_m} \exp\left(-.5\left(Y_{ij}^* - X_{ij}'\beta - (V_{ij} - W_{ij}'\alpha)\delta\right)^2\right) \end{aligned}. \tag{A3}$$

The augmented posterior density is proportional to the product of the prior density, some appropriate indicator functions, and the density of the latent variables given above. It is

$$\phi(V, Y^*, \theta | IPO, \mu, X, W) = C \times \phi_o(\theta) \times 1[\delta \geq 0] \times \prod_{m=1}^{N} \left(1[Y_m^* \in \Gamma_{IPO_m}] \times 1[V_m \in \Gamma_{\mu_m}] \times \phi_m(V_m, Y_m^* | \theta, X_m, W_m) \right), \tag{A4}$$

where $\Gamma_{IPO_m} \equiv \left\{ Y_m^* \mid IPO_{ij} = 1[Y_{ij}^* \geq 0], \forall ij \in \mu_m \right\}$ contains the restriction on the latent outcome variables imposed by the outcome equation. It is analogous to the restriction on the latent valuation variables imposed by Γ_{μ_m}. The conditional densities derived below are all proportional to selected factors in the density in equation (A4).

A. Conditional Distributions of Outcome Variables

The conditional augmented posterior density for each latent outcome variable is proportional to the two terms this variable enters in ϕ in equation (A4). The first is in the indicator function $1[Y_m^* \in \Gamma_{IPO_m}]$, and the second is the corresponding term in ϕ_m from equation (A3). There are two cases. When IPO_{ij} equals one, the indicator function $1[Y_m^* \in \Gamma_{IPO_m}]$ equals one when Y_{ij}^* is nonnegative (conditional on the other outcome variables being in Γ_{IPO_m}), and the density of the conditional distribution is

$$\pi(Y_{ij}^* | V, Y_{-ij}^*, \theta, IPO, \mu, W, X) = C \times 1[Y_{ij}^* \geq 0] \times \exp\left(-.5 \left(Y_{ij}^* - X_{ij}' \beta - (V_{ij} - W_{ij}' \alpha) \delta \right)^2 \right). \tag{A5}$$

This is the normal distribution $N\left(X_{ij}' \beta + (V_{ij} - W_{ij}' \alpha)\delta, 1 \right)$ truncated from below at zero. When IPO_{ij} equals zero, the distribution is the same normal distribution, now truncated from *above* at zero.

B. Conditional Distributions of Valuation Variables

The conditional augmented posterior distribution of V_{ij} depends on whether investor *i* and company *j* are matched or not. When $ij \notin \mu_m$, the density is simply

$$\pi(V_{ij} | V_{-ij}, Y^*, \theta, IPO, \mu, X, W) = C \times 1[V_{ij} \leq \overline{V}_{ij}] \times \exp\left(-.5(V_{ij} - W_{ij}'\alpha)^2\right). \tag{A6}$$

When $ij \in \mu_m$, the outcome of the match is observed. Correlation between the error terms means that the outcome contains additional information about the valuation, and the conditional density is

$$\begin{aligned} &\pi(V_{ij} | V_{-ij}, Y^*, \theta, IPO, \mu, X, W) = C \times 1[V_{ij} \geq \underline{V}_{ij}] \times \\ &\exp\left(-.5\left(V_{ij} - W_{ij}'\alpha - \frac{(Y_{ij}^* - X_{ij}'\beta)\delta}{1+\delta^2}\right)^2 \times (1+\delta^2)\right). \end{aligned} \tag{A7}$$

Both of these distributions are truncated normal distributions. The first is $N\left(W_{ij}'\alpha, 1\right)$ truncated from above at $\overline{V}_{ij}$. The second is $N\left(W_{ij}'\alpha + (Y_{ij}^* - X_{ij}'\beta)\delta/(1+\delta^2), 1/(1+\delta^2)\right)$ truncated from below at $\underline{V}_{ij}$. The expressions for $\overline{V}_{ij}$ and $\underline{V}_{ij}$ are given in equations (5) and (6).

C. Conditional Distributions of Parameters

The conditional distributions of the parameters are normal distributions. The distributions of α and β are not truncated. The distribution of δ is truncated below from at zero to normalize the sign of the valuation equation. Each parameter enters all the terms in ϕ, and the derivation of the distributions requires "completing the square" in a product of normal densities. To illustrate, let γ be a random vector with density

$$\pi(\gamma) = C_1 \times \exp\left(-.5(\gamma' M_\gamma \gamma + 2\gamma' N_\gamma + C_2)\right). \tag{A8}$$

Here, M_γ is a corresponding matrix and N_γ is a corresponding vector (naturally, C_1 and C_2 could be combined into the single normalizing constant $C = C_1 \times \exp(-.5C_2)$). Completing the square in this expression shows that the distribution of γ is the normal distribution $N(-M_\gamma^{-1}N_\gamma, M_\gamma^{-1})$.

Collecting terms in ϕ involving α gives $\pi(\alpha|V,Y^*,\beta,\delta,IPO,\mu,X,W)$. This distribution is determined by M_α and N_α as follows:

$$M_\alpha = \Sigma_\alpha^{-1} + \sum_{m=1}^{N}\left[\sum_{ij\in M_m} W_{ij}W_{ij}' + \sum_{ij\in\mu_m}\delta^2 W_{ij}W_{ij}'\right] \tag{A9}$$

$$N_\alpha = -\Sigma_\alpha^{-1}\overline{\alpha} + \sum_{m=1}^{N}\left[\sum_{ij\in M_m} -W_{ij}V_{ij} + \sum_{ij\in\mu_m}\delta W_{ij}(Y_{ij}^* - X_{ij}'\beta - V_{ij}\delta)\right]. \tag{A10}$$

Similarly, collecting terms in ϕ involving β gives

$$M_\beta = \Sigma_\beta^{-1} + \sum_{m=1}^{N}\sum_{ij\in\mu_m} X_{ij}X_{ij}' \tag{A11}$$

$$N_\beta = -\Sigma_\beta^{-1}\overline{\beta} - \sum_{m=1}^{N}\sum_{ij\in\mu_m} X_{ij}(Y_{ij}^* - V_{ij}\delta + W_{ij}'\alpha\delta) . \tag{A12}$$

Finally, for δ, collecting terms gives

$$M_\delta = \Sigma_\delta^{-1} + \sum_{m=1}^{N}\sum_{ij\in\mu_m}(V_{ij} - W_{ij}\alpha)^2 \tag{A13}$$

$$N_\delta = -\Sigma_\delta^{-1}\overline{\delta} - \sum_{m=1}^{N}\sum_{ij\in\mu_m}(Y_{ij}^* - X_{ij}'\beta)(V_{ij} - W_{ij}'\alpha) . \tag{A14}$$

The conditional distribution of δ is $N\left(-N_\delta/M_\delta, 1/M_\delta\right)$ truncated from below at zero.

References

Ackerberg, Daniel, and Maristella Botticini, 2002, Endogenous matching and the empirical determinants of contract form, Journal of Political Economy 110, 564-591.

Admati, Anat, and Paul Pfleiderer, 1994, Robust financial contracting and the role of venture capitalists, Journal of Finance 49, 371-402.

Albert, James, and Siddhartha Chib, 1993, Bayesian analysis of binary and polychotomous response data, Journal of the American Statistical Association 88, 669-679.

Barry, Christopher, Chris Muscarella, John Peavy, and Michael Vetsuypens, 1990, The role of venture capital in the creation of public companies: Evidence from the going-public process, Journal of Financial Economics 27, 447-472.

Berger, James, 1993, Statistical Decision Theory and Bayesian Analysis (Springer-Verlag, New York).

Berry, Steve, James Levinsohn, and Ariel Pakes, 1995, Automobile prices in market equilibrium, Econometrica 63, 841-890.

Brander, James, Raphael Amit, and Werner Antweiler, 2002, Venture-capital syndication: Improved venture selection vs. the value-added hypothesis, Journal of Economics and Management Strategy 11, 423-452.

Bresnahan, Timothy, 1987, Competition and collusion in the American automobile industry: The 1955 price war, Journal of Industrial Economics 335, 457-482.

Bresnahan, Timothy, and Peter Reiss, 1991, Econometric models of discrete games, Journal of Econometrics 48, 57-81.

Casamatta, Catherine, 2003, Financing and advising: Optimal financial contracts with venture capitalists, Journal of Finance 58, 2059-2085.

Cornelli, Francesca, and Oved Yosha, 2003, Stage financing and the role of convertible securities, Review of Economic Studies 70, 1-32.

Fox, Jeremy, 2006, Estimating Matching Games with Transfers, Working paper University of Chicago

Gale, David, and Lloyd Shapley, 1962, College admissions and the stability of marriage, American Mathematical Monthly 69, 9-15.

Gelfand, Alan, and Adrian Smith, 1990, Sampling based approaches to calculating marginal densities, Journal of the American Statistical Association 85, 398-409.

Geweke, John, 1999, Using simulation methods for bayesian econometric models: Inference, development, and communication (with discussion and rejoinder), Econometric Reviews 18, 1-126.

Geweke, John, Gautam Gowrisankaran, and Robert Town, 2003, Bayesian inference for hospital quality in a selection model, Econometrica 71, 1215-1239.

Geweke, John, Michael Keane, and David Runkle, 1994, Alternative computational approaches to inference in the multinomial probit model, Review of Economics and Statistics 76, 609-632.

Gompers, Paul, 1996, Grandstanding in the venture capital industry, Journal of Financial Economics 42, 133-156.

Gompers, Paul, and Josh Lerner, 1999, The Venture Capital Cycle (MIT Press, Cambridge).

Gompers, Paul, and Josh Lerner, 2000. The Determinants of Corporate Venture Capital Successes: Organizational Structure, Incentives, and Complementarities, in Randall Morck, ed.:Concentrated Corporate Ownership (University of Chicago Press, Chicago).

Gorman, Michael, and William Sahlman, 1989, What do venture capitalists do? Journal of Business Venturing 4, 231-248.

Heckman, James, 1976, The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models, Annals of Economic and Social Measurement 5, 475-492.

Heckman, James, 1979, Sample selection bias as a specification error, Econometrica 47, 153-162.

Hellmann, Thomas, 1998, The allocation of control rights in venture capital contracts, Rand Journal of Economics 29, 57-76.

Hellmann, Thomas, and Manju Puri, 2002, Venture capital and the professionalization of start-up firms: Empirical evidence, Journal of Finance 57, 169-197.

Hochberg, Yael, Alexander Ljungqvist, and Yang Lu, 2006, Venture capital networks and investment performance, Journal of Finance, Forthcoming.

Holmstrom, Bengt, and Jean Tirole, 1997, Financial intermediation, loanable funds, and the real sector, Quarterly Journal of Economics 112, 663-691.

Hsu, David, 2004, What do entrepreneurs pay for venture capital affiliation? Journal of Finance 59, 1805-1844.

Inderst, Roman, and Holger Müller, 2004, The effects of capital market characteristics on the value of start-up firms, Journal of Financial Economics 72, 319-356.

Judd, Kenneth, 1998, Numerical Methods in Economics (MIT Press, Cambridge).

Kaplan, Steven, and Antoinette Schoar, 2005, Private equity performance: Returns, persistence, and capital flows, Journal of Finance 60, 1791-1823.

Kaplan, Steven, Berk Sensoy, and Per Strömberg, 2002, How well do venture capital databases reflect actual investments? Working paper, University of Chicago.

Kaplan, Steven, and Per Strömberg, 2001, Venture capitalists as principals: Contracting, screening, and monitoring, American Economic Review 91, 426-430.

Kaplan, Steven, and Per Strömberg, 2003, Financial contracting meets the real world: An empirical analysis of venture capital contracts, Review of Financial Studies 70, 281-315.

Kaplan, Steven, and Per Strömberg, 2004, Characteristics, contracts, and actions: Evidence from venture capitalist analyses, Journal of Finance 59, 2177-2210.

Lerner, Josh, 1994, The syndication of venture capital investments, Financial Management 23, 16-27.

Lerner, Josh, 1995, Venture capitalists and the oversight of private firms, Journal of Finance 50, 301-318.

Li, Kai, and Nagpurnanand Prabhala, 2006, Self-selection models in corporate finance, in Espen Eckbo, ed.: Handbook of Corporate Finance: Empirical Corporate Finance (Elsevier/North-Holland), Forthcoming.

Megginson, William, and Kathleen Weiss, 1991, Venture capitalist certification in initial public offerings, Journal of Finance 46, 879-903.

Quindlen, Ruthann, 2000, Confessions of a Venture Capitalist: Inside the High-stakes World of Start-up Financing (Warner Books, New York).

Robert, Christian, and George Casella, 2004, Monte Carlo Statistical Methods (Springer-Verlag, New York).

Roberts, Gareth, and Alan Smith, 1994, Simple conditions for the convergence of the Gibbs sampler and Metropolis-Hastings algorithms, Stochastic Processes and Their Applications 49, 207-216.

Roth, Alvin, and Marilda Sotomayor, 1990, Two-sided Matching: A Study in Game-theoretic Modeling and Analysis (Cambridge University Press, Cambridge).

Sahlman, William, 1990, The structure and governance of venture capital organizations, Journal of Financial Economics 27, 473-521.

Sørensen, Morten, 2005, An Economic and Econometric Analysis of Market Sorting with an Application to Venture Capital, Unpublished dissertation (Stanford University).

Tanner, Martin, 1998, Tools for Statistical Inference: Methods for the Exploration of Posterior Distributions and Likelihood Functions (Springer-Verlag, New York).

Tanner, Martin, and Wing Wong, 1987, The calculation of posterior distributions by data augmentation, Journal of the American Statistical Association 82, 528-549.

[1] Geweke, Gowrisankaran, and Town (2003) use this variable to identify a selection model of hospital choice. The identifying assumption is that patients are more likely to choose a closer hospital independent of the severity of their illness. Ackerberg and Botticini (2002) study a closely related matching problem between landlords and tenants in Renaissance Tuscany using tenants' wealth to instrument for risk aversion.

[2] There are two fundamental differences between hospital choice analyzed by Geweke, Gowrisankaran, and Town (2003) and the VC market. First, distance is a valid instrumental variable for hospital choice, that is, distance is related to hospital choice but is unrelated to the outcome. No such instrument is available for the VC market. Second, there is no interaction in hospital choice. Hospital choice is a unilateral decision, and one patient's choice of a particular hospital does not exclude other patients from also going to this hospital. Thus, there is no sorting in hospital choice in the sense analyzed here.

[3] Recent papers emphasize informational asymmetries in venture capital, i.e. Casamatta (2003), Cornelli and Yosha (2003), Hellmann (1998), Kaplan and Strömberg (2001).

[4] The model also fits in the growing literature on self-selection models in corporate finance. See Li and Prabhala (2006) for a recent survey of this literature.

[5] Admati and Pfleiderer (1994), Brander, Amit, and Antweiler (2002) and Lerner (1994) test different theories for the syndication of VC investments.

[6] VCs distinguish between the *pre-money valuation* and the *post-money valuation* of a deal. The valuation here is similar to the pre-money valuation.

[7] This slight abuse of standard set notation is common in the literature. A more correct way of stating this equivalence would be $(i,j) \in \mu \Leftrightarrow \{j\} \in \mu(i) \Leftrightarrow \{i\} = \mu(j)$.

[8] It is assumed that companies always prefer matching with some investor to remaining unmatched, and investors always prefer investing rather than leaving their quota unfilled. Theoretically, relaxing this assumption is straightforward (see Roth and Sotomayor (1990) and Sørensen (2005)), but it complicates the empirical model since it requires an empirical specification of agents' outside options.

[9] One nonstandard feature is that the identification of the sign of α depends on interaction terms in W_{ij}. These are terms that depend on the characteristics of both the investor and the company, for example, the investor's previous experience with other companies in the same industry. Due to data limitations, no interaction terms are included in the specifications estimated below, and the sign of the valuation equation is normalized by restricting the covariance between the error terms in the valuation equation and the outcome equation to be nonnegative.

[10] Formally, the marginal conditional augmented posterior distributions for the Gibbs sampler are either normal or truncated normal, see the Appendix.

[11] In the likelihood function in equation (12), this follows since Γ_μ is not a product set, that is, it cannot be written as $\Gamma_\mu = \Gamma_1 \times \Gamma_2 \times \ldots \times \Gamma_{I+J}$, where $I + J$ is the number of agents.

[12] Young companies are companies with a SDC stage code of 22 or less.

[13] Megginson and Weiss (1991) and Barry, Muscarella, Peavy, and Vetsuypens (1990) use a similar classification of the lead investor. Gompers (1995) defines the lead investor as the VC that has been present on the board the longest.

[14] Early-stage companies have SDC stage codes equal to 11, 12, and 13.

[15] Defining fund size as total committed capital generates similar results.

[16] In a few cases a venture capital firm invests with two different funds in the same market. This creates ambiguity about the fund size for the characteristics of the other potential investments involving this investor. For these unobserved potential investments, the average size of the two funds is used.

[17] The other standard errors are calculated with clustering at the investor level.

[18] Note that the coefficient on this variable is not identified in the valuation equation, since the variable is constant within each market. The exclusion of this variable is only a concern for the estimation of the outcome equation.

[19] Describing Bayesian estimates as statistically significant is a slight abuse of standard terminology. Here, it means that zero is not contained in the corresponding Bayesian credible interval.

[20] The probability that the match *i'j'* is preferred to the match *ij* equals $\Pr(W'_{i'j'}\alpha + \eta_{i'j'} > W'_{ij}\alpha + \eta_{ij})$ $= \Phi\left((W'_{i'j'} - W'_{ij})\alpha / \sqrt{2}\right)$ and *dP/dW* equals $\phi(0)\alpha / \sqrt{2}$. Here, Φ and ϕ denote the distribution function and density of the standard normal distribution, respectively. For binary variables the marginal effect is evaluated for a discrete change.

Table I

Descriptive Statistics

Reported figures are averages over the observed investments in the sample. The sample contains investments in 1,666 companies. The variable *IPO* equals one if the investment results in a public offering and zero otherwise. *YEAR* contains the year of the investment. *FUND_SIZE* is the amount of capital available to the investor, measured as the total amount of investments made by the fund. *STATE* equals zero if the company is located in Massachusetts and equals one if the company is located in California. *STAGE* equals zero if the company is an early-stage company and equals one if it is a late-stage company. *I_XXXX* are six industry controls. *TOTAL_EXPERIENCE* is the experience of the investor, measured as the number of previous investments made by the VC firm.

	OBS	MEAN	STD. DEV.	MIN	MAX
IPO	1666	0.269	0.444	0	1
YEAR	1666	1988.340	4.219	1982	1995
FUND_SIZE ($ MIL)	1666	130.556	210.650	1.4	2113.9
STATE	1666	0.781	0.414	0	1
STAGE	1666	0.178	0.382	0	1
I_COMMUNICATION	1666	0.131	0.338	0	1
I_COMPUTER	1666	0.421	0.494	0	1
I_ELECTONICS	1666	0.101	0.301	0	1
I_BIOTECHNOLOGY	1666	0.062	0.241	0	1
I_MEDICAL	1666	0.134	0.341	0	1
I_OTHER	1666	0.150	0.357	0	1
TOTAL_EXPERIENCE	1666	69.562	66.178	0	443

Table II

Probit Estimates

The table contains estimated coefficients from three specifications of a Probit model. The dependent variable is *IPO*. The dummy variables *Y_XXXX* control for individual years, with omitted base year 1982. The reported coefficients are ML estimates, and dF/dX is the marginal effect of a change in the variable evaluated at the sample average. The marginal effects for binary variables (marked °) are the effects of a discrete change. STD. ERR. and P-VALUE are standard error and statistical significance of marginal effect. The first three columns with standard errors and p-values are evaluated with clustering at the investor level. The last column with p-values is evaluated without clustering. *, **, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

Table II: cont.

	Specification 1			Specification 2			Specification 3			
	dF/dX	STD. ERR.	P-VAL	dF/dX	STD. ERR.	P-VAL	dF/dX	STD. ERR.	P-VAL	P-VAL
TOTAL_EXPERIENCE	0.000657	(0.000250)	0.01 ***				0.000657	(.000248)	0.01 ***	0.00 ***
LOG(EXPERIENCE+1)				0.048717	(0.015269)	0.00 ***				
FUND_SIZE ($ MIL)	0.000139	(0.000071)	0.05 **	0.000128	(0.000064)	0.04 **	0.000141	(0.000069)	0.04 **	0.01 ***
STAGE°	0.0620	(0.0320)	0.04 **	0.0671	(0.0335)	0.04 **	0.0579	(0.0309)	0.05 *	0.05 *
I_COMMUNICATION°	0.2741	(0.0480)	0.00 ***	0.2775	(0.0477)	0.00 ***	0.2682	(0.0486)	0.00 ***	0.00 ***
I_COMPUTER°	0.1487	(0.0428)	0.00 ***	0.1514	(0.0428)	0.00 ***	0.1418	(0.0428)	0.00 ***	0.00 ***
I_ELECTRONICS°	0.2189	(0.0651)	0.00 ***	0.2203	(0.0653)	0.00 ***	0.2052	(0.0643)	0.00 ***	0.00 ***
I_BIOTECH°	0.4359	(0.0554)	0.00 ***	0.4334	(0.0554)	0.00 ***	0.4328	(0.0572)	0.00 ***	0.00 ***
I_MEDICAL°	0.1943	(0.0472)	0.00 ***	0.1973	(0.0473)	0.00 ***	0.1949	(0.0490)	0.00 ***	0.00 ***
STATE°	0.000014	(0.032100)	1.00	-0.000814	(0.031104)	0.98	0.003116	(0.032470)	0.92	0.91
YEAR							-0.0043	(0.0033)	0.20	0.12
Y_1983°	-0.0400	(0.0396)	0.33	-0.0516	(0.0390)	0.21				
Y_1984°	-0.0400	(0.0442)	0.38	-0.0534	(0.0435)	0.24				
Y_1985°	-0.1167	(0.0449)	0.03 **	-0.1246	(0.0432)	0.02 **				
Y_1986°	0.0259	(0.0636)	0.68	0.0129	(0.0633)	0.84				
Y_1987°	-0.0065	(0.0562)	0.91	-0.0202	(0.0549)	0.72				
Y_1988°	0.0239	(0.0534)	0.65	0.0081	(0.0516)	0.88				
Y_1989°	-0.0644	(0.0487)	0.22	-0.0769	(0.0470)	0.13				
Y_1990°	-0.0831	(0.0568)	0.19	-0.0971	(0.0551)	0.12				
Y_1991°	0.0731	(0.0751)	0.31	0.0547	(0.0764)	0.46				
Y_1992°	-0.0972	(0.0406)	0.04 **	-0.1097	(0.0393)	0.02 **				
Y_1993°	0.0286	(0.0631)	0.64	0.0090	(0.0610)	0.88				
Y_1994°	-0.1063	(0.0467)	0.05 **	-0.1209	(0.0425)	0.02 **				
Y_1995°	-0.0906	(0.0500)	0.10 *	-0.1053	(0.0465)	0.04 **				

Table III:

Probit Estimates with Alternative Specifications

The table contains estimated coefficients from four specifications of a Probit model. In all specifications the dependent variable is *IPO*. The variable *TOTAL_EXPERIENCE* contains the number of previous investments by the investor. *STATE_CONTROL* is a binary control for the location of the company (California or Massachusetts). *YEAR_CONTROLS* are controls for individual years. The reported coefficients are ML estimates of a standard Probit model. The marginal effect for variables marked ° is the effect of discrete change in the variable. STD.ERR. are standard errors of the marginal effects. All standard errors are evaluated with clustering at the investor level. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

	Specification 1				Specification 2				Specification 3			
	COEF.	dF/dX	STD. ERR.		COEF.	dF/dX	STD. ERR.		COEF.	dF/dX	STD. ERR.	
TOTAL_EXPERIENCE	0.002242	0.000732	(0.000248)	***	0.002041	0.000657	(0.000250)	***	0.002030	0.000658	(0.000248)	***
FUND_SIZE ($ MIL)	0.000397	0.000130	(0.000073)	*	0.000433	0.000139	(0.000071)	**	0.000435	0.000141	(0.000069)	
YEAR									-0.013	-0.004	(0.003)	
STAGE°					0.186	0.062	(0.032)	**	0.173	0.058	(0.031)	*
I_COMMUNICATION°					0.753	0.274	(0.048)	***	0.735	0.268	(0.048)	***
I_COMPUTER°					0.454	0.149	(0.043)	***	0.431	0.142	(0.043)	***
I_ELECTRONICS°					0.605	0.219	(0.065)	***	0.568	0.205	(0.064)	***
I_BIOTECH°					1.166	0.436	(0.055)	***	1.156	0.433	(0.057)	***
I_MEDICAL°					0.545	0.194	(0.047)	***	0.546	0.195	(0.049)	***
STATE CONTROL	yes				yes				no			
YEAR CONTROLS	yes				yes				no			
N	1666				1666				1666			

Table IV

OLS Estimates

The table presents coefficients from three OLS regressions. The dependent variable is *TOTAL_EXPERIENCE*. *AVG_EXPERIENCE* contains average experience of the investors in each market. *STATE_CONTROL* controls for the location of the company. *YEAR_CONTROLS* controls for individual years. STD. ERR. is the robust standard error of the coefficients. *, **, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

	Specification 1			Specification 2			Specification 3		
	COEF.	STD. ERR.		COEF.	STD. ERR.		COEF.	STD. ERR.	
YEAR	5.084	(0.364)	***	-0.114	(0.640)				
STAGE	19.153	(4.025)	***	18.552	(3.916)	***	18.653	(3.943)	***
I_COMMUNICATION	3.068	(5.787)		2.162	(5.631)		2.132	(5.680)	
I_COMPUTER	-1.246	(4.587)		-1.843	(4.463)		-1.924	(4.532)	
I_ELECTRONICS	5.170	(6.221)		3.436	(6.054)		3.435	(6.108)	
I_BIOTECHNOLOGY	11.960	(7.351)		14.321	(7.156)	**	14.520	(7.242)	**
I_MEDICAL	5.765	(5.738)		4.404	(5.584)		4.498	(5.651)	
AVG_EXPERIENCE				1.000	(0.103)	***	0.997	(0.116)	***
CONSTANT	-10044.3	(722.8)	***	222.4	(1267.5)		-4.5	(13.9)	
STATE CONTROL	no			no			yes		
YEAR CONTROLS	no			no			yes		
R^2	0.12			0.17			0.17		
N	1666			1666			1666		

Table V

Bayesian Estimates of Structural Model

The table presents Bayesian estimates of the parameters in the two equations in the structural model. The dependent variable in the outcome equation is IPO, and the dependent variable in the valuation equation is the latent valuation variable. A detailed description of the construction of all variables is in the main text. MEAN, MEDIAN, and STD. DEV. are, respectively, the mean, median, and standard deviation of the simulated posterior distributions of the parameters. The reported coefficients for the outcome equation are normalized to be comparable to standard Probit estimates. dF/dX is the marginal effect. The marginal effects are evaluated at the mean of the posterior. The marginal effects for variables marked ° are effects of a discrete change. dP/dW is the marginal change in the probability of preferring one of two otherwise identical matches with a change in the variable (the precise definition is in the text). Estimates are based on 1,200,000 simulations of the posterior distribution. The initial 400,000 simulations are discarded for burn-in. *, **, and *** denote that zero is not contained in the 10%, 5%, and 1% credible interval, respectively.

	MEAN	MEDIAN	dF/dX	dP/dW	STD. DEV.	
OUTCOME EQUATION						
TOTAL_EXPERIENCE	0.001385	0.001385	0.000346		(0.000302)	***
FUND_SIZE ($ MIL)	0.000328	0.000328	0.000082		(0.000092)	***
YEAR	-0.000895	-0.000907	-0.000223		(0.000527)	*
STAGE°	0.1843	0.1849	0.0486		(0.0499)	***
I_COMMUNICATION°	0.7101	0.7076	0.2183		(0.0785)	***
I_COMPUTER°	0.4092	0.4082	0.1052		(0.0676)	***
I_ELECTRONICS°	0.5612	0.5603	0.1687		(0.0838)	***
I_BIOTECHNOLOGY°	1.1274	1.1269	0.3859		(0.0890)	***
I_MEDICAL°	0.5238	0.5226	0.1538		(0.0793)	***
CONSTANT	0.1763	0.1931			(1.0458)	
VALUATION EQUATION						
TOTAL_EXPERIENCE	0.0290	0.0281		0.0082	(0.0127)	***
FUND_SIZE ($ MIL)	0.0465	0.0418		0.0131	(0.0251)	***
STAGE°	0.1657	0.1644		0.0466	(0.0330)	***
I_COMMUNICATION°	0.0252	0.0259		0.0071	(0.0434)	
I_COMPUTER°	0.0074	0.0082		0.0021	(0.0360)	
I_ELECTRONICS°	0.0272	0.0276		0.0077	(0.0462)	
I_BIOTECHNOLOGY°	0.2808	0.2801		0.0787	(0.0567)	***
I_MEDICAL°	0.0713	0.0718		0.0201	(0.0446)	
VARIANCE						
COVARIANCE	0.2222	0.2208			(0.0670)	***

Table VI

Bayesian Estimates of Outcome Equation under Alternative Specifications

The table contains Bayesian estimates of the coefficients in the outcome equations for three specifications of the structural model. The dependent variable in the outcome equation is IPO. INV1, INV2, up to INV5 are investor-specific controls for the five most experienced investors in the data. A detailed description of the construction of all variables is in the main text. MEAN and STD. DEV. are, respectively, the mean and standard deviation of the simulated posterior distributions of the coefficients. The reported coefficients for the outcome equation are normalized to be comparable to standard Probit estimates. dF/dX is the marginal effect, evaluated at the MEAN of the posterior. The marginal effect for discrete variables marked ° is the effect of a discrete change. Estimates are based on 1,200,000 simulations of the posterior distribution. The initial 400,000 simulations are discarded for burn-in. *, **, and *** denote that zero is not contained in the 10%, 5%, and 1% credible interval, respectively.

	Specification 1				Specification 2				Specification 3			
	MEAN	dF/dX	STD. DEV.		MEAN	dF/dX	STD. DEV.		MEAN	dF/dX	STD. DEV.	
OUTCOME EQUATION												
TOTAL_EXPERIENCE	0.001808	0.000441	(0.000483)	***	0.001385	0.000346	(0.000302)	***	0.000260	0.000068	(0.000435)	
FUND_SIZE ($ MIL)					0.000328	0.000082	(0.000092)	***	0.000752	0.000196	(0.000179)	***
YEAR					-0.000895	-0.000223	(0.000527)	*	-0.000884	-0.000230	(0.000751)	
STAGE					0.1843	0.0486	(0.0499)	***	0.1747	0.0478	(0.0503)	***
I_COMMUNICATION					0.7101	0.2183	(0.0785)	***	0.6897	0.2176	(0.0785)	***
I_COMPUTER					0.4092	0.1052	(0.0676)	***	0.3761	0.1005	(0.0680)	***
I_ELECTRONICS					0.5612	0.1687	(0.0838)	***	0.5370	0.1657	(0.0847)	***
I_BIOTECHNOLOGY					1.1274	0.3859	(0.0890)	***	1.1319	0.3949	(0.0912)	***
I_MEDICAL					0.5238	0.1538	(0.0793)	***	0.5120	0.1550	(0.0783)	***
INV1									0.5415	0.1716	(0.1287)	***
INV2									-0.0869	-0.0218	(0.1039)	
INV3									0.3094	0.0906	(0.0818)	***
INV4									0.1618	0.0451	(0.1013)	
INV5									-0.7920	-0.1366	(0.3041)	***
CONSTANT	-1.1180		(0.1013)	***	0.1763		(1.0458)		0.2212		(1.6848)	
VARIANCE												
COVARIANCE	0.2499		(0.0732)	***	0.2222		(0.0670)	***	0.1925		(0.0651)	***

Table VII

Bayesian Estimates under Alternative Market Specifications

The table contains Bayesian estimates of the outcome equation of the structural model. The model is estimated with four market specifications. The first specification takes a market to be delineated by a half-year and a state (either CA or MA). The second specification expands the time dimension, and takes a market to be delineated by a whole year and a state. The third specification expands the geographical dimension and takes a market to be delineated by a half-year, but treats CA and MA as one market. The final specification takes a market to be delineated by a whole year and treats CA and MA as one market. The dependent variable is IPO. A description of the variables is presented in Table I and in the main text. MEAN and STD. DEV. are, respectively, the mean and standard deviation of the simulated posterior distributions. The reported coefficients for the outcome equation are normalized to be comparable to standard Probit estimates. Estimates are based on 1,200,000 simulations of posterior distribution. The initial 400,000 simulations are discarded for burn-in. *, **, and *** denote that zero is not contained in the 10%, 5%, and 1% credible interval, respectively.

	NORMAL MKT DEF			EXPANDED YEARS			EXPANDED STATES			EXPANDED YEARS AND STATES		
	MEAN	STD. DEV.		MEAN	STD. DEV.		MEAN	STD. DEV.		MEAN	STD. DEV.	
OUTCOME EQUATION												
TOTAL_EXPERIENCE	0.001385	(0.000302)	***	0.001093	(0.000310)	***	0.001079	(0.000310)	***	0.001322	(0.000300)	***
FUND_SIZE ($ MIL)	0.000328	(0.000092)	***	0.000316	(0.000080)	***	0.000322	(0.000080)	***	0.000370	(0.000098)	***
YEAR	-0.000895	(0.000527)	*	-0.000924	(0.000895)		-0.001410	(0.000751)	*	-0.001089	(0.000506)	***
STAGE	0.184	(0.050)	***	0.186	(0.050)	***	0.187	(0.049)	***	0.175	(0.049)	***
I_COMMUNICATION	0.710	(0.079)	***	0.712	(0.077)	***	0.711	(0.078)	***	0.706	(0.078)	***
I_COMPUTER	0.409	(0.068)	***	0.413	(0.067)	***	0.413	(0.068)	***	0.415	(0.067)	***
I_ELECTRONICS	0.561	(0.084)	***	0.572	(0.083)	***	0.570	(0.083)	***	0.565	(0.083)	***
I_BIOTECHNOLOGY	1.127	(0.089)	***	1.141	(0.091)	***	1.143	(0.090)	***	1.118	(0.089)	***
I_MEDICAL	0.524	(0.079)	***	0.536	(0.079)	***	0.536	(0.078)	***	0.525	(0.078)	***
CONSTANT	0.176	(1.046)		0.065	(1.783)		1.039	(1.500)		0.319	(1.024)	
VARIANCE												
COVARIANCE	0.222	(0.067)	***	0.316	(0.084)	***	0.311	(0.084)	***	0.323	(0.091)	***
N	1666			1666			1666			1666		
Markets	56			28			28			14		

Company

		4 (X_4=4)	3 (X_3=3)	2 (X_2=2)	1 (X_1=1)
Investor	C (X_C=3)	- 40	- -	- -	- -
	B (X_B=2)	**40** -	**30** 30	- 20	- -
	A (X_A=1)	- -	- -	**20** -	**10** 10

Figure 1. Matching market with four companies and two or three investors. Bold numbers indicate outcomes from a matching market with investors A and B. Remaining (not bold) numbers are outcomes from a matching market with investors A, B, and C.

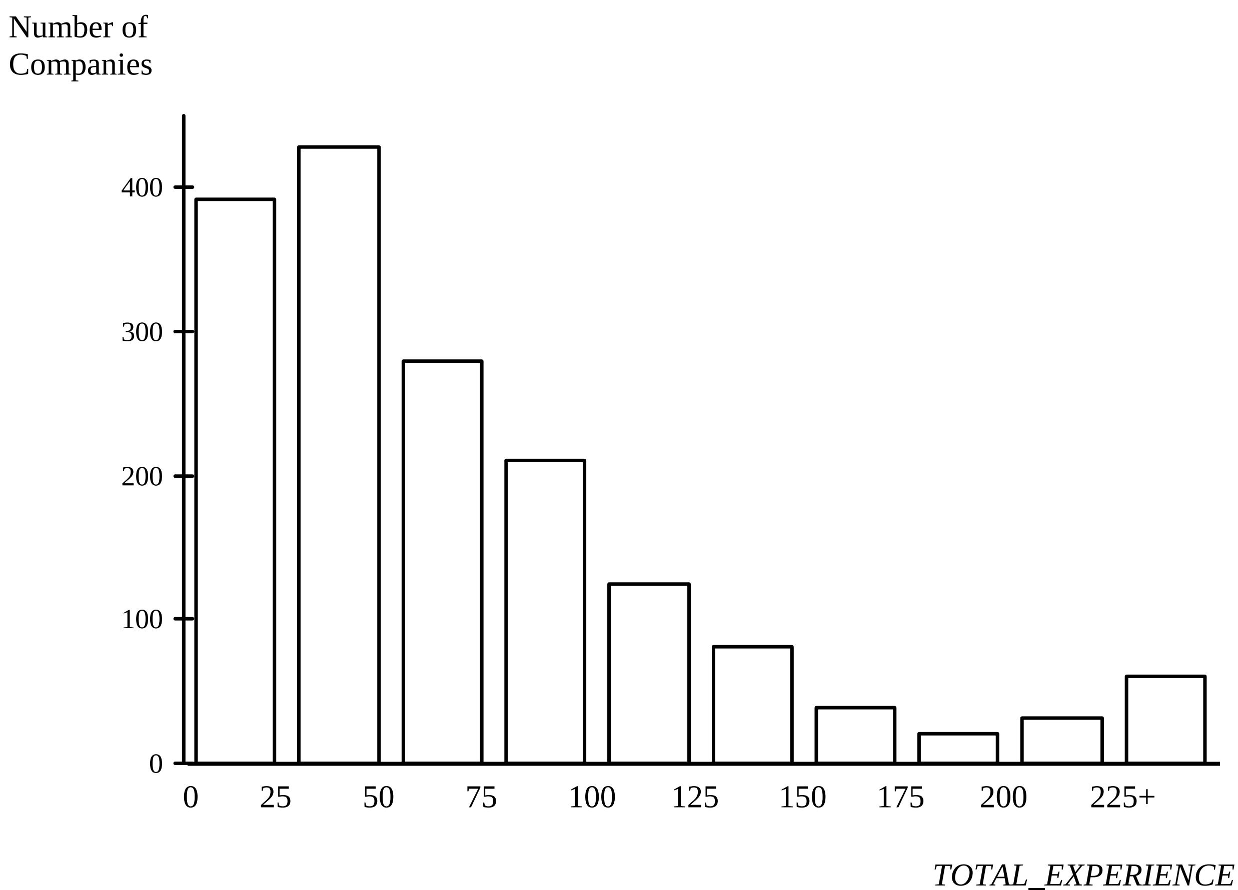

Figure 2. Number of companies in each group. The 1,666 companies are grouped into 10 groups according to the experience of the investor. The figure illustrates the number of companies in each group.

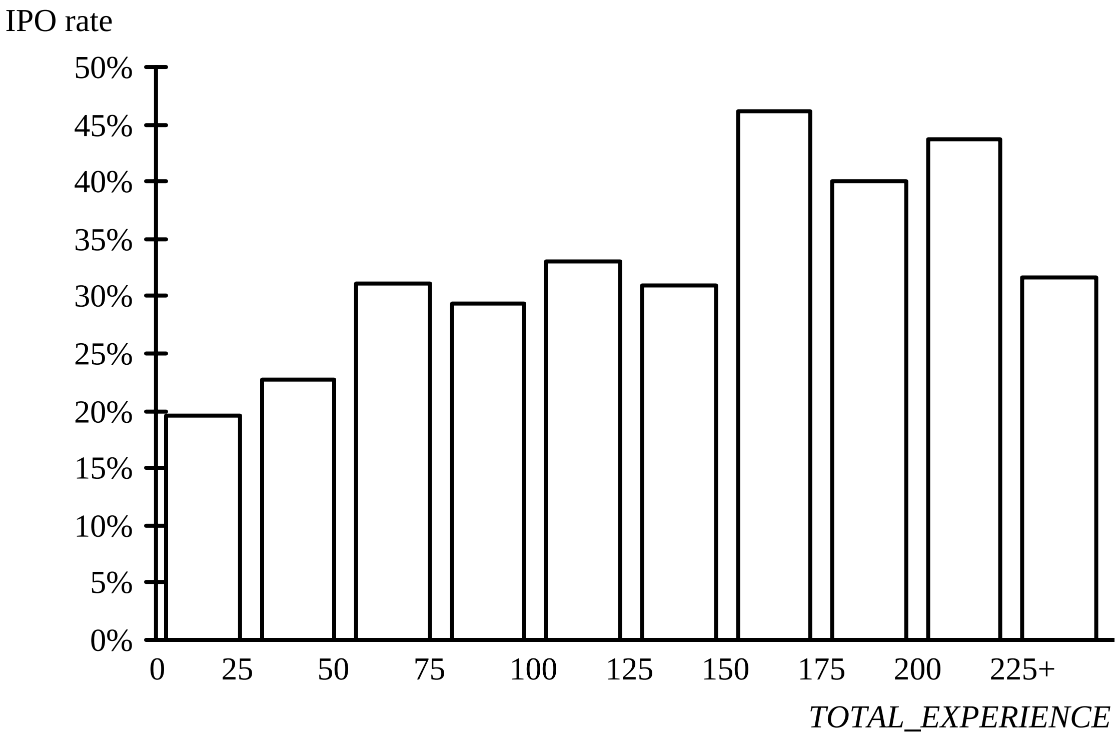

Figure 3. IPO rate in each group. The 1,666 companies are grouped into 10 groups according to the experience of the investor. The figure shows IPO rates of companies in each group.

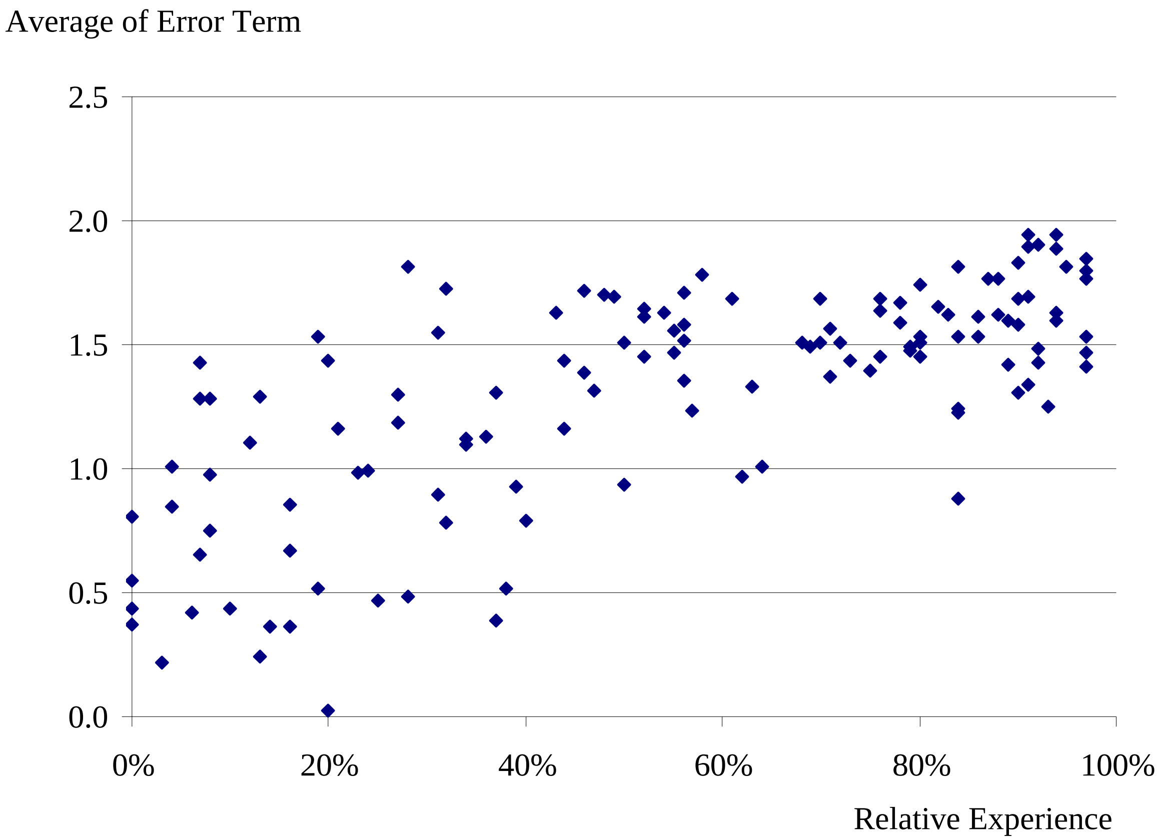

Figure 4. Error terms of matched pairs. For each investor in each market, the average of the error terms in the valuation equation is plotted as a function of the investor's experience relative to the other investors in the same market. Relative experience is the investor's experience as a percentile of the experience of all the investors in the market. For the most experienced investor in the market, the relative experience is one. The average is calculated as the average of the error terms in the matches that are made by the investor. Only investors who participate in two or more matches in the market are included in the figure.

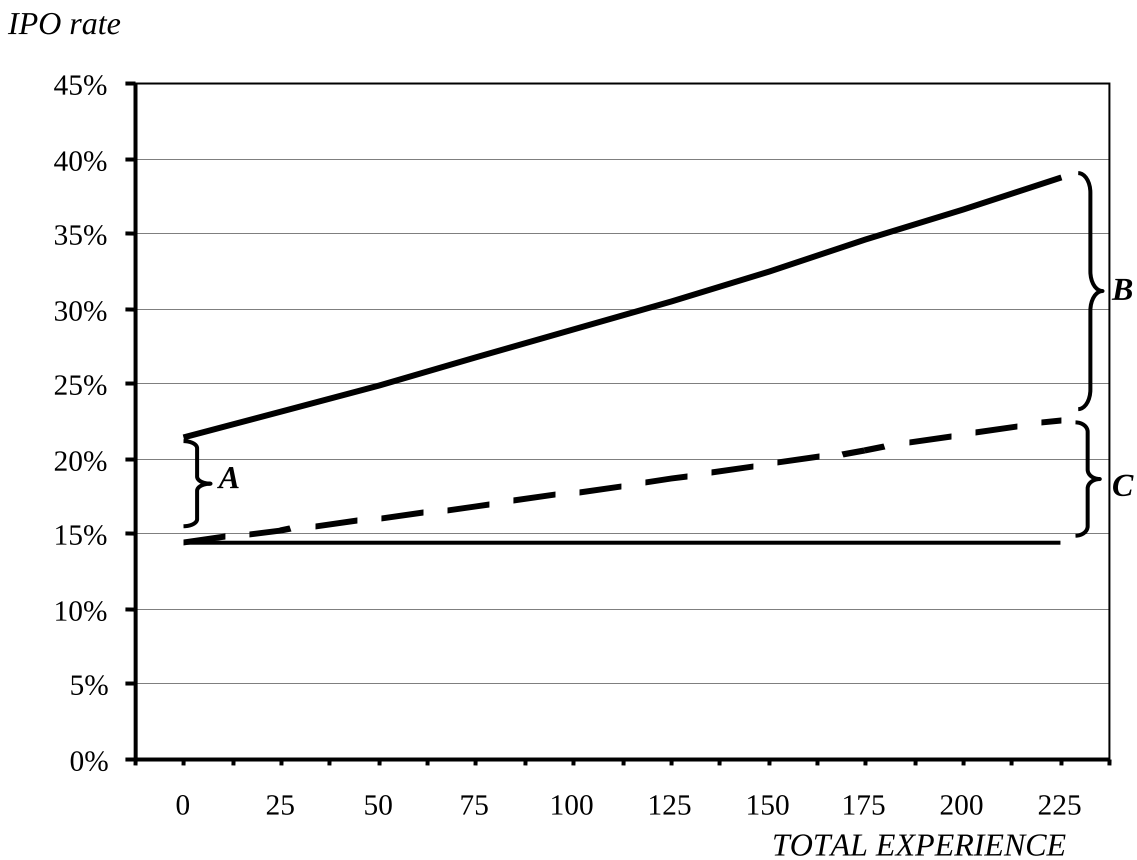

Figure 5. Decomposition of influence and sorting. The figure shows the IPO rates from the Probit model (Specification 1 in Table III) and the outcome equation of the structural model (Table V). The solid line indicates the IPO rate predicted by the Probit model. This rate is the empirical rate of the IPOs observed in the sample. The broken line represents the IPO rate from the outcome equation of the structural model. This is the IPO rate after controlling for the selection of the investments, and it represents the rate that would be observed if an average company were randomly assigned to investors with different degrees of experience.