Informational Hold-up and Performance Persistence in Venture Capital*

Yael V. Hochberg
Northwestern University and NBER

Alexander Ljungqvist
New York University and CEPR

Annette Vissing-Jørgensen
Northwestern University and NBER

July 22, 2010

Abstract

Why don't successful venture capitalists eliminate excess demand for their follow-on funds by aggressively raising their performance fees? We propose a theory of learning that leads to informational hold-up in the VC market. Investors in a fund learn whether the VC has skill or was lucky, whereas potential outside investors only observe returns. This gives the VC's current investors hold-up power when the VC raises his next fund: Without their backing, he cannot persuade anyone else to fund him, since outside investors would interpret the lack of backing as a sign that his skill is low. This hold-up power diminishes the VC's ability to increase fees in line with performance. The model provides a rationale for the persistence in after-fee returns documented by Kaplan and Schoar (2005) and predicts low expected returns among first-time funds, persistence in investors from fund to fund, and over-subscription in follow-on funds raised by successful VCs. Empirical evidence from a large sample of U.S. VC funds raised between 1980 and 2006 is consistent with these predictions.

The performance of 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 performance 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, thereby increasing the dollar fees fund managers can earn?

As Berk and Green (2004) show in the context of mutual funds, if investors supply their capital competitively but fund management skill is scarce, investors' expected excess returns must equal zero, realized returns must be unpredictable from public information, and fund managers will earn economic rents reflecting their skill. This fits the mutual fund industry, where returns do not appear persistent, but not the VC industry. Instead, we argue that to explain performance persistence in VC funds, the investor market must be uncompetitive in some way, forcing VCs to share the rents their skills generate with investors.

A constant level of market power among investors over time is not enough to generate persistence. Suppose there is a permanent shortage of investors willing to tie up their capital for the ten-year duration that is common in VC funds. Market power then implies that investors earn positive expected excess returns, by virtue of sharing in the VC's rents, but these expected returns, though positive, must be equal across funds (holding risk constant). Moreover, realized returns must remain unpredictable from public information or investors could improve their expected returns by reallocating their capital across VCs. Thus, to explain persistence, we need investors' market power to have increased when a VC raises his next fund.

In this paper, we propose a model of learning and informational hold-up in the VC market that can explain performance persistence. We model fund managers (the general partners or GPs) as potentially managing a sequence of two funds, each lasting one period. There is a large set of risk-neutral potential investors (the limited partners or LPs) so that at the beginning of the first period, the LP market is perfectly competitive. Whether the second fund is raised depends on what is learned

1

about GP skill during the first fund. The key ingredient of the model is that investing in a fund gives the LP an opportunity to collect 'soft' information about the GP's skill, whereas the market can observe only 'hard' information such as realized returns. Access to soft information gives the investor an informational advantage over the market when it comes to distinguishing between skill and luck.[1] Soft information is arguably particularly important in the VC industry: VCs invest in risky, unlisted, and hard-to-value companies which they hold for a number of years before eventually selling them or, more often, writing them off. Objective returns thus take many years to materialize, unlike in the mutual fund industry where managers invest in traded securities that can easily and objectively be valued, typically every quarter.[2]

The asymmetric evolution of information in our model enables LPs to hold the GP up when he next raises a fund, because other potential investors in the market would interpret failure to reinvest by the 'incumbent' LP as a negative signal about the GP's skill. Specifically, outside investors face a winner's curse—the better-informed incumbent LPs will outbid them in a follow-on fund whenever the GP has skill—and so withdraw from the market for follow-on funds. This gives incumbent LPs bargaining power when negotiating follow-on investments with GPs and leads to performance persistence: Net of the manager's performance fee (or 'carry'), high LP returns in a first fund predict high LP returns in a follow-on fund as the hold-up problem prevents the manager from raising his carry to the point where investors just break even.

The model also predicts that average returns are lower in first than in follow-on funds. Because the LP market is competitive ex ante, and LPs realize they will enjoy market power ex post, GPs are funded on 'too good' terms in the first fund, but on average 'pay this back' to their LPs later. Effectively, investing in a first fund entails a valuable option to invest in a follow-on fund, which will be exercised only if the GP turns out to be skilled. Thus, the model implies that conditional on a

---

[1] For empirical evidence of the importance of soft information in learning about *corporate* managers' skill, see Cornelli, Kominek, and Ljungqvist (2010).

[2] 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." Lerner and Schoar (2004) argue that LPs typically demand wide-ranging information rights in order to inform their decision whether to reinvest. Chung et al. (2010) use a learning model to calibrate the incentive effects of future fundraising in the VC market.

follow-on fund being raised, LPs who invested in a GP's first fund should reinvest in the GP's next fund. After all, LPs earn a return in excess of their opportunity cost of capital in follow-on funds. The same argument provides a rational explanation for oversubscription in follow-on funds, especially in those raised by GPs whose first funds generated high returns.

We verify these predictions with one of the most comprehensive datasets on U.S. VC funds assembled to date. Unlike Kaplan and Schoar (2005), who have access only to anonymized fund performance data, we know the identity of each fund in our dataset. We are thus able to merge in data from other sources, such as fund characteristics, carry information, and LP identities. While each of these data items has individually been the subject of prior work, we not only put all the pieces of the puzzle together but have also collected data for many more funds than prior work. Like all studies in this area, our data suffer from selective reporting (as funds are not required to disclose data), but arguably less so than prior work.

Our results confirm that VC performance is persistent, as in Kaplan and Schoar (2005). Consistent with our model, we also find that average returns are higher in later funds, that LPs 'chase performance' by reinvesting following higher first-fund returns, and that this leads to oversubscription. However, these results are consistent with *any* model in which the LP market becomes less competitive over time. To get at the economic mechanism at the heart of our model—informational hold-up—we need to provide evidence of asymmetric learning. To this end, we test whether incumbent LPs behave in ways that suggest they have soft information about GP skill.

How to capture soft information? Most VCs, in practice, raise their next fund well before the end of their current fund's life. Thus, their current fund's eventual (or ex post) return is not yet known when they go out fundraising. All the market knows at this point is an interim return number. While the interim return constitutes hard information, by construction it reflects a mixture of objective cash-on-cash returns and subjective unrealized capital gains.[3] (Unlike previous authors, we have access to these interim returns.) Incumbent LPs observe the reported interim return as well, but in

---

[3]Blaydon and Horvath (2002) document that absent agreed valuation standards in the VC industry, different VC funds report radically different valuations for the same portfolio companies at a given point in time.

our model they also possess soft information, say knowledge of whether the GP's unrealized capital gains are likely to materialize or to evaporate. Soft information allows incumbent LPs to learn the GP's skill and thereby helps them predict the current fund's eventual ex post return. Based on this argument, we treat a current fund's future ex post return (which will be revealed many years later) as a proxy for the soft information that incumbent LPs have at the time the GP raises his next fund.

When we include both a prior fund's current interim and future ex post returns in the persistence regressions, we find that the hard information publicly available at the time of fundraising does *not* predict the subsequent performance of the GP's next fund, whereas our proxy for soft information does. Consistent with the informational assumptions of our model, this suggests that performance is not predictable from public information, such as interim returns, but that it *is* predictable from future ex post returns about which incumbent LPs may be better informed. Moreover, we show that the fraction of incumbent LPs who reinvest in the next fund is unrelated to public information (in the form of interim returns) but strongly increases in the as-yet unknown future ex post return. Both patterns are consistent with asymmetric learning: Incumbent LPs reinvest before the return on the GP's current fund is publicly known, and such GPs continue to perform well on their next fund.

A corollary of the assumption that incumbent LPs learn to distinguish between luck and skill is that some low-performing first-time funds should keep their LPs ('unlucky funds') while some high-performing first-time funds should lose their LPs and so close ('lucky funds'). The data support these patterns. Furthermore, persistence in lucky funds that survive and in unlucky funds should be lower than for other funds as their luck runs out or their bad luck ends in the next fund, respectively. Again, our evidence is consistent with this prediction.

Our paper is related to the literature on relationship-banking, which employs similar informational assumptions (e.g. Sharpe (1990), Rajan (1992), von Thadden (2004)), and to the literature on learning more generally in financial markets (see Pastor and Veronesi (2009) for a recent survey). However, unlike in hold-up models in the banking literature, asymmetric learning is efficient in the VC setting. This follows because VC contracts specify both an investment level (fund size) and the division of

the fund's surplus between GP and LPs, and fund size is NPV-maximizing in both first and follow-on funds. Moreover, GPs may even benefit from informational hold-up ex ante, because under certain conditions, first funds can only be raised under asymmetric learning. Such a preference is consistent with the fact that GPs are willing to provide their LPs with considerable amounts of soft information about strategies and performance which cannot credibly be communicated to potential new LPs.[4]

In addition, our paper relates to the literatures on VC performance and the relationship between LPs and GPs. Kaplan and Schoar (2005) are the first to report evidence of performance persistence in VC funds. In addition, they document a positive and concave relation between performance and future fund-raising, which we show to be consistent with both symmetric and asymmetric learning. Jones and Rhodes-Kropf (2003) provide empirical evidence in support of the hypothesis that VCs need to be compensated for bearing idiosyncratic risk through higher expected returns. Ljungqvist and Richardson (2003) analyze the cash flow, return, and risk characteristics of private equity funds. Cochrane (2005), Korteweg and Sørensen (2010), Quigley and Woodward (2003), and Gottschalg and Phalippou (2009) 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 when investing in private equity and suggest that LPs vary in their level of sophistication.

Finally, in contemporaneous work, Glode and Green (2010) provide a model to explain performance persistence in the hedge fund industry. The Glode and Green (2010) model also exhibits learning that generates ex ante competitive and ex post uncompetitive markets. In their model, which captures a primary concern of hedge fund managers, incumbent LPs learn about the profitability of a 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 in the VC industry as complementary to that of Glode and Green in the hedge fund industry. In

---

[4]Intuitively, if the average NPV across GP types is negative, first funds cannot be funded with symmetric learning. Asymmetric learning (the provision of soft information) serves as a commitment device for GPs who turn out to have high 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 invest in first funds even if they do not earn their opportunity cost of capital on first funds.

our model, which emphasizes the role of soft information, GPs *cannot* credibly reveal their type to outside LPs (and may optimally set up their 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, depending on the institutional setting. If GP skill refers to a comparative advantage in screening or monitoring investments (as the VC literature tends to assume), it is unlikely that LPs can 'steal' this skill. If, on the other hand, GP skill refers to knowledge of which industries are worth investing in or of what trading strategies are profitable (as in hedge funds), LPs might well steal that knowledge.

The remainder of the paper is structured as follows. Section I presents our model of learning and informational hold-up. Section II presents the sample and data. Section III presents the empirical analysis, and Section IV discusses and concludes.

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.[5] 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 $\left[\mu^L, \mu^H\right]$.[6] We abstract from

---

[5]This is similar to Berk and Green's (2004) assumption for mutual funds and consistent with the evidence reported for private equity funds in Lopez de Silanes, Phalippou, and Gottschalg (2010).

[6]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.

agency problems by assuming that GPs manage their funds in their LPs' best interest.[7]

Limited partners: There is a large set of ex ante identical, risk-neutral investors so that the LP market is perfectly competitive at $t=0$. To capture the possibility of absence of perfect competition at $t=1$, we assume that a single LP is sufficient to finance each fund.[8] This assumption is stronger than required and our results go through with multiple LPs financing each fund, as long as they do not compete all the rents away at $t=1$. As discussed in the introduction, with perfect competition among LPs ex post, there cannot be any persistence.[9] We distinguish between the 'incumbent' LP who has invested in the GP's first fund, and 'outside' LPs who have not. 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 soft information that allows them to infer the GP's type, μ^i, perfectly. Outside LPs only observe hard information, in the form of the first fund's cash flows, C_1.[10] We refer to this setup as asymmetric learning, meaning that the incumbent LP learns the GP's type faster than do outside investors. We distinguish this setup from one with symmetric learning in which both incumbent and outside LPs learn the GP's type perfectly at $t=1$.

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 $NPV_2\left(\mu^i\right) = 0$, i.e. for which the NPV at the NPV-maximizing investment level is zero. The NPV-maximizing investment-level for a given value of μ^i is $I_2\left(\mu^i\right) = \frac{e^{\mu^i}}{1+r} - 1$, so $NPV_2\left(\mu^i\right) = \frac{e^{\mu^i}\left[\mu^i - \ln(1+r)\right]}{1+r} - \left[\frac{e^{\mu^i}}{1+r} - 1\right]$. Therefore, $NPV_2\left(\mu^i\right)$ (and $I_2\left(\mu^i\right)$) equals zero for $\mu^* = \ln\left(1+r\right)$. Assume that $\mu^L < \mu^* < \mu^H$.

---

[7]For a model of agency problems among fund managers in a learning setting, see Ljungqvist, Richardson, and Wolfenzon (2007).

[8]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.

[9]In practice, in our data, funds have fewer than nine LPs on average, and based on the empirical results of Lerner, Schoar, and Wongsunwai (2007), only a minority of these are likely to have the experience and sophistication to learn the GP's type. The remaining LPs may be thought of as a competitive fringe.

[10]This setting corresponds to 'passive monitoring' by investors in Tirole's (2006) terminology.

Payoff functions: To characterize the division of a fund's surplus, we assume the following contract, which is a standard feature in nearly all VC funds. The LP receives all cash flows up to the amount he invested; additional cash flows are divided according to a linear sharing rule, with the GP receiving a fraction f, representing his performance fee or 'carry.'[11]

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. Asymmetric Learning

Under asymmetric learning, the LP market is perfectly competitive at $t=0$ but not at $t=1$. Because outside LPs do not learn the GP's type, incumbent LPs have an informational advantage over outside investors when the GP attempts to raise a 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. In order to identify the source of LPs' hold-up power, we model the bargaining game explicitly.

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 could represent 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

---

[11]In practice, VC contracts also include an annual management fee (often 2% of fund size) that does not vary with performance. Our model does not have enough elements to pin down both an optimal carry and an optimal management fee. For simplicity, we therefore set the management fee to zero, but we would obtain similar results if we allowed for a positive fee. This would simply reduce the optimal carry to yield the same expected payoff to LPs. For details of VC fees, see Metrick and Yasuda (2009).

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.

Appendices A and B show that sequential bargaining results in the LP accepting the GP's first offer. For a given GP type μ^i, both the GP's and the LP's offer sets fund size at the NPV-maximizing level, since it is in neither party's interest to reduce the size of the surplus available to be shared:

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

Only funds with $\mu^i > \mu^*$ are raised. As the shrinkage factor $p \to 0$, 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 Appendices A and B. In what follows, we focus on the case where $p \to 0$.

Equal division of the fund NPV is without loss of generality. As in Binmore, Rubinstein, and Wolinsky (1986), an unequal division would result if the parties had different discount rates, if they had different beliefs about the shrinkage factor p, or if one party was able to make offers faster than the other party.

Our assumption that a single LP is sufficient to finance each fund allows us to solve the model using standard bilateral bargaining tools. In practice, VC funds have several LPs. Still, the assumption of a representative LP is not restrictive as long as the presence of multiple incumbent LPs does not lead to the informational rents being competed away in the style of Bertrand competition. Empirically, this

is, of course, indirectly testable, for if the rents *were* competed away, the implications of our model should be rejected in the data. Conceptually, the standard argument against the Bertrand outcome is the Edgeworth solution (see, for example, Tirole (1988), Section 5.2). The Edgeworth solution shows that with capacity constraints, the outcome of competition in an oligopolistic market will not be marginal-cost pricing but instead will involve agents earning economic rents. Our model could easily be extended in this way, for example by relaxing the assumption that LPs have sufficient wealth to finance the entire fund or more generally by assuming that LPs are institutionally prohibited from financing the entire fund, perhaps for diversification reasons. As in the Edgeworth solution, either of these would reduce LPs' incentive to compete on price (i.e., carry) in follow-on funds.

A second argument against multiple LPs competing away their rents comes from the literature on bargaining coalitions. Chae and Heidhues (2004) derive conditions under which an individual (LP) is better off joining a bargaining coalition than bargaining in competition with other individuals (LPs), and Vidal-Puga (2005) generalizes the Nash bargaining solution to games involving such coalitions. The upshot of these studies is that multiple LPs working in coalition can extract nonzero rents from the GP in equilibrium. Our representative-LP framework can be viewed as a reduced-form of these more extensive bargaining games.

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

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

(5)

Thus, expected LP returns net of carries vary with GP skill and so are *not* equalized across GP types in follow-on funds. This contrasts with the symmetric learning case modeled in Berk and Green (2004).[12] Furthermore, the average expected LP return in follow-on funds is

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

(6)

---

[12]In the symmetric learning case, f_2 is set to ensure that LPs earn an expected return of r in all follow-on funds.

Investment and carry in first-time funds: As no learning has taken place yet, the LP market is perfectly competitive at $t=0$. The GP can therefore offer any LP a contract $\left(f_1, I_1^{opt}\right)$ where I_1^{opt} is the NPV-maximizing fund size:

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

By the uniform distribution of GP types, $E_i \left(e^{\mu^i} \right) = \frac{1}{\mu^H - \mu^L} \left[e^{\mu^H} - e^{\mu^L} \right]$, and f_1 is chosen 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 $\mu^i \geq \mu^*$. Thus, f_1 solves:

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

$$f_1 = \frac{E_i \left(e^{\mu^i} \right) \ln \left(1 + I_1^{opt} \right) - (1 + r) I_1^{opt}}{E_i \left(g_1 \left(\mu^i \right) \right)} + \frac{\frac{1}{2} E_i \left(NPV_2 \left(\mu^i \right) \right)}{E_i \left(g_1 \left(\mu^i \right) \right)} \tag{8}$$

with $NPV_2 \left(\mu^i \right)$ calculated at the NPV-maximizing investment level. For notational simplicity, we denote $E(max(0, C_1 - I_1(\mu^i))|\mu = \mu^i)$ by $g_1(\mu^i)$. Closed-form expressions for $g_1(\mu^i)$ and f_1 can be found in Appendices A and C, respectively.

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

$$1 + E_i \left(E \left(r_1 | \mu = \mu^i \right) \right) = \frac{E_i \left(E \left(X_1^{LP} | \mu = \mu^i \right) \right)}{I_1^{opt}} = \frac{E_i \left(e^{\mu^i} \right) \ln \left(1 + I_1^{opt} \right) - f_1 E_i \left(g_1 \left(\mu^i \right) \right)}{I_1^{opt}}. \tag{9}$$

C.   Optimality of Asymmetric Learning

Learning is valuable whether it happens symmetrically (with incumbent and outside LPs learning about GP skill at the same speed) or asymmetrically (with incumbent LPs learning faster than outside LPs). It ensures that skilled GPs receive more capital in follow-on funds and that low-skill GPs exit the industry. This increases the overall value created by the 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).

Can *asymmetric* learning among incumbent and outside LPs lead to more efficient investment outcomes than *symmetric* learning? In our setting it might, namely if the expected NPV in the population of GPs, at the optimal investment level, is negative (i.e., if $\frac{E_i\left(e^{\mu^i}\right)\ln\left(1+I_1^{opt}\right)}{1+r} - I_1^{opt} < 0$). In this case, a GP would not be able to raise a first-time fund (nor any follow-on funds) if learning was symmetric. However, with asymmetric learning, LPs earn informational rents in follow-on funds and this may be sufficient to make up for the expected losses on first-time funds. This will be the case if there is enough dispersion in GP skill. Effectively, with asymmetric learning, an investment in a first fund gives the LP an option to invest in a follow-on fund, and the value of this option increases in uncertainty about GP skill. If the option value exceeds the expected loss on first funds, LPs will invest in first funds despite a negative expected NPV.[13]

Soft information about skill effectively commits GPs to sharing the follow-on fund NPV with their LPs, 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 in the banking literature such as Sharpe (1990), but there the gains from long-term contracting must be weighed against inefficient 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). The fact that contracts between GPs and LPs provide exclusive informational rights to incumbent LPs while prohibiting LPs from sharing such information is consistent with GPs recognizing that subjecting themselves to informational hold-up may be value-increasing for GPs.

Of course, even if subjecting themselves to informational hold-up is value-increasing ex ante, it is clear that GPs who subsequently learn that they have skill will have an incentive to signal their type to outside investors prior to raising a follow-on fund. In practice, skilled GPs do try to signal their type, but they are unlikely to do so with sufficient precision to eliminate the information asymmetry

---

[13]This efficiency argument mirrors Tirole's (2006), though Tirole considers a setting where a borrower's profitability improves over time, while we consider a setting where average profitability improves over time as low-skill GPs exit.

between incumbent and outside investors. For example, one way that skilled GPs try to signal is by taking portfolio firms public earlier than may otherwise be optimal. This phenomenon is known as grandstanding (Gompers (1996)). Grandstanding is unlikely to fully reveal the GP's type, however, since the number of IPOs is unlikely to be fully informative about skill.

Finally, perhaps long-term contracting should be possible even with symmetric learning, since hard information is verifiable, and so courts could 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.

D.   Empirical Implications

If LPs learn GP skill in an asymmetric way, our model yields the following empirical implications:

Implication A1: 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., $E\left(r_2|r_1, \mu^i > \mu^*\right)$. increases in r_1.

Implication A2: 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 A3: 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 A4: 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.

Implication A5: Fundraising. Some GPs with poorly performing first-time funds will be able to raise follow-on funds, and some GPs with high-performing first-time funds will not be able to raise follow-on funds.

Implication A1 is, of course, what the model is designed to capture. It is proved in Appendix D. One might think that outside LPs could simply invest in all follow-on funds with a high realized

r_1, thus expecting to earn a high value of $E\left(r_2|r_1\right)$. But our model makes it clear why this is not possible. The winner's curse problem described earlier implies that outside LPs would only be able 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 eliminating performance persistence breaks down in the VC setting when there is asymmetric learning.

Implications A2 and A4 follow directly from the fact that incumbent LPs share in the follow-on fund's NPV and thus earn a return in excess of their opportunity cost of capital at $t = 1$. Implication A3 follows from equations (8) and (9). Carries are set such that incumbent LPs expect to break even across the two funds. Since they earn informational rents in the follow-on fund, they must expect losses in first funds. Formally, the second term in (8) implies that f_1 is higher if learning is asymmetric, as the GP captures half the follow-on fund's NPV through a higher first-fund carry. Thus, (9) implies that $E_i\left(E\left(r_1|\mu = \mu^i\right)\right) < r$. Furthermore, since LPs earn half of the follow-on fund NPV, $r < E_i\left(r_2|\mu^i \geq \mu^*\right)$ and thus $E_i\left(E\left(r_1|\mu = \mu^i\right)\right) < E_i\left(r_2|\mu^i \geq \mu^*\right)$. Implication A5 follows from the fact that follow-on fund raising is based on whether μ^i exceeds the zero-NPV cutoff μ^*. Since r_1 is merely a noisy indicator of μi, some GPs will have $\mu^i \geq \mu^*$ despite having realized low r_1. Similarly, some GPs will have $\mu^i < \mu^*$ despite having realized high r_1. Incumbent LPs learn μ^i perfectly, and therefore will reinvest (or not) accordingly.

While we are particularly interested in the implications of asymmetric learning, we note that any model of learning, whether symmetric or asymmetric, would generate the following five implications about fund-raising and the first and second moments of fund size and carry:

Implication S1: Fund-raising. The probability that a GP raises a follow-on fund is increasing in the LP return of the GP's first fund: $P\left(\mu^i > \mu^*|r_1\right)$ increases in r_1.

Implication S2: 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: $E\left(I_2|r_1, \mu^i > \mu^*\right)$ increases in r_1.

Implication S3: 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: $E\left(f_2|r_1, \mu^i > \mu^*\right)$ increases in r_1.

Implication S4: 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 S5: Cross-fund standard deviation of GP carry. The cross-fund standard deviation of GP carry is higher among follow-on than among first funds: $SD_i\left(f_2|\mu^i > \mu^*\right) > SD_i\left(f_1^{opt}\right)$.

These implications are independent of how the GP and LP split the surplus and so hold even when learning is symmetric. Implications S1, S2, and S3 are not surprising, and we provide proofs in Appendix D. Implications S4 and S5 follow immediately from the fact that the cross-fund standard deviation of both fund size and carry in the model is zero in first-time funds and positive in follow-on funds.

E.   Extension: Multiple Follow-on Funds

How does persistence evolve over time if a GP can raise 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. If GP types are constant over time, as we have assumed so far, persistence should eventually diminish for sufficiently high fund sequence numbers as hard information (such as audited returns) accumulates, eventually allowing outside LPs to infer the GP's type perfectly.

Alternatively, GP type may be less than perfectly positively correlated over time. This would be the case, for example, if there is time variation in the team of partners, say when one of the partners retires or leaves to start a new fund or join a competing fund. This scenario implies quite different dynamics of persistence across a VC firm's funds. Some persistence would remain even for high fund sequence numbers, as incumbent LPs retain an informational advantage over outside LPs by receiving

soft information about the evolution of the team's type.[14] The following prediction summarizes:

Implication A6: Evolution of return persistence. Return persistence should initially increase in fund sequence. Then either (a) or (b) should occur:

(a) If GP type is constant over time, return persistence eventually disappears for sufficiently high fund sequence numbers. The R^2 of a regression of fund size on lagged fund size (with both adjusted for vintage effects to capture changes in optimal fund size due to changing investment opportunities for given GP type) should then be close to 1 for sufficiently high fund sequence numbers.

(b) If GP type is positively, but not perfectly, correlated over time, some return persistence remains for high fund sequence numbers and return persistence may not even diminish with fund sequence. The R^2 of a regression of fund size on lagged fund size (with both adjusted for vintage effects) should stay substantially below 1 even for high fund sequence numbers.

II.   Sample and Data

To examine whether the implications of our model are consistent with empirical patterns in the VC industry, we construct a sample of U.S. VC funds obtained from two databases, Thomson Reuters' Venture Economics (VE) and Private Equity Intelligence (PREQIN).[15] As Table 1 details, our sample contains 2,790 funds raised by 1,164 VC firms between 1980 and 2006. Of these, 783 funds are in both VE and PREQIN, 44 appear only in PREQIN, and the remaining 1,963 appear only in VE. The number of funds raised per year averages 64 in the 1980s, 137 in the 1990s, and 119 between 2001 and 2006. The average (median) sample fund raised \$124.7 million (\$49.9 million) in nominal dollars. Average fund size increased from \$30.1 million in 1980 to \$44.3 million in 1990, \$202.7 million in 2000, and \$217.2 million in 2006.

---

[14]Of course, if GP type is uncorrelated over time, learning is irrelevant and we should see no return persistence for any sequence number. Our model of performance persistence is relevant only if learning about GP type is important.

[15]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 such as *Pratt's Guide*, *CapitalIQ*, *Galante's*, and a web search. We screen out funds of funds, buyout funds, hedge funds, venture leasing funds, and evergreen funds (i.e., funds without a predetermined dissolution date).

VC funds are under no obligation to disclose performance data publicly. Based on data disclosed voluntarily by GPs and/or LPs, VE and PREQIN report IRRs, calculated net of fees and carries, for a subsample of funds.[16] 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). This realized return is the IRR measure commonly used in the empirical VC literature. Since VC funds typically have a ten-year life, this IRR measure accurately reflects ultimate, 'ex post' performance only in the case of ten-year-old sample funds, i.e. those raised between 1980 and 1996. For more recent sample funds, the reported IRR is liable to change as investments are subsequently exited or written off after the end of our sample period. Thus, whenever we use ex post IRRs, we restrict the sample to vintage years 1980-1996.

While PREQIN reports only realized IRRs, VE also reports 'interim' IRRs for each year between a fund's inception and the earlier of its liquidation or 2006. This allows us to track performance as it evolves over a fund's life, or more specifically, as it is reported to LPs over time.[17] Recall that interim IRRs reflect a mixture of objective cash-on-cash returns and unrealized capital gains which VCs can compute in a subjective fashion. Interim IRRs are available for 651 funds. The average (median) interim IRR in the entire sample is 27.7% (13.7%); funds raised between 1980 and 1996 reported lower average (median) interim IRRs of 9% (7.3%), averaged over their ten-year lives.

As Table 1 shows, we have ex post IRRs for 1,007 of the 2,790 funds. The average realized IRR is 14.1%, though this includes recent funds that have yet to switch from making investments to exiting them. The average IRR for the 601 funds from the 1980-1996 vintages is 18.8%, with a lower median of 10.3%.[18] Realized returns in the VC industry have varied considerably over time. Average ex post 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

---

[16]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.

[17]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.

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

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.

To investigate LPs' reinvestment decisions, we compile a large sample of LP fund holdings, using data obtained from Venture Economics, VentureOne, PREQIN, and *CapitalIQ*. None of the four data sources provides complete coverage of any given LP's investments, or of the LPs in any given fund.[19] Combining the four sources gives us 12,491 observations, where each observation is a pairing of an LP with a fund the LP invested in. We identify 1,878 distinct LPs investing in 1,526 sample funds. The average LP invests in 6.7 sample funds. Conditional on LP data being available, the average sample fund has 8.2 LPs.

III. Empirical Analysis

The main focus of our empirical analysis is on the role of asymmetric learning and soft information in explaining VC fund performance persistence and incumbent LPs' reinvestment decisions. We first replicate the motivating fact of our paper, namely that VC fund performance is persistent. We then ask what type of information—publicly available hard information or privately available soft information—predicts VC fund returns and find it to be the latter. We show that incumbent LPs tend to reinvest, as the model predicts, and that there are instances where a GP cannot raise a follow-on fund despite high realized returns. We test the follow-on hypothesis that apparently lucky but unskilled GPs and apparently unlucky but skilled GPs do not exhibit persistence, presumably because their luck eventually runs out or their bad luck eventually turns, and find support in the data. We also show that LPs' reinvestment decisions are sensitive to a proxy for soft information but not to publicly available hard information, consistent with learning. Finally, we examine whether first-time funds indeed have lower average returns than follow-on funds and what drives fund oversubscription.

Two of the implications that are generic to learning (in the sense that they apply whether learning is asymmetric or symmetric) have already been shown to hold in prior studies. At the end of this

---

[19]This drawback is also noted by Lerner, Schoar, and Wongsunwai (2007), who use VE data in a related exercise.

section, we briefly replicate existing tests of these implications, S1 and S2. We also provide novel evidence in support of Implications S3, S4, and S5, which have not previously been tested.

A.   Persistence, Learning, and Soft Information

A.1.   Performance Persistence

We begin by replicating Kaplan and Schoar's (2005) persistence test in our larger dataset. In column (1) of Table 2, 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 differences in risk-taking across funds, we also include a dummy variable that equals one for funds classified as investing in early-stage companies.[20] Like Kaplan and Schoar, we find that fund performance increases with fund size and—consistent with Implication A1—prior-fund performance ($p<0.001$).

One concern regarding the persistence result is selection bias: Not every VC fund 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 estimate a persistence regression with exit rates as the dependent variable instead of IRRs. Hochberg, Ljungqvist, and Lu (2007) define exit rates 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. Exit rates can thus be computed for all funds. As the estimates in column (2) show, we continue to find persistence using this alternative performance measure.

How long does persistence persist for? According to Implication A6, persistence should initially increase in fund sequence. To test for this, column (3) interacts the previous fund's ex post IRR with an indicator for fourth and higher-numbered funds. (The cut-off is arbitrary but not selective.) We find significant persistence among fourth and later funds but not among earlier funds, which is consistent with Implication A6. Our dataset contains too few higher-numbered funds with reported IRRs to reliably test whether persistence eventually diminishes.[21] We leave this prediction to be

---

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

[21]There is some evidence suggesting substantial persistence even among the highest fund sequence numbers. The R^2 from a regression of log fund size on previous log fund size (both adjusted for vintage effects by subtracting vintage means) is far below 1 even for eighth and higher-numbered funds (0.34) and is fairly similar to the R^2 obtained for the

tested in future data as IRRs become available for funds with still higher fund sequence numbers.

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. This, of course, implies that the previous fund's ex post return is not yet publicly known when the current fund is raised. To mitigate this problem, they suggest including the two prior funds' ex post IRRs. When we do so, in column (4), we find that only the immediately preceding fund's ex post IRR significantly predicts the current fund's future ex post IRR ($p<0.001$).

A.2. What Type of Information Predicts Returns?

The patterns in column (4) are interesting: The ex post IRR of the immediately preceding fund—which will not be known until several years after the current fund has been raised—predicts the current fund's future IRR, while the second prior fund's IRR—which is typically already known at the time the current fund is raised—does not. This hints at the possibility that future fund returns are better predicted using information about prior-fund performance that is not yet publicly available at the time of fundraising. This would be consistent with our informational assumptions as outside investors then could not 'chase returns,' while incumbent LPs might chase returns based on privately obtaining signals that are informative about the future performance of the GP's previous fund. We now ask whether incumbent LPs do learn soft information and whether this soft information indeed predicts returns.

To this end, we run a horse race between the prior fund's current interim and future ex post returns. Interim IRRs are usually audited and so constitute hard information. They are publicly available in real time to both incumbent LPs and, during fundraising, to the LP market at large. The ex post IRR, on the other hand, will not be publicly known until the prior fund has come to the end of its ten-year life, i.e., a few years after the current fund is raised. To the extent that incumbent LPs obtain soft information about the GP's skill, they should be in a better position to forecast the

---

same regression run for fourth or fifth funds (0.43) or for sixth and seventh funds (0.32). These relatively low R^2 are consistent with GP types that are positively correlated over time but not constant. Thus, Implication A6(b) implies that there should be substantial return persistence even for the highest fund sequence numbers.

prior fund's future ex post IRR. Thus, future ex post IRRs may correlate with what incumbent LPs currently know, allowing incumbent LPs to make better informed reinvestment decisions. Of course, it is an empirical question whether future ex post IRRs are a useful proxy for incumbent LPs' private information.

In column (5) of Table 2, we augment the column (1) specification by including both the prior fund's interim IRR (measured as of the year-end prior to the year the GP raised the current fund) and its future ex post return.[22] When we do so, we find that the hard information available to outsiders and incumbents at the time of fundraising—the prior fund's interim IRR—does *not* predict the next fund's performance (p=0.515), whereas the prior fund's as-yet-unrealized future ex post IRR does (p=0.007). This pattern is consistent with the informational assumptions of our model: Information not yet publicly known at the time of fundraising (i.e., ex-post IRRs) predicts returns on follow-on funds above and beyond hard information known at the time of fundraising (i.e., interim IRRs). We will shortly provide further evidence consistent with this interpretation when we model LP reinvestment decisions.

A.3.   Soft Information, Luck, and Skill

Implication A5 captures a corollary of our model. If incumbent LPs do obtain soft information that enables them to better distinguish between skill and luck, as we have assumed, we should see two patterns in the data. First, we should see some high-return first-time GPs who lose the backing of their LPs and therefore cannot raise a follow-on fund. This would happen if incumbent LPs came to the conclusion that an unskilled GP had simply been lucky. Second, we should see some low-performing funds whose LPs, despite poor performance, reinvest in a follow-on fund. This would be the case when a skilled GP had simply been unlucky.

To test Implication A5, we sort all first-time funds into five quintiles based on their ex-post

---

[22] 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.

IRR and then compute the fraction of GPs in each quintile who are able to raise a follow-on fund. Consistent with the notion that LPs employ soft information in distinguishing between skill and luck in making their reinvestment decisions, we find that 49% of the GPs with the worst performing first funds are able to raise follow-on funds, while 6% of the top performing funds are unable to raise a follow-on fund. These patterns support Implication A5.

A follow-on prediction is that persistence should be lower both among GPs who lose LPs after good performance ('lucky' GPs) and among GPs who keep LPs despite bad performance ('unlucky' GPs). This follows because skill should be a stronger predictor of future performance than either good luck or bad luck. Using the LP data described earlier, we interact LP reinvestment rates with realized returns to construct proxies for which high-performing funds were likely to have been 'lucky' rather than skilled and which low-performing funds were likely simply 'unlucky' rather than unskilled. Specifically, we independently double-sort GPs' current funds into 3x3 bins based on their prior funds' IRRs and the fraction of their current LPs who also invested in their previous fund. We denote funds in the top performance tercile and bottom reinvestment tercile as 'lucky' while those in the bottom performance tercile and top reinvestment tercile are denoted 'unlucky.'

In column (6) of Table 2, we include in the persistence regression interactions between the prior fund's ex post IRR and indicators for 'lucky' and 'unlucky' funds. Consistent with our prediction, we find that both lucky and unlucky funds experience significantly lower performance persistence than do other funds. In fact, we cannot reject the hypothesis that persistence for these two groups is zero. (The p-values are 0.40 and 0.51, respectively, for the test that the sum of the coefficients of the prior fund's IRR and the relevant interaction term is zero.) Thus, the prior performance of lucky and unlucky GPs is uninformative about their future performance, as predicted. Since our identification of such GPs is based on incumbent LPs' reinvestment behavior, this provides indirect support for our model.

A.4. The Role of Soft Information in LP Reinvestment Decisions

The evidence in Table 2 shows that future fund returns can be predicted using prior funds' future ex post IRRs, which will not be known until some years after fund-raising, while they cannot be predicted using publicly available information available at the time of fund-raising, namely prior funds' interim IRRs. To examine whether future ex post IRRs are a useful proxy for soft information, as these findings suggest, we now test for similar patterns in LP reinvestment decisions.

On average, 50.7% of LPs in a first fund continue to invest in the GP's next fund, falling to 45.2% in later funds. This is a greater reinvestment rate than chance alone would predict in observational data and so supports Implication A2.[23] An important caveat is that our LP data are incomplete. For example, it is possible that an LP listed as an investor in fund 2 but not in fund 1 (or vice versa) actually invested in both funds, leading us to underestimate LP persistence. An alternative way to gauge LP persistence is to ask how often a given LP chooses to reinvest when the GP raises a follow-on fund. We can compute this for a subset of LPs whose investment record is well-known (and so relatively complete), due to Freedom of Information Act suits. The LPs in question are the California Public Employees Retirement System, the California State Teachers Retirement System, and the endowments of the Universities of California, Michigan, and Texas. These five LPs reinvested on average in 74.3% of follow-on funds, suggesting that there is a significant degree of persistence in LP composition across funds, as predicted.

In Table 3, we model the determinants of the LP reinvestment rate, i.e., the fraction of prior fund LPs who invest in the next fund. As the dependent variable is a fraction with support on [0,1] and positive mass at both 0 and 1, we estimate fractional logit models. In column (1), the independent variables are the prior fund's future ex-post IRR, the natural logarithm of fund size, and an indicator for whether the previous fund was the GP's first, as well as untabulated vintage-year effects. We find that a significantly larger fraction of LPs reinvest in cases where the previous fund will *eventually*

---

[23]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 five or more (out of ten) LPs who were investors in the previous fund is 0.001% (applying the hypergeometric pdf).

report a higher realized return (p=0.001). Economically, the sensitivity is large: A one-standard deviation increase in the prior fund's future ex post IRR increases incumbent LPs' reinvestment rate by 9.7 percentage points from the unconditional mean of 45.2%.

In column (2), we add our measure of publicly available hard information, the prior fund's interim IRR, which is publicly known at the time LPs decide whether to reinvest in the GP's next fund. Its coefficient estimate return is both statistically and economically zero while that of the prior fund's future ex post IRR remains positive and significant. In column (3), we add another piece of hard information available to all investors at the time of fund-raising, namely the ex post IRR of the GP's fund prior to the previous fund. This piece of information also does not affect reinvestment decisions. These patterns mirror those for returns in Table 2 and are consistent with the notion that future ex post IRRs are a potentially useful proxy for incumbent LPs' private information.

Combining these findings with the persistence result in Table 2 suggests that incumbent LPs reinvest in follow-on funds raised by GPs whose current funds will *eventually* exhibit good performance, and they do so *before* the return on the GP's current fund is publicly known. Such GPs then continue to perform well on their next fund. A plausible explanation for these findings is that ex post IRRs correlate with incumbent LPs' private (soft) information. In other words, incumbent LPs appear to know something that is not captured by publicly available interim performance measures and which allows them to make reinvestment decisions that resemble the return-chasing behavior seen in mutual funds—except that the returns being chased are not yet publicly known.

B. Performance of First-time Funds

Implication A3 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 A3, the difference is not statistically significant (the t-statistic is 1.49). Table 4 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

24

(1), a dummy identifying first-time funds is not statistically significant, contrary to Implication A3, though we find that larger funds have higher IRRs, which is consistent with Implication A3 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. Many of the funds coded as follow-on funds in column (1) were, in practice, raised well 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. This likely corresponds better to the model's distinction between funds for which GP type is unknown and funds for which GP type is (to a large extent) known to GPs and incumbent investors. 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 ($p=0.05$). These patterns are consistent with Implication A3.

C. Oversubscription in Follow-on Funds

Implication A4 predicts that oversubscription is concentrated in follow-on funds and is more severe, the better the GP's previous fund performed. To our knowledge, there is no prior evidence regarding the concentration of oversubscription in follow-on funds or its relation to prior-fund performance, likely because demand data are hard to obtain in venture capital. To proxy for excess demand, 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. These data are available from 1991. Of course, we do not observe investor demand for fund units (i.e., how much capital investors are willing to invest) separately from the supply of fund units (i.e., how much capital a GP is willing to accept). But as long as supply responds to excess demand, excess demand should correlate with the ratio of final to

target fund size.[24]

The ratio of final fund size to target fund size (the subscription ratio) 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 A4. 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 5.

Column (2) tests whether oversubscription is related to prior fund performance. Restricting the sample to follow-on funds, we regress the subscription ratio on the previous fund's IRR. Controlling for log target size and vintage-year effects, we find that a one-standard deviation increase in previous returns is associated with a 4.8 percentage point increase in the subscription ratio ($p<0.001$), consistent with Implication A4.

D. Effect of Learning on Fund-Raising and GP Compensation

We end our empirical analysis by briefly discussing tests of Implications S1-S5. Because these predictions do not require learning to be asymmetric, they should hold using interim returns (i.e., publicly available hard information) when conditioning on prior performance.[25]

Kaplan and Schoar (2005) report evidence consistent with Implication S1, which states that the probability of raising a follow-on fund increases in the return the LPs earned in the first fund. We replicate this finding in our sample using a Cox hazard model with time-varying covariates, which can capture how changes in reported interim IRRs affect the probability that a VC firm raises a new fund the following year. Column (1) of Table 6 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 interim 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 25.1% higher likelihood

---

[24]In practice, GPs usually set a minimum and maximum fund size target and allow the final fund size to depend on investor demand. This provides support for our assumption that supply responds (to some extent) to excess demand.

[25]This implies that prior funds' future ex post IRRs should have no significant explanatory power in these models. This is indeed the case, though their inclusion affects the precision of the estimate of interest in one of the specifications discussed next. These results are available on request.

of raising a follow-on fund in year τ.

According to Implication S2, the size of a follow-on fund increases in the return LPs earned in the previous fund. This implication is also consistent with results reported in Kaplan and Schoar (2005). We replicate this finding using a Tobit estimator to control 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*.[26] The dependent variable equals the log fund size if the firm raises a follow-on fund and zero if it does not. The results are presented in column (2). Like Kaplan and Schoar (2005), we find that good prior performance begets larger follow-on funds. A one-standard deviation increase in previous fund's interim IRR is associated with a 35.4% or \$20.5 million increase in fund size, from the unconditional mean in the estimation sample of \$57.8 million ($p<0.001$).

Implication S3 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: 22.7%). It is an open question whether carries increase in prior-fund returns. Column (3) suggests that they do. A one-standard deviation increase in the previous fund's ex ante IRR is associated with a 2.07 percentage point increase in carry on the next fund ($p<0.001$). These estimates could be biased if poor performance prevents a VC firm from raising a follow-on fund (left-censoring). In column (4), 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 4.19 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.[27]

Implication S4 states that the standard deviation of fund size should be higher in follow-on funds

---

[26]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.

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

than in first-time funds. Using the set of funds raised between 1980 and 2006 for which we have performance information, we find statistically significant support for Implication S4: The standard deviation of fund size for first-time funds is significantly smaller, at \$80.8m versus \$345.7m for follow-on funds (p=0.01).

Similarly, Implication S5 predicts higher variation in carry among follow-on funds than among first-time funds. The data again support this prediction: The standard deviation of GP carry for first-time funds is 1.6%, versus 4.0% for follow-on funds, and the difference is statistically significant at the 1% level. 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.

IV. Discussion and Conclusion

Performance in the VC market appears persistent, suggesting VCs have skill. But why then do successful VCs not eliminate excess demand for their next funds by raising their carries? We propose a model of learning and informational hold-up that can explain performance persistence in the VC market. We argue that persistence requires that the LP market is perfectly competitive when a GP raises his first fund and that his investors subsequently gain market power. We propose that the source of their market power is learning: Investing in a fund gives an investor the opportunity to collect soft information about the GP's skill, while outside investors can observe only hard information such as realized returns. Thus, incumbent investors have an informational advantage when the GP raises his next fund. This imposes a winner's curse on outside investors—the better-informed incumbent LPs will outbid them whenever the GP has skill—and enables incumbent LPs to hold the GP up when he next raises a fund. Performance is persistent because the hold-up problem prevents the GP from raising his carry to the point where investors simply break even.

In addition to persistence, the model also predicts that expected returns are lower in first than in follow-on funds; that LPs who invested in a GP's first fund should invest in his next fund; and that

follow-on funds will rationally be oversubscribed, especially following high returns in first funds. We verify these predictions with one of the most comprehensive datasets on U.S. VC funds assembled to date. Our results confirm that VC performance is persistent (as in Kaplan and Schoar (2005)) and that average returns are higher in later funds. LPs 'chase performance' by reinvesting following higher first-fund returns, leading to oversubscription.

Importantly, we document that incumbent LPs appear to behave in a manner that suggests that they learn and that such learning gives them an informational advantage. Incumbent LPs reinvest in follow-on funds raised by GPs whose current funds will *eventually* exhibit good performance, and they do so *before* the return on the GP's current fund is publicly known. Such GPs then continue to perform well on their next fund. Information that is publicly available to all investors at the time a GP raises his next fund cannot be used to predict how well the GP will perform in future, nor does it explain his incumbent LPs' decision whether to invest in his next fund. Though the inference is necessarily indirect, these patterns point to incumbent LPs obtaining private information about GP skill and so are at least consistent with asymmetric learning.

Our model thus provides a simple unifying framework that explains both performance persistence and lower returns for first-time funds and that generates a rich set of empirical predictions consistent both with patterns uncovered by past studies of VC performance and with patterns unique to the importance of soft information in distinguishing between GP skill and luck.

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.

Blaydon, C., and M. Horvath, 2002, "What's a Company Worth? Depends Which GP You Ask," *Venture Capital Journal (May)*, 40-41.

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

Chae, S., and P. Heidhues, 2004, "A Group Bargaining Solution," *Mathematical Social Sciences* 48, 37-53.

Chung, J.W., B.A. Sensoy, L.H. Stern, and M.S. Weisbach, 2010, "Incentives of Private Equity General Partners from Future Fundraising," Unpublished working paper, Ohio State University.

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

Cornelli, F., Z. Kominek, and A. Ljungqvist, 2010, "Monitoring managers: Does it matter?," Unpublished working paper, NYU.

Glode, V., and R. Green, 2010, "Information Spill-Overs and Performance Persistence for Hedge Funds," Unpublished working paper, Carnegie-Mellon University.

Gompers, P.A., 1996, "Grandstanding in the Venture Capital Industry," *Journal of Financial Economics* 43, 133-156.

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.

Gottschalg, O., and L. Phalippou, 2009, "The Performance of Private Equity Funds," *Review of Financial Studies* 22, 1747-1776.

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

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.

Korteweg, A., and M. Sørensen, 2010, "Risk and Return Characteristics of Venture Capital-Backed Entrepreneurial Companies," *Review of Financial Studies*, forthcoming.

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.

Ljungqvist, A., M. Richardson, and D. Wolfenzon, 2007, "The Investment Behavior of Buyout Fund Managers," Unpublished working paper, NYU.

Lopez de Silanes, F., L. Phalippou, and O. Gottschalg, 2010, "Giants at the Gate: Diseconomies of Scale in Private Equity," Unpublished working paper, University of Amsterdam.

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

Metrick, A., and A. Yasuda, 2009, "The Economics of Private Equity Funds," *Review of Financial Studies*, forthcoming.

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.

Pastor, L., and P. Veronesi, 2009, "Learning in Financial Markets," *Annual Review of Financial Economics* 1, 361-381.

Phalippou, L., 2010, "Venture Capital Funds: Performance Persistence and Flow-Performance Relation," *Journal of Banking and Finance*, forthcoming.

Quigley, J., 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.

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., 1988, *The Theory of Industrial Organization*, MIT Press.

Tirole, J., 2006, *The Theory of Corporate Finance*, Princeton University Press.

Vidal-Puga, J., 2005, "A Bargaining Approach to the Owen Value and the Nash Solution With Coalition Structure," *Economic Theory* 25, 679-701.

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 $g_1\left(\mu^i\right)$ and $g_2\left(\mu^i\right)$

This section defines two quantities, $g_1\left(\mu^i\right)$ and $g_2\left(\mu^i\right)$, which will be used in the derivations in Appendix B and C below.

Define $g_1\left(\mu^i\right)$ 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,

$$\ln\left(C_1/I_1\right)=\ln\left(e^{A_1^i}\ln\left(1+I_1\right)/I_1\right)=A_1^i+\ln\left(\frac{\ln\left(1+I_1\right)}{I_1}\right)\sim\Phi\left(\mu_i+\ln\left(\frac{\ln\left(1+I_1\right)}{I_1}\right)-0.5\sigma^2,\sigma^2\right).$$

Using equations (18.30), (18.24), and (12.2a, 12.2b) in McDonald (2003), this implies that

$$
\begin{aligned}
g_1\left(\mu^i\right) &= E\left[C_1|C_1>I_1^{opt},\mu=\mu^i\right]P\left(C_1>I_1^{opt}|\mu=\mu^i\right)-I_1^{opt}P\left(C_1>I_1^{opt}|\mu=\mu^i\right)\\
&= \left[e^{\mu_i}\ln\left(1+I_1^{opt}\right)\Phi\left(d_1\left(\mu^i\right)\right)-I_1^{opt}\Phi\left(d_1\left(\mu^i\right)-\sigma\right)\right]
\end{aligned}
$$

where

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

Similarly, define $g_2\left(\mu^i\right)$ 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

$$\ln\left(C_2/I_2\right)\sim\Phi\left(\mu_i+\ln\left(\frac{\ln\left(1+I_2\right)}{I_2}\right)-0.5\sigma^2,\sigma^2\right)$$

and

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

where

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

B. 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, $f_2^{LP}(\mu^i)$, $f_2^{GP}(\mu^i)$ such that the incumbent LP is indifferent between the contract $\left(f_2^{GP}(\mu^i), I_2(\mu^i)\right)$ now and the contract $\left(f_2^{LP}(\mu^i), I_2(\mu^i)\right)$ in the next round of bargaining and such that the GP is indifferent between the contract $\left(f_2^{LP}(\mu^i), I_2(\mu^i)\right)$ now and the contract $\left(f_2^{GP}(\mu^i), I_2(\mu^i)\right)$ in the next round of bargaining. These carries are given by

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

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:

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

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

Combining the two expressions, we get

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

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

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

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

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

and $f_2^{LP}(\mu^i)=f_2^{GP}(\mu^i)$. Thus, the GP's expected payoff, $f_2^{GP}(\mu^i)g_2(\mu^i)$, 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.

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

$$E_i\left(e^{\mu^i}\mu^i|\mu^i > \mu^*\right) = \frac{1}{\mu^H - \mu^*}\int_{\mu^*}^{\mu^H} e^{\mu^i}\mu^i d\mu^i = \frac{1}{\mu^H - \mu^*}\left[e^{\mu^i}\left(\mu^i - 1\right)\right]_{\mu^*}^{\mu^H}$$

$$= \frac{1}{\mu^H - \mu^*}\left[e^{\mu^H}\left(\mu^H - 1\right) - e^{\mu^*}\left(\mu^* - 1\right)\right]$$

$$E_i\left(e^{\mu^i}|\mu^i > \mu^*\right) = \frac{1}{\mu^H - \mu^*}\left[e^{\mu^H} - e^{\mu^*}\right]$$

$$E_i\left(g_1\left(\mu^i\right)\right) = \frac{1}{\mu^H - \mu^L}\int_{\mu^L}^{\mu^H} g_1\left(\mu^i\right) d\mu^i.$$

D. Proof of Implications A1, S1, S2 and S3

Proof of Impliation S1

Implication A1 states that $E\left(r_2|r_1, \mu^i > \mu^*\right)$ is increasing in r_1. We prove that $E\left(r_2|\ln C_1, \mu^i > \mu^*\right)$ is increasing in $\ln C_1$. Since $r_1 = \frac{X_1^{LP}}{I_1} - 1 = \frac{C_1 - \max(0, f_1(C_1 - I_1))}{I_1} - 1 = \frac{\exp(\ln C_1) - \max(0, f_1(\exp(\ln C_1) - I_1))}{I_1} - 1$ is an increasing function of $\ln C_1$, this implies that $E\left(r_2|r_1, \mu^i > \mu^*\right)$ will be increasing in r_1.

Start by rewriting equation (5) as

$$E\left(r_2|\mu^i\right) = \frac{\frac{1}{2}e^{\mu^i}\ln\left(1 + I_2\left(\mu^i\right)\right) + \frac{1}{2}(1+r)\,I_2\left(\mu^i\right)}{I_2\left(\mu^i\right)} - 1$$

$$= \frac{\frac{1}{2}e^{\mu^i}\left(\mu^i - \ln(1+r)\right) + \frac{1}{2}(1+r)\left[\frac{e^{\mu^i}}{1+r} - 1\right]}{\left[\frac{e^{\mu^i}}{1+r} - 1\right]} - 1 = h\left(\mu^i\right)$$

since $I_2\left(\mu^i\right) = \frac{e^{\mu^i}}{1+r} - 1$, and with the function $h(.)$ defined by the last equality.

We then observe that

$$
\begin{aligned}
E\left(r_2|\ln C_1, \mu^i > \mu^*\right) &= \int_{\mu^*}^{\mu^H} E\left(r_2|\mu^i, \ln C_1\right) f\left(\mu^i|\ln C_1, \mu^i > \mu^*\right) d\mu^i \\
&= \int_{\mu^*}^{\mu^H} E\left(r_2|\mu^i\right) f\left(\mu^i|\ln C_1, \mu^i > \mu^*\right) d\mu^i \\
&= \int_{\mu^*}^{\mu^H} h\left(\mu^i\right) f\left(\mu^i|\ln C_1, \mu^i > \mu^*\right) d\mu^i.
\end{aligned}
$$

and thus

$$
\frac{dE\left(r_2|\ln C_1, \mu^i > \mu^*\right)}{d\ln C_1} = \int_{\mu^*}^{\mu^H} h\left(\mu^i\right) \frac{df\left(\mu^i|\ln C_1, \mu^i > \mu^*\right)}{d\ln C_1} d\mu^i.
$$

We can write $f\left(\mu^i|\ln C_1, \mu^i > \mu^*\right)$ as

$$
f\left(\mu^i|\ln C_1, \mu^i > \mu^*\right) = f\left(\ln C_1|\mu^i, \mu^i > \mu^*\right) \frac{f\left(\mu^i|\mu^i > \mu^*\right)}{f\left(\ln C_1|\mu^i > \mu^*\right)}.
$$

Since $\ln C_1 = A_1^i + \ln\left(\ln\left(1 + I_1\right)\right) \sim N\left(\mu^i - \frac{1}{2}\sigma^2 + \ln\left(\ln\left(1 + I_1\right)\right), \sigma^2\right)$, we have that

$$
f\left(\ln C_1|\mu^i, \mu^i > \mu^*\right) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\left[\left\{\ln C_1 - \left(\mu^i - \frac{1}{2}\sigma^2 + \ln(\ln(1+I_1))\right)\right\}^2/(2\sigma^2)\right]} = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\left[\left(\mu^i - k\right)^2/(2\sigma^2)\right]},
$$

with $k = \left(\ln C_1 + \frac{1}{2}\sigma^2 - \ln\left(\ln\left(1 + I_1\right)\right)\right)$. Furthermore, μ^i is uniform over the interval from μ^* to μ^H, so $f\left(\mu^i|\mu^i > \mu^*\right) = 1/(\mu^H - \mu^*)$.

It follows that

$$
\begin{aligned}
f\left(\ln C_1|\mu^i > \mu^*\right) &= \int_{\mu^*}^{\mu^H} f\left(\ln C_1|\mu^i\right) f\left(\mu^i|\mu^i > \mu^*\right) d\mu^i \\
&= \int_{\mu^*}^{\mu^H} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\left[\left(\mu^i - k\right)^2/(2\sigma^2)\right]} \frac{1}{\mu^H - \mu^*} d\mu^i \\
&= \frac{1}{\mu^H - \mu^*} \left[\Phi\left(\frac{\mu^H - k}{\sigma}\right) - \Phi\left(\frac{\mu^* - k}{\sigma}\right)\right]
\end{aligned}
$$

where Φ is the CDF of the standard normal distribution. Therefore,

$$
\begin{aligned}
f\left(\mu^i|\ln C_1, \mu^i > \mu^*\right) &= f\left(\ln C_1|\mu^i, \mu^i > \mu^*\right) \frac{f\left(\mu^i|\mu^i > \mu^*\right)}{f\left(\ln C_1|\mu^i > \mu^*\right)} \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\left[\left(\mu^i - k\right)^2/(2\sigma^2)\right]} \frac{1}{\left[\Phi\left(\frac{\mu^H - k}{\sigma}\right) - \Phi\left(\frac{\mu^* - k}{\sigma}\right)\right]}
\end{aligned}
$$

The above then implies that

$$
\frac{df\left(\mu^i|\ln C_1, \mu^i > \mu^*\right)}{d\ln C_1} = f\left(\mu^i|\ln C_1, \mu^i > \mu^*\right) \frac{1}{\sigma} \left\{\frac{\left(\mu^i - k\right)}{\sigma} + \frac{\phi\left(\frac{\mu^H - k}{\sigma}\right) - \phi\left(\frac{\mu^* - k}{\sigma}\right)}{\Phi\left(\frac{\mu^H - k}{\sigma}\right) - \Phi\left(\frac{\mu^* - k}{\sigma}\right)}\right\}.
$$

$f\left(\mu^i|\ln C_1,\mu^i>\mu^*\right)\frac{1}{\sigma}$ is positive for all values of μ^i. $\left\{\frac{\left(\mu^i-k\right)}{\sigma}+\frac{\phi\left(\frac{\mu^H-k}{\sigma}\right)-\phi\left(\frac{\mu^*-k}{\sigma}\right)}{\Phi\left(\frac{\mu^H-k}{\sigma}\right)-\Phi\left(\frac{\mu^*-k}{\sigma}\right)}\right\}$ is increasing in μ^i (since

$\frac{\phi\left(\frac{\mu^H-k}{\sigma}\right)-\phi\left(\frac{\mu^*-k}{\sigma}\right)}{\Phi\left(\frac{\mu^H-k}{\sigma}\right)-\Phi\left(\frac{\mu^*-k}{\sigma}\right)}$ does not depend on μ^i). Since $\int_{\mu^*}^{\mu^H}\frac{df\left(\mu^i|\ln C_1,\mu^i>\mu^*\right)}{d\ln C_1}=0$, there thus exists a value of μ^i,

call it μ^{**} (which will depend on $\ln C_1$), such that $\frac{df\left(\mu^i|\ln C_1,\mu^i>\mu^*\right)}{d\ln C_1}$ is negative for $\mu^i<\mu^{**}$ and positive for

$\mu^i>\mu^{**}$.

Furthermore, $h\left(\mu^i\right)>r$ for $\mu^i>\mu^*$ (where $\mu^*=1+r$) since

$$h\left(\mu^i\right)>r\Leftrightarrow e^{\mu^i}\ln\left(1+I_2\left(\mu^i\right)\right)>(1+r)I_2\left(\mu^i\right)\Leftrightarrow NPV_2\left(\mu^i\right)>0$$

which is true for $\mu^i>\mu^*$. Also, $h\left(\mu^i\right)$ is an increasing function of μ^i since

$$\frac{dh\left(\mu^i\right)}{d\mu^i}=\frac{(1+r)e^{\mu^i}}{2\left[e^{\mu^i}-(1+r)\right]^2}\left\{e^{\mu^i}-(1+r)\left[\left(1+\mu^i\right)-\ln(1+r)\right]\right\}$$

is positive if $\frac{e^{\mu^i}}{1+r}>1+\mu^i-\ln(1+r)\Leftrightarrow\mu^i-\ln(1+r)>\ln\left(1+\mu^i-\ln(1+r)\right)$, which is true for all μ^i of

funds raised, since only funds with $\mu^i>\mu^*=\ln(1+r)$ are raised and $x>\ln(1+x)$ for all $x>0$.

Therefore, $\frac{dE\left(r_2|\ln C_1,\mu^i>\mu^*\right)}{d\ln C_1}=\int_{\mu^*}^{\mu^H}h\left(\mu^i\right)\frac{df\left(\mu^i|\ln C_1,\mu^i>\mu^*\right)}{d\ln C_1}d\mu^i$ is positive (for all values of $\ln C_1$) since

$\int_{\mu^*}^{\mu^H}\frac{df\left(\mu^i|\ln C_1,\mu^i>\mu^*\right)}{d\ln C_1}=0$ and $h\left(\mu^i\right)$ is positive and increasing, implying that in $\int_{\mu^*}^{\mu^H}h\left(\mu^i\right)\frac{df\left(\mu^i|\ln C_1,\mu^i>\mu^*\right)}{d\ln C_1}d\mu^i$

the positive values of $\frac{df\left(\mu^i|C_1\right)}{dC_1}$ are multiplied by a larger positive number than the negative values of $\frac{df\left(\mu^i|C_1\right)}{dC_1}$

are.

Proof of Impliation S1

Implication S1 states that $P\left(\mu^i>\mu^*|r_1\right)$ is increasing in r_1. We prove that $P\left(\mu^i>\mu^*|\ln C_1\right)$ is increasing

in $\ln C_1$. Since $r_1=\frac{X_1^{LP}}{I_1}-1=\frac{C_1-\max(0,f_1(C_1-I_1))}{I_1}-1=\frac{\exp(\ln C_1)-\max(0,f_1(\exp(\ln C_1)-I_1))}{I_1}-1$ is an increasing

function of $\ln C_1$, this implies that $P\left(\mu^i>\mu^*|r_1\right)$ will be increasing in r_1.

We start by observing that

$$P\left(\mu^i>\mu^*|\ln C_1\right)=\int_{\mu^L}^{\mu^H}h\left(\mu^i\right)1\left(\mu^i>\mu^*\right)f\left(\mu^i|\ln C_1\right)d\mu^i$$

where $h\left(\mu^i\right)$ is an indicator function equation to 1 if $\mu^i>\mu^*$ and 0 otherwise. This implies

$$\frac{dP\left(\mu^i>\mu^*|\ln C_1\right)}{d\ln C_1}=\int_{\mu^L}^{\mu^H}h\left(\mu^i\right)\frac{df\left(\mu^i|\ln C_1\right)}{d\ln C_1}d\mu^i.$$

By the same arguments as in the proof of Implication A1

$$f\left(\mu^i|\ln C_1\right)=f\left(\ln C_1|\mu^i\right)\frac{f\left(\mu^i\right)}{f\left(\ln C_1\right)}=\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\left[\left(\mu^i-k\right)^2/(2\sigma^2)\right]}\frac{1}{\left[\Phi\left(\frac{\mu^H-k}{\sigma}\right)-\Phi\left(\frac{\mu^L-k}{\sigma}\right)\right]}$$

37

and therefore

$$\frac{df\left(\mu^i|\ln C_1\right)}{d\ln C_1} = f\left(\mu^i|\ln C_1\right)\frac{1}{\sigma}\left\{\frac{\left(\mu^i-k\right)}{\sigma} + \frac{\phi\left(\frac{\mu^H-k}{\sigma}\right) - \phi\left(\frac{\mu^L-k}{\sigma}\right)}{\Phi\left(\frac{\mu^H-k}{\sigma}\right) - \Phi\left(\frac{\mu^L-k}{\sigma}\right)}\right\}.$$

$f\left(\mu^i|\ln C_1\right)\frac{1}{\sigma}$ is positive for all values of μ^i. $\left\{\frac{\left(\mu^i-k\right)}{\sigma} + \frac{\phi\left(\frac{\mu^H-k}{\sigma}\right) - \phi\left(\frac{\mu^L-k}{\sigma}\right)}{\Phi\left(\frac{\mu^H-k}{\sigma}\right) - \Phi\left(\frac{\mu^L-k}{\sigma}\right)}\right\}$ is increasing in μ^i.

Since $\int_{\mu^L}^{\mu^H}\frac{df\left(\mu^i|\ln C_1\right)}{d\ln C_1} = 0$, there thus exists a value of μ^i, call it μ^{**} (which will depend on $\ln C_1$), such that $\frac{df\left(\mu^i|\ln C_1\right)}{d\ln C_1}$ is negative for $\mu^i < \mu^{**}$ and positive for $\mu^i > \mu^{**}$.

$h\left(\mu^i\right)$ is zero for $\mu^i \leq \mu^*$ and one for $\mu^i > \mu^*$. Therefore, in $\frac{dP\left(\mu^i>\mu^*|\ln C_1\right)}{d\ln C_1} = \int_{\mu^L}^{\mu^H} h\left(\mu^i\right)\frac{df\left(\mu^i|\ln C_1\right)}{d\ln C_1}d\mu^i$ a larger fraction of the positive values of $\frac{df\left(\mu^i|\ln C_1\right)}{d\ln C_1}$ will be multiplied by one than the fraction of the negative values of $\frac{df\left(\mu^i|\ln C_1\right)}{d\ln C_1}$ that are multiplied by one (if $\mu^* > \mu^{**}$, some of the positive and none of the negative values of $\frac{df\left(\mu^i|\ln C_1\right)}{d\ln C_1}$ are multiplied by 1, and if $\mu^* < \mu^{**}$, all of the positive and only some of the negative values of $\frac{df\left(\mu^i|\ln C_1\right)}{d\ln C_1}$ are multiplied by 1). Therefore $\frac{dP\left(\mu^i>\mu^*|\ln C_1\right)}{d\ln C_1} = \int_{\mu^L}^{\mu^H} h\left(\mu^i\right)\frac{df\left(\mu^i|\ln C_1\right)}{d\ln C_1}d\mu^i > 0.$

Proof of Implication S2

Implication S2 states that $E\left(I_2|r_1,\mu^i > \mu^*\right)$ is increasing in r_1. The proof of Implication S2 is identical to the proof of Implication A1 with the function $h\left(\mu^i\right)$ now defined as $I_2\left(\mu^i\right) = \frac{e^{\mu^i}}{1+r} - 1$. For $\mu^i > \mu^*$, $h\left(\mu^i\right)$ is a positive function which is increasing in μ^i. Therefore, $\frac{dE\left(I_2|\ln C_1,\mu^i>\mu^*\right)}{d\ln C_1} = \int_{\mu^*}^{\mu^H} h\left(\mu^i\right)\frac{df\left(\mu^i|\ln C_1,\mu^i>\mu^*\right)}{d\ln C_1}d\mu^i$ is positive using the same argument as in the proof of Implication A1.

Proof of Implication S3

Implication S3 states that $E\left(f_2|r_1,\mu^i > \mu^*\right)$ is increasing in r_1. The proof of Implication S3 is identical to the proof of Implication A1 with the function $h\left(\mu^i\right)$ now defined as $f_2\left(\mu^i\right) = \frac{\frac{1}{2}\left[e^{\mu^i}\ln\left(1+I_2\left(\mu^i\right)\right) - (1+r)I_2\left(\mu^i\right)\right]}{g_2\left(\mu^i\right)}$. For $\mu^i > \mu^*$, $h\left(\mu^i\right)$ is a positive function which is increasing in μ^i.

Therefore, $\frac{dE\left(f_2|\ln C_1,\mu^i>\mu^*\right)}{d\ln C_1} = \int_{\mu^*}^{\mu^H} h\left(\mu^i\right)\frac{df\left(\mu^i|\ln C_1,\mu^i>\mu^*\right)}{d\ln C_1}d\mu^i$ is positive using the same argument as in the proof of Implication A1.

$f_2\left(\mu^i\right)$ is positive for $\mu^i > \mu^*$ since $e^{\mu^i}\ln\left(1+I_2\left(\mu^i\right)\right) - (1+r)I_2\left(\mu^i\right) > 0 \Leftrightarrow NPV_2\left(\mu^i\right) > 0$ which is true for $\mu^i > \mu^*\left(= \ln\left(1+r\right)\right).$

The argument for why $f_2\left(\mu^i\right)$ is increasing in μ^i for $\mu^i > \mu^*$ is as follows. Using the expression for $g_2\left(\mu^i\right)$ in Appendix A $f_2\left(\mu^i\right)$ can be written as

$$
\begin{aligned}
f_2\left(\mu^i\right) &= \frac{\frac{1}{2}\left[e^{\mu^i}\ln\left(1+I_2\left(\mu^i\right)\right) - (1+r)I_2\left(\mu^i\right)\right]}{e^{\mu^i}\ln\left(1+I_2\left(\mu^i\right)\right)\Phi\left(\frac{1}{\sigma}\mu^i + \frac{1}{\sigma}\ln\left(\frac{\ln\left(1+I_2\left(\mu^i\right)\right)}{I_2\left(\mu^i\right)}\right) - \frac{1}{2}\sigma\right)} \\[2mm]
&= \frac{\frac{1}{2}\left[e^{\mu^i}\left(\mu^i - \ln\left(1+r\right)\right) - \left(e^{\mu^i} - (1+r)\right)\right]}{e^{\mu^i}\left(\mu^i - \ln\left(1+r\right)\right)\Phi\left(\frac{1}{\sigma}\mu^i + \frac{1}{\sigma}\ln\left(\frac{\left(\mu^i-\ln\left(1+r\right)\right)}{\frac{e^{\mu^i}}{1+r}-1}\right) - \frac{1}{2}\sigma\right)} \\[2mm]
&= \frac{\frac{1}{2}\left[e^{\mu^i}\left(\mu^i - \ln\left(1+r\right)\right) - \left(e^{\mu^i} - (1+r)\right)\right]}{e^{\mu^i}\left(\mu^i - \ln\left(1+r\right)\right)\Phi\left(\frac{1}{\sigma}\mu^i + \frac{1}{\sigma}\ln\left(\mu^i - \ln\left(1+r\right)\right) - \frac{1}{\sigma}\ln\left(\frac{e^{\mu^i}}{1+r}-1\right) - \frac{1}{2}\sigma\right)}
\end{aligned}
$$

Using mathematical software, we verify that the derivative of $f_2\left(\mu^i\right)$ is positive for $\mu^i > \mu^*$.

Table 1. Descriptive Statistics.

		Number of sample funds			Fund size ($m)		Performance							
			of which				All funds				First-time funds		Follow-on funds	
		only in	only in	in			no. of funds with	mean IRR	sd IRR	median IRR	no. of funds with	mean IRR	no. of funds with	mean IRR
vintage	all	VE	Preqin	both	mean	median	IRR data	(%)	(%)	(%)	IRR data	(%)	IRR data	(%)
1980	40	34	4	2	30.1	20.0	17	13.0	12.7	12.9	10	10.9	7	16.1
1981	46	39	1	6	23.1	20.0	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	77	65	1	11	31.7	20.2	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	56	38	0	18	53.0	21.6	34	6.8	7.4	6.1	16	5.2	18	8.3
1987	83	66	1	16	35.9	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	74	42	0	32	83.9	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	164	109	1	54	85.4	56.4	63	42.4	76.5	6.3	16	35.0	47	44.9
1998	169	111	1	57	136.5	70.5	73	23.9	105.9	2.2	7	16.9	66	24.7
1999	258	178	2	78	173.5	100.0	72	-6.7	24.8	-9.8	12	16.5	60	-11.4
2000	334	228	5	101	202.7	100.0	93	-7.0	12.4	-6.5	12	-0.9	81	-7.9
2001	176	124	2	50	217.0	70.0	47	-4.3	12.4	-5.8	3	-19.9	44	-3.2
2002	75	43	4	28	139.1	49.3	22	-1.4	8.7	-0.9	3	-10.0	19	-0.1
2003	69	38	5	26	129.5	37.6	19	-0.7	15.3	-0.3	3	1.5	16	-1.2
2004	109	77	1	31	137.4	50.0	15	-4.2	18.4	-5.2	4	3.7	11	-7.1
2005	131	96	1	34	248.4	107.7	2	-19.3	17.3	-19.3	1	-7.0	1	-31.5
2006	153	122	5	26	217.2	107.3								
1980-2006	2,790	1,963	44	783	124.7	49.9	1,007	14.1	48.3	4.5	253	14.7	754	13.9
1980-1996	1,152	837	17	298	51.4	28.2	601	18.8	39.3	10.3	192	15.3	409	20.4

Table 2. VC Fund Performance Persistence.

The dependent variable in columns (1) and (3) through (6) is a fund's ex post IRR, net of carry and fees, measured at the end of the fund's ten-year life. The sample is accordingly restricted to funds that are at least ten years old as of summer 2007 (that is, funds raised between 1980 and 1996). In column (2), 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. To test for persistence of performance across funds managed by the same VC firm, we regress the ex post performance of fund N on the performance of the fund manager's previous funds and controls for fund size and risk. Columns (1) and (2) condition on the previous fund's performance (i.e., fund N-1). In column (3), we interact the ex post IRR of fund N-1 with an indicator for being a later-sequence fund (defined as fund sequence 4 or higher). In column (4), we condition on the ex post IRRs of the fund manager's two previous funds, N-1 and N-2. In column (5), we run a horse race between the interim IRR of fund N-1, measured as of the year-end prior to the year the GP raises fund N, and the ex post IRR of fund N-1. (Where a VC firm operates multiple funds in parallel as of the prior year-end, we compute the maximum interim IRR.) In column (6), we interact the ex post IRR of fund N-1 with indicators for whether fund N-1 was "lucky" or "unlucky". "Lucky" is defined as being in the top tercile of return performance and bottom tercile of LP reinvestment, and "unlucky" is defined as being in the bottom tercile of return performance and top tercile of LP reinvestment. 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.

	Ex post performance of fund N					
Performance measure:	IRR	Exit rate	IRR			
	(1)	(2)	(3)	(4)	(5)	(6)
log size of fund *N-1*	0.050***	0.025***	0.065***	0.070***	0.065***	0.076***
	0.018	*0.009*	*0.018*	*0.019*	*0.031*	*0.028*
dummy =1 if fund *N* has early-stage focus	0.071**	0.061***	0.079**	0.108**	0.082	0.143**
	0.034	*0.019*	*0.032*	*0.051*	*0.052*	*0.056*
ex post IRR or exit rate of fund *N-1*	0.457***	0.379***	0.196	0.750***	0.367***	1.092***
	0.134	*0.063*	*0.150*	*0.187*	*0.135*	*0.435*
… x (sequence number = 4 or higher)			0.323*			
			0.196			
ex post IRR of fund *N-2*				0.165		
				0.254		
interim IRR of fund *N-1* as of previous year					0.251	
					0.384	
ex post IRR of fund *N-1* x "lucky"						-0.894***
						0.310
ex post IRR of fund *N-1* x "unlucky"						-1.299***
						0.322
Vintage year FE	yes	yes	yes	yes	yes	yes
Wald test: all coeff. = 0	5.7***	8.1***	6.0***	5.4***	4.3***	7.7***
Adjusted R^2	27.2%	20.3%	22.2%	26.6%	20.0%	34.4%
No. of observations	318	480	315	165	175	123

Table 3. LP Reinvestment Decisions.
We model incumbent LPs' decision to invest in a GP's next fund as a function of a proxy for the soft information known to incumbent LPs (but not to outside investors) as well as the hard information known publicly to all investors at the time the next fund is raised. Specifically, based on the results in Table 2, we use the ex post IRR of the GP's previous fund (*N-1*) as a proxy for the incumbent LPs' soft information. At the time a GP raises his next fund, the ex post IRR of his previous fund has not yet been publicly announced; in fact, it will not be known publicly for many years. If incumbent LPs learn the GP's skill while investing in the previous fund, their information should correlate with the ex post IRR when it is eventually realized. We take any information that is publicly available at the time a fund is raised as hard information. Hard information in the models reported in this table includes the ex post IRR of the GP's previous fund but two (*N-2*) as well as the interim IRR of the previous fund (*N-1*) as reported at the end of the year before fund *N* is raised. We also control for an indicator for whether the previous fund was the GP's first, the natural logarithm of the fund's size, and vintage-year effects (not reported). The dependent variable is the fraction of LPs from the GP's previous fund (*N-1*) that reinvest in the next fund (*N*). This variable has support on [0,1] and positive mass at both 0 and 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 $E(y|x)=\exp(x\beta)/(1+\exp(x\beta))$. In column (3), the previous fund can never be a first fund as the specification includes performance data from the GP's two prior funds. Therefore, the first-fund dummy is excluded. Heteroskedasticity-consistent standard errors are shown in italics. We use ***, **, and * to denote significance at the 1%, 5%, and 10% level (two-sided), respectively.

	Fraction of LPs in previous fund (*N-1*) that reinvest in GP's next fund (*N*)		
	(1)	(2)	(3)
log fund size	0.030	-0.065	-0.105
	0.063	*0.082*	*0.090*
dummy=1 if fund *N-1* is a first fund	-0.026***	-0.036***	
	0.006	*0.009*	
ex post IRR of fund *N-1*	0.796***	0.642***	0.564***
	0.239	*0.237*	*0.246*
ex post IRR of fund *N-2*			0.160
			0.190
interim IRR of fund *N-1* as of previous year		-0.007	-0.197
		0.276	*0.332*
Vintage year FE	yes	yes	yes
Pseudo-R^2	30.2%	39.1%	16.9%
No. of observations	452	316	236

Table 4. Performance of First-time Funds Relative to Follow-on Funds.
We test for differences in average performance of first-time and follow-on funds using OLS regressions of fund IRRs that control for log fund size, fund stage focus, and vintage-year fixed effects (not reported). The dependent variable in each column is a fund's ex post IRR, net of carry and fees, measured over its ten-year life. The sample is accordingly restricted to funds that are at least ten years old as of summer 2007 (that is, 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.

	Ex post IRR of fund N		
	(1)	(2)	(3)
log size of fund N	0.043*	0.030	0.031
	0.022	*0.023*	*0.023*
dummy =1 if fund N has early-stage focus	0.086***	0.080***	0.079***
	0.028	*0.028*	*0.027*
dummy =1 if fund N is a follow-on fund	-0.014		
	0.036		
dummy =1 if fund N was raised at least 10 years after the 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

Table 5. 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. The models are estimated as OLS regressions with vintage-year fixed effects (not shown). 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 so that we can condition on the performance of the previous fund. Heteroskedasticity-consistent standard errors are shown in italics. We use ***, **, and * to denote significance at the 1%, 5%, and 10% level (two-sided), respectively.

	Subscription ratio (= amount raised / fund N's target amount)	
	All funds (1)	Follow-on funds (2)
log target size of fund N	0.031^{**}	0.006
	0.013	*0.023*
dummy =1 if fund N is a follow-on fund	0.068^{**}	
	0.033	
interim IRR of fund N-1 as of previous year		0.128^{***}
		0.035
Vintage year FE	yes	yes
Wald test: all coeff. = 0	5.0^{***}	3.4^{***}
Adjusted R^2	6.2%	8.2%
No. of obs	901	366

Table 6. Effect of Learning on Fund-raising and GP Compensation.

This table tests three implications of our hold-up model that result from learning but that do not require incumbent LPs to have an informational advantage over outside LPs. These are Implications S1, S2, and S3. Because these predictions do not require learning to be asymmetric, they should hold using interim returns (i.e., publicly available hard information) when conditioning on prior performance. In column (1), 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 on the interim 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 interim IRR.) Thus, the hazard model uses only information that was available at the time of fund-raising. We ignore the first four years of interim 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. The hazard model includes 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. The dependent variable in column (2) 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 126 funds raised by VC firms that later became defunct. The model is estimated using Tobit. The variable of interest in column (2) is the interim IRR of the previous fund measured as of the year-end prior to the year the GP raises the current fund. If no follow-on fund is raised, the IRR of the previous fund is measured ex post (i.e., as of year ten.) The dependent variable in columns (3) and (4) is the GP's performance fee or "carry." The variable of interest is again the interim IRR of the previous fund measured as of the year-end prior to the year the GP raises the current fund. Since we condition on the performance of the previous fund, the estimation sample in column (3) is restricted to follow-on funds and the model is 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 (4), 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 again use the 362 defunct VC firms identified from *CapitalIQ*; performance and carry data are available for 81 funds raised by VC firms that later became defunct. The estimation samples in all four specifications include all observations through 2006 for which an interim IRR is available. Standard errors are shown in italics. They are heteroskedasticity-consistent in columns (1) and (3); note that the Tobit estimator does not support a heteroskedasticity correction in columns (2) and (4). We use ***, **, and * to denote significance at the 1%, 5%, and 10% level (two-sided), respectively.

	Prob(follow-on fund raised)	Log size of follow-on fund	GP carry in follow-on fund	
	(1)	(2)	(3)	(4)
log fund size	0.391***	0.925***	0.001	0.057***
	0.042	*0.068*	*0.003*	*0.006*
interim IRR of fund *N-1* as of previous year-end	0.224***	0.594***	0.039***	0.084***
	0.064	*0.156*	*0.006*	*0.015*
Vintage year FE	n.a.	yes	yes	yes
Wald test: all coeff. = 0	130.3***	328.4***	8.5***	170.9***
Pseudo-R^2 / adjusted R^2	2.7%	10.4%	23.3%	43.7%
No. of observations	3,721	724	195	270
No. of VC firms	262			
No. of funds raised	621			
Model estimated	Hazard	Tobit	OLS	Tobit