Informational Hold-up and Performance Persistence in Venture Capital[*]

Yael V. Hochberg
Northwestern University

Alexander Ljungqvist
New York University and CEPR

Annette Vissing-Jørgensen
Northwestern University and NBER

November 10, 2008

Abstract

We propose and test a theory of learning and informational hold-up in the venture capital market. The model predicts that higher returns on the current fund increase the probability that a VC will raise a follow-on fund, the size of the follow-on fund, and the performance fee investors are charged in the follow-on fund. If learning is asymmetric, such that incumbent investors learn more about fund manager skill than potential new investors, the model also predicts persistence in returns, poor performance among first-time funds, persistence in investors from fund to fund, and over-subscription in follow-on funds raised by successful fund managers. Our empirical evidence is consistent with these predictions. The model provides a unified framework for understanding a series of empirical facts about the venture capital industry.

Performance in venture capital (VC) funds appears highly persistent across a sequence of funds managed by the same manager (Kaplan and Schoar (2005)). This contrasts with evidence for mutual funds (Malkiel (1995)), and raises an interesting question: Why do successful VCs not raise their fees, effectively auctioning off the stakes in their follow-on funds to the highest bidder? Alternatively, why do successful funds not grow to the point where their return on fund capital equals investors' outside option, thus increasing fund managers' dollar fees?

We propose and test a model of learning and informational hold-up in the VC market which provides a rationale for VC performance persistence and generates a rich set of empirical predictions concerning fund returns and fund-raising patterns. The model formalizes the interaction between fund managers, known in the industry as general partners (GPs), and the investors in their funds, known as limited partners (LPs). It exploits logic similar to that used in the relationship-banking literature (Sharpe (1990), Rajan (1992), von Thadden (2004)): LPs who invest in a GP's fund learn more about the GP's skill than do other investors. In particular, they can distinguish between skill and luck. This asymmetric evolution of information enables 'incumbent' LPs to hold up the GP when he next raises a fund, because other potential investors will interpret failure to reinvest by incumbent LPs as a negative signal about the GP's skill.

Our informational assumptions are consistent with the importance the literature attaches to 'soft' information in the VC industry. Lerner, Schoar, and Wongsunwai (2007) note that "Reinvestment decisions by LPs are particularly important in the private equity industry, where information about the quality of different private equity groups is more difficult to learn and often restricted to existing investors." VC fund investors typically require wide-ranging information rights which enable them to monitor a fund's performance, primarily in order to inform the reinvestment decision (Lerner and Schoar (2004)).

The existence of asymmetric information between incumbent LPs and potential new investors is the key difference between our model of VC performance and models of mutual fund performance, such as Berk and Green (2004). It is the reason why arguments in such models against performance persistence in the mutual fund industry do not apply to the VC industry.

In a setting with learning about GP skill, we derive the optimal fund size and division of surplus between GPs and LPs. We first derive a set of testable predictions that result from a setting with symmetric learning by incumbent LPs and potential new LPs. We then introduce asymmetric learning, assuming that incumbent LPs learn more about the GP's skill. Asymmetric learning generates informational hold-up of high-skill GPs which in turn generates persistence in the net-of-fees returns LPs earn, even when both fund size and the division of surplus are derived endogenously. While the predictions of the symmetric learning model also hold with asymmetric learning, we derive additional testable predictions that hold only under asymmetric learning.

We model GPs with a sequence of two funds, each lasting one period. Whether the second fund is raised depends on what is learned about GP skill during the first fund. GPs differ in investment skill. There is a large set of risk-neutral potential LPs so that at the beginning of the first period, the LP market is perfectly competitive. LPs are wealthy and only one LP is required to fund a given fund in our baseline setting. At the end of period one, the GP and the incumbent LP learn the GP's type. With symmetric learning, outside LPs also learn the GP's type perfectly, while with asymmetric learning, outside LPs can only update their beliefs about GP skill based on observing the return of the GP's first fund.

Under symmetric learning, the LP market is perfectly competitive in both periods, so the GP receives the entire NPV in each fund. Similar to Berk and Green (2004), the GP chooses the NPV-maximizing fund size and share of the surplus (the carried interest, or "carry") such that the

2

LP earns his outside option. This yields the following predictions. The probability of raising a follow-on fund increases in the return to LPs in the GP's first fund. GPs with higher LP returns in their first fund raise larger follow-on funds and charge a higher carry. Fund size and carry are more variable in follow-on funds than in first funds.

Under asymmetric learning, the LP market is perfectly competitive at $t=0$ but not at $t=1$. The incumbent LP has an informational advantage when a GP raises a follow-on fund, allowing him to extract part of the follow-on fund's NPV.[1] This yields additional predictions. Most importantly, there will be persistence in net-of-fee returns to LPs. Intuitively, LPs earn higher expected returns in follow-on funds run by high-skill GPs (since they extract part of the follow-on fund NPV, which is higher for high-skill GPs) and high-skill GPs tend to have delivered high returns to LPs on their first fund. Second, the average return to LPs is lower in first-time funds than in follow-on funds. GPs receive funding on "too good" terms in the first fund, but on average pay this back to their LPs in follow-on funds. Finally, conditional on a follow-on fund being raised, LPs who invested in a GP's first fund should also invest in the GP's follow-on fund. That is, there should be persistence in the LP composition across funds, because LPs earn a return in excess of their opportunity cost of capital in follow-on funds.

If we relax the assumption that incumbent LPs have sufficient capital to fully fund follow-on funds, additional LPs are needed. Because all LPs in a fund earn the same return, LPs who get follow-on fund allocations earn a return in excess of their opportunity cost of capital. Our model thus provides a rational explanation for oversubscription in follow-on funds, especially in those raised by GPs with high first-time fund returns.

---

[1]How much they extract depends on the GP's and the incumbent LP's relative bargaining power when raising the follow-on fund. Without loss of generality, we assume a sequential-offer game. This results in symmetric Nash bargaining, implying that LPs and GPs share equally in the NPV.

Unlike in standard models of informational hold-up such as Sharpe (1990), asymmetric learning is efficient in the VC setting. VC contracts specify both an investment level (fund size) and the division of the fund's surplus, and investment (fund size) is NPV-maximizing in both first-time funds and follow-on funds. Moreover, it is possible that GPs strictly prefer asymmetric to symmetric learning, because under certain conditions, first-time funds can only be funded under asymmetric learning.[2] Such a preference is consistent with the fact that GPs are willing to provide their LPs with considerable amounts of soft information about fund strategies and performance which cannot credibly be communicated to potential new LPs.

Using a large sample of U.S. VC funds raised between 1980 and 2006, we find support for the predictions of our model, which suggests that asymmetric learning and informational hold-up play a role in the VC industry. While parts of our empirical evidence overlap with published work, our evidence relating GP carry to fund returns and our evidence relating oversubscription to fund sequence and fund returns is to our knowledge new. Furthermore, we believe we are the first to avoid look-ahead bias in conditioning only on performance measures known ex ante.[3]

In addition to the relationship-banking literature, our paper relates to the literature on VC performance and the relationship between LPs and GPs. Kaplan and Schoar (2005) find evidence of performance persistence in VC funds raised by the same GP and document a positive and concave relation between performance and future fund-raising. Jones and Rhodes-Kropf (2003) provide empirical evidence in support of the hypothesis that VCs need to be compensated through higher expected returns for bearing idiosyncratic risk. Ljungqvist and Richardson (2003) analyze the

---

[2]Intuitively, if the average NPV across GP types is negative, first-time funds cannot be funded with symmetric learning since LPs cannot earn a return equal to or above their outside option. Asymmetric information (the provision of soft information) serves as a commitment device for GPs who turn out to have skill to give up part of the follow-on fund NPV to LPs. The resulting rents earned by LPs in follow-on funds make them willing to fund first-time funds even if they do not earn their opportunity cost of capital on first funds.

[3]In contemporaneous work, Phalippou (2008) also analyzes performance persistence using ex ante performance measures.

cash flow, return, and risk characteristics of private equity funds. Cochrane (2005), Korteweg and Sørensen (2008),Quigley and Woodward (2003), and Gottschalg and Phalippou (2007) estimate the risk and return of VC investments. Lerner, Schoar, and Wongsunwai (2007) find large heterogeneity in the returns that different classes of institutional investors earn in private equity and suggest that LPs vary in their level of sophistication.

Finally, in contemporaneous theoretical work, Glode and Green (2008) provide an alternative explanation for performance persistence. In their model, incumbent LPs learn about the profitability of a given GP's strategy. LPs can 'steal' the strategy by revealing it to another GP. This increases the outside option of LPs who have invested in successful funds and enables them to extract part of the follow-on fund surplus, generating performance persistence. We view our explanation for persistence as complementary to that of Glode and Green. In our model, which emphasizes the role of soft information, GPs *cannot* credibly reveal their type to outside LPs (and may optimally set up funds specifically to avoid such revelation). In Glode and Green, GPs *do not want* their type revealed for competitive reasons. Both elements are likely relevant in practice.

I.   A Model of Learning About GP Skill

A.   Setup

General partners and funds: At $t=0$, risk-neutral GPs raise funds of size I_1 lasting one period. GPs may raise a second fund of size I_2 at $t=1$, after the return of fund 1 is known. GPs differ in their investment skill, and this heterogeneity affects expected payoffs. For a GP of type i, fund $k = 1, 2$ returns a cash flow of $C_k^i = e^{A_k^i} \ln (1 + I_k)$ at $t = k$. The log function captures decreasing returns to scale. For a given GP, the cash flows of funds 1 and 2 are drawn independently but from

the same distribution, with $A_k^i \sim N\left(\mu^i - \frac{1}{2}\sigma^2, \sigma^2\right)$. There is a continuum of GP types characterized by μ^i. For simplicity, we assume that μ^i is distributed uniformly over the interval $\left[\mu^L, \mu^H\right]$.[4]

Limited partners: There is a large set of identical, risk-neutral investors. We assume each LP has sufficient wealth so that each fund only requires one LP.[5] We distinguish between the 'incumbent' LP who has invested in the GP's first fund, and 'outside' LPs who have not. The LP market is perfectly competitive at $t=0$. LPs can earn a (risk-adjusted) return of r outside the VC industry.

Learning about GP type: At $t=0$, no-one knows the GP's type. At $t=1$, the GP and the incumbent LP learn the GP's type perfectly, and under symmetric learning, so do outside LPs. Under asymmetric learning, outside LPs only observe the cash flows of the first fund, C_1.[6] Once μ^i is known, the NPV of fund 2, as of $t=1$, is

$$NPV_2\left(\mu^i\right) = \frac{E_1\left(C_2^i\right)}{1+r} - I_2 = \frac{e^{\mu^i}\ln\left(1+I_2\right)}{1+r} - I_2. \tag{1}$$

Define μ^* as the value of μ^i for which $NPV_2\left(\mu^i\right)$ at the optimal investment level equals zero. If the NPV is zero, the optimal investment I_2 is zero also, and thus $\mu^* = \ln\left(1+r\right)$. Assume that $\mu^L < \mu^* < \mu^H$.

Payoff functions: To characterize the division of a fund's surplus, we assume the following contract. The LP receives all cash flows up to the amount invested. Additional cash flows are

---

[4]The log-normal distribution of cash flows and the uniform distribution of GP types allow us to solve the model in closed form but do not drive our results. The more important choice is the functional form of the relation between cash flows and investment. To generate performance persistence, we need a functional form where C_2/I_2 is increasing in GP type even when I_2 is chosen optimally to reflect GP skill.

[5]For simplicity, we assume that GPs have no investable wealth. In practice, LPs typically contribute 99% of a fund's capital, with GPs providing the remainder.

[6]This corresponds to 'passive monitoring' in Tirole's (2005) terminology. We do not incorporate a role for active monitoring, thus implicitly assuming that the GP always exerts effort. While active monitoring could play a central role in the relationship between GPs and the entrepreneurs they back, we believe it to be less important in the relationship between GPs and LPs given the fairly high-powered incentives provided to the GPs via their carry. We discuss the nature of the GP-LP contract below.

divided according to a linear sharing rule, with the GP receiving a fraction f, which is his carry.[7]

For a fund of size I with carry f, the payoffs to the GP and the LP at the end of the fund's life are

$$X^{GP} = \max\left(0, f\left(C - I\right)\right) = \begin{cases} 0 \text{ when } C \leq I \\ f\left(C - I\right) \text{ when } C > I. \end{cases} \tag{2}$$

$$X^{LP} = C - \max\left(0, f\left(C - I\right)\right) = \begin{cases} C \text{ when } C \leq I \\ C - f\left(C - I\right) \text{ when } C > I. \end{cases} \tag{3}$$

B. Symmetric Learning

Under symmetric learning, the LP market is perfectly competitive at both $t{=}0$ and $t{=}1$ so that the GP receives the entire NPV of the fund. Thus, the GP chooses the NPV-maximizing investment level I_2 and sets the carry f_2 such that the LP earns the outside option.

Investment in follow-on funds: For follow-on funds, assuming symmetric learning, the GP's type is known and a GP of type μ^i picks I_2 to maximize $NPV_2|\mu = \mu^i$:

$$\max_{I_2} \frac{e^{\mu^i} \ln\left(1 + I_2\right)}{1 + r} - I_2 \iff I_2\left(\mu^i\right) = \frac{e^{\mu^i}}{1 + r} - 1. \tag{4}$$

Only funds with $\mu^i \geq \mu^*$ are raised.

Carry in follow-on funds: The carry is set such that the LP earns the outside option:

$$E\left(X_2^{LP}|\mu = \mu^i\right) = \left(1 + r\right) I_2\left(\mu^i\right) \iff$$

$$E\left(C_2|\mu = \mu^i\right) - f_2\left(\mu^i\right) E\left[\max\left(0, C_2 - I_2\left(\mu^i\right)\right)|\mu = \mu^i\right] = \left(1 + r\right) I_2\left(\mu^i\right) \iff$$

$$f_2\left(\mu^i\right) = \frac{e^{\mu_i} \ln\left(1 + I_2\left(\mu^i\right)\right) - \left(1 + r\right) I_2\left(\mu^i\right)}{g_2\left(\mu^i\right)} \tag{5}$$

where we denote $E\left[\max\left(0, C_2 - I_2\left(\mu^i\right)\right)|\mu = \mu^i\right]$ by $g_2\left(\mu^i\right)$ for notational simplicity and provide

---

[7]In practice, VC contracts also include a management fee, as a percent of fund size. For simplicity, we set this fee to zero. For details of VC fees, see Metrick and Yasuda (2007).

the closed-form expression for $g_2\left(\mu^i\right)$ in Appendix A.

LP return in follow-on funds: The LP's expected return for a GP of type μ^i is given by

$$1 + E\left(r_2|\mu = \mu^i\right) = \frac{E\left(X_2^{LP}|\mu = \mu^i\right)}{I_2\left(\mu^i\right)} = 1 + r. \tag{6}$$

With symmetric learning, expected LP returns are equalized across all GP types in follow-on funds.

Thus, there is no performance persistence in the cross-section of first-time and follow-on funds.

First-fund investment, carry, and expected returns are derived in the same way as for follow-on

funds except that the GP's type is unknown and so expectations are taken over GP types i.

Investment in first-time funds: Each GP picks I_1 to maximize $E_i\left(NPV_1(\mu_i)\right)$:

$$\max_{I_1} E_i\left[\frac{e^{\mu^i}\ln\left(1 + I_1\right)}{1 + r} - I_1\right] \iff I_1^{opt} = \frac{E_i\left(e^{\mu^i}\right)}{1 + r} - 1 \tag{7}$$

where, by the uniform distribution of GP types, $E_i\left(e^{\mu^i}\right) = \frac{1}{\mu^H - \mu^L}\left[e^{\mu^H} - e^{\mu^L}\right]$. If the average

NPV across GPs, at the optimal investment level, is negative (i.e., if $\frac{E_i\left(e^{\mu^i}\right)\ln\left(1 + I_1^{opt}\right)}{1 + r} - I_1^{opt} < 0$),

a GP cannot raise a first-time fund (nor any follow-on funds) when learning is symmetric.

Carry in first-time funds: The carry is again set such that the LP earns the outside option:

$$E_i\left(E\left(X_1^{LP}|\mu = \mu^i\right)\right) = (1 + r)\, I_1^{opt} \iff$$

$$E_i\left(E\left(C_1|\mu = \mu^i\right)\right) - f_1 E_i\left(E\left[\max\left(0, C_1 - I_1\left(\mu^i\right)\right)|\mu = \mu^i\right]\right) = (1 + r)\, I_1^{opt} \iff$$

$$f_1 = \frac{E_i\left(e^{\mu_i}\right)\ln\left(1 + I_1^{opt}\right) - (1 + r)\, I_1^{opt}}{E_i\left(g_1\left(\mu^i\right)\right)} \tag{8}$$

where we denote $E\left[\max\left(0, C_1 - I_1\left(\mu^i\right)\right)|\mu = \mu^i\right]$ by $g_1\left(\mu^i\right)$ for notational simplicity and provide

the closed-form expression for $g_1\left(\mu^i\right)$ in Appendix A.

LP return in first-time funds: The LP's expected return (given that GP type is not known at $t=0$) is given by

$$1 + E_i \left(E \left(r_1 | \mu = \mu^i \right) \right) = \frac{E_i \left(E \left(X_1^{LP} | \mu = \mu^i \right) \right)}{I_1^{opt}} = 1 + r. \tag{9}$$

B.1. Testable Implications under Symmetric Learning

Implication 1: Fund-raising. The probability that a GP raises a follow-on fund is increasing in the LP return of the GP's first-time fund: $P \left(\mu^i > \mu^* | r_1 \right)$ increases in r_1.

Implication 2: Evolution of fund-size. In the cross-section of GPs with follow-on funds, a high return to the LP in the first fund predicts a larger second fund: $E \left(I_2 | r_1 \right)$ increases in r_1.

Implication 3: Evolution of GP carry. In the cross-section of GPs with follow-on funds, a high first-fund return predicts a larger GP carry in the second fund: $E \left(f_2 | r_1 \right)$ increases in r_1.

Implication 4: Cross-fund standard deviation of fund size. The cross-fund standard deviation of fund sizes is higher among follow-on than among first funds: $SD_i \left(I_2 | \mu^i > \mu^* \right) > SD_i \left(I_1^{opt} \right)$.

Implication 5: Cross-fund standard deviation of GP carry. The cross-fund standard deviation of GP carry is higher among follow-on than among first funds: $SD_i \left(f_2 | \mu^i > \mu^* \right) > SD_i \left(f_1^{opt} \right)$.

We verify Implications 1, 2, and 3 numerically in Appendix B. Implications 4 and 5 follow immediately from the fact that the cross-fund standard deviation of both fund size and carry is zero in first-time funds and positive in follow-on funds.

C. Asymmetric Learning

Under asymmetric learning, the LP market is perfectly competitive at $t=0$ but not at $t=1$. When outside LPs do not learn the GP's type, incumbent LPs have an informational advantage in the follow-on fund. This allows incumbent LPs to extract part of the follow-on fund's NPV. How much

they extract depends on how GPs and LPs are assumed to bargain.

Investment and carry in follow-on funds: We assume the GP and the incumbent LP take turns making offers consisting of a proposed carry and fund size, (f_2, I_2). The GP goes first. In each round, if an offer is rejected, the NPV of the fund shrinks by a factor $1 - p$. This represents lost deal flow due to the delayed start of the fund.

If no agreement is reached, each party receives its outside option. For the incumbent LP, this equals r (i.e., NPV=0). The GP's outside option depends on what outside LPs are willing to offer if no agreement is reached with the incumbent LP. We assume that outside LPs cannot see (or cannot verify) the bids made prior to breakdown of bargaining with the incumbent LP. We furthermore assume that the incumbent LP can counter any offer an outside LP makes. The GP's outside option is then zero, because outside LPs face a winner's curse. Say, an outside LP observes an attractive return on the GP's first fund. The outside LP knows that if she offers an investment and carry based on assuming the GP's type μ^i exceeds μ^*, the incumbent LP will counter with an offer that is more attractive to the GP only when μ^i in fact exceeds μ^*. Thus, the outside LP never makes a positive NPV investment and rationally withdraws from the market.

We show in Appendix C that sequential bargaining results in the LP accepting the GP's first offer. For a given GP type μ^i, investment in follow-on funds $I_2\left(\mu^i\right)$ is as under symmetric learning. As $p \to 0$, the follow-on fund carry $f_2\left(\mu^i\right)$ is set such that the follow-on fund NPV is divided equally between the GP and the incumbent LP. The following proposition summarizes.

Proposition 1: *As $p \to 0$, the equilibrium outcome of the sequential bargaining game is immediate agreement with equal division of the fund's NPV between the GP and the incumbent LP.*

Proof: See Appendix C. In what follows, we focus on the case where $p \to 0$.

LP return in follow-on funds: The LP's expected return is

$$1 + E\left(r_2 | \mu = \mu^i\right) = \frac{E\left(X_2^{LP} | \mu = \mu^i\right)}{I_2\left(\mu^i\right)} = \frac{e^{\mu_i} \ln\left(1 + I_2\left(\mu^i\right)\right) - \frac{1}{2}\left[e^{\mu_i} \ln\left(1 + I_2\left(\mu^i\right)\right) - (1+r) I_2\left(\mu^i\right)\right]}{I_2\left(\mu^i\right)}.$$

(10)

This implies expected LP returns are no longer equalized across GP types in follow-on funds, unlike under symmetric learning. Furthermore, the average expected LP return in follow-on funds is

$$1 + E_i\left(E\left(r_2 | \mu = \mu_i\right) | \mu_i > \mu^*\right) = \frac{1}{\mu^H - \mu^*} \int_{\mu^*}^{\mu^H} \left(1 + E\left(r_2 | \mu = \mu^i\right)\right) d\mu^i.$$

(11)

Investment and carry in first-time funds: With asymmetric learning, the LP market remains perfectly competitive at $t=0$ (as no learning has taken place yet). The GP can therefore offer any LP a contract $\left(f_1, I_1^{opt}\right)$ where I_1^{opt} is the NPV-maximizing investment derived in the symmetric learning case (equation (7)) and where f_1 is such that the LP earns a fair return (i.e., a zero overall NPV) across the current fund and the follow–on fund that will be raised if $\mu^i \geq \mu^*$. We thus solve for f_1 as the value which sets the LP's overall NPV across the two funds to zero. Then f_1 solves

$$0 = \left[E_i\left(NPV_1\left(\mu^i\right)\right) - \frac{f_1 E_i\left(g_1\left(\mu^i\right)\right)}{1+r}\right] + \frac{1}{1+r} \frac{E_i\left(NPV_2\left(\mu^i\right)\right)}{2} \Longleftrightarrow$$

$$f_1 = \frac{E_i\left(e^{\mu_i}\right) \ln\left(1 + I_1^{opt}\right) - (1+r) I_1^{opt}}{E_i\left(g_1\left(\mu^i\right)\right)} + \frac{\frac{1}{2} E_i\left(NPV_2\left(\mu^i\right)\right)}{E_i\left(g_1\left(\mu^i\right)\right)}$$

(12)

with $NPV_2\left(\mu^i\right)$ calculated at the NPV-maximizing investment level. (The exact expression for f_1 can be found in Appendix D.) Since LPs earn positive NPV in follow-on funds, it is possible that first-time funds can be raised even if they have negative expected NPV, i.e., if $\frac{E_i\left(e^{\mu^i}\right) \ln\left(1 + I_1^{opt}\right)}{1+r} - I_1^{opt} < 0$. This will be the case if there is enough dispersion in skill to ensure that the expected NPV in follow-on funds is sufficient to compensate LPs for the expected loss on first funds.

11

LP return in first-time funds: The LP's expected return (given that GP type is unknown at $t=0$) becomes

$$1 + E_i \left(E \left(r_1 | \mu = \mu^i \right) \right) = \frac{E_i \left(E \left(X_1^{LP} | \mu = \mu^i \right) \right)}{I_1^{opt}} = \frac{E_i \left(e^{\mu^i} \right) \ln \left(1 + I_1^{opt} \right) - f_1 E_i \left(g_1 \left(\mu^i \right) \right)}{I_1^{opt}}. \quad (13)$$

C.1. Testable Implications under Asymmetric Learning

Implications 1-5 do not depend on how the follow-on fund NPV is divided and so hold regardless of whether learning is symmetric or not. In addition, we have the following testable implications:

Implication 6: Persistence. In the cross-section of GPs with follow-on funds, a high return to the LP in fund 1 predicts a high return to the LP in fund 2, i.e., $E \left(r_2 | r_1 \right)$ increases in r_1.

Implication 7: Performance of first-time versus follow-on funds. The average return to LPs is lower in first funds than in follow-on funds, i.e., $E_i \left(E \left(r_1 | \mu = \mu^i \right) \right) < E_i \left(r_2 | \mu^i \geq \mu^* \right).$

Implication 8: Persistence in LPs. Conditional on a follow-on fund being raised, LPs who invested in a GP's first fund should invest in that GP's follow-on fund.

Implication 9: Oversubscription in follow-on funds. Oversubscription is concentrated in follow-on funds and is more severe for follow-on funds with higher first-fund returns.

We verify Implication 6 numerically in Appendix B. Implication 7 follows from equations (12) and (13). The second term in (12) implies that f_1 is higher if learning is asymmetric. Since with symmetric learning, $E_i \left(E \left(r_1 | \mu = \mu^i \right) \right) = r$, it follows from (13) that with asymmetric learning, $E_i \left(E \left(r_1 | \mu = \mu^i \right) \right) < r$. Furthermore, since LPs earn half of the follow-on fund NPV, $r < E_i \left(r_2 | \mu^i \geq \mu^* \right)$ and thus $E_i \left(E \left(r_1 | \mu = \mu^i \right) \right) < E_i \left(r_2 | \mu^i \geq \mu^* \right)$. Implications 8 and 9 follow directly from the fact that LPs in follow-on funds earn half of the fund's NPV and thus earn a return in excess of their opportunity cost of capital.

An obvious question when faced with performance persistence is why outside LPs could not simply invest in all follow-on funds with a high realized r_1, thus expecting to earn a high value of $E(r_2|r_1)$. Our model makes it clear why this is not feasible. The winner's curse problem described above implies that outside LPs would only get to invest with those GPs for whom their offers implied negative NPV to investors. This implies that the "return-chasing" behavior emphasized by Berk and Green (2004) as the mechanism for elimination of performance persistence breaks down in the VC setting.

C.2. Extending the asymmetric learning model

Extension 1 – Adding further follow-on funds: How does persistence evolve over a sequence of more than two funds? In practice, most GPs raise a second fund well before the end of the first fund's ten-year life. At that point, incumbent LPs may be only slightly better informed than outside LPs, so one might expect only weak performance persistence when regressing fund 2 returns on fund 1 returns. Then, as soft information accumulates, the information asymmetry between incumbent and outside LPs increases, leading to more performance persistence in higher-sequence funds. Finally, persistence decreases as sufficient hard information (such as audited returns stated in a fundraising prospectus) accumulates, allowing outside LPs to infer the GP's type.

The following quote from Metrick (2006) suggests incumbent LPs retain an informational advantage even for some of the industry's most experienced GPs. Arguing why successful GPs, such as Sequoia Capital, do not increase their carry to eliminate over-subscription, Metrick states that if a firm were to increase its carry, "... [its] mix of LPs would be different, and some of the long-serving LPs would be gone. The new LPs, lacking the long-standing relationship, are less likely to remain loyal if the firm has a poor performing fund." One way to interpret investor loyalty is as a proxy for

the LP's information about the GP. Recall that outside LPs only observe hard information, while the incumbent LP additionally observes soft information that cannot be credibly communicated to outside LPs. Thus, if faced with one bad fund return, the incumbent LP will update his belief of GP skill less than outside LPs will.

The following prediction summarizes:

Implication 10: Evolution of persistence. Persistence should be hump-shaped in fund sequence and some persistence should remain even for funds raised by the most experienced GPs.

Extension 2 – Multiple LPs per fund: The relationship-banking literature suggests entrepreneurs reduce informational hold-up by borrowing from multiple banks; see Ongena and Smith (1998) and Boot (2000) for recent surveys. We argue in the next section that GPs ex ante have no incentive to reduce hold-up in the VC setting because asymmetric learning and informational hold-up likely lead to more efficient investment outcomes than with symmetric learning. Why? A pre-commitment to sharing future fund NPV enables GPs to raise first funds even if first funds have a negative expected NPV.

GPs do, of course, accept capital from multiple LPs, maybe because they want to reduce dependence on any single LP to avoid fund-raising difficulties in future funds should an LP experience a negative liquidity shock. Alternatively, it may allow LPs to diversify idiosyncratic GP risk and thus effectively increase the size of the pie (in a risk-adjusted sense).

Whatever the motive for having multiple LPs, the question is whether competition among incumbent LPs reduces their bargaining power and thus the amount of persistence implied by the model. For example, could the GP not offer each of a set of N LPs the chance to invest 1/N of the optimal follow-on fund size at a carry that ensures that their expected return equals their outside option r? This may not happen, for two reasons. First, LPs might collude against the GP. Collusion

is possible if LPs value future access to funds sufficiently. In our model, LPs expect to earn r (across first-time and follow-on funds) when investing with future GPs, but suppose more realistically that there is a finite supply of LPs who are good at learning. Since learning is valuable, such LPs would earn more than r in equilibrium. If the LPs collude, they continue to earn more than r forever. If a given LP competes on carry, he may obtain a larger allocation of the (high-skilled) GP's next fund. But other LPs can punish such behavior by not coinvesting with a deviating LP in future funds. If future returns are sufficiently valuable relative to a larger allocation in today's fund, LPs will refrain from competing on carry. Second, it is not necessarily in the GP's interest to make an aggressive offer to LPs because he may lose the ability to raise funds for uncertain future lines of business.[8]

D.   Optimality of Asymmetric Learning

Learning is valuable in both the symmetric and asymmetric learning settings. It ensures that high-skill GPs receive more capital in follow-on funds and that low-skill GPs (those with negative NPV) do not raise follow-on funds. This increases the overall value created by the VC industry. In expectation, GPs earn the full NPV of both their first-time and follow-on funds and thus prefer learning to no learning (not yet knowing their skill level).

Does *asymmetric* learning lead to better investment outcomes (i.e., a more efficient outcome) than *symmetric* learning? In our model it might. As we argued in Section C, when the expected NPV of first-time funds is negative, GPs can still raise first-time funds if learning is asymmetric and there is sufficient dispersion in GP skill that LPs can make up first-time fund losses from infor-

---

[8]Think of Kleiner, Perkins starting a series of funds investing in China. If the first China fund is expected to have negative NPV, these funds cannot be started unless Kleiner, Perkins is able to convince LPs that they will share in some of the later China funds' positive NPV.

mational rents on follow-on funds.[9] Soft information effectively serves as a commitment device for GPs to share some of the follow-on fund NPV with investors, and thus enables long-term contracting which in turn leads to better investment outcomes. This is also the case in standard models of informational hold-up such as Sharpe (1990), but there the gains from long-term contracting must be weighed against a distortion in investment in each period caused by distorted interest rates. This distortion is not present in the VC setting because VC contracts specify both an investment level (fund size) and the division of the fund's surplus. We have shown above that this yields first-best investment levels in each period (i.e., NPV-maximizing fund sizes).

One could argue that long-term contracting should be possible even with symmetric learning, since hard information is verifiable by courts which could thus enforce long-term contracts. In practice, contracts do not give LPs explicit rights to invest particular amounts with particular carries should a follow-on fund be raised, suggesting enforcement problems. Besides such problems, it may also be difficult or cost-inefficient for LPs to learn the GP's type only through the collection of hard information. In that sense, asymmetric learning may lead to more efficient outcomes by leading to more learning.

II. Sample and Data

We test the model using a sample of U.S. VC funds from Thomson Financial's Venture Economics (VE) and from Private Equity Intelligence (PREQIN). Both databases provide data on fund characteristics (such as a fund's vintage year and size) and, for a sub-sample, fund performance. Given our focus on GP-LP relations, we screen out entities that are not structured as limited partnerships.

---

[9]This efficiency argument is similar to Tirole's (2005, section 9.4.5), though Tirole considers a setting where profitability improves over time for a given borrower, while we consider a setting where average profitability improves over time as low-skill GPs fail to raise follow-on funds.

Using secondary sources such as *Pratt's Guide*, *CapitalIQ*, *Galante's*, and a web search, we exclude universities, state or local pension plans, government-sponsored entities, and vehicles representing angel or other individual investors. We also screen out funds of funds, buyout funds,[10] hedge funds, and evergreen funds (i.e., funds without a predetermined dissolution date).

Table 1 describes our sample. It contains 2,812 funds raised by 1,164 VC firms between 1980 and 2006. Of these, 782 funds are in both VE and PREQIN, 44 appear only in PREQIN, and the remaining 1,986 appear only in VE. The number of funds raised per year averages 64 in the 1980s, 138 in the 1990s, and 121 between 2001 and 2006. The average (median) sample fund raised $124.5 million ($50.0 million) in nominal dollars. Average fund size increased from $30.1 million in 1980 to $44.3 million in 1990, $201.4 million in 2000, and $215.9 million in 2006.

VC funds are under no obligation to disclose performance data publicly. Based on cash flow data disclosed voluntarily by GPs and/or LPs, VE and PREQIN calculate IRRs net of fees and carries.[11] VE provides two types of IRRs. The first is a single number per fund, reflecting a fund's performance as measured from its inception to the earlier of the fund's liquidation date or the date we downloaded the data (summer 2007). As funds typically have a ten-year life, this single number will reflect ultimate, "ex post" performance in the case of funds raised between 1980 and 1996. For more recent funds, the reported IRR is liable to change as investments are exited or written off. Thus, whenever we use ex post IRRs, we restrict the sample to vintage years 1980-1996.

While PREQIN reports only this single ex post number, VE in addition reports annual IRRs for each year between a fund's inception and the earlier of its liquidation or 2006. This allows us to

---

[10]We define as VC funds all funds listed in VE or PREQIN as focusing on start-up, early-stage, late-stage, or expansion investments, as well as those listed as "venture (general)" or "balanced" funds. In cases where VE and PREQIN classify a fund differently, we verify fund type using secondary sources.

[11]VE and PREQIN also report DVPI (the ratio of distributed to invested capital) and TVPI (the ratio of fund value to invested capital, which is based on both realized cash returns and subjective valuations of unrealized investments). Our results are qualitatively similar using these performance measures.

track performance as it evolves over a fund's life. We can thus avoid the look-ahead bias of relating a VC firm's ability to raise a follow-on fund in, say, year four of its current fund to that fund's ultimate, ex post performance. Instead, we can condition on "ex ante" available information.[12]

As Table 1 shows, we have ex post IRRs for 1,009 of the 2,812 funds. The average IRR is 14.2%. However, this includes recent funds that have yet to switch from making investments to exiting them. Focusing on the 601 funds from the 1980-1996 vintages, the average IRR is 18.8%. The median IRR is considerably lower, at 10.3%.[13] IRRs vary considerably over time. Average IRRs were in the single digits for funds raised between 1981 and 1987, in the mid to high teens between 1988 and 1990, in the twenties between 1991 and 1994, 44.3% for 1995 vintage funds, and 63.8% in the 1996 vintage. For the 1980-1996 vintages, IRRs average 15.3% for first-time funds and 20.4% for follow-on funds.

III.   Testing the Theory's Predictions

The VC literature already contains evidence to support some of our predictions. Kaplan and Schoar's (2005) and Gottschalg and Phalippou's (2007) findings, using ex post IRRs, support Implications 1, 2, and 6 (dependence of fund-raising, follow-on fund size, and persistence). Our purpose in re-testing these implications is to examine whether these findings are robust to using ex ante IRRs. Our evidence in support of Implications 3 (dependence of follow-on carry on prior-fund returns), 4 (more variable fund size for follow-on funds), and 9 (more oversubscription for follow-on funds) is to our knowledge novel. As for the evidence on Implications 5 (more variable carries for follow-on funds), 7 (lower average performance of first-time funds), and 8 (LP persistence), our work

---

[12]As Ljungqvist and Richardson (2003) show, over a fund's life, performance follows a 'J-curve', in the sense that IRRs tend to be negative in the first few years as the fund is mainly in investment mode and then turn positive after five or six years as the fund begins to exit its investments through IPOs or M&A transactions.

[13]While this suggests some positive outliers, winsorizing the data does not materially affect our results.

confirms earlier findings by Gompers and Lerner (1999), Kaplan and Schoar (2004), Gottschalg and Phalippou (2008), and Lerner, Schoar, and Wongsunwai (2007).

A. Implication 1: Fund-raising

Implication 1 states that the probability of raising a follow-on fund increases in the return the LPs earned in the first fund. Column 1 of Table 2 presents a probit model where the dependent variable equals one if the GP raises a second fund and zero otherwise. The main variable of interest is the ex post IRR of the GP's first fund. Since we require ten years of data to measure ex post IRRs, we restrict the sample to first funds raised in 1980-1996. Note that we allow the follow-on fund to be raised at any point after the first fund, up until the end of 2006. Following Kaplan and Schoar (2005), we also control for the log of the first fund's size and for vintage-year fixed effects.[14]

The coefficient estimated for prior-fund IRR is reliably positive (p=0.001), consistent with Implication 1. At the means of the other covariates, a one-standard deviation increase in first-fund IRR increases the probability of raising a follow-on fund by 27 percentage points. In column 2, we widen the sample to include all funds (not just first-time funds) to see if the performance-sensitivity is attenuated for later funds. The results suggest that the most recent fund's performance is less important if the VC has a longer track record: While still statistically significant (p=0.043), the coefficient estimated for prior-fund performance decreases, so that a one-standard deviation increase in IRR is associated with only a nine point increase in the likelihood of raising another fund.

A potential shortcoming of using ex post IRRs, which we share with Kaplan and Schoar (2005) and Gottschalg and Phalippou (2007), is that it implies perfect foresight, in the sense that LPs can predict, when raising a fund, what the previous fund's ultimate performance will be before the end

---

[14]The model does not address variation in first-time fund size. The variation we see in the data could reflect investors having some (imperfect) ex ante knowledge about which GPs have higher skill or differences in optimal fund size across industries or geographic areas, so we follow the literature and control for fund size.

of its investment life. To condition only on information that is known ex ante, we estimate a Cox hazard model with time-varying covariates. This relates the probability that a VC firm raises a new fund in year τ to the interim (i.e., ex ante) IRR of its previous fund as reported at the end of year $\tau - 1$ and the log size of its previous fund.[15] Since we use ex ante rather than ex post IRRs in the estimation, we can include all available vintages through 2006. As VC firms have a non-zero probability of raising further funds after 2006, we adjust the likelihood function for right-censoring.

Column 3 of Table 2 reports the coefficient estimates. Controlling for the fact that VC firms with larger funds are more likely to raise another fund, we find that higher LP returns on the previous fund significantly increase the hazard of raising a new fund ($p<0.001$). A unit increase in IRR in year $\tau - 1$ (e.g., from 0 to 100%) is associated with a 25.9% higher likelihood of raising a follow-on fund in year τ. This mirrors the probit results in columns 1 and 2.

B.   Implication 2: Evolution of Fund Size

According to Implication 2, the size of a follow-on fund increases in the return LPs earned in the previous fund. Table 3 provides evidence consistent with this implication. We use a Tobit estimator to model the evolution of fund size, controlling for left-censoring in the size variable as a result of a firm being unable to raise a follow-on fund (presumably due to poor performance). To code failure to raise a follow-on fund, we identify 362 defunct VC firms in *CapitalIQ*.[16] The dependent variable equals the log fund size if the firm raises a follow-on fund and zero if it does not. Like Kaplan and Schoar (2005), we regress this on the prior fund's IRR and log size as well as vintage-year effects.

---

[15]We ignore the first four years of IRR data as the IRR of a fund that is mainly investing and not yet generating returns is not meaningful. In practice, this affects only first-time funds as VC firms with later funds nearly always have a prior fund that is at least four years old. For the purpose of the hazard model, first-time funds are treated as left-censored during their first four years, and the likelihood function is adjusted accordingly.

[16]Defunct VC firms are those *CapitalIQ* labels "out of business", "dissolved", "liquidating", "no longer investing", or "reorganizing." We also assume that firms that haven't raised a fund since 1996 are defunct.

Controlling for the fact that GPs that have managed larger funds tend to raise larger funds in the future, column 1 shows that fund size is positively related to the previous fund's ex post IRR (p=0.001). However, this specification assumes perfect foresight. In column 2, we instead condition on the previous fund's ex ante IRR measured as of the year-end prior to the GP raising the current fund. Unlike in column 1, which restricts the previous fund to those raised in 1980-1996 to ensure we have a final IRR, in column 2 we can include all funds with available data.[17] The results are virtually unchanged. At the means of the other covariates, a one-standard deviation increase in the ex ante performance of the previous fund is associated with a 35.3% or $20.6 million increase in fund size, from the unconditional mean in the estimation sample of $58.3 million ($p$<0.001).

C. Implication 3: Evolution of GP Carry

Implication 3 states that GPs increase their carry following high returns on their previous funds. We hand-collect carry data for 367 funds from GPs and public sources (including the *Venture Capital Journal*, press reports in Factiva, and various Harvard Business School case studies). Consistent with Gompers and Lerner (1999) and Litvak (2008), first-time funds in our data have a lower carry (mean: 20.6%) than do follow-on funds (mean: 22.6%).

In Table 4, we relate the GP carry to the ex post (column 1) or ex ante (column 2) IRR of the GP's previous fund, controlling for fund size and vintage-year effects. In each case, carries increase significantly in the previous fund's IRR. In column 2, for instance, a one-standard deviation increase in the previous fund's ex ante IRR is associated with a 1.98 percentage point increase in carry on the next fund (p<0.001).

These estimates could be biased to the extent that poor performance prevents a VC firm from

---

[17]Again, we ignore data for the first four years of a fund's life.

raising a follow-on fund (left-censoring). In column (3), we estimate a Tobit model where we set the carry equal to zero if the firm fails to raise a follow-on fund. This increases the performance-sensitivity of the GP carry to 4.2 percentage points ($p<0.001$). To our knowledge this is the first set of results to systematically document a relation between follow-on fund carry and prior fund performance.[18]

D.   Implications 4 and 5: Cross-fund Standard Deviation of Fund Size and Carry

Table 5 reports the results of tests comparing the standard deviation of fund size and carry in three different samples: The full set of funds raised between 1980 and 2006; the set of funds raised between 1980 and 1996; and the set of 1980-1996 funds for which we have performance information. In each sample, we find statistically significant support for Implication 4; see Panel A. Panel B shows significantly greater variation in carry among more established VC firms, as predicted, in the first two samples. This is driven by the fact that 88% of first-fund carries cluster at 20% (with the remainder mostly at 25%), while 60% of follow-on funds have a 20% carry, 23% have a 25% carry, and 16% have a 30% carry. (In the set of 1980-1996 funds for which we have IRR data, there is no variation in first-fund carries, so the test statistic cannot be computed.)

E.   Implications 6 and 10: Persistence in Returns to LPs

We first replicate Kaplan and Schoar's (2005) persistence test. In column 1 of Table 6, Panel A, we regress a fund's ex post IRR on log fund size, the ex post IRR of the VC firm's previous fund, and vintage-year effects. As a crude control for difference in risk-taking across funds, we also include a dummy variable that equals one for funds classified as investing in early-stage companies.[19] Like

---

[18]Neither Gompers and Lerner (1999) nor Litvak (2008) condition on prior performance. Metrick and Yasuda (2007) show that carry per individual partner increases in fund sequence number, but do not relate carry to prior fund performance.

[19]Kaplan and Schoar (2005) show that the persistence result is robust to this and other proxies for risk.

Kaplan and Schoar, we find that fund performance increases with fund size and—consistent with Implication 6—prior performance ($p<0.001$). Kaplan and Schoar note that a VC firm's current and previous funds will tend to overlap in time, as they are usually raised fewer than 10 years apart. To mitigate this problem, they suggest including the two prior funds' IRRs. When we do so, in column 2, we find that only the previous fund's IRR is significantly related to the current fund's IRR ($p<0.001$). In column 3, we control for the performance of the second-prior fund only, to reduce overlap. The persistence effect remains significant ($p=0.04$), suggesting that Kaplan and Schoar's result is not a spurious artifact of the overlapping nature of the data.

Since we have access to ex ante performance data, we can improve on Kaplan and Schoar's (2005) persistence test by conditioning on the previous fund's IRR measured as of the year-end prior to the year the GP raised the current fund. This yields a somewhat larger persistence coefficient ($p=0.035$). A one percentage point increase in the ex ante IRR of the VC firm's previous fund is associated with a 0.924 percentage point increase in the ex post IRR of the next fund. In contemporaneous work, Phalippou (2008) also finds that persistence is present based on performance data known at the time of follow-on fund raising.

According to Implication 10, persistence should first increase and then decrease across a sequence of funds; in early funds, it may even be absent. To test this prediction, we interact the ex ante IRR measure from column 4 with a dummy variable identifying fourth and higher-numbered funds. (The cut-off is arbitrary but not selective.) We find no evidence of persistence among second and third funds and significant persistence among fourth and later funds. (We have too few higher-numbered funds in the dataset to reliably test whether persistence subsequently diminishes.)

One concern regarding the persistence result is selection bias: Not every fund in our sample reports an IRR, and it is possible that those that do are those that experience persistent good

performance. To explore the extent of this bias, we repeat the analyses in Panel A with exit rates as the dependent variable instead of IRRs; see Hochberg, Ljungqvist, and Lu (2007). Exit rates are defined as the fraction of a fund's investments that were exited through an IPO or an M&A transaction over the course of the fund's ten-year life and can be computed for nearly every fund. As Panel B shows, we continue to find persistence using this alternative measure of performance.

F. Implication 7: Performance of First-time Funds

Implication 7 predicts that follow-on funds outperform first-time funds on average. For vintage years 1980-1996, average IRRs among follow-on funds are 5.1 percentage points higher (at 20.4%) compared to first funds (at 15.3%); see Table 1. Though consistent with Implication 7, the difference is not statistically significant (t-stat = 1.49). Table 7 presents estimates from IRR regressions that control for fund size, fund stage focus (as a crude proxy for risk), and vintage-year effects. In column 1, a dummy identifying first-time funds is not statistically significant, though we find that larger funds have higher IRRs, which is consistent with Implication 7 insofar as follow-on funds tend to be larger. In column 2, we explore an alternative way of splitting funds into 'early' and 'later' funds. Until a fund has liquidated all its positions, its reported IRR includes unrealized capital gains and so contains a subjective element. Many of the funds coded as follow-on funds in column 1 were, in practice, raised before the tenth anniversary of the GP's first fund. (In our sample, the average second fund is raised after 3.2 years.) In column 2, we define follow-on funds as those raised at least 10 years after the first fund. Such funds perform significantly better than earlier funds, by 8.5 percentage points on average (p=0.05). In column 3, we use a different functional form and regress IRRs on the VC firm's age (measured in log years since it raised its first fund). We find that IRRs increase over a VC firm's lifetime, by about 3.5 percentage points for a one-standard

deviation increase in VC firm age ($p=0.05$).

G. Implication 8: Exit of LPs

Implication 8 predicts that if a GP successfully raises a follow-on fund, investors in the first fund should also invest in the follow-on fund. Using a large sample of LP data compiled from Venture Economics and VentureOne,[20] we find that on average, 42.4% of LPs in a first fund continue to invest in the GP's next fund, falling to 39.9% in later funds. How likely is it that we observe this level of persistence purely by chance? Say the average fund has ten LPs and there are 250 possible LPs to choose from in the average year. The probability of randomly choosing four or more (out of ten) LPs who were investors in the previous fund is 0.00025 (applying the hypergeometric pdf). We conclude that there is a significant degree of persistence in LP composition across funds.

H. Implication 9: Oversubscription in Follow-on Funds

There is no prior evidence regarding the concentration of oversubscription in follow-on funds and its relation to prior-fund performance. We compile data on target and final fund sizes from January issues of the *Private Equity Analyst* in the three years centered on a fund's vintage year.[21] These data are available from 1991. The ratio of final to target fund size averages 101.4%, with a standard deviation of 35.8% and a range from 6.7% to 310%. For first and follow-on funds, the subscription ratio averages 94.9% and 103.1%, respectively, consistent with Implication 9. This difference continues to hold when we regress the subscription ratio on a follow-on fund indicator, log target fund size, and vintage-year effects; see column 1 of Table 8.

Column 2 tests whether oversubscription is related to prior fund performance. Restricting the

---

[20]Neither VE nor VentureOne covers every LP, and even when combined do not provide a complete list of LPs in a fund. This drawback is also noted by Lerner, Schoar, and Wongsunwai (2007), who use LP data from VE.

[21]Some funds remain open for investment for longer.

sample to follow-on funds, we regress the subscription ratio on the previous fund's IRR measured as of the year-end before the GP raised the current fund. Controlling for log target size and vintage-year effects, we find that a one-standard deviation increase in prior performance is associated with a 4.7 percentage point increase in subscription ($p<0.001$), consistent with Implication 9.

IV.   Discussion and Conclusion

We propose and test a model of the interaction between VC fund managers (GPs) and their investors (LPs) in a setting where investors learn about GP skill over time. We exploit logic similar to that presented in the relationship-banking literature: LPs who invest in a GP's fund learn more about that GP's skills than do other LPs who are not invested in the fund. This asymmetric evolution of information creates a hold-up problem, and this in turn leads to performance persistence.

Our empirical evidence is consistent with asymmetric learning. We find performance persistence in after-fee returns across a GP's funds (see also Kaplan and Schoar (2005)), even when we depart from the literature by conditioning only on ex ante available information. Follow-on funds outperform first funds after fees. There is persistence in LPs from fund to fund, and follow-on funds are oversubscribed, especially following high returns. We also document empirical evidence consistent with the predictions of both asymmetric and symmetric learning: The probability that a GP raises another fund increases in the first fund's return; better performing GPs raise larger follow-on funds and charge higher performance fees; and the cross-fund standard deviation of fund size and fees is higher in later funds.

It may be possible to extend our informational hold-up framework to explain other features of the data. For example, suppose LPs differ in their ability to learn about GP skill. If there is a shortage of LPs who are able to learn, and if learning is valuable (as it is in our model), then

LPs who are better at learning will in equilibrium be compensated by the GP for the services they provide. This prediction is consistent with the large heterogeneity in the performance experienced by different classes of LPs shown in Lerner, Schoar, and Wongsunwai (2007). In particular, if university endowments such as the Harvard Management Company or the Yale Investments Office are particularly good at learning about a GP's skill, the model can provide an explanation for the limited partner puzzle presented in Lerner, Schoar, and Wongsunwai, who find that endowments earn higher returns from VC investing than do other institutional investor types.

References

Berk, J., and R. Green, 2004, "Mutual Fund Flows and Performance in Rational Markets", *Journal of Political Economy* 112, 1269-1295.

Binmore, K., M. J. Osborne, and A. Rubinstein, 1992, "Non-Cooperative Models of Bargaining", Handbook of Game Theory, Volume 1, Edited by R. J. Aumann and S. Hart.

Boot, A. W. A., 2000, "Relationship Banking: What Do We Know?", *Journal of Financial Intermediation*, 9, 7-25.

Cochrane, J., 2005, "The Risk and Return of Venture Capital", *Journal of Financial Economics* 75, 3-52.

Glode, V., and R. Green, 2008, "Information Spill-Overs and Performance Persistence in Private Equity Partnerships", Unpublished working paper, Carnegie-Mellon University.

Gompers, P.A., and J. Lerner, 1999, "An Analysis of Compensation in the U.S. Venture Capital Partnership," *Journal of Financial Economics* 51, 3-44.

Gompers, P.A., and J. Lerner, 2001, *The Money of Invention: How Venture Capital Creates New Wealth*, Harvard Business School Publishing, Boston.

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

Jones, C., and M. Rhodes-Kropf, 2003, "The Price of Diversifiable Risk in Venture Capital and Private Equity", Unpublished working paper, Columbia University.

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

Kaplan, S.N., B. Sensoy, and P. Strömberg, 2002, "How Well Do Venture Capital Databases Reflect Actual Investments?", Unpublished working paper, University of Chicago.

Korteweg, A., and M. Sørensen, 2008, "Estimating Risk and Return of Infrequently Traded Assets: A Bayesian Selection Model of Venture Capital," Unpublished working paper, Columbia University.

Lerner, J., and A. Schoar, 2004, "The Illiquidity Puzzle: Evidence from Private Equity Partnerships," *Journal of Financial Economics* 72, 3-40.

Lerner, J., A. Schoar, and W. Wongsunwai, 2007, "Smart Institutions, Foolish Choices?: The Limited Partner Performance Puzzle", *Journal of Finance* 62, 731-764.

Litvak, K., 2008, "Are Venture Capitalists Paid Like Bureaucrats? Analysis of the Pay-Performance Link in Venture Capital Partnership Agreements," *University of Chicago Law Review*, forthcoming.

Ljungqvist, A., and M. Richardson, 2003, "The Cash Flow, Return, and Risk Characteristics of Private Equity", NBER Working Paper No. 9454.

Malkiel, B.G., 1995, "Returns from Investing in Equity Mutual Funds 1971-1991," *Journal of Finance* 50, 549-572.

McDonald, R.L., 2003, *Derivatives Markets*, Addison Wesley.

Metrick, A., 2006, *Venture Capital and the Finance of Innovation*, Wiley.

Metrick, A., and A. Yasuda, 2007, "The Economics of Private Equity Funds", Unpublished working paper, University of Pennsylvania.

Ongena S., and D.C. Smith, 1998, "Bank Relationships: A Review", in Zenios, S. A., and P. Harker (eds.), *Performance of Financial Institutions*, Cambridge University Press.

Phalippou, L., 2008, "Venture Capital Funds: Performance Persistence and Flow-Performance Relation", Unpublished working paper, University of Amsterdam.

Quigley, R., and S. Woodward, 2003, "An Index for Venture Capital", Unpublished working paper, University of California, Berkeley.

Rajan, R., 1992, "Insiders and Outsiders: The Choice between Informed and Arm's Length Debt", *Journal of Finance* 47, 1367-1400.

Rubinstein, A., 1982, "Perfect Equilibrium in a Bargaining Model", *Econometrica* 50, 97-109.

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

Sharpe, S.A., 1990, "Asymmetric Information, Bank Lending and Implicit Contracts: A Stylized Model of Customer Relationships", *Journal of Finance* 45, 1069-1087.

Tirole, J., 2005, "The Theory of Corporate Finance", Princeton University Press.

Von Thadden, E.-L., 2004, "Asymmetric Information, Bank Lending and Implicit Contracts: The Winner's Curse", *Finance Research Letters* 1, 11-23.

Appendix: Derivations and Proofs for the Theoretical Model

A. Expressions for $g_2\left(\mu^i\right)$ and $g_1\left(\mu^i\right)$

$g_2\left(\mu^i\right)$ is defined as $E\left[\max\left(0, C_2 - I_2\left(\mu^i\right)\right) | \mu = \mu^i\right]$. From the normality of A_2^i it follows that

$$\ln\left(C_2/I_2\right) = \ln\left(e^{A_2^i}\ln\left(1+I_2\right)/I_2\right) = A_2^i + \ln\left(\frac{\ln\left(1+I_2\right)}{I_2}\right) \sim N\left(\mu_i + \ln\left(\frac{\ln\left(1+I_2\right)}{I_2}\right) - 0.5\sigma^2, \sigma^2\right)$$

Using equations (18.30), (18.24), and (12.2a, 12.2b) in McDonald (2006), this implies that

$$
\begin{aligned}
g_2\left(\mu^i\right) &= E\left[C_2 | C_2 > I_2\left(\mu^i\right), \mu = \mu^i\right] P\left(C_2 > I_2\left(\mu^i\right) | \mu = \mu^i\right) - I_2\left(\mu^i\right) P\left(C_2 > I_2\left(\mu^i\right) | \mu = \mu^i\right) \\
&= \left[e^{\mu_i}\ln\left(1+I_2\left(\mu^i\right)\right) N\left(d_2\left(\mu^i\right)\right) - I_2\left(\mu^i\right) N\left(d_2\left(\mu^i\right) - \sigma\right)\right]
\end{aligned}
$$

where

$$d_2\left(\mu^i\right) = \frac{\mu_i + \ln\left(\frac{\ln\left(1+I_2\left(\mu^i\right)\right)}{I_2\left(\mu^i\right)}\right)}{\sigma} + \frac{1}{2}\sigma.$$

$g_1\left(\mu^i\right)$ is defined as $E\left[\max\left(0,\left(C_1 - I_1^{opt}\right)\right) | \mu = \mu^i\right]$. From the normality of A_1^i,

$$\ln\left(C_1/I_1\right) \sim N\left(\mu_i + \ln\left(\frac{\ln\left(1+I_1\right)}{I_1}\right) - 0.5\sigma^2, \sigma^2\right).$$

Therefore,

$$g_1\left(\mu^i\right) = \left[e^{\mu_i}\ln\left(1+I_1^{opt}\right) N\left(d_1\left(\mu^i\right)\right) - I_1^{opt} N\left(d_1\left(\mu^i\right) - \sigma\right)\right]$$

where

$$d_1\left(\mu^i\right) = \frac{\mu_i + \ln\left(\frac{\ln\left(1+I_1^{opt}\right)}{I_1^{opt}}\right)}{\sigma} + \frac{1}{2}\sigma.$$

B. Verifying Implications 1, 2, 3, and 6

We verify Implications 1, 2, 3, and 6 numerically using two simulated datasets, one for the symmetric learning version of the model and one for the asymmetric learning version. The construction of each of the two data sets is as follows. Pick values for r, μ^L, μ^H, and σ. Define a grid of N equidistant μ^i values between μ^L and μ^H. For each value of μ^i, draw a vector of M values of r_1

and of r_2, respectively, and define an $M \times 1$ vector in which all elements are μ^i, an $M \times 1$ vector in which all elements are $I_2(\mu^i)$, and an $M \times 1$ vector in which all elements are $f_2(\mu^i)$.[22] Stack the M r_1-vectors on top of each other, and do likewise for the r_2-vectors, the μ-vectors, the I_2-vectors, and the f_2-vectors. Align the stacked vectors to generate an $NM \times 5$ dataset of values of r_1, r_2, μ^i, $I_2(\mu^i)$ and $f_2(\mu^i)$.

We use the following parameter values: $r = e^{0.14} - 1$ (i.e., $\mu^* = 0.14$), $\mu^L = 0.04$, $\mu^H = 0.4$, and $\sigma = 0.1$. We use $M = 10,000$ and $N = 361$, corresponding to a grid of GP types with grid size 0.001. The qualitative results are not sensitive to the exact parameter values we choose. To verify Implication 1, we graph the fraction of funds for which $\mu^i \geq \mu^*$ in each of 25 r_1-buckets in Figure 1, Panel A. Consistent with Implication 1, the probability of raising a follow-on fund increases in r_1. Graphing mean fund size I_2 in Panel B, conditional on $\mu^i \geq \mu^*$, confirms Implication 2 that follow-on fund size increases in first-fund return. Mean carry in Panel C increases in r_1, consistent with Implication 3. To verify Implication 6, Figure 2 graphs the mean of r_2 within a set of r_1-buckets, using only those observations for which $\mu^i \geq \mu^*$. The graph confirms that performance is persistent.[23] Note also that consistent with Implication 7, the mean of r_1 in the simulated data is $E_i(E(r_1|\mu = \mu^i)) = 0.078$ while the mean of r_2 for $\mu^i \geq \mu^*$ is $E_i(r_2|\mu^i \geq \mu^*) = 0.188$.

With symmetric learning, our chosen parameters imply $I_1 = 0.089$ and $f_1 = 0.24$. With asymmetric learning, I_1 is still 0.089 whereas f_1 increases to 0.59. While the values depend on the parameters chosen for the simulation, the fact that f_1 is higher when learning asymmetric is general. Intuitively, the first-time fund carry must be high enough that the GP earns the full NPV of both the first fund and the follow-on fund (since the LP market is competitive ex ante). A carry of 59% is unrealistic. To temper the required increase in carry, we could introduce a management fee. Alternatively, we could use a non-uniform distribution of GP types or introduce asymmetric bargaining with a higher fraction of follow-on fund NPV going to the GP.[24] None of these changes

---

[22] $I_2(\mu^i)$ and $f_2(\mu^i)$ are not defined and not needed for $\mu^i < \mu^*$, so just plug in zeros in the $I_2(\mu^i)$-vectors and $f_2(\mu^i)$-vectors corresponding to $\mu^i < \mu^*$.

[23] The spread in $E(r_2)$ is substantial, ranging from around 16% in the bottom bucket to 21% in the top. The range of $E(r_2)$ (and thus the amount of performance persistence) is wider the wider is the range of GP types. To get a sense of whether the assumed parameter values are reasonable, consider a simple calculation. The optimal investment at the midpoint of the range of GP types ($\mu = 0.22$) is $I_2 = 0.0833$. At this level of investment, the lowest GP type generates returns with a mean of $E(C_2/I_2) = 1$ and a standard deviation of $\sqrt{V(C_2/I_2)} = 0.10$. The highest GP type generates $E(C_2/I_2) = 1.43$ with $\sqrt{V(C_2/I_2)} = 0.14$. The highest GP type is thus 43% more effective than the lowest GP type for the same I_2, with somewhat higher risk. While it is not possible to map our simple model directly to the data (as GP types are unlikely to be uniformly distributed), this amount of dispersion in GP skill seems reasonable and is sufficient to generate a wide spread in expected returns on follow-on funds.

[24] Binmore, Rubinstein, and Wolinsky (1986) provide alternating-offer bargaining games that converge to Nash bargaining with unequal bargaining powers. One of their settings assumes that one party is faster at making counter-

would alter our qualitative conclusions.

C. Outcome of Sequential Bargaining

We derive the equilibrium strategies and outcome of the sequential bargaining game between the GP and the incumbent LP at the start of the follow-on fund. This is done both for general p and for $p \to 0$. We omit the proof that the proposed equilibrium is the unique perfect equilibrium; it follows Rubinstein (1982) and Binmore, Osborne, and Rubinstein (1992), section 2.1.

Strategies: For a given value of p, the following constitutes a set of equilibrium strategies:

1. All offers propose the NPV-maximizing investment level $I_2\left(\mu^i\right) = \frac{e^{\mu_i}}{1+r} - 1$.

2. For a GP of type μ^i, there exists a single pair of proposed carries, $f_2^{LP}\left(\mu^i\right)$, $f_2^{GP}\left(\mu^i\right)$ such that the incumbent LP is indifferent between the contract $\left(f_2^{GP}\left(\mu^i\right), I_2\left(\mu^i\right)\right)$ now and the contract $\left(f_2^{LP}\left(\mu^i\right), I_2\left(\mu^i\right)\right)$ in the next round of bargaining and such that the GP is indifferent between the contract $\left(f_2^{LP}\left(\mu^i\right), I_2\left(\mu^i\right)\right)$ now and the contract $\left(f_2^{GP}\left(\mu^i\right), I_2\left(\mu^i\right)\right)$ in the next round of bargaining. These carries are given by

$$
\begin{aligned}
f_2^{GP}\left(\mu^i\right) &= \frac{p\left[e^{\mu^i}\ln\left(1+I_2\left(\mu^i\right)\right) - (1+r)I_2\left(\mu^i\right)\right]/\left[1-(1-p)^2\right]}{g_2\left(\mu^i\right)} \\
f_2^{LP}\left(\mu^i\right) &= (1-p)\,f_2^{GP}\left(\mu^i\right).
\end{aligned}
$$

Proof: The indifference conditions for the GP and the LP are that the expected payoff from accepting equals the expected payoff from waiting and having your own offer accepted in the next round:

$$
f_2^{LP}\left(\mu^i\right)g_2\left(\mu^i\right) = (1-p)\,f_2^{GP}\left(\mu^i\right)g_2\left(\mu^i\right) \iff f_2^{LP}\left(\mu^i\right) = (1-p)\,f_2^{GP}\left(\mu^i\right)
$$

$$
\begin{aligned}
&\frac{E\left(C_2|\mu=\mu^i\right) - f_2^{GP}\left(\mu^i\right)g_2\left(\mu^i\right)}{1+r} - I_2\left(\mu^i\right) \\
&= (1-p)\left[\frac{E\left(C_2|\mu=\mu^i\right) - f_2^{LP}\left(\mu^i\right)g_2\left(\mu^i\right)}{1+r} - I_2\left(\mu^i\right)\right].
\end{aligned}
$$

offers than the other party. Another setting assumes that the parties have different beliefs about the probability of a breakdown in negotiations.

Combining the two expressions, we get

$$
\begin{aligned}
f_2^{GP}\left(\mu^i\right) &= \frac{p\left[\frac{E\left(C_2|\mu=\mu^i\right)}{1+r} - I_2\left(\mu^i\right)\right] / \left[1-(1-p)^2\right]}{g_2\left(\mu^i\right)/(1+r)} \\
&= \frac{p\left[e^{\mu^i}\ln\left(1+I_2\left(\mu^i\right)\right) - (1+r)I_2\left(\mu^i\right)\right] / \left[1-(1-p)^2\right]}{g_2\left(\mu^i\right)}.
\end{aligned}
$$

3. The equilibrium strategies are that the GP always offers $\left(f_2^{GP}\left(\mu^i\right), I_2\left(\mu^i\right)\right)$ and always rejects offers with $f_2 < f_2^{LP}\left(\mu^i\right)$, and the incumbent LP always offers $\left(f_2^{LP}\left(\mu^i\right), I_2\left(\mu^i\right)\right)$ and always rejects offers with $f_2 > f_2^{GP}\left(\mu^i\right)$.

Outcome: Since the GP makes the first offer, the equilibrium outcome is $\left(f_2^{GP}\left(\mu^i\right), I_2\left(\mu^i\right)\right)$, agreed to in the first round of bargaining. If the incumbent LP made the first offer, it would be $\left(f_2^{LP}\left(\mu^i\right), I_2\left(\mu^i\right)\right)$, agreed to in the first round of bargaining.

Limit as $p \to 0$: Taking p to zero, and using l'Hôpital's rule:

$$
f_2^{GP}\left(\mu^i\right) \to \frac{\frac{1}{2}\left[e^{\mu^i}\ln\left(1+I_2\left(\mu^i\right)\right) - (1+r)I_2\left(\mu^i\right)\right]}{g_2\left(\mu^i\right)}
$$

and $f_2^{LP}\left(\mu^i\right) = f_2^{GP}\left(\mu^i\right)$. Thus, the GP's expected payoff, $f_2^{GP}\left(\mu^i\right)g_2\left(\mu^i\right)$, equals half the fund's NPV, which is the same as what results from Nash bargaining with equal bargaining power and outside options of zero.

D. Expression for f_1 under Asymmetric Learning

f_1 solves

$$
f_1 = \frac{E_i\left(e^{\mu_i}\right)\ln\left(1+I_1^{opt}\right) - (1+r)I_1^{opt}}{E_i\left(g_1\left(\mu^i\right)\right)} + \frac{\frac{1}{2}E_i\left(NPV_2\left(\mu^i\right)\right)}{E_i\left(g_1\left(\mu^i\right)\right)}.
$$

$NPV_2\left(\mu^i\right)$ is calculated at the NPV-maximizing investment level $I_2\left(\mu^i\right) = \frac{e^{\mu^i}}{1+r} - 1$ and thus given by

$$
NPV_2\left(\mu^i\right) = \frac{e^{\mu^i}\ln\left(1+I_2\left(\mu^i\right)\right)}{1+r} - I_2\left(\mu^i\right) = \frac{e^{\mu^i}\left(\mu^i - \ln(1+r)\right)}{1+r} - \left[\frac{e^{\mu^i}}{1+r} - 1\right]
$$

which implies

$$f_1 = \frac{\left[E_i\left(e^{\mu^i}\right)\ln\left(1+I_1^{opt}\right) - (1+r)I_1^{opt} + \frac{1}{2}\left(\frac{E_i\left(e^{\mu^i}\mu^i|\mu^i>\mu^*\right)}{(1+r)} - \frac{E_i\left(e^{\mu^i}|\mu^i>\mu^*\right)[\ln(1+r)+1]}{(1+r)} + 1\right)\right]}{E_i\left(g_1\left(\mu^i\right)\right)}.$$

with $I_1^{opt} = \frac{\frac{1}{\mu^H-\mu^L}\left[e^{\mu^H}-e^{\mu^L}\right]}{1+r} - 1$, $\mu^* = \ln\left(1+r\right)$. Furthermore, from the uniform distribution of GP types it follows that:

$$
\begin{aligned}
E_i\left(e^{\mu^i}\mu^i|\mu^i>\mu^*\right) &= \frac{1}{\mu^H-\mu^*}\int_{\mu^*}^{\mu^H}e^{\mu^i}\mu^i d\mu^i = \frac{1}{\mu^H-\mu^*}\left[e^{\mu^i}\left(\mu^i-1\right)\right]_{\mu^*}^{\mu^H} \\
&= \frac{1}{\mu^H-\mu^*}\left[e^{\mu^H}\left(\mu^H-1\right)-e^{\mu^*}\left(\mu^*-1\right)\right] \\
E_i\left(e^{\mu^i}|\mu^i>\mu^*\right) &= \frac{1}{\mu^H-\mu^*}\left[e^{\mu^H}-e^{\mu^*}\right] \\
E_i\left(g_1\left(\mu^i\right)\right) &= \frac{1}{\mu^H-\mu^L}\int_{\mu_L}^{\mu_H}g_1\left(\mu^i\right)d\mu^i.
\end{aligned}
$$

Table 1. Descriptive Statistics.

| | Number of sample funds | | | | Fund size ($m) | | Performance | | | | | | | |
| | | of which | | | | | All funds | | | | First-time funds | | Follow-on funds | |
vintage	all	only in VE	only in Preqin	in both	mean	median	no. of funds with IRR data	mean IRR (%)	sd IRR (%)	median IRR (%)	no. of funds with IRR data	mean IRR (%)	no. of funds with IRR data	mean IRR (%)
1980	40	34	4	2	30.1	20.0	17	13.0	12.7	12.9	10	10.9	7	16.1
1981	47	40	1	6	22.6	19.6	18	7.8	8.3	9.9	8	6.3	10	9.1
1982	66	57	0	9	23.5	15.1	29	3.1	8.9	5.9	19	2.9	10	3.3
1983	78	66	1	11	31.4	20.1	42	8.6	11.4	7.8	20	9.3	22	7.9
1984	90	79	0	11	30.4	22.0	52	4.4	7.9	3.5	23	5.5	29	3.6
1985	62	45	1	16	40.1	20.0	33	9.7	10.8	9.6	9	9.9	24	9.6
1986	57	39	0	18	52.5	21.6	34	6.8	7.4	6.1	16	5.2	18	8.3
1987	84	67	1	16	36.1	24.2	57	7.1	15.0	7.2	21	4.5	36	8.6
1988	57	36	2	19	66.2	32.8	39	13.7	14.2	10.3	7	11.4	32	14.2
1989	77	46	1	30	66.3	30.5	51	16.3	31.8	12.2	10	26.2	41	13.9
1990	50	37	2	11	44.3	35.0	20	16.2	21.4	10.4	2	-0.8	18	18.1
1991	35	24	1	10	40.6	30.0	15	21.9	17.2	22.6	3	27.8	12	20.4
1992	50	29	0	21	75.3	44.0	26	23.6	28.8	13.8	4	2.7	22	27.4
1993	71	44	2	25	55.8	35.7	37	26.4	33.3	15.3	8	15.8	29	29.4
1994	75	43	0	32	83.5	45.0	42	23.5	33.1	17.0	6	7.6	36	26.2
1995	118	82	1	35	71.0	43.0	51	44.3	57.2	28.1	16	34.6	35	48.8
1996	100	74	0	26	69.2	50.0	38	63.8	103.8	23.8	10	98.5	28	51.4
1997	165	110	1	54	85.6	56.7	64	42.7	76.5	6.3	16	35.0	48	45.3
1998	169	111	1	57	136.5	70.5	73	23.9	105.9	2.2	7	16.9	66	24.7
1999	259	179	2	78	173.0	100.0	72	-6.7	24.8	-9.8	12	16.5	60	-11.4
2000	338	234	5	99	201.4	97.5	93	-7.0	12.4	-6.5	12	-0.9	81	-7.9
2001	177	125	2	50	215.8	69.1	47	-4.3	12.4	-5.8	3	-19.9	44	-3.2
2002	77	45	4	28	138.6	49.3	22	-1.4	8.7	-0.9	3	-10.0	19	-0.1
2003	70	39	5	26	127.9	36.6	19	-0.7	15.3	-0.3	3	1.5	16	-1.2
2004	113	80	1	32	137.6	50.0	16	-0.8	18.4	-5.2	4	3.7	12	-2.2
2005	133	98	1	34	248.0	107.7	2	-19.3	17.3	-19.3	1	-7.0	1	-31.5
2006	154	123	5	26	215.9	107.0								
1980-2006	2,812	1,986	44	782	124.5	50.0	1,009	14.2	48.3	4.5	253	14.7	756	14.0
1980-1996	1,157	842	17	298	51.3	28.1	601	18.8	39.3	10.3	192	15.3	409	20.4

Table 2. Fund-raising.

In columns (1) and (2), we estimate probit models with vintage-year fixed effects. The dependent variable is an indicator variable set equal to one if the VC firm raises another fund, and zero otherwise. The variable of interest is the previous fund's ex post IRR. As this can only be measured once the fund's ten-year life is over, we include only funds from the 1980-1996 vintages. Column (1) restricts the sample to first-time funds and thus asks what determines whether the VC firm raises a second fund. In column (2), we widen the sample to include all funds (not just first-time funds) to see if the performance-sensitivity is attenuated for later funds. In column (3), we estimate a Cox semi-parametric hazard model with time-varying covariates using annual data. This models the hazard (i.e., the instantaneous probability) that a VC firm raises a new fund in year t. We allow a VC firm to raise multiple funds in succession (i.e., we estimate a "multiple-failure" hazard model). The hazard model conditions not on the ex post IRR of the previous fund, but on the interim (i.e., ex ante) IRR as reported at the end of year t-1. (Where a VC firm operates multiple funds in parallel as of the prior year-end, we use the maximum IRR.) Thus, unlike the probit models in columns (1) and (2), the hazard model uses only information that was available at the time of fund-raising. We ignore the first four years of reported IRRs over a fund's life as the IRR of a fund that is mainly investing and not yet generating returns is not meaningful. Practically, this affects only first-time funds as VC firms with later funds nearly always have a prior fund that is at least four years old. For the purposes of the hazard model, first-time funds are treated as left-censored during their first four years, and the likelihood function is adjusted accordingly. Since we do not require ex post IRRs for estimating the hazard model, we can include all available vintages through 2006. Since VC firms have a non-zero probability of raising further funds after our data end in 2006, the hazard model adjusts for right-censoring. Heteroskedasticity-consistent standard errors are shown in italics. We use ***, **, and * to denote significance at the 1%, 5%, and 10% level (two-sided), respectively.

	Prob(follow-on fund raised)		
	First-time funds	All funds	All funds
	(1)	(2)	(3)
log fund size	-0.012	0.266***	0.389***
	0.137	*0.518*	*0.042*
fund IRR (ex post)	5.861***	1.392**	
	1.770	*0.687*	
fund IRR (ex ante)			0.230***
			0.065
Vintage year FE	yes	yes	n.a.
Wald test: all coeff. = 0	30.4***	44.8***	131.2***
Pseudo-R^2	32.6%	11.6%	2.7%
No. of observations	191	598	3,721
No. of VC firms			262
No. of funds raised			620
Model estimated	Probit	Probit	Hazard

Table 3. Evolution of Fund Size.

The dependent variable in each column is the log of the size of the follow-on fund (in $m) if the firm raises a follow-on fund and zero if it does not. To code failure to raise a follow-on fund, we identify 362 defunct VC firms in *CapitalIQ*. Performance data is available for 71 and 126 funds raised by defunct VC firms in columns (1) and (2), respectively. The models are estimated using Tobit. The variable of interest in column (1) is the ex post IRR of the previous fund, which can either be a first-time and or a follow-on fund. As this can only be measured once the fund's ten-year life is over, we include only data for previous funds from the 1980-1996 vintages. In column (2), we replace this variable with the interim IRR of the previous fund measured as of the year-end prior to the year the GP raises the current fund. (Where a VC firm operates multiple funds in parallel as of year-end prior, we take the maximum IRR.) If no follow-on fund is raised, the IRR of the previous fund is measured ex post (i.e., as of year ten.) In column (2), it is not necessary to restrict the vintages. Standard errors are shown in italics. We use ***, **, and * to denote significance at the 1%, 5%, and 10% level (two-sided), respectively. Note that the Tobit estimator does not support a heteroskedasticity correction.

	Log size of follow-on fund	
	(1)	(2)
log size of previous fund	0.844***	0.892***
	0.069	*0.068*
IRR of previous fund (ex post)	0.553***	
	0.172	
IRR of previous fund (ex ante)		0.593***
		0.156
Vintage year FE	yes	yes
Wald test: all coeff. = 0	222.2***	322.8***
Pseudo-R^2	10.4%	10.2%
No. of observations	534	726
Model estimated	Tobit	Tobit

Table 4. Evolution of GP Carry.

The dependent variable in each column is the GP's carried interest ("carry"). The variable of interest in column (1) is the ex post IRR of the previous fund. As this can only be measured once the fund's ten-year life is over, we include only data for previous funds from the 1980-1996 vintages. In column (2), we replace this variable with the interim IRR of the previous fund measured as of the year-end prior to the year the GP raises the current fund. (Where a VC firm operates multiple funds in parallel as of year-end prior, we take the maximum IRR.) In this case, it is not necessary to restrict the vintages. Since we condition on the performance of the previous fund, the estimation samples in columns (1) and (2) are restricted to follow-on funds and the models are estimated using OLS. The OLS results could be biased to the extent that poor performance results in a VC firm being unable to raise a follow-on fund (left-censoring). In column (3), we estimate a Tobit model where we set the dependent variable equal to zero if the firm fails to raise a follow-on fund. To code failure to raise a follow-on fund, we identify 362 defunct VC firms in *CapitalIQ*. Performance data is available for 81 raised by defunct VC firms. Standard errors are shown in italics; in the case of the OLS specifications, they are heteroskedasticity-consistent. We use ***, **, and * to denote significance at the 1%, 5%, and 10% level (two-sided), respectively.

	GP carry in follow-on fund		
	(1)	(2)	(3)
log size of current fund	-0.001	0.000	0.054***
	0.003	*0.003*	*0.000*
IRR of previous fund (ex post)	0.013**		
	0.005		
IRR of previous fund (ex ante)		0.037***	0.085***
		0.006	*0.000*
Vintage year FE	yes	yes	yes
Wald test: all coeff. = 0	1.8*	8.1***	163.1***
Adjusted R^2	7.2%	21.5%	n.a.
No. of obs	86	197	272
Model estimated	OLS	OLS	Tobit

Table 5. Cross-fund Standard Deviation of Fund Size and GP Carry.
The table reports the results of variance ratio tests comparing the standard deviation of fund size (Panel A) or fund carry (Panel B) for first and follow-on funds. We use three different samples: The full set of funds raised between 1980 and 2006; the set of funds raised between 1980 and 1996; and the set of 1980-1996 funds for which we have performance information. We use ***, **, and * to denote significance at the 1%, 5%, and 10% level (two-sided), respectively.

	All funds raised in 1980-2006		All funds raised in 1980-1996		All funds raised in 1980-1996 with IRR information	
	No. of funds	St.dev.	No. of funds	St.dev.	No. of funds	St.dev.
Panel A: Fund size ($m)						
First-time funds	901	81.0	400	63.6	191	40.7
Follow-on funds	1,906	344.1	751	127.1	407	159.0
F-statistic		18.0***		4.0***		15.2***
Panel B: GP carry (%)						
First-time funds	50	1.6	14	1.5	8	0.0
Follow-on funds	317	4.0	59	3.2	47	3.0
F-statistic		6.6***		4.5***		n.a.

Table 6. Performance Persistence.

In Panel A, the dependent variable in each column is a fund's ex post IRR, net of fees, measured over its ten-year life. Accordingly, the sample is restricted to funds raised between 1980 and 1996. To test for persistence of performance across funds managed by the same VC firm, we regress the ex post IRR of fund N on the ex post IRR of fund N-1 (column 1), the ex post IRRs of funds N-1 and N-2 (column 2), and the ex post IRR of fund N-2 (column 3). In column (4), we use the interim IRR of fund N-1 measured as of the year-end prior to the year the GP raises the current fund. (Where a VC firm operates multiple funds in parallel as of the prior year-end, we take the maximum IRR.) In Panel B, we measure performance using exit rates, defined as the fraction of a fund's investments that were exited through an IPO or an M&A transaction over the course of the fund's ten-year life. All models are estimated using OLS. Heteroskedasticity-consistent standard errors are shown in italics. We use ***, **, and * to denote significance at the 1%, 5%, and 10% level (two-sided), respectively.

Panel A	IRR of current fund (ex post)				
	(1)	(2)	(3)	(4)	(5)
log size of previous fund	0.065***	0.071***	0.058**	0.120***	0.103***
	0.014	*0.019*	*0.027*	*0.036*	*0.037*
dummy =1 if early-stage focus	0.100***	0.114**	0.083	0.128**	0.118**
	0.033	*0.052*	*0.058*	*0.058*	*0.057*
IRR of fund -1 (ex post)	0.797***	0.741***			
	0.147	*0.188*			
IRR of fund -2 (ex post)		0.184	0.671**		
		0.259	*0.324*		
IRR of fund -1 as of previous year (ex ante)				0.924**	0.214
				0.435	*0.329*
… X (sequence number = 4 or higher)					1.011*
					0.512
Vintage year FE	yes	yes	yes	yes	yes
Wald test: all coeff. = 0	5.1***	5.9***	6.4***	2.2**	2.1**
Adjusted R^2	33.8%	26.7%	16.9%	13.6%	15.6%
No. of observations	318	163	177	189	189

Panel B	Exit rates (ex post)		
	(1)	(2)	(3)
log size of previous fund	0.022**	0.027*	0.037**
	0.010	*0.014*	*0.015*
dummy =1 if early-stage focus	0.051**	0.053**	0.054**
	0.020	*0.025*	*0.026*
exit rate of fund -1 (ex post)	0.378***	0.267***	
	0.062	*0.095*	
exit rate of fund -2 (ex post)		0.236**	0.331***
		0.114	*0.111*
Vintage year FE	yes	yes	yes
Wald test: all coeff. = 0	6.3***	5.7***	4.7***
Adjusted R^2	18.0%	21.7%	12.7%
No. of observations	486	269	285

Table 7. Performance of First-time Funds Relative to Follow-on Funds.
We test for differences in average performance by first-time versus follow-on funds using OLS regressions of fund IRR that control for log fund size, fund stage focus, and vintage-year fixed effects. The dependent variable in each column is a fund's ex post IRR, net of fees, measured over its ten-year life. Accordingly, the sample is restricted to funds raised between 1980 and 1996. Heteroskedasticity-consistent standard errors are shown in italics. We use ***, **, and * to denote significance at the 1%, 5%, and 10% level (two-sided), respectively.

|	IRR of fund (ex post)		
	(1)	(2)	(3)
log size of fund	0.043*	0.030	0.031
	0.022	*0.023*	*0.023*
dummy =1 if early-stage focus	0.086***	0.080***	0.079***
	0.028	*0.028*	*0.027*
dummy =1 if follow-on fund	-0.014		
	0.036		
dummy for funds raised at least 10 years after VC firm's first fund		0.085**	
		0.043	
log years since VC firm raised its first fund			0.036**
			0.018
Vintage year FE	yes	yes	yes
Wald test: all coeff. = 0	5.9***	6.1***	6.1***
Adjusted R^2	16.1%	16.7%	15.9%
No. of observations	598	598	598
Model estimated	OLS	OLS	OLS

Table 8. Oversubscription.
We obtain data on target fund sizes and final amounts raised per fund by searching January issues of the *Private Equity Analyst* in the three years centered on each fund's vintage year, as reported by VE or PREQIN. The *Private Equity Analyst* provides this information from 1991, so the sample is restricted to the 1991-2006 vintages. The dependent variable is the subscription ratio, that is, the ratio of the final amount raised and the original target fund size. These models are estimated as OLS regressions with vintage-year fixed effects. Column (1) uses all funds for which data on actual and target fund size can be found in the *Private Equity Analyst*. Column (2) restricts the sample to follow-on funds. Heteroskedasticity-consistent standard errors are shown in italics. We use ***, **, and * to denote significance at the 1%, 5%, and 10% level (two-sided), respectively.

|	Amount raised / target amount	
	(1)	(2)
log target fund size	0.030**	0.006
	0.013	*0.023*
dummy =1 if follow-on fund	0.069**	
	0.033	
IRR of fund -1 as of previous year (ex ante)		0.125***
		0.035
Vintage year FE	yes	yes
Wald test: all coeff. = 0	5.0***	3.4***
Adjusted R^2	6.2%	8.0%
No. of obs	908	367
Model estimated	OLS	OLS

Figure 1. Symmetric Learning – Probability of Follow-On Fund, Mean Size of Follow-On Fund, and Mean Carry of Follow-On Fund

The figures illustrate Implications 1, 2 and 3 in the case with symmetric learning. Panel A shows the probability that a follow-on fund is raised as a function of the realized return on the first-time fund, $P(\mu^i > \mu^*|r_1)$. Panel B shows expected (i.e. mean) size of the follow-on fund as a function of the realized return on the first-time fund, $E(I_2|r_1, \mu^i > \mu^*)$. Panel C shows the expected (i.e. mean) size of the carry in follow-on funds as a function of the realized return on the first-time fund, $E(f_2|r_1, \mu^i > \mu^*)$. The model parameters used are $r = e^{0.14} - 1$ (and thus $\mu^* = 0.14$), $\mu^L = 0.04$, $\mu^H = 0.4$, and $\sigma = 0.1$.

Panel A. Probability of Raising a Follow-On Fund
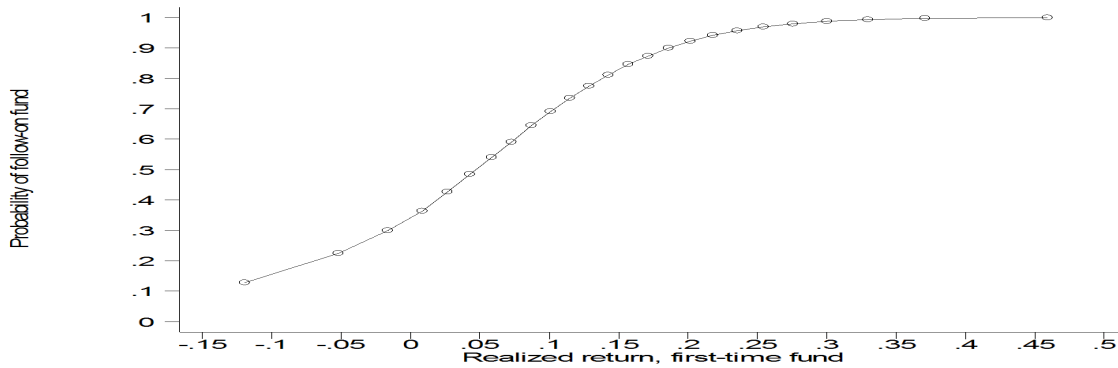
Panel B. Mean Size of Follow-On Fund
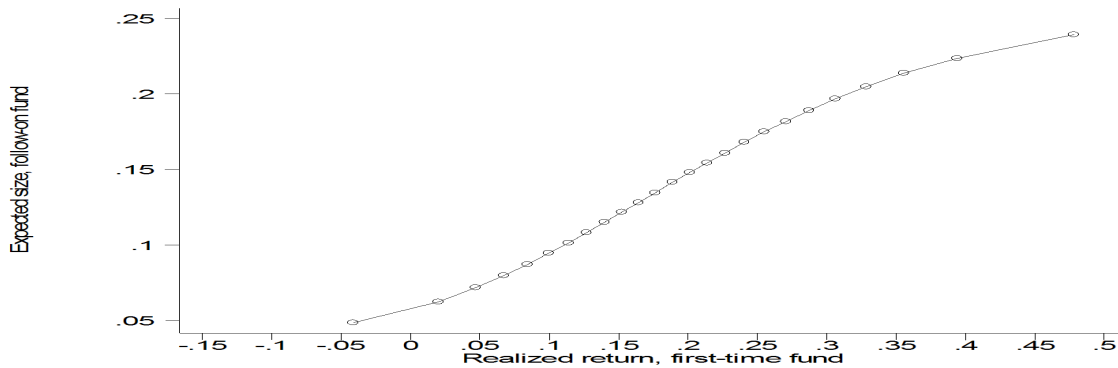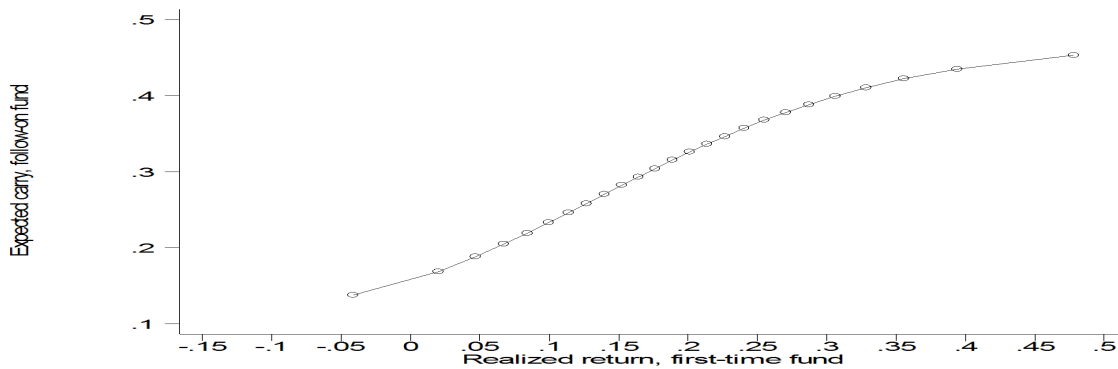

Panel C. Mean Carry of Follow-On Fund

Figure 2. Asymmetric Learning – Performance Persistence

The figure illustrates Implication 6. It plots the expected (i.e. mean) return in follow-on funds as a function of the realized return on the first-time fund. The model parameters used are $r = e^{0.14} - 1$ (and thus $\mu^* = 0.14$), $\mu^L = 0.04$, $\mu^H = 0.4$, and $\sigma = 0.1$.