The Size and Specialization of Direct Investment Portfolios*

Yael V. Hochberg
Northwestern University

Mark M. Westerfield
University of Southern California

February 25, 2010

Preliminary and Incomplete

Abstract

We investigate the optimal size and scope of a portfolio of direct investments in a model in which investment quality is heterogeneous and the portfolio manager observes a limited pool of investment opportunities. We use an order statistics argument to show that the number and specialization of investments within a portfolio are substitutes: while specialization by the portfolio manager may increase the cash flows from any given set of ideas, it restricts the pool of potential ideas that he can undertake. Since potential ideas are heterogeneous in quality, the ability to choose from a large pool increases the returns to investment and is necessary to support a large number of investments. Variables that increase (decrease) the set of available projects or returns to investment activity, such as skill (competition) cause the portfolio to be larger (smaller) and more generalized (specialized). Increasing the extent to which specialization restricts the pool of potential projects causes the portfolio to be smaller and less specialized. We test and confirm the predictions of our model in a setting with highly skewed payoffs and in which access to a wide range of ideas is critical: the U.S. venture capital industry.

*We are grateful to Harry DeAngelo, Wayne Ferson, Mike Fishman, Gareth James, Jiro Kondo, Alexander Ljungqvist, Tim Loughran, Ramana Nanda, Oguzhan Ozbas, Mitchell Petersen, Josh Rauh, Paola Sapienza, Berk Sensoy, and seminar participants at USC, Notre Dame, the University of Florida, Northwestern University, Harvard Business School, Tel Aviv University and the HKUST Finance Symposium for helpful discussions and suggestions. Vineet Bhagwat and Amol Deshpande provided excellent research assistance. Hochberg gratefully acknowledges funding from the Heizer Center for Private Equity and Venture Capital and the Zell Center for Risk Research at the Kellogg School of Management. Address correspondence to y-hochberg@kellogg.northwestern.edu (Hochberg) or mwesterf@usc.edu (Westerfield).

1 Introduction

The ideal size and scope of investment activity has long been of interest to economists, and much of the research in this area has focused either on the boundaries of the firm or on the investment behavior of firms.[1] Investment within the firm is usually understood to be both an extension of and determined by the ongoing activities of the firm. In particular, returns to scale and scope result from an interaction of current business activities, their organizational form, and potential opportunities. Instead, we wish to know how the size and scope of a portfolio of investments should optimally be structured without the attached firm. Many investment vehicles and financial intermediaries, such as real estate funds, hedge funds, venture capital, and private equity, have such a structure. In addition, isolating project selection from ongoing business activities may shed light on forces that guide project selection within firms. In this paper, we ask how ideas are optimally found and exploited in a portfolio of direct investments.

We answer this question with a model of direct investment with endogenous choices of size and specialization (scope). The model takes as key ingredients both that potential investments are heterogeneous in quality and that investors observe only a limited pool of possible investment opportunities. Returns to investment are determined in part by the choice of specialization because of two competing effects on project quality. First, specialization can increase the cash flows from any given project or improve the average quality of a pool of projects. Second, specialization restricts the set of potential projects the investor can choose from because some of the potential projects are outside of either the investor's expertise or his area of focus. If some projects are more valuable than others, then one must be able to sift through many potential projects to find the most desirable. Since potential projects ideas are assumed to be heterogeneous in quality, the ability to choose from a large pool of potential projects is necessary to extract a high value from investments. We show that when opportunities are sufficiently heterogeneous and investors are sufficiently constrained in how much they can observe, the number and specialization of investments within a portfolio are substitutes.

Our model does not assume the existence of decreasing marginal returns to capital,[2] but instead

[1]See e.g. Stein (1997), Maksimovic and Phillips (2002), Rajan, Servaes, and Zingales (2000), Scharfstein and Stein (2000), or Scharfstein and Mullainathan (2002). See Maksimovic and Phillips (2007) for a recent survey of the internal capital markets and investments literature.

[2]Berk and Green (2004) provide a model of investments that assumes decreasing returns to scale and successfully

provides a micro-foundation based on optimal choices for size and scope. If an investor can see or access n potential investments and has the funding to carry through with m of them, then he will pick the best m. The investor thus faces endogenously decreasing marginal returns to capital since the $m+1$th investment must be worse than the first m. The *rate* at which returns to capital decline is determined by the distribution of projects and by the size of the pool of projects available to the investor. When the pool of potential projects (n) is relatively small, returns to capital are small and decline rapidly. Because specialization restricts the pool of potential projects, it causes returns to capital to decline more rapidly, making size and specialization *substitutes*. In addition, the decreasing marginal returns to capital effect becomes stronger when the distribution has a large right tail (has a high skewness or variance). In that case, the average difference in quality between the $(m+1)$th and mth projects will be larger, especially when the initial pool of potential projects is smaller. In total, the decreasing returns effect will dominate any direct positive effect of specialization on cash flow when the distribution has sufficiently large tails.

In this setting, variables that increase access to projects or to the returns to investment activities, such as skill, cause the investor's portfolio to be larger and more generalized. Conversely, variables that decrease access to projects or reduce the returns to investment activity, such as competition, cause the portfolio to be smaller and more specialized. Finally, increasing the extent to which specialization restricts the pool of potential projects the portfolio manager can undertake reduces both the size and the specialization of the portfolio.

To test the predictions of our model, we desire a setting in which access to a large pool of ideas or projects (deal flow) is a significant constraint and in which investments are highly heterogeneous in quality. Venture Capital (VC) thus provides an ideal laboratory for testing our model. First, access to deal flow is considered a critical asset within the VC industry and the shadow price of access to a high quality pool of ideas is high. VCs cannot usually observe all startup companies that are seeking investment capital, and Sorensen (2007) shows that the ability to select projects contributes on the order of 60% of VC returns, with the remainder attributed to value-added activities. Second, the cash flows from VC investment opportunities are highly variable and skewed. Sahlman (1990) describes the returns to venture capital investments and finds that 34.5% of money

describes the mutual fund industry. Chen, Hong, Huang, and Kubik (2004) empirically document the existence of decreasing returns in the mutual fund industry.

invested returned at best a partial loss, while 6.8% of money invested was responsible for nearly 50% of the final VC portfolio value. Ljungqvist and Richardson (2003) document that nearly three-quarters of VC portfolio companies are written off completely, suggesting that a small proportion of VC investments account for the majority of venture fund returns. Finally, the industry and geographic specialization of a VC fund is often formalized in a governing partnership agreement that may use covenants to limit investment shares outside of particular industries or geographical areas (Lerner, Hardymon, and Leamon (2007)).

Using a large dataset of U.S. VC funds raised during the period 1980 to 1999 (with investments between 1980 and 2003), we find five main empirical results. Each is consistent with the predictions of our model but often counter to common intuitions regarding investment. First, size and specialization are strong substitutes. Figure 1 displays a simple scatter plot and linear fit of the relationship between VC portfolio size and specialization. It demonstrates that VC funds likely do not raise large funds based on expertise or specialization within a particular single industry or geography, because large specialized funds are not common.

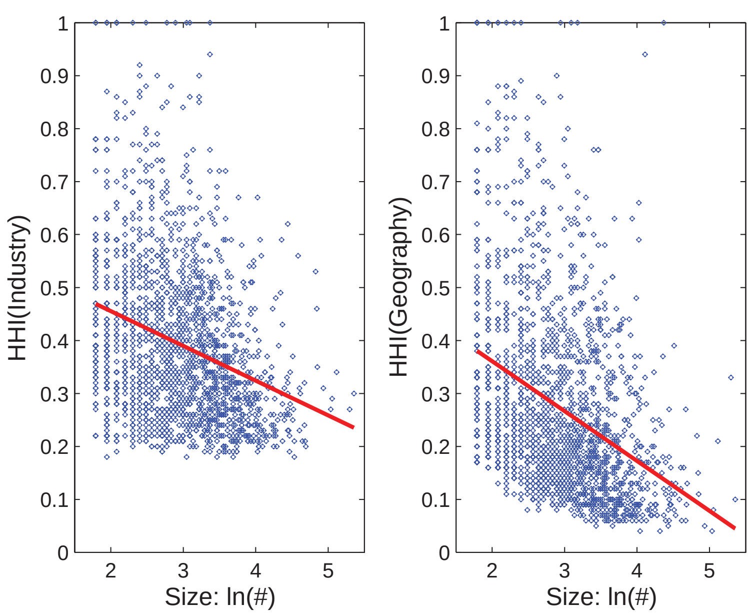

Figure 1: The industry and geographic specialization of venture capital portfolios, plotted as function of size. Specialization is measured by the Herfindahl-Hirschman Index (HHI) of a VC fund's investments across industries (left panel) and CMSAs (right panel). Size is measured as the log of the number of investments made. Exact specifications for variable construction can be found in section 4.

Second, the most experienced VC firms raise funds that are more generalized. This runs directly counter to the standard learning-by-doing intuition that experience in a given industry allows for specialization and enhances productivity; productivity should be necessary to support a larger fund in the presence of decreasing returns to capital. Instead, we show that the cost of having a more narrow pool of ideas from which to select investments overcomes the productivity enhancement resulting from specialization.

Third, earlier stage startup companies receive investments from funds that are smaller and more specialized. This is surprising, because early stage investments are generally smaller and riskier than later stage companies for whom some uncertainty has been resolved, and so the specialized VC funds that invest in these companies are giving up an obvious benefit to diversification.

Fourth, when aggregate VC inflows increase, VC portfolios have *fewer* investments and are more specialized. One might think that since greater aggregate inflows result in a larger dollar size for portfolios (Kaplan and Schoar (2005)), they would also result in portfolios with more numerous investments. However, our model predicts that the result of the increased competition from increased aggregate VC funding ("money chasing deals"[3]) produces portfolios of investments that are smaller in number and more specialized.

Finally, when specialization is more costly to the VC in terms of restricting available potential deal flow, VC portfolios become smaller and less specialized.

Our work is related to the large literature on internal capital markets and firm investments, although our question is very different. Papers such as Stein (1997), Rajan, Servaes, and Zingales (2000), Scharfstein and Stein (2000), and Ozbas (2005) focus on the ability of headquarters to allocate capital across already existing units and the associated agency costs. Agency costs and existing assets/forms play a central role in these papers and thus provide an understanding or organizational form and capital flows based on ongoing business activities. In contrast, we focus on investment choices independently of the ongoing business activities, and generate decreasing marginal returns to capital directly through project selection rather than through agency costs. Other papers, such as Inderst, Mueller, and Münnich (2007) and Fulghieri and Sevilir (2008), examine the development and agency considerations of direct investment projects (such as venture capital), but they

[3] Gompers and Lerner (2000) show that as inflows into the VC industry increase, valuations for portfolio companies are bid up. There is an increase in competition within the industry for a limited number of attractive investment opportunities, often referred to as "money chasing deals."

do not examine the role of project selection in the fashion explored in this paper. Rather, their projects are taken to all be ex-ante identical in quality. Maksimovic and Phillips (2002) provides a neo-classical benchmark for conglomeration based on inherent productivity differences and exogenously decreasing returns to scale. Unproductive firms diversify because they are unproductive in their primary activity. In our model, the fund jointly chooses size and specialization and faces an investment frontier based on those choices.

While there is very little existing work analyzing the size and specialization of real estate funds or hedge funds, our empirical findings also contribute to an emerging literature on VC fund and firm organization and specialization. Hochberg, Mazzeo, and McDevitt (2009) examine the competitive structure of venture capital markets and the effect of venture firm specialization on competition within markets. Gompers, Kovner, and Lerner (2009) examine the relationship between specialization of human capital and success, focusing in particular on the difference between generalist firms who employ generalist individuals and generalist firms that employ specialist individuals. Our model does not make an unambiguous prediction about the relationship between performance and specialization, controlling for opportunities, because each VC firm will optimally choose size and specialization. We discuss this in more detail in Section 3.

The remainder of the paper is organized as follows: Section 2 lays out the model and the comparative statics. Section 3 describes the empirical setting and describes our empirical proxies and tests. Section 4 describes the data and the variable construction. Section 5 describes the empirical results. Section 6 concludes. Appendix A contains all proofs and some examples using specific distributional assumptions.

2 The Model

In this section, we will model the project selection for a single investment vehicle, such as a real estate fund, hedge fund, or private equity fund. The portfolio manager will choose both the size and scope of the fund's activities: how many projects to finance and to what extent to specialize the fund's activities. The key intuition is twofold. Specialization increases a fund's ability to develop a particular project or improves the average quality of a pool of potential projects. However, specialization reduces the set of projects that a fund can undertake because some of those projects will be outside the capabilities or focus of the specialized portfolio manager. Since the fund must find profitable projects, heterogeneity in project quality implies that the size of the *set of potential projects* is the key to the quality of the projects that the fund actually finances.

2.1 Portfolio Managers and Projects

There is a pool of potential projects or ideas that require funding and a portfolio manager with financial and human capital. The pool of projects is large and each one has an associated base quality Δ_i drawn from the same distribution with cdf F: $\Delta_i \sim F(\Delta)$ with $F(0) = 0$ and no atoms.[4] We will assume that the Δ_i are normalized to the portion of the project's payoffs that can be claimed by the portfolio manager, as opposed to by those who supply him with capital or ideas/projects. Δ_i may represent the investment's true quality or the portfolio manager's expected payoff conditional on some signal gained by examining the project in detail.

Before identifying any projects, the portfolio manager must raise funds and choose which human capital to develop and employ. The timing of fund raising in our model is important: a portfolio manager must raise funds and choose human capital *before* he identifies the specific projects to undertake.[5] We assume that there is an un-modeled information asymmetry problem between investment funds and the capital markets, so that raising capital takes time. Thus, a desirable project will be found and funded by someone else before the initial fund can return with additional money.

The portfolio manager makes two choices. First, by raising funds, a portfolio manager chooses how many projects can be funded. Since all projects are identical in size, we say the portfolio manager can finance M projects. We will assume that the cost of raising this capital is equal to $M\theta$.[6]

Second, a portfolio manager chooses his human capital by choosing the level of specialization for his fund: $\phi \in [0, 1]$. A specialized portfolio manager is better able to both apply human capital

[4] $F(0) = 0$ implies that $\Delta_i > 0$ and so projects cannot produce negative gross cash flows to portfolio managers. The most a portfolio manager can lose is the initial investment.

[5] When the portfolio manager chooses what human capital to employ while raising money, it is equivalent to choosing, for example, a strategy or asset class for hedge funds, an industry or region for venture capital, or a region and property class for a real estate fund. The projects are then the specific assets to be purchased/developed within the chosen class.

[6] θ can be interpreted both as a return required by the capital markets or as a transactions cost to raising capital. For investment funds in which the portfolio manager's skill is hard to determine (for example, venture capital or hedge funds), this might also represent the cost of raising money from adverse selection.

to prospective projects and also to increase the average quality of the pool of projects he evaluates. Thus, a successful project pays off $\phi\eta + \Delta_i$ where η captures both the value added to the project from specialized human capital and the upward shift in the underlying quality of the pool of projects. The shift in the quality of the underlying pool might be a result of the specialist being better able to find good projects or a market structure advantage in which better projects match more easily with specialists. For any combination of the three reasons, specialization increases the payoff to making any particular investment.[7]

An additional benefit to specialization is that the portfolio manager is able to recover all of his or her specific human capital ($\phi\eta$) if a project fails, whereas he can only recover a fraction $\mu < 1$ of the base value of the project. μ can be taken to represent both recoverable human capital and the management fees that are paid to the investment fund on all capital raised by the fund, thus capturing the benefit in fee income from raising capital for a larger portfolio.[8]

Not all projects are successful: given funding, the probability of achieving a positive payoff is $\alpha \in (0, 1]$.[9] The base ability of an portfolio manager is ψ, and so the gross payoff to the portfolio manager given a particular value for Δ_i for successful funded project is $\psi\left(\phi\eta + \Delta_i\right)$.[10] If the project fails, then the portfolio manager is still able to recover $\psi\left(\phi\eta + \mu\Delta_i\right)$ in value. Thus, the expected gross payoff to a project, given Δ_i, is

$$\psi\left(\phi\eta + (\alpha + (1-\alpha)\mu)\,\Delta_i\right). \tag{1}$$

Next, we consider the pool of projects that the portfolio manager can potentially finance. We assume that a portfolio manager sees N different projects. However, the portfolio manager can only access (evaluate and undertake) at most $\llcorner(1-\lambda\phi)N\lrcorner$ of these projects. $\lambda < 1$ measures the effect of specialization on the size of the pool, and the $\llcorner X \lrcorner$ notation indicates the greatest integer

[7] One can think of η as having two components, η_1 and η_2. The change in the average quality of the pool of projects is an upward shift of the distribution so that the base quality of the project becomes $\phi\eta_1 + \Delta_i$ instead of just Δ_i. The second component is the value added to the project because of the specialized human capital: $\phi\eta_2$. Thus the total project payoff is $\phi\eta + \Delta_i = \phi\eta_1 + \phi\eta_2 + \Delta_i$.

[8] Additionally, μ can also capture the transaction fees earned on each investment made by the fund in industries such as leveraged buyouts.

[9] Alternately, α might represent the size of the atom at zero in the payoff distribution. Then Δ represents the payoff conditional on being positive. We assume for simplicity that α is independent of the choice of ϕ.

[10] ψ may also serve as a proxy for market conditions. We abstract from structurally modeling market conditions, but those conditions are implicitly part of the distribution of Δ or parameters like ψ. For example, more intense competition among VCs might mean that entrepreneurs have more bargaining power, and so the total surplus available to any given VC on a project will decline. We discuss our empirical proxies in more detail in section 3.

below X. Because the specialist investor enhances his ability to succeed in one area while giving up knowledge other areas, the advantage of specialization comes at the cost of breadth: the more specialized a portfolio manager is, the more projects are outside his capabilities. Alternately, just as a specialist might have a matching advantage within his area of focus, he may have a matching disadvantage outside that area. $\llcorner(1-\lambda\phi)N\lrcorner$ represents the portfolio manager's *pool of projects*.

Upon evaluating the $\llcorner(1-\lambda\phi)N\lrcorner$ projects, the portfolio manager will choose the M best projects to undertake. Thus, we are interested in order statistics on Δ. Denote by $\mathrm{E}\left[\Delta_{n,m}\right]$ the expected value of the mth highest value of Δ picked from a total of n i.i.d choices. The portfolio manager's expected payoff is

$$\psi\left(M\phi\eta+(\alpha+(1-\alpha)\mu)\sum_{j=1}^{M}\mathrm{E}\left[\Delta_{\llcorner(1-\lambda\phi)N\lrcorner,j}\right]\right)-M\theta. \tag{2}$$

2.2 Order Statistics on Δ

While closed form solutions for order statistics can be messy, they follow certain basic rules that we can exploit. In doing so, we will provide the mechanism by which choices and size and specialization determine the shape of the investment frontier and the rate at which the returns to capital decrease.

We will label the decreasing returns portion of the investor's objective function by the function G and determine its properties:

Definition 1 (Returns Function) *The function $G(n,m)$ is the sum of project values when the best m projects are pulled from a group of n potential choices:*

$$G(n,m)=\sum_{j=1}^{m}\mathrm{E}\left[\Delta_{\llcorner n\lrcorner,j}\right]. \tag{3}$$

Proposition 1 (Cumulative Order Statistics) *Assume that the expectations in (3) exist. Then, $G(n,m)$ is*

- *increasing and concave in m: $(G(n,m)-G(n,m-1))$ is positive and declining in m.*
- *increasing in n: $(G(n+1,m)-G(n,m))$ is positive.*

- *has increasing differences (is super-modular) in* (n, m)[11]*:*

$$[(G(n+1,m) - G(n+1,m-1)) - (G(n,m) - G(n,m-1))] > 0. \quad (4)$$

Proof. See Appendix A. ■

Asssumption 1 (Cumulative Order Statistics) *We will assume that* $G(n,m)$ *is concave in* n*:* $(G(n+1,m) - G(n,m))$ *is declining in* n*.*[12]

Intuitively, the properties of G provided by Proposition 1 and Assumption 1 can be understood in relatively simple fashion. A portfolio manager picking m projects from a pool of n choices will always pick the m best projects. Thus, the $(m+1)$th project to be added is always worse than the first m. Since projects will not produce a negative gross cash flow to the portfolio manager (before accounting for the cost of capital), each additional project adds to the gross payoff, but at a decreasing rate. Thus, the total returns to capital will be increasing and concave in the number of investments made.

Similarly, we can consider adding one potential project to a pool of n projects, from which the portfolio manager will pick the best m. If this new $(n+1)$th project is worse than the m already selected, then the portfolio manager gains nothing. However, if this project is better, than the portfolio manager benefits from an expanded pool by substituting the worst existing project for the new one. As the pool of potential choices becomes very large, only the very best projects are undertaken (for a fixed m), and so the probability that any new choice will be good enough becomes very small. Thus, each additional choice adds to the expected gross payoff with a declining marginal value. We illustrate this distributional shift effect in Figure 2 for the exponential distribution.

The final property, the supermodularity of G (4), is a cross effect between n and m, and it is the key to understanding the tradeoff between size and scope. It says that as the choice set (n) of the portfolio manager increases, the marginal value of each new project undertaken (m) also increases.

[11] See, for example, Athey (2002). Increasing differences (or super-modularity) means that the gains from increasing n increase in m, and vice versa. For differentiable functions, super-modularity is equivalent to $\frac{\partial^2}{\partial n \partial m} G(n,m) > 0$.

[12] We are not able to provide a general proof that G is concave in n for all distributions for which the relevant expectations exist and $F''(\Delta) < 0$. However, for every distribution we have explicitly calculated, G is concave in n. As examples, Appendix A contains closed form solutions for the function $G(n,m)$ for the the uniform, exponential, and power law distributions.

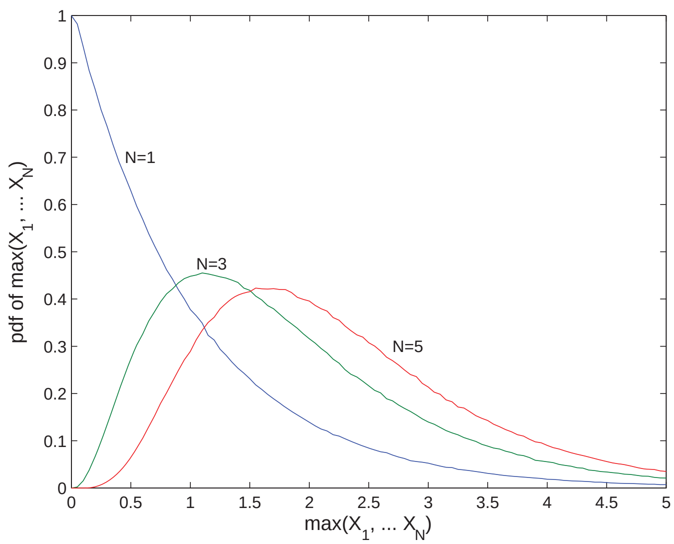

Figure 2: The probability density function of $\max(X_1, ...X_N)$ where the X_i are independent standard exponential variables. The pdfs are simulated from one to five groups of 2,000,000 draws from a standard exponential distribution.

When the total number of choices is higher, the total number of good choices is also higher, and so the value of the mth project must increase in expectation.

One can also understand the structure of order statistics through an option intuition. Consider a portfolio manager with a pool of n projects from which he will pick the m best. When the size of the pool expands, the portfolio manager has an option: if this new project is better the mth best, then the portfolio manager finances the new one, otherwise he ignores it. The option payoff is based on the project's underlying value, and the marginal value to financing an additional project is higher when the pool (the number of "options", n) is larger.

The complementarity in G between m and n controls the complementarity in payoffs between the number of projects a portfolio manager wishes to examine (n) and the number of projects he wishes to finance (m). This complementarity in G is in turn controlled by the tails of the underlying distribution of project quality (F). For example, Appendix A shows that for the exponential distribution, the complementarity in (4) is equal to $\beta\frac{1}{n+1}$: the complementarity in G is directly proportional to the scale of the distribution, β. We also derive similar results for the uniform and power law distributions. Intuitively, this means that the right tail is large enough that when a portfolio manager examines an additional project, that project has a high enough probability of being larger than the mth best existing project. Using the option intuition, the marginal value to financing an additional project increases faster in the number of options when there is a high probability that the options are "in the money": when the distribution has large tails. We illustrate the complementarity of G in Figure 3 for the exponential distribution. The gains to increasing m (left panel) are higher for larger n and the gains from increasing n (right panel) are higher for larger m.

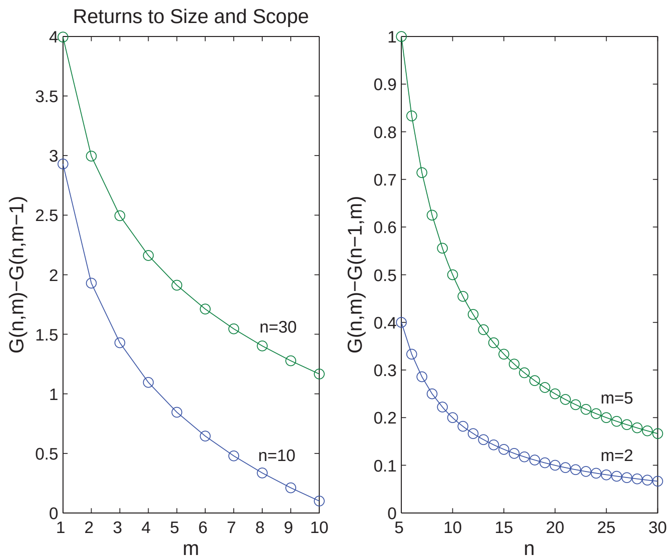

Figure 3: The left plot is the marginal value to funding an additional project, $G(n,m)-G(n,m-1)$, as a function of the number of projects already funded. The right plot is the marginal value to an additional potential project, $G(n,m)-G(n-1,m)$, as a function of the number of potential projects. The plot is generated using the exponential distribution with scale parameter $\beta = 1$.

2.3 Comparative Statics

We now have an expression for the total profits of an investment fund:

$$\pi(\phi, M) = \psi \left[M\phi\eta + (\alpha + (1-\alpha)\mu)\, G\left((1-\lambda\phi)N, M\right)\right] - M\theta \tag{5}$$

To gain intuition for how the choice of specialization (ϕ) affects the portfolio manager's investment opportunities, we will treat G as if it were a differentiable function, so that we can take "approximate derivatives", which we will denote using the standard derivative notation G_x.

$M\phi\eta + (\alpha + (1-\alpha)\mu)\, G\left((1-\lambda\phi)N, M\right)$ represents the portfolio manager's project pool quality

– the total expected value of the M best projects, including the benefits of specialization – and is increasing and concave in the number of deals examined and the number of deals undertaken. This value is then multiplied by the portfolio manager's skill (ψ) to obtain the gross payoff from the portfolio or projects.

ϕ represents specialization, which is a choice of production (or idea flow) technology. The advantage of a higher level of specialization is that the base profitability of any given project increases by $\phi\eta$, and this value is recoverable even if the project does not pay off. However the disadvantage of choosing high ϕ is that the portfolio manager faces more rapidly decreasing returns to capital. The marginal value of an additional project is $\psi\left[\phi\eta + (\alpha + (1-\alpha)\mu)\, G_m\right] - \theta$, and so the change in marginal product as ϕ increases is proportional to

$$\eta - (\alpha + (1-\alpha)\mu)\, \lambda N G_{n,m} \tag{6}$$

This cross effect – the change in the marginal value of a new project as ϕ increases – can be positive or negative depending on the size of the cross effect in G. When that cross effect in G is strong, the marginal value of additional projects is declining in specialization (ϕ), and so size and specialization must be substitutes.

When size and specialization are substitutes, an increase in any parameter that makes investments as a whole more attractive must lead portfolio managers to invest in more projects. Following the logic of substitution, they must become generalists to gather the deal flow necessary to support a high level of activity. Thus, we expect size to increase and specialization to decrease when projects are more likely to pay off (α), when failed projects are easier to recover (μ), and when portfolio managers are more skilled (ψ). We expect that portfolio managers that face a high cost of capital (θ) will choose more specialized and fewer projects.

While each of the parameters above affects investments as a whole, we also have two parameters, λ and N, that affect the investment curvature through the order statistics and G. Thus, without making specific distributional assumptions, we can only sign the partial effect of λ and N on M^* and ϕ^*. When the cost of specialization on breadth of project choice (λ) increases, this makes specialization most costly and also reduces the marginal value to funding an additional project. Thus, we expect that the direct effect of increasing λ is to decrease both specialization (ϕ) and size

(M). However, decreasing specialization has an indirect effect of increasing size, and we cannot sign the sum without further assumptions. Put differently, we have an income effect (on M) and a substitution effect (on M through ϕ) that go in opposite directions. However, we can say that holding ϕ constant, the effect of λ on M will be negative. Similarly for the effect of λ on ϕ and N on M. However, sign the partial effect of N on ϕ.

More formally:

Proposition 2 *There is a unique optimal size and specialization choice for the portfolio manager, (M^*, ϕ^*).*

If the the cross difference in G exceeds

$$\left[\begin{array}{c} (G(n+1,m) - G(n+1,m-1)) - ... \\ ...\,(G(n,m) - G(n,m-1)) \end{array} \right] > \frac{\eta}{(\alpha + (1-\alpha)\mu)\,\lambda(n+1)}. \tag{7}$$

then ϕ and M are substitutes: increasing M reduces the marginal returns to ϕ, while increasing ϕ reduces the marginal returns to M.

In addition, we have two sets of comparative statics. First:

$$\frac{d}{d\psi}\phi^* \leq 0 \qquad \frac{d}{d\psi}M^* \geq 0 \tag{8}$$
$$\frac{d}{d\alpha}\phi^* \leq 0 \qquad \frac{d}{d\alpha}M^* \geq 0 \tag{9}$$
$$\frac{d}{d\theta}\phi^* \geq 0 \qquad \frac{d}{d\theta}M^* \leq 0 \tag{10}$$
$$\frac{d}{d\mu}\phi^* \leq 0 \qquad \frac{d}{d\mu}M^* \geq 0 \tag{11}$$

Second, if we hold ϕ fixed, then M^* is increasing in N and decreasing in λ. If we hold M fixed, then ϕ^* is decreasing in λ.

The $\frac{\eta}{(\alpha+(1-\alpha)\mu)\lambda(n+1)}$ term in (7) ensures that the cross effect in G – the change in the marginal value of a new project as the pool of potential projects increases – stays above a minimum positive level, as opposed to zero in (4). Loosely, as described above, this condition means that the distribution of projects has a right tail that is "large enough".

Our model makes one additional point: the constraint on access to additional ideas alone does not justify a negative relationship between size and specialization. Instead, it is the *interaction* of limited access to ideas or projects with a payoff distribution that has a large right tail that generates a high shadow price for new potential ideas. To see the importance of the distribution, consider a case in which there was little or no variation in Δ. Then the best and worst projects would be roughly equivalent, there would be less gain to seeking a larger project pool, and the largest investment funds would have the most to gain from specialization. If all potential projects are similar, then the value of properly exploiting a project is much higher than the value of finding the right project to undertake. However, if some projects are vastly more valuable than others, then one must be able to sift through potential projects to find the most desirable. Large tails make selecting from a large set of projects the key to value.

3 An Empirical Setting: The Venture Capital Industry

To bring our model to the data, we require a direct investment setting that closely matches the features and assumptions of our model and for which sufficient data is available to determine both the size and specialization of portfolios, as well as proxies for other parameters of the model. We desire a setting in which access to ideas, deals, or projects is critical and constrained, and where investments are heterogeneous in quality. While our model potentially addresses a number of different settings, such as real estate investment funds, private equity, and hedge funds, for many of these industries, data on portfolio composition is sparse or largely unavailable. A notable exception is the venture capital industry, for which detailed data on portfolio investments and fund composition is available.

3.1 The VC Industry

The venture capital industry plays a vital economic function by identifying, funding, and nurturing promising entrepreneurs. VCs invest funds provided by institutional investors via fixed pools of capital, or funds, that are raised in advance of investment. Most VC funds are structured as closed-end, often ten-year, limited partnerships, and are not traded. The VC fund is closed with a specified amount of committed capital, which is filed with the SEC under Regulation D as a sale of securities. Funds are often marketed specifically as being specialized to specific industry or geography segments of the investment spectrum. This specialization is often formalized in the Limited Partnership Agreement that governs the fund, via covenants that limit the VC fund from investing more than a specific share of the fund in any industry or geography outside of the declared area(s) of specialization (Lerner, Hardymon, and Leamon (2007)). While the exact number of portfolio companies the fund will invest in is not specified, the fund's stated investment focus and preferred stage of investment, combined with the total dollar amount raised, defines to a great extent the number of portfolio companies the fund will be able to invest in. Empirically, VC portfolios vary widely in size, with some VCs choosing to invest in many startups, and other choosing to keep their portfolios small. Some VC firms choose to specialize in a particular industry or geographic region, while others choose to generalize across industries or invest across wider geographical boundaries.[13][14]

Further consistent with the setting of our model, access to deal flow is considered a critical asset within the VC industry (Fenn, Liang, and Prowse (1995)).[15] VCs cannot usually observe all startup companies that are seeking investment capital, and any completed investment results from a two-sided matching process (Sorensen (2007)), in which the startup must also be prefer to take funds from the particular VC rather than one of his competitors (Hsu (2004)). Due to these deal flow access constraints, VCs often engage in quid pro quo sharing of deal flow to increase the pool of potential investments they have access to (Lerner (1994), Hochberg, Ljungqvist, and Lu (2007)).[16]

Finally, the cash flows from VC investment opportunities are highly variable and skewed.

[13] For example, consider Sequoia Capital XI, a large VC fund raised in 2003 by one of Silicon Valley's premier VC firms. Sequoia XI successfully invested in both shoe stores and network security firms (Zappos.com, sold to Amazon in 2009 for about $800 million, and Sourcefire, IPOed in 2007 with a market value of about $350 million). The same fund also invested in fabless semi-conductors (Xceive), network control technology (ConSentry), airline IT and services (ITA) and social networking websites (LinkedIn). In contrast, consider Longitude Venture Partners, a smaller VC fund raised in 2008. Longitude focuses on biotechnology investments, and its portfolio consists primarly of drug development companies.

[14] The specialization of VC funds in a particular industry or geography does not appear to be merely segregation of a larger pool of capital into multiple parallel funds investing in different specialization areas. The mean time between funds raised by the same firm is 2.87 years, and the vast majority of firms do not switch areas of specialization between sequential funds, though they may generalize away from a particular area of specialization. Indeed, only 55 of the funds in our sample of 1820 funds are raised in the same vintage year that their firm raised a fund in a different reported area of specialization. Thus specialized funds do not appear to reflect a choice of internal organization as either a firm with one generalist fund or a firm with multiple diversified specialist funds.

[15] The VC literature typically attributes two important activities to VCs: deal selection and value-added support. Sorensen (2007) shows that the ability to select projects contributes on the order of 60% of VC returns, with the remainder attributed to value-added activities. Hochberg, Ljungqvist, and Lu (2007) additionally emphasize the importance of access to deal flow for VC returns. Hellmann and Puri (2000), Hellmann and Puri (2002), and Lindsey (2008) present evidence on value-added support activities by VCs.

[16] Note that in addition to limiting potential deal flow through the loss of deals outside the area of specialization, a specialist must also be concerned about loss of potential deal flow due to potential conflicts of interest when investing in closely related deals. Product market competitors will often seek out different VCs in order to avoid information spill-overs and potential cannibalization.

Sahlman (1990) describes the returns to venture capital investments from 1969 through 1985. He notes that 34.5% of money invested went to investments that were at best a partial loss and that 6.8% of money invested was responsible for nearly 50% of the final value of investments. Ljungqvist and Richardson (2003), using a datset from a later period, document that nearly three-quarters of VC portfolio companies are written off completely, while only 13% or so of investments return a multiple of three or more times invested capital over the ten to twelve year life of the fund, further supporting the notion that a small proportion of VC investments account for the majority of venture fund returns. Indeed, it is commonly accepted in the VC community that fund returns are often driven by merely one or two portfolio company successes out of an entire portfolio of investments.

3.2 From Predictions to Empirical Tests

To derive predictions and tests of our model in the VC setting, we will first look across parameters and see how they induce the choice variables for specialization and portfolio size, ϕ and M, to change. We will then construct proxy variables and convert our statements about parameters of the model into specific empirical tests.

In choosing our empirical proxies, we will take advantage of some ambiguity in the model. Most of the parameters of the model that deliver unambiguous comparative statics can be said to be broadly "good" or "bad" for a given venture capital fund. An increase in any of $\phi, \alpha, -\theta, \mu$ – higher skill, higher probability of success, lower cost of capital, higher redeployability of capital, more potential deals – result in a fund choosing to become a large generalist. In contrast, a reduction in any of $\phi, \alpha, -\theta, \mu$ will result in a fund choosing to become a small specialist. Thus, the model can deliver unambiguous predictions about the effect of proxy variables even if the parameter the variable is proxying for is ambiguous. For example, the experience of a VC might be a proxy for higher skill, higher success probability, or lower cost of capital, but our model still makes an unambiguous prediction: more experienced VCs should have larger and more generalized funds.

3.3 Parameters

3.3.1 Across VCs

For our first three predictions, we will follow the logic of substitution. First, from Proposition 1 above, we have

Prediction 1 *Generalists VCs will have larger portfolios.*

Furthermore, anything that makes VC investments more attractive or profitable will induce VCs to invest in more projects. Then, the VC must be a generalist in order to have the deal flow required to find a large number of good projects. Thus, looking across skill (ψ), we have:

Prediction 2 *More skilled VCs have less specialized and larger portfolios.*

Conversely, when all investments are constrained because a VC or the VC industry faces a high cost of capital, then the VC will choose to take on only a small number of projects and to specialize as a result. Thus, looking across cost of capital (θ), we have:

Prediction 3 *VCs with a higher cost of capital have more specialized and smaller portfolios.*

3.3.2 Across Investments

While the first set of predictions looks across VC firms, we can also look potential investments and investment conditions. Not all investments have the same characteristics: some may have higher probabilities of failure $(1-\alpha)$ than others. This makes the investment less desirable. Following the logic of substitution, fewer investments will be undertaken and the VCs that undertake riskier investments will be specialists.[17] Looking across the probability of failure $(1-\alpha)$, we find

Prediction 4 *When the probability of failure is higher, VCs will choose to specialize more and to undertake fewer projects.*

[17] In the model α is not a choice variable – since α is the probability of success, all VCs would pick $\alpha = 1$ if they could. However, one could alter the model to add a relationship between project success probability and payoffs. Assume, for example, that choosing a project with failure probability $(1-\alpha)$ increased the payoff conditional on success by a $(1-\xi\alpha)$. Then, (5) becomes $\pi(\phi, M, \alpha) = \psi \left[M\phi\eta + M\alpha(1-\xi\alpha) + (\alpha + (1-\alpha)\mu)\, G\left((1-\lambda\phi)N, M\right)\right] - M\theta$. This is concave in α, so a unique optimum in α exists. As long as ξ is not too high (analogous to Condition 7), π is super-modular in $(M, -\phi, \alpha)$, so, following the proof of Proposition 2, α^* will be complement to M^* and a substitute for ϕ^*. Thus, even endogenously, the specialists will choose investments with a higher likelihood of failure.

Similarly, if capital (human or physical) from failed investments is more redeployable (higher μ), then more investments will be undertaken. Following the logic of substitution, this will lead VCs to generalize:

Prediction 5 *When capital can be more easily re-deployed from failed projects, VCs will choose to generalize more and to undertake more projects.*

3.3.3 Across Opportunities

Different venture capital markets have different scales. For example, the number of potential entrepreneurs varies in time, geographically, and by industry. In doing our analysis, we can look across these heterogeneous markets. Markets with less prospective deal flow (smaller N) should have smaller VC fund portfolios. However, the model does not provide unambiguous comparative statics for the effect of N on specialization. Thus, we have:

Prediction 6 *When prospective deal flow is lower, VCs undertake fewer projects.*

Additionally, some choices of specialization are more costly to the VC in terms of the manner in which they affect the set of potential projects he can then undertake. In doing our analysis, we can look across different areas of specialization to capture heterogeneity in the cost of specialization. specialization that is more costly to deal flow (higher λ) should lead to smaller VC portfolios, and to *less* specialization:

Prediction 7 *When the cost of specialization is higher, VCs undertake fewer projects and have less specialized portfolios.*

3.4 Proxies and Tests

The first prediction is that, unconditionally, larger VCs should be more generalized. This is directly testable:

Test 1 *Generalist VCs have larger portfolios.*

3.4.1 Experience

Traditionally, the VC literature has used VC experience as a proxy for skill: in order to be able to continue participating as an investor in the industry, the VC firm must be able to raise sequences of overlapping funds, and the ability to raise a follow-on fund is increasing in the VC's past performance. Thus, to become experienced, a VC must have some level of skill (Hochberg, Ljungqvist, and Vissing-Jorgensen (2009)). With the addition of learning-by-doing, VC parent firm experience, as well as the sequence number of the fund, represent plausible proxies for VC skill. Greater experience and higher fund sequence numbers suggest the VC has performed well enough, over a sequence of funds, that he is likely of high(er) skill.

Similarly, experience and fund sequence number can also be taken as plausible proxies for cost of capital: there is likely to be less uncertainty about the skill level of a VC firm that has been in existence for some time and has a long track history of returns. Thus, the asymmetric information problem facing investors in an experienced VC firm is likely to be less severe.

Test 2 *More experienced VCs have larger and less specialized portfolios.*

Notice that this prediction is directly opposite of the standard learning-by-doing argument regarding experience. In those models, experience in a specific area – such as internet security – should make the VC more skilled *in that area*. This skill then enhances productivity, which should support larger portfolios in the presence of decreasing returns to capital. Here, our prediction is that while expertise indeed enhances productivity, this enhancement is not sufficient to overcome the cost of having a more narrow pool of ideas from which to select investments.

3.4.2 Stage of Investment

While we cannot observe the underlying ex-ante probability of success for any individual portfolio company investment, the literature provides us with a number of possible ways to empirically examine this prediction. Portfolio companies differ in systematic ways: some may be early-stage investments, where uncertainty of outcome is high, while others are later-stage investments, where much of the uncertainty has been resolved. Investments in earlier-stage rounds are more speculative and so the probability of failure $(1-\alpha)$ for any given attempt is higher. Thus, we predict that[18]

Test 3 *VCs that invest primarily in seed or early stage investments will have smaller and more specialized portfolios.*

This prediction runs counter to a common diversification argument for early investment. Since early stage investments are generally smaller and riskier than later stage companies for whom some uncertainty has been resolved, the specialized VC funds that invest in these companies are giving up an obvious benefit to diversification.

3.4.3 VC Inflows

Gompers and Lerner (2000) show that as inflows into the VC industry increase, valuations for portfolio companies are bid up. There is an increase in competition within the industry for a limited number of attractive investment opportunities, often referred to as "money chasing deals." Hochberg, Ljungqvist, and Lu (2007) further show that higher flows to the VC industry in a year a fund is raised lead to a lower rate of successful exit companies in the fund's portfolio. This suggests that as money chases deals, not only are prices for attractive projects bid up, but lower quality projects are funded as well.

Thus, VC inflows may proxy for time-variation in the probability of a given project's success (lower α) or for overall project profitability (lower ψ). In either case, we predict

Test 4 *As inflows into the VC industry increase, VCs will have smaller and more specialized portfolios.*

Notice that this prediction is directly opposite of the obvious mechanical relationship. As total inflows increase, the dollar size of VC funds increases (Kaplan and Schoar (2005)). One might therefore expect that the effect would be to increase the number of projects undertaken in any potential VC portfolio. Our model leads us to predict the reverse: more money in the industry overall should lead to a smaller average portfolio size.

[18] While seed and early stage investments are generally smaller in dollar amount, we measure portfolio size in count of portfolio companies, not committed (or invested) capital. Thus, we do not create a large mechanical relationship between size and stage of investment.

3.4.4 High Deal Flow and Cost of Specialization

While we cannot directly observe either the pool of potential projects or the cost of specialization to deal flow, certain well-known industry features allow us to proxy for these variables. An area of geographic specialization may proxy for the general pool of potential deals available to the VC (N). For example, a VC who is located in a known cluster of entrepreneurial activity, such as Silicon Valley or the Route 128 area around Boston, will likely see a larger pool of potential projects than a VC located in a less active area. Additionally, the area of geographic specialization can provide us with a proxy for the cost of specialization on deal flow (λ). For example, a VC fund specializing geographically in Silicon Valley reduces the available number of deals much less than a similar fund specializing geographically in the Lehigh Valley area of Pennsylvania. While both VCs, ex-ante, may see the same set of potential deals from across the country, different geographic specialization choices have different effects on the final pool of investments from which VCs select – the VCs differ in the value of λ.

Thus, we can employ the MSA in which a VC fund invests the largest portion of their capital to proxy for the cost of geographic specialization in that area. Funds that make the largest portion of their investments in the San Francisco Bay Area or Boston-Route 128 area, the two largest clusters of entrepreneurial activity in the United States, presumably face a lower cost of geographic specialization (λ) and a greater ex-ante pool of potential deal flow (N). Both lower λ and higher N in the model lead to larger portfolios, and thus we can predict:

Test 5 *VCs who concentrate the largest portion of their investments in the SF Bay Area or Boston Metro area should have larger portfolios.*

The model also unambiguously predicts that lower λ leads to more specialized portfolios, though it does not provide an unambiguous prediction for N. However, we can further test whether indeed:

Test 6 *VCs who concentrate the largest portion of their investments in the SF Bay Area or Boston Metro area have more geographically specialized portfolios.*

3.4.5 Performance

Our model does not provide predictions for fund performance, neither in terms of returns nor in terms of exit rates. The Δ employed in our model refer not to returns from project success to those who provide the capital (LPs), but rather to the profits of the portfolio manager, which may differ substantially. More formally, the optimization problem presented in the model is over payoffs to GPs, not to the limited parters or entrepreneurs. Thus, it is difficult to translate from the model to the typical net-of-fee returns to limited partner investors or to the entrepreneur success rates typically employed in the VC literature.

Let us first consider returns. Specialization and portfolio size are choice variables, taken as a function of parameters, in order to maximize the total value of the VC's profits. Thus, if we were able to perfectly measure and control for all of the relevant parameters, interior choices of size and specialization would contain no additional information for performance. Any non-zero regression coefficient would simply be the result an interaction of omitted variable bias and parameter measurement error. In any event, we cannot examine performance of this sort empirically, as the given optimization problem is over payoffs to the VC general partner, not to the limited partners or to the entrepreneur, and we do not observe VC general partner profits.

Next, consider exit rates, a common measure of VC fund performance in the finance literature. Let us assume that the portfolio firm is able to achieve exit (via IPO or successful trade sale) if its realized value is above a certain threshold.[19] Our model assumes that more generalist VCs will have the additional deal flow that allows them to pick more projects that are likely to be above the threshold, while specialists have the developmental ability to push marginal projects above the threshold. We cannot know which effect will dominate without making specific assumptions about the distribution of underlying projects, the severity of the exit threshold, and the relationship between the VC's payoff and project value. That said, recent empirical work by Gompers, Kovner, and Lerner (2009), which finds that specialist VCs enjoy higher exit rates than generalists, suggests that the additional developmental ability of specialists may be the dominating force.

[19]We interpret this to mean that the project must be successful (which happens with probability α) and that the cash flows ($\phi\eta + \Delta$) must be above a certain level. The effect of specialization on the fraction of projects that will exceed the threshold is ambiguous. Generalists will have higher Δ on average, while specialists have higher $\phi\eta$.

3.4.6 Life-cycle

Our model can also be consistent with a life-cycle model of specialization, whereby new VCs who enter the industry with lower skill and higher cost of capital will have smaller and more specialized first funds. As the VC gains skill over time and is able to lower his cost of capital, he will become larger and less specialized. While our model does not make any life-cycle predictions as is, as illustrated in by the proxies discussed above, the parameters of the model can be interpreted to capture the distinguishing features of younger or more established VC firms.

4 Data

The data for our analysis come from Thomson Financial's Venture Economics database. Venture Economics began compiling data on venture capital investments in 1977, since backfilled to the early 1960s. Most VC funds are structured as closed-end, often ten-year, limited partnerships. They are not usually traded, nor do they disclose fund valuations. The typical fund spends its first three or so years selecting companies to invest in, and then nurtures them over the next few years. In the second half of the fund's life, successful portfolio companies are exited via IPOs or sales to other companies generating capital inflows that are distributed to the fund's investors. At the end of the fund's life, any remaining portfolio holdings are sold or liquidated and the proceeds distributed to investors.

Our data contains the vast majority of U.S. VC investments made between the years 1975 and 2003. Owing to the VC investment cycle, relatively recent funds have not yet operated for long enough to fully observe the breadth of their investment types and determine the extent to which they are specialized or generalized. To allow our measure of portfolio size and specialization to include the first four years of a fund's life, when investments are made, we exclude all funds raised after 1999.[20] Our results are robust to including funds of later vintages. We further exclude funds raised before 1980, both because the reliability of the Venture Economics data pre-1980 is lower, and because venture capital as an asset class that attracts institutional investors has only existed since 1980.[21]

[20] Closing the sample period at year end 1999 provides at least four years of investment activity for the youngest funds, using November 2003 as the latest date of investment data in our sample.

[21] The institutionalization of the VC industry is commonly dated to three events: The 1978 Employee Retirement

We concentrate solely on investments by U.S. based VC funds, and exclude angel and buy-out funds. We exclude all VC funds that are not independent (structured as limited partnerships with overlapping sequences of funds), since corporate and banking VCs often have strategic goals that determine their level of specialization. In addition, we exclude all funds with fewer than five unique portfolio companies. This ensures that if we see a fund whose investments are primarily concentrated in a single industry, it is likely due to intent, rather than chance.

Our final data-set includes 1820 funds managed by 879 VC firms. Table I describes our sample funds. The average sample fund had \$87 million of committed capital, with a range from \$0.1 million to \$5 billion. (Fund size is unavailable for 33 of the 1,820 sample funds.) Fund sequence numbers denote whether a fund is the first, second and so forth fund raised by a particular VC management firm. The average sample fund is a third fund, and the median is a second fund, though sequence numbers are missing in Venture Economics for 258 of the sample funds. 30% of funds are identified as first-time funds.

4.1 Funds Versus Firms

We distinguish between funds and firms. While VC funds have a limited (usually ten-year) life, the VC firms that manage the funds have no predetermined lifespan. Success in a first-time fund often enables the VC firm to raise a follow-on fund (Kaplan and Schoar (2005)), resulting in a sequence of funds raised a few years apart. We assume that experience and contacts acquired in the running of one fund carry over to the firm's next fund and so measure VC experience at the parent firm rather than the fund level. We aggregate round-by-round investments in portfolio companies to calculate fund-specific variables.

For the purposes of determining portfolio size and specialization, we focus on the VC *fund*. As noted in Section 3, the decision regarding fund size and specialization is made at the time a VC raises the fund: the fund is closed with a specified amount of committed capital, which is filed with the SEC under Regulation D as a sale of securities. Partnership agreements are often drafted with covenants that limit funds with stated areas of focus from investing more than a certain fraction

Income Security Act (ERISA) whose "Prudent Man" rule allowed pension funds to invest in higher-risk asset classes; the 1980 Small Business Investment Act which redefined VC fund managers as business development companies rather than investment advisers, so lowering their regulatory burdens; and the 1980 ERISA "Safe Harbor" regulation which sanctioned limited partnerships which are now the dominant organizational form in the industry.

of the committed capital outside of those areas of focus. Finally, the managing members of the general partner vehicle of the fund are defined at fund-raising as well, and typically remain constant over the life of the fund.

4.2 Portfolio Size and Specialization

As our measure of portfolio size, we compute the number of unique portfolio companies in which a given fund invests over the course of its life. The average portfolio for our sample funds consists of approximately 23 unique portfolio companies, while the median fund portfolio consists of 17 unique portfolio companies.

To measure specialization, we compute two concentration measures. As a measure of industry specialization, we compute the Herfindahl-Hirschman Index (HHI) of investment by industry for each fund, based on the number of investments made by the fund in each industry. Venture Economics uses six industries: biotechnology, communications and media, computer related, medical/health/life science, semiconductors/other electronics, and non-high-technology. As a measure of geographic specialization, we compute the HHI by CMSA based on the number of investments made by the fund in each of the 287 Consolidated Metropolitan Statistical Areas in our data-set. (For example, San Francisco, Oakland, and San Jose are considered a single location). All our reported results are robust to employing HHI computed using dollar investment amounts in each industry or CMSA instead of number of portfolio companies. The median fund in our sample has an industry HHI measure of 0.36, with a range from 0.18 to 1, and a geography HHI measure of 0.22, with a range of 0.04 to 1.

4.3 VC Firm Experience

We derive four direct proxies for the experience of the VC parent firm. These are the age of the VC firm (the number of days since the VC firm's first investment); the number of rounds the firm has participated in; the cumulative amount the firm has invested; and the number of portfolio companies it has backed. Each measure is calculated using data from the VC firm's creation through the year the fund in question was raised. To illustrate, by the time Sequoia Capital raised Fund IX in 1999, it had been active for 24 years and had participated in 888 rounds, investing a total of $1,275 million in 379 separate portfolio companies. As an alternative measure of experience, we use the fund's sequence number. In the interest of brevity, we present univariate sorts and regression results using only the cumulative number of days since the VC's first-ever investment, and the fund's sequence number, though we obtain similar results using any of the alternative measures.

4.4 Investment Stage

We calculate the proportion of deals in which the fund has invested that were reported to be at seed or early stage of development at the time the fund first invested in them. We define a seed or early stage dummy variable as taking the value of one if the fund first invested in its portfolio companies at the seed or early stage with greatest frequency. 13.3% of funds are thus defined as primarily investing in seed or early-stage investment opportunities.

4.5 Money Chasing Deals

We compute the aggregate inflows into VC funds in the year a sample fund was raised. Table I shows that the average fund in our sample was raised in a year in which $23.4 billion flowed into the VC industry.

4.6 Deal Flow and Cost of Specialization

As discussed in Section 3, we proxy for the cost of specialization in restricting deal flow and the size of the overall pool of potential deal flow using the geographic area in which the fund invests the largest portion of its capital. For each fund, we identify the MSA in which the fund invested in the most companies. We define the geographic area to be high deal flow / low cost if the fund's primary MSA of investment is either the San Francisco-Oakland-San Jose consolidated MSA or the Boston-Worcester-Lawrence consolidated MSA. Roughly 45.5% of the funds in our sample make the largest fraction of their investments in the San Francisco Bay Area, while 13.9% of funds make the largest fraction of thier investments in the Boston/Route 128 area. In total, 59.5% of our funds are attributed high deal flow / low cost of specialization based on our geography-based proxy.

5 Testing the Model's Predictions

5.1 Correlations and Univariate Sorts

5.1.1 Size and Specialization

The main implication of our model is that portfolio size and specialization are substitutes. As Table II shows, portfolio size and specialization have a significant negative unconditional correlation between -0.26 and -0.29, depending on the dimension of specialization.

The relationship between size and specialization can also be illustrated in univariate sorts. Panel A of Table III presents sorts of portfolio size over quartiles of fund specialization, while Panel B presents sort of fund specialization measures across quartiles of portfolio size. The negative relationship between portfolio size and specialization is striking: regardless of specialization measure, portfolio size increases sharply as we move from the highest quartile of specialization to the lowest quartile. The differences between portfolio size for the most generalist funds and most specialist funds is significant at the 1% level. In addition, the magnitude is very large: average portfolio size in the most generalist quartile is between 176% and 201% of portfolio size in the most specialist quartile.For example, firms in the bottom quartile of industry specialization – the industry generalists – have a mean portfolio size of approximately 30 companies, while firms in the top quartile of industry specialization – the industry specialists – have a mean portfolio size of approximately 17 companies. We find the same pattern when we reverse the order of the univariate sorts, sorting fund specialization measures across quartiles of portfolio size. As we move from the lowest quartile of portfolio size to the highest quartile of portfolio size, specialization decreases significantly, regardless of the dimension on which it is measured. Once again, these differences are significant at the 1% level.

5.1.2 Experience

As noted in Section 3, our model implies that greater experience should be associated with larger, less-specialized portfolios. As Table II shows, the unconditional correlation between experience and size is positive regardless of the measure of experience employed. The unconditional correlations between portfolio size and experience range from 0.18 to 0.26 and are all significant at the 1%

level. As an alternative measure of experience, we can also consider fund sequence number. The unconditional correlation between fund sequence number and portfolio size is also positive and significant, at 0.10.

Panels C and D of Table III presents univariate sorts of portfolio size and specialization measures over quartiles of fund experience. (The table employs experience based on number of days since parent firm's first investment, though similar results obtain over the other three experience measures, or fund sequence number.) Firms in the highest quartile of experience are on average 155% larger than those in the lowest quartile of experience, with a difference in mean portfolio size between the highest and lowest quartiles of experience of approximately 10 portfolio companies, and the difference is significant at the 1% level.[22] Similarly, sorts of specialization over quartiles of experience demonstrate that more experienced firms create funds that are less specialized, with the difference in fund specialization between the highest and lowest quartiles of firm experience significant at the 1% level for both dimensions of specialization.

5.1.3 Investment Stage

As stated in Section 3, our model implies that VC funds investing in earlier rounds should be smaller and more specialized. As Table II shows, the correlation between the early stage indicator variable and portfolio size is -0.07, while the correlation between the early stage indicator and our measures of specialization is positive, ranging from 0.07 to 0.20. All the reported correlations are statistically significant at the 1% level.

Funds investing primarily in seed or early stage deals exhibit a mean portfolio size of 19.4 companies, versus a mean portfolio size of 23.6 portfolio companies for funds that do not invest primarily in seed or early stage deals; the difference is significant at the 1% level. Early-stage funds also exhibit higher levels of specialization. Early-stage funds have mean industry HHIs of 0.42, versus 0.39 for later-stage focused funds, and geography HHIs of 0.37, versus 0.26 for later-stage funds; both these differences are significant at the 1% level.

[22]This is also consistent with findings on dollar size of funds and experience documented in Kaplan and Schoar (2005), Gottschalg and Phalippou (2007) and Hochberg, Ljungqvist, and Vissing-Jorgensen (2009).

5.1.4 Money Chasing Deals

As stated in Section 3, our model implies that inflows into the entire VC industry should lead funds to be smaller and more specialized. As Table II shows, the correlation between inflows and portfolio size is -0.13, and the correlations between inflows and the four measures of specialization are positive, ranging from 0.16 to 0.29. All the reported correlations are statistically significant at the 1% level.

Panel C of Table III presents univariate sorts of portfolio size and specialization measures over quartiles of total inflows. The portfolio size of funds raised in years in which inflows into the VC industry were highest are significantly smaller than the portfolio size of funds raised in years in which inflows into the VC industry were lowest, with the difference in mean portfolio size between the highest and lowest quartiles of inflows is 7.25 firms. Similarly, sorts of specialization over quartiles of total demonstrate that funds raised in years in which inflows into the VC industry were highest are significantly more specialized than funds raised in years in which inflows into the VC industry were lowest, with the difference in specialization between the highest and lowest quartiles of firm experience significant at the 1% level for both geography and industry specialization measures.

5.1.5 Deal Flow and Cost of Specialization

As noted in section 3, our model provides an unambiguous prediction that funds investing primarily high deal flow / low cost of specialization areas should have larger portfolios, holding the optimal choice of specialization constant. The prediction for the effect of our proxy on specialization is less clear. The model provides an unambiguous prediction for the effect of the cost of specialization (λ) on the specialization of the portfolio, holding optimal portfolio size constant – funds with lower costs of specialization to deal flow should be more specialized. However the model is ambiguous on the equivalent effect of having a larger ex-ante pool of potential deal flow (N). Given that our proxy may capture both a larger pool of ex-ante deal flow and a lower cost of specialization, we therefore do not have an unambiguous prediction for the effect of the proxy on the geographic specialization of the portfolio, though we nevertheless explore the resulting empirically observed relationship.

As the predictions resulting from these comparative statics require us to hold the optimal level of specialization and/or portfolio size constant, we do not have univariate predictions for the relationship between our deal flow / cost of specialization proxy and portfolio size or specialization. That said, we note that funds investing primarily in the San Francisco Bay Area or Boston-Route 128 area – our proxy for higher ex-ante deal flow and lower cost of geographic specialization – exhibit a mean portfolio size of 26.7 companies, versus a mean portfolio size of 17.6 portfolio companies for funds that invest primarily in other geographic areas; the difference is significant at the 1% level. In contrast, there are no significant univariate differences in geographic specialization between the two sub-samples (0.27 and 0.28).

Both the correlations and the patterns in the univariate sorts presented above are consistent with the predictions of the model. We now turn to analyzing the relationships between our key variables of interest in a multivariate setting.

5.2 Multivariate Models

We begin by examining the tests that result from our first set of comparative statics: the effect of proxies for $\{\psi, \alpha, -\theta, \mu\}$ on size and specialization.

Our model produces two first order conditions, and so it is tempting, at first glance, to view the proper empirical model as one of simultaneous equations. This, however, is not correct.

Consider a simultaneous equations system (SEM) of the form

$$\text{portfolio size} = f(\text{specialization}, \text{exogenous parameters}) + \epsilon \tag{12a}$$

$$\text{specialization} = f(\text{portfolio size}, \text{exogenous parameters}) + \mu \tag{12b}$$

As Wooldridge (2002) demonstrates, in order for a set of equations to be an SEM, each equation must be autonomous, i.e., "have an economic meaning in isolation from the other equations in the system." For example, in a supply and demand SEM, the decision to supply a given good and the decision to demand that good come from the distinct optimization problems of two separate agents, and thus the supply and demand curves have precise ceteris paribus interpretations. Here, the two choice variables are the result of a single optimization problem for one economic agent (the VC fund). Thus, there is no meaning in asking how an exogenous shock to specialization affects size (or vice versa) because both are determined jointly as a function of the underlying parameters. Our interest is in testing whether, other factors held fixed, there exists a tradeoff between portfolio size and specialization. To do so, Wooldridge (2002) argues that we should simply estimate equation (12a) and/or equation (12b) as stand-alone equations.[23]

We thus present two sets of multivariate models analyzing the relationship between portfolio size, specialization and our parameter proxies. Our first set of models uses specialization as the dependent variable, while our second set uses portfolio size as the dependent variable. Table IV presents models in which the dependent variable is the industry specialization (HHI) of the fund's portfolio, and the independent variables are the proxies for the variables of interest from our model, described above, as well as year dummies that provide additional controls for changing conditions over the course of our sample. Table V presents similar models, this time employing geography specialization as the dependent variable.

Our specialization measures have support on [0,1] and positive mass on 1. To avoid the resulting well-known biases of OLS in this situation, we estimate fractional logit models using quasi-MLE (see Papke and Wooldridge (1996)).[24] This involves modeling the conditional mean, $E(y|x) = e^{\beta x}/(1+e^{\beta x})$. All models are estimated included year controls (not reported), and standard errors are heteroscedasticity-consistent and clustered by VC parent firm.

In each table, we present five models. In column (1), we present a simple model of specialization as a function of portfolio size. In columns (2) and (3), we model specialization as function of our other independent variables of interest, experience, the indicator for early-stage focus, and VC inflows alone, once using the natural logarithm of days since first startup investment by the parent firm as a measure of experience, and once using the natural logarithm of fund sequence number as our measure of experience[25] In columns (4) and (5), we estimate models with the full range of variables of interest from our model.

In all five of the models estimated, for both dimensions of specialization, we observe a clear, sta-

[23]Note that even if we thought (12a) and (12b) made sense as an SEM, we would not be able to estimate the parameters. Estimating such a system requires two exogenous variables, one per equation, each of which affects the LHS endogenous variable but not the RHS endogenous variables. Our model implies the existence of such a variable for portfolio size: both skill and cost of capital appear only in the first-order condition for size. However, every other variable appears in both first order conditions, and so we cannot have an instrument for specialization.

[24]All our reported results are robust to employing either Tobit estimations bounded from above by one or naive OLS estimations.

[25]Fund sequence number and fund experience are highly correlated, on the order of 0.75, and thus we include them in our empirical models separately, rather than together. Our results are robust to employing the three other direct measures of experience described above as well.

tistically significant negative relationship between portfolio size and specialization. The magnitude of the associated relationship is substantial: holding all other variables at their means, a one standard deviation increase in fund portfolio size is associated with a reduction in industry HHI that ranges from -0.033 to -0.045, depending upon the exact specification estimated. (This compares to the unconditional mean industry HHI of 0.40.) Similarly, holding all other variables at their means, a one standard deviation increase in fund portfolio size is associated with a reduction in geography HHI that ranges from -0.056 to -0.066, depending upon the model estimated (compared to the unconditional mean geography HHI of 0.28).

As our model predicts, the estimates from Tables IV and V indicate that experience and sequence number are associated with less specialization: holding all other variables at their means, a one standard deviation increase in fund parent firm experience (sequence number) is associated with a reduction in industry HHI of approximately -0.021 (-0.022), and a reduction in geography HHI of approximately -0.012 (-0.017). Increases in investment in earlier stage companies or in total VC inflows are associated, as predicted by the model, with increased specialization of the fund's portfolio. Holding all other variables at their means, focusing on early-stage portfolio companies is associated with an increase of 0.016 to 0.025 in industry specialization of the portfolio, and 0.085 to 0.089 in geography specialization. A one standard deviation increase in total VC inflows into the industry is associated with an increase in industry HHI in the range of 0.037 to 0.050, and an increase in geography HHI of 0.018 to 0.034.

In, Table VI, we reverse the designation of dependent variable, and analyze the relationship between the exogenous variables in our model and portfolio size. We take two approaches to analyzing portfolio size. In Panel A of the table, the dependent variable is the natural logarithm of portfolio size. In Panel B, to demonstrate robustness by avoiding the issues related to the estimation of count variables, we estimate poisson models of portfolio size (the count of unique number of portfolio companies). In the first two columns of each panel, we estimate simple models of portfolio size as a function of specialization (industry or geography HHI). In the third column, we model size as a function of the proxies for the exogenous parameters in our model. In the fourth and fifth columns, we estimate models of size as a function of industry specialization in addition to the proxies for exogenous parameters. In the sixth column we model size as a function of geography specialization and our proxies.

In both panels, similar patterns emerge. In all five of the models that include a specialization measures on the RHS, we observe a positive association between specialization and size, consistent with the predictions of our model. We observe a positive association between experience and size, and a negative association between early stage or VC inflows and size consistent with the predictions of the model. Here too, the magnitudes of the effects are large. Holding all other variables at their means, a one standard deviation increase in industry (geography) specialization is associated with a decrease in portfolio size of 4.2 to 5.6 (5.32 to 6.6)companies, depending on the model estimated. This compares to an unconditional mean portfolio size of 23 companies. A one standard deviation increase in experience is associated with an increase in portfolio size of 4.3 to 4.4 companies, a one standard deviation increase in total VC inflows is associated with a decrease in portfolio size of 2.2 to 3.5 companies, and focusing on early stage companies is associated with a decrease in portfolio size of 1.6 to 3.4 companies, depending on the model. With the exception of the coefficients on the indicator for early-stage investment focus in two of the twelve models, all the coefficients are statistically significant at conventional levels, and the vast majority are significant at the 1% level.

Our final set of models extends the estimations conducted above to examine the effects of our proxy for the technology parameters $\{-\lambda, N\}$. The comparative statics for these parameters lead to predictions on portfolio size and specialization that hold only when either the level of specialization or portfolio size is held constant. Here, it is once again useful to note that estimation of an SEM-type model would be doubly inappropriate, as we specifically wish to test the direct effects of a change in these parameters on portfolio size (specialization), rather than the sum of the direct and indirect effects through specialization (size).

Table VII augments the models estimated in Tables IV through VI with our proxy for high deal flow / low cost of specialization, the indicator variable taking the value of one if the fund primarily invests in either the San Francisco Bay Area or the Boston-Route 128 area. Columns (1) and (2) present estimates from OLS models where the dependent variable is the natural logarithm of the fund's portfolio size measured in number of unique firms. Columns (3) and (4) present estimates from poisson regressions where the dependent variable is the portfolio size measured in number of unique firms. In all four models, the independent variables are as in the models in Table VI, with the addition of the indicator for primarily investing in San Francisco Bay Area or Boston-Route 128 area – our proxy for higher ex-ante deal flow and lower cost to geographic specialization. All models are estimated including year controls (not reported), and standard errors are heteroscedasticity-consistent and clustered by VC parent firm.

Consistent with the predictions arising from our comparative statics on λ and N, we observe a positive and significant relationship between being located in a high deal flow / low cost of geographic specialization area – having the San Francisco Bay Area or Boston-Route 128 area as a primary geography of investment – and portfolio size, holding geographic specialization constant. The magnitude of the effect is significant. Holding all other variables at their means, investing primarily in SF/BOS is associated with a increase in portfolio size of 7.2 to 7.7 companies, depending on the model estimated. In all four of the models, we observe (as in the models in Table VI) a positive association between specialization and size, a positive association between experience and size, and a negative association between early stage or VC inflows and size consistent with the predictions of the model. The magnitudes of these effects remain similar to those in the models in Table VI, with the exception of the indicator for investing primarily in seed or early stage investments.

While our model does not provide an unambiguous prediction for the effect of higher N on specialization, it provides a prediction for the effect of lower λ on specialization, holding portfolio size constant. Our proxy – primarily investing in San Francisco Bay Area or Boston-Route 128 area – can proxy for both λ and N, and thus it is interesting to examine what empirical effect is observed when relating the proxy to geographic specialization, when portfolio size is held constant. To that end, columns (5) and (6) of Table VII present estimates of fractional logit models where the dependent variable is the geographic specialization (HHI) of the fund, and the independent variables are as in the models estimated in Table V, augmented by the indicator for primarily investing in San Francisco Bay Area or Boston-Route 128 area – our proxy for higher ex-ante deal flow and lower cost to geographic specialization. Both models include year controls (not reported), and standard errors are heteroscedasticity-consistent and clustered by VC parent firm. In both models, we observe a positive and significant relationship between being located in a high deal flow / low cost of geographic specialization area – having the San Francisco Bay Area or Boston-Route 128 area as a primary geography of investment – and geographic specialization, consistent with the model's predictions for the effect of lower λ on specialization. Holding all other variables at their means, investing primarily in SF/BOS is associated with a increase in portfolio geography specialization of roughly 0.04 regardless of model estimated. The coefficients on the remaining independent variables remain similar to those estimated in Table V.

Overall, the empirical relationships and patterns documented in Tables II through VII are consistent with the main predictions of our model, and are often different than the patterns generally presumed in the investments literature with regards to size and scope. These patterns emphasize the importance of accounting for access to deal flow when examining the choice of specialization and size in a setting where deal selection is critical.

6 Conclusion

There is a general presumption in economics and finance that specialization enhances productivity and is thus an important driver of value. In this paper, we present and test a model of project selection that generates endogenously decreasing marginal returns to capital, so that the investment frontier is a function of the specialization choice of the project manager. Specialization can increase the cash flows from any given idea or improve the average quality of a pool of ideas. However, specialization restricts the pool of potential ideas the investor can choose from because some of the potential projects are outside of either the investor's expertise or his area of focus. The second effect dominates when investments are highly heterogeneous in quality, and so size and specialization are substitutes. In addition, positive factors, such as higher skill will induce portfolio managers to become larger in size and scope, and negative factors, such as competition induce portfolio managers to become smaller and more specialized. We test our model with data from the venture capital industry and find empirical support in the form of several results that run counter to common intuitions regarding project selection. As a result, our model and empirical work can provide insight into how investment funds are structured, and, by implication, how investment vehicles without ongoing business activities differ from firms in the way they make choices.

References

Berk, J., and R. Green, 2004, Mutual Fund Flows and Performance in Rational Markets, *Journal of Political Economy* 112, 129–1295.

Chen, J. S., H. Hong, M. Huang, and J. D. Kubik, 2004, Does Fund Size Erode Mutual Fund Performance? The Role of Liquidity and Organization, *American Economic Review* 94, 1276–1302.

Fenn, G., N. Liang, and S. Prowse, 1995, The Economics of Private Equity Markets, Staff Study 168, Board of Governors of the Federal Reserve System.

Fulghieri, P., and M. Sevilir, 2008, Size and Focus of a Venture Capitalist's Portfolio, *Review of Financial Studies* Forthcoming.

Gompers, P. A., A. Kovner, and J. Lerner, 2009, Specialization and Success: Evidence from Venture Capital, *Journal of Economics, Management and Strategy* 18, 817–844.

Gompers, P. A., and J. Lerner, 2000, Money Chasing Deals? The Impact of Fund Inflows on Private Equity Valuations, *Journal of Financial Economics* 55, 281–325.

Gottschalg, O., and L. Phalippou, 2007, The Performance of Private Equity Funds, *Review of Financial Studies* p. forthcoming.

Hellmann, T., and M. Puri, 2000, The Interaction between Product Market and Financing Strategy: The Role of Venture Capital, *Review of Financial Studies* 13, 959–984.

Hellmann, T., and M. Puri, 2002, Venture Capital and the Professionalization of Start-up Firms: Empirical Evidence, *Journal of Finance* 57, 169–197.

Hochberg, Y. V., A. Ljungqvist, and Y. Lu, 2007, Whom You Know Matters: Venture Capital Networks and Investment Performance, *Journal of Finance* 62, 251–301.

Hochberg, Y. V., A. Ljungqvist, and A. Vissing-Jorgensen, 2009, Informational Hold-up and Performance Persistence in Venture Capital, Working paper, Northwestern University and New York University.

Hochberg, Y. V., M. Mazzeo, and R. McDevitt, 2009, Market Structure, Competition and Specialization in Venture Capital, Working paper, Northwestern University.

Hsu, D. H., 2004, What do Entrepreneurs Pay for Venture Capital Affiliation?, *Journal of Finance* 59, 1805–1844.

Inderst, R., H. M. Mueller, and F. Münnich, 2007, Financing a Portfolio of Projects, *Review of Financial Studies* 20, 1289–1325.

Kaplan, S. N., and A. Schoar, 2005, Private Equity Performance: Returns, Persistence and Capital Flows, *Journal of Finance* 60, 1791–1823.

Lerner, J., 1994, The Syndication of Venture Capital Investments, *Financial Management* 23, 16–27.

Lerner, J., F. Hardymon, and A. Leamon, 2007, Note on Private Equity Partnership Agreements, Harvard Business School Publishing, 9-294-084.

Lindsey, L. A., 2008, Blurring the Boundaries: The Role of Venture Capital in Strategic Alliances, *Journal of Finance* 63, 1137–1168.

Ljungqvist, A., and M. Richardson, 2003, The Investment Behavior of Private Equity Fund Managers, Working paper.

Maksimovic, V., and G. Phillips, 2002, Do Conglomerate Firms Allocate Resources Inefficiently Across Industries? Theory and Evidence, *Journal of Finance* 57, 721–767.

Maksimovic, V., and G. Phillips, 2007, Conglomerates and Internal Capital Markets, in B. Espen Eckbo, eds.: *Handbook of Corporate Finance: Empirical Corporate Finance, Vol. 1* (Elsevier/North Holland, Amsterdam, Netherlands).

Ozbas, O., 2005, Integration, Organizational Processes, and Allocation of Resources, *Journal of Financial Economics* 75, 201–242.

Papke, L. E., and J. M. Wooldridge, 1996, Econometric Methods for fractional response variables with an application to 401(k) plan participation rates, *Journal of Applied Econometrics* 11, 619–632.

Rajan, R., H. Servaes, and L. Zingales, 2000, The Cost of Diversity: The Diversification Discount and Inefficient Investment, *Journal of Finance* 55, 35–80.

Rényi, A., 1953, On the theory of order statistics, *Acta Mathematica Hungarica* 4, 191–231.

Sahlman, W., 1990, The Structure and Governance of Venture Capital Organizations, *Journal of Financial Economics* 27, 473–521.

Scharfstein, D., and S. Mullainathan, 2002, Do Firm Boundaries Matter?, *American Economic Review* 91, 195–199.

Scharfstein, D., and J. Stein, 2000, The Dark Side of Internal Capital Markets: Divisional Rent-Seeking and Inefficient Investment, *Journal of Finance* 55, 2537–2564.

Sorensen, M., 2007, How Smart is Smart Money? An Empirical Two-Sided Matching Model of Venture Capital, *Journal of Finance* 62, 2725–2762.

Stein, J., 1997, Internal Capital Markets and the Competition for Corporate Resources, *Journal of Finance* 52, 111–133.

Topkis, D. M., 1998, *Supermodularity and Complementarity*. Princeton University Press Princeton, NJ.

Wooldridge, J. M., 2002, *Econometric Analysis of Cross Section and Panel Data*. MIT Press Cambridge, MA.

A Order Statistics and G

A.1 A Representation of Order Statistics

We will repeatedly use the Rényi representation for order statistics given in Rényi (1953). However, we will reverse the ordering, so that $\zeta_{n,j}$ is the jth *highest* draw from a sample of n i.i.d. variables, i.e. $\zeta_{n,1} = \max_{1 \le i \le n} \zeta_i$. If the ζ_i have a standard exponential distribution, then,

$$\zeta_{n,j} \stackrel{d}{=} \sum_{i=j}^{n} \frac{e_i}{i}$$

where the e_i are independent standard exponential random variables.

If F is a general distribution function, we follow the method of Rényi (1953) and look at a sample of n i.i.d. variables with distribution function (cdf) $F(\xi)$. We set[26] $\zeta_i = -\ln(1 - F(\xi_i))$ and observe that the ζ_i are exponentially distributed and independent. Since the transformation is strictly increasing, the ordering of the sample is maintained. Then we can write

$$\xi_{n,j} \stackrel{d}{=} F^{-1}\left\{1 - \exp\left[-\left(\sum_{i=j}^{n} \frac{e_i}{i}\right)\right]\right\} \tag{13}$$

A.2 Proof of Proposition 1

From the definition of $G(n, m)$ (3), $G(n, m) - G(n, m-1) = \mathrm{E}[\Delta_{n,m}]$. This expectation is positive because $F(\Delta = 0) = 0$ and declining because $\Delta_{n,m}$ is defined to be larger than $\Delta_{n,m+1}$.

Similarly, $G(n+1, m) - G(n, m) = \sum_{j=1}^{m} (\mathrm{E}[\Delta_{n+1,j}] - \mathrm{E}[\Delta_{n,j}])$. This sum is positive because (13) shows that $\Delta_{n+1,j}$ first order stochastically dominates $\Delta_{n,j}$.

Finally, $[(G(n+1, m) - G(n+1, m-1)) - (G(n, m) - G(n, m-1))] = \mathrm{E}[\Delta_{n+1,m}] - \mathrm{E}[\Delta_{n,m}]$. This expectation is positive because (13) shows that $\Delta_{n+1,m}$ first order stochastically dominates $\Delta_{n,m}$.

A.3 Proof of Proposition 2

Because G is shown to be concave in m (Proposition 1) and assumed to be concave in n (Assumption 1), it is also the case that $\pi(\phi, M)$ (Equation 5) is concave in ϕ and M. In addition, both ϕ and M are bounded. Thus, the portfolio manager's problem has a unique solution.

Maximizing (5) is equivalent to maximizing

$$[M\phi\eta + (\alpha + (1-\alpha)\mu)\, G((1-\lambda\phi)N, M)] - M\frac{\theta}{\psi}. \tag{14}$$

Condition (7), Assumption 1, and the results of Proposition 1, imply that (14) is super-modular in $(M, -\phi, \psi, -\theta)$, $(M, -\phi, \alpha, \mu)$, $(\phi, -\lambda)$, and $(M, N, -\lambda)$. Topkis's theorem (Topkis (1998)) then proves the comparative statics.[27]

[26] We use a slightly different transformation here then Rényi here (equation 1.10 in Rényi (1953)). For the equivalence, observe that $F(\xi_i)$ and $1 - F(\xi_i)$ are both uniformly distributed. The remaining differences follow from our rank ordering from highest to lowest, rather than the reverse.

[27] Super-modularity of $f(x, y)$ in (x, y) means that the return to increasing x goes up with y. Intuitively, if the gains from x go up with y, then an optimizing agent should do more x when y is plentiful.

As an example, assume f is continuous. Then super-modularity is equivalent to stating $f_{xy}(x, y) > 0$. To see that

A.4 Distributional Examples

In all three examples, β is a measure of the scale of the distribution. When β is larger, the distribution has its mass pushed into the tail. Thus, a large β means that the best project will likely be drawn from further out in the now larger tail.

A.4.1 The Exponential Distribution

Assume that the Δ_i are distributed exponentially: $F(\Delta) = 1 - e^{-\frac{1}{\beta}\Delta}$, with scale parameter β. Then the Rényi representation theorem on order statistics (13) implies that $\mathrm{E}[\Delta_{n,m}] = \beta \sum_{k=m}^{n} \frac{1}{k}$, and so

$$G(n,m) = \sum_{j=1}^{m} \mathrm{E}[\Delta_{n,j}] = \beta \sum_{j=1}^{m} \sum_{k=j}^{n} \frac{1}{k}$$

Then we have

$$\begin{aligned}(G(n,m) - G(n,m-1)) &= \beta \sum_{k=m}^{n} \frac{1}{k} \\ (G(n+1,m) - G(n,m)) &= \beta \frac{m}{n+1} \\ [(G(n+1,m) - G(n+1,m-1)) - (G(n,m) - G(n,m-1))] &= \beta \frac{1}{n+1}\end{aligned}$$

Thus G is concave in n as assumed, and if $\beta > \frac{\eta}{(\alpha+(1-\alpha)\mu)\lambda}$, then the required condition on G in Proposition 2 is met.

A.4.2 The Uniform Distribution

Assume that the Δ_i are distributed uniformly: $F(\Delta) = \frac{\Delta}{\beta}$ for $\Delta \in [0, \beta]$. Then the Rényi representation theorem on order statistics (13) implies that $\mathrm{E}[\Delta_{n,m}] = \beta \frac{n-m+1}{n+1}$, and so

$$G(n,m) = \sum_{j=1}^{m} \mathrm{E}[\Delta_{n,j}] = \beta \left(m - \frac{m(m+1)}{2(n+1)} \right)$$

Then we have

$$\begin{aligned}(G(n,m) - G(n,m-1)) &= \beta \frac{n-m+1}{n+1} \\ (G(n+1,m) - G(n,m)) &= \beta \frac{m(m+1)}{2(n+1)(n+2)} \\ [(G(n+1,m) - G(n+1,m-1)) - (G(n,m) - G(n,m-1))] &= \beta \frac{m}{(n+1)(n+2)}\end{aligned}$$

Thus G is concave in n as assumed, and if $\beta > \frac{\eta}{(\alpha+(1-\alpha)\mu)\lambda} (N+2)$, then the required condition on G in Proposition 2 is met.

---

this implies that optimal choice of x is increasing in y, we examine the first order condition on x: $f_x(x^*(y), y) = 0$. Using the implicit function theorem, $\frac{\partial}{\partial y} x^*(y) = -\frac{f_{xy}}{f_{xx}}$. Since the denominator must be negative for an optimum to exist, $f_{xy}(x,y) > 0$ implies $x^*(y)$ is increasing in y. Topkis's theorem proves this result when f is defined only on a (possibly discrete) set.

A.4.3 The Power Law Distribution

Assume that the Δ_i are distributed according to a power law: $F(\Delta) = 1 - x^{-\frac{1}{\beta}\Delta}$ for $\beta < 1$, defined on $\Delta \in [1, \infty)$. (If $\beta \geq 1$, then the required expectations fail to exist because the distribution has no mean). Then $\ln(\Delta)$ has an exponential distribution with scale parameter β. The Rényi representation theorem on order statistics (13) implies that

$$\mathrm{E}\left[\Delta_{n,m}\right] = \mathrm{E}\left[\exp\left(\ln\left(\Delta_{n,m}\right)\right)\right] = \mathrm{E}\left[\exp\left(\beta \sum_{k=m}^{n} \frac{e_k}{k}\right)\right] = \prod_{k=m}^{n} \mathrm{E}\left[\exp\left(\beta \frac{e_k}{k}\right)\right] = \prod_{k=m}^{n} \frac{k}{k-\beta}$$

where the e_k are i.i.d. standard exponential variables. Then,

$$G(n,m) = \sum_{j=1}^{m} \mathrm{E}\left[\Delta_{n,j}\right] = \sum_{j=1}^{m} \prod_{k=j}^{n} \frac{k}{k-\beta}. \tag{15}$$

Then we have

$$\begin{aligned} (G(n,m) - G(n,m-1)) &= \prod_{k=m}^{n} \frac{k}{k-\beta} \\ (G(n+1,m) - G(n,m)) &= \left(\frac{\beta}{n+1-\beta}\right) G(n,m) \\ \left[(G(n+1,m) - G(n+1,m-1)) - (G(n,m) - G(n,m-1))\right] &= \left(\frac{\beta}{n+1-\beta}\right) \prod_{k=m}^{n} \frac{k}{k-\beta} \end{aligned}$$

Thus G is concave in n as assumed. The cross difference achieves its minimum at $m = n$, so the required condition on G for Proposition 2 is met if $\left(\frac{\beta}{n+1-\beta}\right)\left(\frac{n}{n-\beta}\right) > \frac{\eta}{(\alpha+(1-\alpha)\mu)\lambda(n+1)}$. Since the left hand side is increasing in β from 0 to ∞, we can define unique β^* as the value of β for which the condition is met with equality. Then, if $\beta > \beta^*$ and $\beta^* < 1$, the required cross difference condition on G is met.

Table I. Descriptive Statistics

The sample consists of 1820 independent venture capital funds headquartered in the U.S. that were started between 1980 and 1999 (the "vintage years") and make at least five investments over the course of their lives. Fund portfolio size (#) is the number of unique portfolio companies the fund invested in over the course of its life. Fund assets ($) is the amount of committed capital reported in the Venture Economics database. Sequence number denotes whether a fund is the first, second and so forth fund raised by a particular VC management firm. A fund is defined as investing primarily in seed or early stage deals if the largest fraction of the fund's investments were invested in at the seed or early stage. A fund is defined as primarily investing in the greater San Francisco (SF) Bay Area or Boston (BOS) /Route 128 area if the largest fraction of the fund's investments were located either in the San Francisco-Oakland-San Jose CMSA or in the Boston-Worcester-Lawrence CMSA. The four measures for the investment experience of a sample fund's parent (management) firm are based on the parent's investment activities measured between the parent's creation and the fund's vintage year. By definition, the experience measures are zero for first-time funds. The VC inflows variable is the aggregate amount of capital raised by other VC funds in the sample fund's vintage year. Specialization measures are derived using the investments made by the sample fund over its lifetime. Industry HHI (#) is the Herfindahl-Hirschman Index of the fund's investments across industries, using the number of unique portfolio companies invested in by the fund in each Venture Economics industry category. Venture Economics uses six industries: biotechnology, communications and media, computer related, medical/health/life science, semiconductors/other electronics, and non-high-technology. Industry HHI ($) is the Herfindahl-Hirschman Index of the fund's investments across industries, using the total dollar values invested by the fund in each industry. Geography HHI is the Herfindahl-Hirschman Index of the fund's investments across Metropolitan Statistical Areas (MSAs), using the number of unique portfolio companies invested in by the fund in each of the 287 US MSAs represented in the dataset.

	No.	Mean	Std. dev.	Min	Median	Max
Fund characteristics						
fund portfolio size (# companies)	1820	23.03	19.31	6	17	212
fund committed capital ($m)	1789	87.33	195.08	0.1	36	5000
sequence number	1562	3.54	3.78	1	2	31
first fund (fraction, %)	1820	29.8				
primarily seed or early stage (fraction, %)	1820	13.3				
primarily invest in SF or BOS (fraction, %)	1814	59.5				
Fund specialization						
industry HHI (# companies)	1820	0.40	0.15	0.18	0.36	1
industry HHI ($ value)	1819	0.44	0.17	0.17	0.4	1
geography HHI (# companies)	1814	0.28	0.18	0.04	0.22	1
Fund parent's experience (as of vintage year)						
days since parent's first investment	1820	2195.59	2380.06	0	1279	9130
no. of rounds parent has participated in so far	1820	108.32	227.14	0	19	2292
aggregate amount parent has invested so far ($m)	1820	101.16	306.41	0	14.84	6563.61
no. of portfolio companies parent has invested in so far	1820	42.05	70.91	0	13	601
Money chasing deals						
VC inflows in fund's vintage year ($bn)	1820	23.38	27.87	2.29	75.13	84.63

Table II. Correlations

The sample consists of 1820 independent venture capital funds headquartered in the U.S. that were started between 1980 and 1999 (the "vintage years") and make at least five investments over the course of their lives. Fund portfolio size (#) is the number of unique portfolio companies the fund invested in over the course of its life. Fund assets ($) is the amount of committed capital reported in the Venture Economics database. Sequence number denotes whether a fund is the first, second and so forth fund raised by a particular VC management firm. A fund is defined as investing primarily in seed or early stage deals if the largest fraction of the fund's investments were invested in at the seed or early stage. The four measures for the investment experience of a sample fund's parent (management) firm are based on the parent's investment activities measured between the parent's creation and the fund's vintage year. By definition, the experience measures are zero for first-time funds. The VC inflows variable is the aggregate amount of capital raised by other VC funds in the sample fund's vintage year. Specialization measures are derived using the investments made by the sample fund over its lifetime. Industry HHI (#) is the Herfindahl-Hirschman Index of the fund's investments across industries, using the number of unique portfolio companies invested in by the fund in each Venture Economics industry category. Venture Economics uses six industries: biotechnology, communications and media, computer related, medical/health/life science, semiconductors/other electronics, and non-high-technology. Industry HHI ($) is the Herfindahl-Hirschman Index of the fund's investments across industries, using the total dollar values invested by the fund in each industry. Geography HHI is the Herfindahl-Hirschman Index of the fund's investments across Metropolitan Statistical Areas (MSAs), using the number of unique portfolio companies invested in by the fund in each of the 287 US MSAs represented in the dataset. The table presents pair-wise correlations between variables of interest and fund portfolio size and specialization measures. We use ***, **, and * to denote significance at the 1%, 5%, and 10% level (two-sided), respectively.

	fund portfolio size	industry HHI (#)	industry HHI ($)	geography HHI (#)
Fund characteristics				
fund portfolio size (# companies)	1.00	-0.26***	-0.29***	-0.29***
fund assets ($m)	0.31***	0.02	-0.03	-0.10***
sequence number	0.10***	-0.09***	-0.06***	-0.10***
first fund (fraction, %)	-0.01	0.01***	0.07***	0.09***
primarily seed or early stage (fraction, %)	-0.07***	0.07***	0.07***	0.20***
Fund specialization				
industry HHI (# companies)		1.00	0.81***	0.18***
industry HHI ($ value)			1.00	0.16***
geography HHI (# companies)				1.00
Fund parent's experience (as of vintage year)				
days since parent's first investment	0.18***	-0.09***	-0.09***	-0.11***
no. of rounds parent has participated in so far	0.25***	-0.10***	-0.08***	-0.11***
aggregate amount parent has invested so far ($m)	0.25***	-0.03***	-0.05***	-0.11***
no. of portfolio companies parent has invested in so far	0.26***	-0.10***	-0.09***	-0.12***
Money chasing deals				
VC inflows in fund's vintage year ($bn)	-0.13***	0.29***	0.22***	0.16***

Table III. Univariate Sorts

The sample consists of 1820 independent venture capital funds headquartered in the U.S. that were started between 1980 and 1999 (the "vintage years") and make at least five investments over the course of their lives. Fund portfolio size (#) is the number of unique portfolio companies the fund invested in over the course of its life. Specialization measures are derived using the investments made by the sample fund over its lifetime. Industry HHI (#) is the Herfindahl-Hirschman Index of the fund's investments across industries, using the number of unique portfolio companies invested in by the fund in each Venture Economics industry category. Venture Economics uses six industries: biotechnology, communications and media, computer related, medical/health/life science, semiconductors/other electronics, and non-high-technology. Industry HHI ($) is the Herfindahl-Hirschman Index of the fund's investments across industries, using the total dollar values invested by the fund in each industry. Geography HHI is the Herfindahl-Hirschman Index of the fund's investments across Metropolitan Statistical Areas (MSAs), using the number of unique portfolio companies invested in by the fund in each of the 287 US MSAs represented in the dataset. Panel A presents univariate sorts of specialization by quartile of fund portfolio size and of fund portfolio size by quartile of specialization. Panel B presents univariate sorts of specialization by quartile of fund portfolio size and of fund portfolio size by quartile of specialization. Panel C presents univariate sorts of specialization by quartile of fund portfolio size and of fund portfolio size by quartile of specialization. We use ***, **, and * to denote significance at the 1%, 5%, and 10% level (two-sided), respectively.

Panel A. Portfolio size by quartile of fund specialization

	Q1: Least Specialized	Q2	Q3	Q4: Most Specialized	Q1-Q4
industry HHI (# companies)	29.88	25.35	19.61	16.91	12.96***
industry HHI ($ value)	31.88	23.53	19.86	16.01	15.87***
geography HHI (# companies)	34.07	21.20	19.19	16.91	17.15***

Panel B. Fund specialization by quartile of fund portfolio size

	Q1: Smallest Portfolio	Q2	Q3	Q4: Largest Portfolio	Q1-Q4
industry HHI (# companies)	0.45	0.41	0.39	0.32	0.12***
industry HHI ($ value)	0.51	0.45	0.42	0.35	0.16***
geography HHI (# companies)	0.36	0.29	0.25	0.20	0.16***

Panel C. Portfolio size by quartile of fund experience

	Q1: Least Experience	Q2	Q3	Q4: Most Experience	Q1-Q4
fund portfolio size (# companies)	17.66	21.35	25.64	27.50	-9.83***

Panel D. Fund specialization by quartile of fund experience

	Q1: Least Experience	Q2	Q3	Q4: Most Experience	Q1-Q4
industry HHI (# companies)	0.42	0.41	0.37	0.38	0.03***
industry HHI ($ value)	0.46	0.46	0.41	0.41	0.04***
geography HHI (# companies)	0.31	0.30	0.25	0.25	0.05***

Table III. Univariate Sorts (Continued).

Panel E. Portfolio size by primary stage of investment

	Seed/Early	Expansion/Late	Difference
fund portfolio size (# companies)	19.39	23.63	4.26***

Panel F. Fund specialization by primary stage of investment

	Seed/Early	Expansion/Late	Difference
industry HHI (# companies)	0.42	0.39	-0.12***
geography HHI (# companies)	0.37	0.26	-0.08***

Panel G. Portfolio size by quartile of $ inflows into VC

	Q1: Low Inflows	Q2	Q3	Q4: High Inflows	Q1-Q4
fund portfolio size (# companies)	26.35	25.31	20.80	19.11	7.25***

Panel H. Fund specialization by quartile of $ inflows into VC

	Q1: Low Inflows	Q2	Q3	Q4: High Inflows	Q1-Q4
industry HHI (# companies)	0.36	0.35	0.42	0.47	-0.12***
industry HHI ($ value)	0.40	0.39	0.46	0.50	-0.10***
geography HHI (# companies)	0.25	0.26	0.28	0.32	-0.08***

Table IV. Industry Specialization

The sample consists of 1820 independent venture capital funds headquartered in the U.S. that were started between 1980 and 1999 and invested in at least five portfolio companies. The dependent variable is Industry HHI (#), the Herfindahl-Hirschman Index of the fund's investments across industries, using the number of unique portfolio companies invested in by the fund in each Venture Economics industry category. Venture Economics uses six industries: biotechnology, communications and media, computer related, medical/health/life science, semiconductors/other electronics, and non-high-technology. These dependent variables have support on [0,1] and positive mass at 1. To avoid the resulting well-known biases of OLS in this situation, we estimate fractional logit models using quasi-MLE; see Papke and Wooldridge (1996). This involves modeling the conditional mean $E(y|x)=\exp(x\beta)/(1+\exp(x\beta))$. Independent variables are as described in Table I. Year controls are included but not reported. Intercepts are not shown. Heteroskedasticity-consistent standard errors (clustered on parent VC firm) are shown in parentheses. We use ***, **, and * to denote significance at the 1%, 5%, and 10% level (two-sided), respectively.

	Industry HHI (# companies)				
	1	2	3	4	5
ln fund portfolio size	-0.278***			-0.205***	-0.215***
	0.023			*0.024*	*0.024*
ln days since parent's first investment		-0.093**		-0.064***	
		0.013		*0.013*	
ln fund sequence number			-0.014***		-0.107***
			0.022		*0.020*
=1 if primarily invests in seed or early stage		0.133***	0.092***	0.105**	0.070
		0.050	*0.049*	*0.049*	*0.048*
ln VC inflows in funding year		0.164***	0.182***	0.134***	0.149***
		0.023	*0.025*	*0.022*	*0.022*
No. of observations	1,820	1,678	1,561	1,678	1,561

Table V. Geography Specialization

The sample consists of 1820 independent venture capital funds headquartered in the U.S. that were started between 1980 and 1999 and invested in at least five portfolio companies. The dependent variable is Geography HHI, the Herfindahl-Hirschman Index of the fund's investments across Metropolitan Statistical Areas (MSAs), using the number of unique portfolio companies invested in by the fund in each of the 287 US MSAs represented in the dataset. These dependent variables have support on [0,1] and positive mass at 1. To avoid the resulting well-known biases of OLS in this situation, we estimate fractional logit models using quasi-MLE; see Papke and Wooldridge (1996). This involves modeling the conditional mean $E(y|x)=\exp(x\beta)/(1+\exp(x\beta))$. Independent variables are as described in Table I. Year controls are included but not reported. Intercepts are not shown. Heteroskedasticity-consistent standard errors (clustered on parent VC firm) are shown in parentheses. We use ***, **, and * to denote significance at the 1%, 5%, and 10% level (two-sided), respectively.

	Geography HHI (# companies)				
	1	2	3	4	5
ln fund portfolio size	-0.492***			-0.421***	-0.453***
	0.039			*0.038*	*0.042*
ln days since parent's first investment		-0.101***		-0.044**	
		0.019		*0.018*	
ln fund sequence number			-0.171***		-0.102**
			0.036		*0.033*
=1 if primarily invests in seed or early stage		0.476***	0.450***	0.427***	0.412***
		0.071	*0.077*	*0.049*	*0.073*
ln VC inflows in funding year		0.177**	0.148***	0.119***	0.080***
		0.034	*0.039*	*0.035*	*0.041*
No. of observations	1,814	1,675	1,556	1,675	1,556

Table VI. Portfolio Size

The sample consists of 1820 independent venture capital funds headquartered in the U.S. that were started between 1980 and 1999 and invested in at least five portfolio companies. The dependent variable is the fund's portfolio size. Panel A presents OLS models where the dependent variable is the natural logarithm of the fund's portfolio size (number of unique firms). Panel B presents poisson models where the dependent variable is the fund's portfolio size (count of unique firms). Independent variables are as described in Table I. Year controls are included but not reported. Intercepts are not shown. Heteroskedasticity-consistent standard errors (clustered on parent VC firm) are shown in parentheses. We use ***, **, and * to denote significance at the 1%, 5%, and 10% level (two-sided), respectively.

Panel A. OLS models	*ln* portfolio size					
	1	2	3	4	5	6
industry HHI (# companies)	-1.296***			-1.012***	-1.014***	
	0.125			*0.136*	*0.134*	
geography HHI (# companies)		-1.309***				-1.127***
		0.100				*0.104*
ln days since parent's first investment			0.141**	0.119***		0.119***
			0.014	*0.013*		*0.014*
ln fund sequence number					0.121***	
					0.031	
=1 if primarily invests in seed or early stage			-0.141***	-0.109**	-0.088*	-0.028
			0.047	*0.045*	*0.046*	*0.045*
ln VC inflows in funding year			-0.148***	-0.108***	-0.113**	-0.109***
			0.036	*0.035*	*0.045*	*0.036*
R^2	0.08	0.12	0.11	0.16	0.10	0.19
No. of observations	1,820	1,814	1,678	1,678	1,561	1,675

Table VI. Portfolio Size (Continued).

Panel B. Poisson models	portfolio size (count)					
	1	2	3	4	5	6
industry HHI (# companies)	-1.632***			-1.217***	-1.271***	
	0.165			*0.174*	*0.179*	
geography HHI (# companies)		-1.629***				-1.332***
		0.167				*0.169*
ln days since parent's first investment			0.161**	0.135***		0.137***
			0.020	*0.017*		*0.017*
ln fund sequence number					0.122***	
					0.033	
=1 if primarily invests in seed or early stage			-0.189***	-0.160***	-0.134**	-0.073
			0.056	*0.054*	*0.055*	*0.053*
ln VC inflows in funding year			-0.172***	-0.126***	-0.127***	-0.134***
			0.035	*0.034*	*0.042*	*0.035*
No. of observations	1,820	1,678	1,678	1,678	1,561	1,675

Table VII. High Deal Flow / Low Cost to Specialization

The sample consists of 1820 independent venture capital funds headquartered in the U.S. that were started between 1980 and 1999 and invested in at least five portfolio companies. The dependent variable in columns (1) and (2) is the natural logarithm of the fund's portfolio size (number of unique firms), and the models estimated are OLS. The dependent variable in columns (3) and (4) is the fund's portfolio size (count of unique firms) and the models estimated are poisson regressions. The dependent variable in columns (5) and (6) is the geographic specialization of the fund, measured by the Herfindahl-Hirschman Index of the fund's investments across Metropolitan Statistical Areas (MSAs), using the number of unique portfolio companies invested in by the fund in each of the 287 US MSAs represented in the dataset. These dependent variables have support on [0,1] and positive mass at 1. To avoid the resulting well-known biases of OLS in this situation, we estimate fractional logit models using quasi-MLE; see Papke and Wooldridge (1996). This involves modeling the conditional mean $E(y|x)=\exp(x\beta)/(1+\exp(x\beta))$. Independent variables are as described in Table I. Year controls are included but not reported. Intercepts are not shown. Heteroskedasticity-consistent standard errors (clustered on parent VC firm) are shown in parentheses. We use ***, **, and * to denote significance at the 1%, 5%, and 10% level (two-sided), respectively.

	ln portfolio size (*OLS*)		portfolio size (*Poisson*)		geography HHI (*fractional logit*)	
	1	2	3	4	5	6
geography HHI (# companies)	-1.165***	-1.111***	-1.141***	-1.134***		
	0.104	*0.101*	*0.171*	*0.166*		
ln fund portfolio size					-0.449***	-0.483***
					0.038	*0.042*
ln days since parent's first investment	0.099***		0.113***		-0.053***	
	0.014		*0.017*		*0.018*	
ln fund sequence number		0.083***		0.082***		-0.125***
		0.030		*0.032*		*0.032*
=1 if primarily invests in seed or early stage	-0.011	-0.010	-0.053	-0.030	0.434***	0.423***
	0.014	*0.043*	*0.051*	*0.052*	*0.068*	*0.071*
ln VC inflows in funding year	-0.090***	-0.105**	-0.109***	-0.120***	0.127***	0.088***
	0.035	*0.045*	*0.035*	*0.042*	*0.035*	*0.041*
=1 if primarily invests in SF or BOS	0.282***	0.285***	0.342***	0.352***	0.199***	0.204***
	0.035	*0.036*	*0.039*	*0.041*	*0.051*	*0.054*
R^2 (OLS models only)	0.23	0.25				
No. of observations	1,675	1,556	1,675	1,556	1,675	1,556