Skill and Luck in Private Equity Performance

Arthur Korteweg

Morten Sorensen

February 2014

Abstract

We evaluate the performance of private equity ("PE") funds, using a variance decomposition model to separate skill from luck. We find a large amount of long-term persistence, and skilled PE firms outperform by 7% to 8% annually. But this performance is noisy, with a large amount of luck, so top-quartile performance does not necessarily imply top-quartile skills, making it difficult for investors ("LPs") to identify skilled PE firms. Buyout ("BO") firms show the largest skill differences, implying the greatest long-term persistence. Venture capital ("VC") performance is the most noisy, making good VC firms hardest to identify, and implying the smallest amount of investable persistence.

The persistence and predictability of returns is a central topic in finance. Studies of individual stocks, mutual funds, and hedge funds generally find that returns are unpredictable and that investors cannot consistently outperform the market. An important exception is private equity ("PE"), including venture capital ("VC"), buyout ("BO") and other types of PE firms. A PE firm typically manages a sequence of PE funds, and Kaplan and Schoar [2005] find that the performance of fund number $N-1$ predicts the performance of the subsequent fund N. A natural interpretation is that PE firms differ in their skills and abilities, and that funds that are managed by skilled PE firms consistently outperform. A puzzle is then why this outperformance is not competed away by the investors ("LPs") in these funds, e.g., by driving up the fees that more skilled PE firms charge their LPs.

Kaplan and Schoar [2005], and subsequent studies,[1] measure persistence as a positive and statistically significant β coefficient in the regression:

$$y_{i,N} = \alpha + \beta \times y_{i,N-1} + \varepsilon_{i,N}, \tag{1}$$

where $y_{i,N}$ is the performance of fund number N managed by PE firm i. This regression is motivated by a cross-sectional intuition. In the cross-section, some funds have better performance, and if these funds follow previous funds by the same PE firm that also had better performance, this is evidence of persistence. Formally, however, equation (1) is a time-series AR(1) model, and there is a tension between this cross-sectional intuition and its time-series properties. The AR(1) model does not distinguish skill from luck. If an unskilled PE firm is lucky and its fund outperforms, which happens occasionally when performance is random, then the AR(1) model implies that this PE firm is now considered a skilled firm, and its next fund is also expected to outperform. Similarly, a skilled firm with an unlucky fund is immediately considered unskilled. In the limit, after all PE firms have undergone a number of such transitions, the AR(1) model implies that their performance will converge to the same limit distribution, with $E[y] = \frac{\alpha}{1-\beta}$.[2] Hence, the AR(1) model is an empirical model of performance persistence that implies no long-term performance

---

[1]Including Phalippou and Gottschalg [2009], Hochberg, Ljungqvist, and Vissing-Jorgensen [2010], Phalippou [2010], Robinson and Sensoy [2011], Chung [2012], Braun, Jenkinson, and Stoff [2013], and Harris, Jenkinson, Kaplan, and Stucke [2013].

[2]This convergence fails when $\beta \geq 1$ and the time series is non-stationary ($\beta = 1$ implies a unit root).

differences, which seems undesirable.

Overview We present a new variance-decomposition model of PE performance that better captures the cross-sectional intuition. Our model explicitly models skill and luck, so an unskilled PE firm with a lucky fund does not become a skilled firm. We model heterogeneous skills of PE firms, and our model allows some PE firms to consistently outperform. It does not imply that all firms converge in the limit. Separating skill from luck also leads to a natural distinction between two types of persistence, which we term *long-term* and *investable* persistence: Long-term persistence reflects the average outperformance of more skilled PE firms, and it captures how heterogeneous skills affect PE performance. In contrast, investable persistence captures whether investors ("LPs") can identify the skilled PE firms. When performance is random, top-quartile performance may be due to luck, and it does not necessarily imply top-quartile skills. This distinction matters. We find a large amount of long-term persistence, and skilled PE firms outperform by 7% to 8% annually, across all fund types. This performance is noisy, though, and we find only a small amount of investable persistence, particularly for venture capital ("VC") firms. VC performance is mostly due to luck, and an LP needs to observe the performance of an excessive number of past funds (on the order of 25 to 50 past funds) to identify a VC firm as having top-quartile skills with reasonable certainty.

Comparing different subsamples, we find that smaller funds have greater persistence than larger funds. Particularly large VC funds have weak long-term persistence and worse signal-to-noise ratios. Comparing locations of PE firms, we find the least persistence for PE firms located in the US, followed by Europe, and the greatest persistence for PE firm located in the rest of the world ("ROW"), although these PE firms also have more volatile performance. We confirm that persistence has declined in the 2000s relative to the 1990s. This decline is largest for VC firms, and we find that Buyout and Other funds still show substantial long-term persistence, even post 2000.

Our finding of large long-term persistence but little investable persistence has several implications: First, it explains LPs' increasing focus on obtaining more detailed information about PE firms and their past funds (such as the PE firms' internal organization and

culture, internal compensation and alignment of incentives, processes and deal sourcing) to help them attribute past performance (e.g., Ewens and Rhodes-Kropf [2013]). Our results show that such detailed information is necessary for LPs to identify top PE firms. Information about past fund performance, by itself, is insufficient. Second, our results may explain why outperformance is not competed away. PE skills are scarce, but when performance is noisy, LPs with the ability to identify skilled PE firms may also be scarce, and those LPs should earn rents (Lerner, Schoar, and Wongsunwai [2007] and Berk, Wang, and Weisbach [2013] study heterogeneous LP skills). Last, our findings confirm the economic realities behind the common saying among VCs that "I'd rather be lucky than smart."

Model   For our analysis, we develop a variance decomposition model, or *hierarchical linear model*, which generalizes the classical *analysis of variance* ("ANOVA") methods. We estimate our model with a Bayesian procedure, as described in the appendix. The model has several advantages relative to the AR(1) model:[3] First, as mentioned, our model explicitly separates skill from luck. When performance is noisy, top-quartile performance does not imply top-quartile skills, and this distinction leads to two different notions of persistence: Long-term persistence arises from the difference between skilled and unskilled PE firms. Investable persistence, in contrast, reflects the difference between PE firms with good and bad past performance. Second, our model explicitly captures the timing of the funds, and it does not rely on the numbering of these. This is important for simultaneous funds, where it is arbitrary which one is labeled N and $N-1$, and our model distinguishes situations where fund N follows fund $N-1$ by a few months from those they are years apart. Moreover, our

---

[3]Persistence is also sometimes studied by estimating transition probabilities across fund quartiles (Billingsley [1961] surveys the statistical issues that arise when estimating parameters and testing hypotheses in Markov chains). If funds' performances are *i.i.d.*, then the probability that a top-quartile fund remains top quartile is 25%; and more generally, $P[y_{i,N} \in Q | y_{i,N-1} \in Q] = 25\%$ when Q contains the performance for any quartile. Hence, the empirical finding that $P[y_{i,N} \in Q | y_{i,N-1} \in Q] > 25\%$ implies that performance cannot be *i.i.d.*, which is sometimes interpreted as evidence of persistence, but this interpretation is tenuous. For example, let there be two types of PE firms, with an equal number of each. The first type determines the return of each of its funds by flipping a dime, and it is either +10% or -10%, with equal probability. The second type flips a quarter, and its returns are either +25% or -25%. Hence, the returns for the four quartiles are: +25%, +10%, -10%, and -25%. For each quartile, the transition probability is $P[y_{i,N} \in Q | y_{i,N-1} \in Q] = 50\%$, so returns are not *i.i.d.* (obviously), but there is no persistence in the conventional sense. Conversely, finding transition probabilities of $P[y_{i,N} \in Q | y_{i,N-1} \in Q] = 25\%$, by itself, does not imply an absence of persistence. It is neither a necessary nor sufficient condition. Hence, the economic magnitudes and statistical significance of persistence are difficult to evaluate using transition probabilities across quartiles.

estimates include firms that only raise a single fund, which then has no fund $N - 1$. Third, our model has a non-parametric component and does not impose normality on fund returns. Instead, we use a general mixtures-of-normals distribution and use a Bayes factor test for the number of mixtures. This flexibility is especially important for VC funds, which have highly skewed returns. Fourth, our Bayesian approach is computationally efficient, and it provides accurate small sample inferences for the estimated parameters, which is important when the parameters of interest are variances (and a ratio of variances in the signal-to-noise ratio), which have non-standard asymptotic distributions.

Literature   Following Kaplan and Schoar [2005], a number of studies have investigated the persistence of PE performance. Phalippou and Gottschalg [2009] consider persistence after correcting for potential biases in reported interim NAVs. Chung [2010] and Phalippou [2010] find weaker effects when regressing $y_{i,N}$ on $y_{i,N-2}$, and they argue that persistence is short lived. Hochberg, Ljungqvist, Vissing-Jorgensen [2010] model performance persistence that arises from asymmetric information between the LP and GP. Recently, Harris, Jenkinson, Kaplan, and Stucke [2013] find that persistence has declined post 2000 for BO firms, and this finding is confirmed by Braun, Jenkinson, and Stoff [2013], using deal-level data. Separating skill from luck is a general question in economics and finance, and our analysis may be useful for other applications, such as the persistence of the performance of serial entrepreneurs (e.g., Bengtson [2013]; Gompers, Kovner, Lerner, and Scharfstein [2010]).

Outline   The paper proceeds as follows. In Section I we present the data. Section II presents our empirical model. Section III presents our results and discusses the evidence for "long-term" persistence in private equity performance. Section IV estimates a learning model and evaluates the "investable persistence." Section V analyzes various subsamples of the data, and Section VI concludes. We provide a detailed description of the Bayesian estimation procedure in the Appendix.

I Data

This paper uses an extensive dataset with PE firms, the funds they manage, and the performance and other information for these funds. The data are obtained from Preqin, a commercial data provider that started collecting performance data using freedom of information act ("FOIA") requests to public investors and later extended the scope of its data collection to other public filings and voluntary reporting by some GPs and LPs. For each fund, Preqin only reports aggregate fund performance, such as the IRR and Total Value to Paid-in Capital multiple ("TVPI"). We do not have individual cash flows between the LPs and GPs. One limitation of these data is that they do not contain the public market equivalent ("PME") measure of fund performance, which has advantages when evaluating risk-adjusted performance (see Sorensen and Jagannathan [2013] and Korteweg and Nagel [2013]).

Harris, Jenkinson, and Kaplan [2013] compare several datasets with PE fund performance. Most of these data are from commercial data providers (Preqin, Burgiss, and Cambridge Associates) and one is from a large anonymous LP (studied by Robinson and Sensoy [2011]). For buyout ("BO") funds, they find that Preqin contains the largest total number of funds in the 1990s and 2000s (but not in the 1980s). For venture capital ("VC") funds, Preqin has slightly weaker coverage in the 1980s and 1990s, but it is the most comprehensive dataset in the 2000s. Importantly, of all the datasets, the Preqin data contain performance information for the greatest number of both BO and VC funds. Moreover, they find no evidence that Preqin's performance data are biased relative to the performance data from other data sources. Hence, when analyzing the performance and persistence of PE funds, the Preqin data are among the best data sets currently available.

The Preqin data contain information about each fund's type. The two main fund types are buyout ("BO") and venture capital ("VC") funds, but Preqin also classifies funds as: real-estate, fund-of-funds, infrastructure, turn-around, special situations, co-investment, and venture debt funds, which we collectively refer to as "Other" funds. The majority of "Other" funds are real-estate and infrastructure funds, and while these two fund types are quite different, we find that they have (surprisingly) similar performance and persistence, and we combine all of these fund types for most of our analysis. We define a fund's

geographical location by the location of its GP. This location may differ from the locations of the portfolio companies, but we obtain very similar results when we instead define location in terms of the fund's geographical investment focus.

Sample We restrict our sample to funds with available performance information. Our model explicitly captures the timing of the individual funds; it does not hinge on the numbering of funds, and our estimates are valid even when performance data are missing for some (randomly chosen) funds. We avoid concerns raised in recent studies about funds' self-reported intermediate IRRs, TVPIs and NAVs ("net-asset values"), by restricting our sample to fully liquidated funds. Finally, we restrict our sample to funds with at least USD 5M of committed capital (in 1990 dollars) to exclude smaller idiosyncratic funds. Our final sample contains 1,924 funds, raised between 1969 and 2001, and managed by 891 firms. There are 842 venture capital ("VC") funds, 562 buyout ("BO") funds, and the remaining 518 funds are classified as Other funds. Table I shows summary statistics for our final sample. Panel B shows sub-classifications of VC and Other funds.

** TABLE I: SUMMARY STATISTICS **

Internal Rate of Return Preqin reports each fund's internal rate of return ("IRR"). The IRR is the annualized return to the limited partners ("LPs") in the fund, net of performance fees ("carried interest" or "carry") and management fees. While the IRR has well-known limitations, it is the most widely available fund performance measure and commonly used in studies of fund performance.

** TABLE II: FUND IRRs BY VINTAGE YEAR **

Table II reports the average IRR for each year. These IRRs are plotted in Figure 2 for VC, BO, and Other funds. For VC funds, we see strong performance during the dot-com bubble in the late 1990s, with average (annualized) IRRs as high as 45.2%, followed by the sharp drop after the bursting of the dot-com bubble. Each fund has a ten-year life, and the indicated year is the fund's year of inception ("vintage year"), so funds with vintage years well before 2000 were exposed to the bubble and show lower performance. BO

performance has been more stable, and it has recently shown a strong recovery relative to VC and Other funds. The performance of Other funds has been even more stable, showing an earlier but more modest decline in the late 1990s, followed by a corresponding recovery.

** FIGURE 1: IRRs BY VINTAGE YEAR **

Our analysis uses total log-returns (or "continuously compounded" returns) rather than annualized IRRs, which is reported by Preqin. The total (log-)return for fund u is denoted y_{iu}, and it is calculated by compounding the fund's IRR over its ten year life, as follows:

$$y_{iu} = 10 \cdot \ln(1 + IRR_{iu}).$$ (2)

This calculation fails for two funds that have IRRs of -100% (one is a 2001 VC fund and the other is a 1998 BO fund). Our analysis excludes these two funds, but our results are robust to including them with IRRs set equal to the first (lowest) percentile of the IRR distribution.

II Variance Decomposition

Model Introduction   For our empirical analysis, we use a *hierarchical linear model*, which generalizes the classical *analysis of variance* ("ANOVA") decomposition. Hierarchical models, using Bayesian estimators that exploit advances in numerical computing (Markov-Chain Monte Carlo, Gibbs sampling, and posterior augmentation), have recently been extensively developed and applied. These models were initially used for educational measurement, because they capture the hierarchical structure that arises when, for example, one observes individual students, who are grouped into classrooms, in different schools, in different districts, etc. (For introductions to hierarchical models and more applications see Raudenbush and Iryk [2008] and de Leeuw and Meijer [2008].) This hierarchical structure also arises for PE when individual PE funds are managed by different PE firms and span different time periods (with data for individual deals, as in Braun, Jenkinson, and Stoff [2013] or with LPs' holdings of PE funds, as in Sensoy, Wang, and Weisbach [2013], our model extends to data at these additional levels as well).

Modeling the hierarchical structure avoids the *unit of analysis* problem (Burstein et al. [1980]). When studying the persistence of PE performance, we are interested in differences between PE firms, so the the unit of analysis is a PE firm, but the unit of observation is the underlying funds, which are *repeated measures* of the PE firm's quality. Increasing the number of funds per firm improves the estimate of each firm's quality but not the number of firms that are compared. With few firms but many funds per firm, observing even more funds per firm becomes uninformative, because the main sampling error arises from the sampling of the firms' qualities, not the sampling of the funds observed for each firm. In contrast, increasing the number of observed firms always improves the estimates. It is difficult for classical regression models, using PE firm-fixed effects ("FEs"), to address this problem, because these models only consider the sampling of the funds for a given set of PE firms (i.e., a given set of PE-firm FEs), not the sampling of the PE firms themselves (i.e., the sampling of the observed FEs from a larger population of potential FEs).

Economic Intuition To illustrate the intuition behind our variance decomposition, consider 60 PE firms. Each firm makes two investments (or manages two funds), each of which either succeeds or fails. For the resulting 120 investments, say, we observe that one half fails and the other half succeeds, so the unconditional success probability is 50%. If the individual investments were statistically independent, each investor would have 25% probability of zero successful investments, 50% probability of a single success, and 25% chance of two successes. We would then see 15 of the 60 PE firms with no successes, 30 with a single one, and the remaining 15 firms with two successful investments. Imagine instead that the observed successes are evenly distributed among the 60 PE firms, so 20 have zero, 20 have one, and 20 PE firms have two successes. In other words, the performance variation *between* PE firms exceeds the amount of variation that is implied by the investments *within* PE firms, if the investments were independent. In this case, the investments cannot be independent, obviously, so some PE firms must have higher (and lower) success probabilities. In other words, some PE firms persistently show better (and worse) performance. For example, the even distribution of success among PE firms is consistent with each PE firm's success probability being drawn from the uniform distribution on $[0, 1]$.

If p_i denotes firm i's success probability, then the expected probability of two successes is $E[p_i^2] = 33\%$ when $p_i \sim U[0,1]$. Based on this intuition, we define and measure persistence by comparing the performance variability *within* funds to the performance variability *between* PE firms. When there is excess variation between firms, as in this example, it implies persistence.

This intuition also leads to a natural distinction between PE firms with high skill and high performance. With $p_i \sim U[0,1]$, using Bayes rule, conditional on observing two successes, the posterior density of p_i is $f(p_i|SS) = 3p_i^2$. Hence, the probability that a firm with top-tercile performance (two successes) has top-tercile skill is $\Pr(p_i \in [0.66;1]|SS) = 70\%$. And the expected success probability for a subsequent fund by a firm with top-tercile performance is only $E[p_i|SS] = 75\%$, whereas the success probability for a firm with actual top-tercile skill is $E[p_i|p_i > 66\%] = 83\%$.

Note that performance is a noisy indicator of skill even when the skill distribution is perfectly known (i.e., $p_i \sim U[0,1]$ is known). When the skill distribution is estimated, additional uncertainty arises due to estimation error. Our model, which we discuss next, incorporates this parameter uncertainty as well.

Formal Model Let PE firms be indexed by i. Each PE firm manages a sequence of underlying PE funds, indexed by u. Each PE fund is managed by a single PE firm, so a fund uniquely identifies its firm. Each observation contains the fund's performance and other characteristics of the firm and fund. We specify the ten-year total log-return of fund u as:

$$y_{iu} = X'_{iu}\beta + \sum_{\tau=t_{iu}}^{t_{iu}+9} (\alpha_i + \eta_{i\tau}) + \varepsilon_u. \tag{3}$$

Here, X_{iu} contains time fixed effects for the timing of the fund (formally, the model is then a *mixed-effects model*). The sum runs over the ten years where the fund is alive and active, with year t_{iu} denoting the fund's first year of operations ("vintage year"). The three terms α_i, $\eta_{i\tau}$, and ε_u are three random effects that define the variance-covariance structure across the funds' performances. Our model cannot determine when a given fund's return is earned during its life, because we only observe each fund's ultimate performance. The model can determine, however, how much of the variation in this performance that is due to each of

10

the three random effects.

Statistical Properties   The random effects in equation (3) decompose the variation in fund performance into three parts: First, α_i is the PE-firm effect, reflecting long-term persistence. For each PE firm, it is distributed $\mathcal{N}\left(0,\sigma_\alpha^2\right)$, and it is constant for all funds managed by the same firm. We interpret a PE firm with a high α_i as having greater skills (corresponding to a higher success probability, p_i, in the example). The model is parameterized with α_i inside the sum in equation (3), so each fund "earns" α_i ten times, and α_i is the annualized return to the PE firm's skill. Second, $\eta_{i\tau}$ is the PE firm-time effect. For each firm and year, $\eta_{i\tau}$ is distributed *i.i.d.* $\mathcal{N}\left(0,\sigma_\eta^2\right)$. Two partially overlapping funds that are managed by the same firm will share an $\eta_{i\tau}$ term for each year of overlap, which introduces correlation between partially overlapping funds that are managed by the same firm. Third, ε_u is an error term, capturing the residual idiosyncratic variation in each fund's performance. Because fund performance is highly skewed, we allow ε_u to be distributed as a mixture of normals, which is considerably more flexible than the normal distribution.[4]

The sum in equation (3) contains the same α_i term ten times, and ten *i.i.d.* $\eta_{i\tau}$ terms, so the total variance of y_u is:

$$\sigma_y^2 = 100\sigma_\alpha^2 + 10\sigma_\eta^2 + \sigma_\varepsilon^2. \tag{4}$$

Economic Motivation   The three random effects are motivated as follows: First, some PE firms may have particular investment or management skills that improve the performance of all their funds. Such long-term persistence is captured by the α_i term, the variation in α_i across PE firms captures differences in skills across PE firms. When there is little variation in α_i, corresponding to a small σ_α^2, then PE firms are similar, and there are few persistent differences in their performance. When σ_α^2 is large, more of the performance difference is due to heterogeneous skills of the PE firms.

Second, PE firms typically manage several contemporaneous funds, and they make simultaneous management and investment decisions across funds. At any time, a PE firm may be attracted to a particular technology, industry, geography, or management practice.

---

[4]Using Bayes factors to test model specifications, we find that VC performance requires a mixture of three normals whereas the performance of Buyout and Other funds are captured by mixtures of one or two normal distributions.

For example, a PE firm that manages two funds with vintage years 1999 and 2001 may be focusing on investments in emerging markets for both funds. Hence, these two funds will be exposed to similar shocks during 2001–09. In this case, a regression of the performance of the latter funds on the performance of the former funds would result in a positive coefficient. This coefficient, however, is not evidence of persistence, as usually defined, and it does not imply that the PE firm's past performance predicts its future performance. Instead, the coefficient arises from the spurious correlation due to the unobserved common component shared by the funds. In our model, these shared components are captured by $\eta_{i\tau}$. All overlapping funds that are managed by the same PE firm share an $\eta_{i\tau}$ term for each year of overlap. These shared terms capture the increasing correlation between funds with greater overlaps. When the estimated σ_η^2 is large, this overlap effect is large. Formally, the covariance between two funds that are managed by the same PE firm, with N years of overlap, is:

$$\text{COV}(y_{iu}, y_{iv}) = 100\sigma_\alpha^2 + N\sigma_\eta^2. \tag{5}$$

This covariance relationship is plotted in Figure 2, and this figure illustrates the identification of the model. The main parameters of interest are the variances of the three random effects, σ_α^2, σ_η^2, and σ_ε^2. In Figure 2, the intercept is σ_α^2 and the slope is σ_η^2, so these two variances are identified by comparing the covariances of funds with increasing amounts of overlap. Given σ_α^2 and σ_η^2, and observing total variance, σ_y^2, the residual variance in equation (4) identifies σ_ε^2.

** FIGURE 2: OVERLAP AND COVARIANCE **

III   Results

A   IRR Regressions

We first confirm the original findings by Kaplan and Schoar [2005] using our data. Table III reports coefficients from OLS regressions of $IRR_{i,N}$ on $IRR_{i,N-1}$, and the reported coefficients show that the previous fund's performance strongly predicts the performance of the subsequent fund. In Specification I, the positive and significant coefficient of 0.125

suggests that a VC fund with a 1% higher IRR predicts a 0.125% higher IRR for the subsequent fund. Specification II suggests that this effect is even stronger when controlling for the performance of fund $N - 2$, although the coefficient on this second fund's performance is negative. For BO funds we find similar positive and significant effects. For Other funds, however, the coefficient is initially positive and significant, but it becomes smaller and insignificant after including fund $N - 2$. The weaker statistical results may be due to the smaller sample size. Moreover, fund $N - 2$ may still be partially overlapping with fund N, so the positive coefficients may also reflect this overlap rather than actual persistence. In Panel B of Table II, we reduce the sample to funds that are entirely non-overlapping, which reduces the sample size substantially, leaving no remaining signs of persistence, but it is difficult to disentangle this weaker result from the low statistical power due to the small sample size. Note also that none of the specifications in the two panels suggest that performance is systematically related to fund size, and there is some weak evidence that a higher sequence number is associated with better performance.

<center>** TABLE III: IRR REGRESSIONS **</center>

A natural interpretation of these results is that BO funds have the most persistence (largest coefficients and R^2, and still significant with fund $N - 2$), followed by VC funds (smaller coefficients and R^2 than BO funds, but still significant with fund $N - 2$), and that Other funds show the least, if any, performance persistence (smallest coefficients and R^2, and insignificant with fund $N - 2$). This analysis and interpretation, however, does not distinguish skill from luck, and it does not distinguish long-term from investable persistence.

B Long-Term Persistence

Table IV reports the estimated parameters of our model. Panel A shows the magnitudes of the three random effects as measured by their standard deviations (σ_α, σ_η and σ_ε).[5] The

---

[5]We use a Bayesian estimator, but we report results using standard frequentist terminology: The "point estimate" is the mean of the posterior distribution, and the "standard error" is the the standard deviation of the posterior distribution. A parameter is "statistically significant," at a given level, when zero is not contained in the corresponding symmetric credible interval, as usually defined in Bayesian statistics. Our Bayesian estimator produces exact small-sample inference, even for non-linear transformations of the estimated parameters, and all reported inference is calculated this way. We do not rely on any asymptotic approximations.

decomposition of the variances ($100 \times \sigma_\alpha^2$, $10 \times \sigma_\eta^2$, σ_ε^2 and σ_y^2) is easier to interpret, and is reported in Panel B. Next, we discuss the interpretation of these estimates in detail.

** TABLE IV: PARAMETER ESTIMATES **

Buyout   For BO funds, Specification I of Table IV shows a total unconditional variance (σ_y^2) of 2.428. This variance can be decomposed into three effects, with 0.361 due to long-term persistence ($100 \times \sigma_\alpha^2$), 0.216 due to the overlap effect ($10 \times \sigma_\eta^2$), and the remaining 1.852 due to idiosyncratic variance (σ_ε^2). The long-term persistence effect, as measured by σ_α, is statistically significant,[6] consistent with the earlier findings using the AR(1) regression.

To evaluate the economic magnitude of the long-term persistence, note that the annual contribution of a PE firm's skill is α_i, which is distributed $\mathcal{N}\left(0, \sigma_\alpha^2\right)$. For notation, let $q_\alpha(\cdot)$ denote the percentiles of the α_i distribution, calculated using σ_α^2. For example, if α_i were distributed standard normal, i.e., $\alpha \sim \mathcal{N}(0,1)$, then $q_\alpha(50\%) = 0$ and $q_\alpha(97.5\%) = 1.96$. With the point estimate of σ_α of 0.060, the marginal (worst) top-quartile BO firm has an α_i of $q_\alpha(75\%) = 4.05\%$, annually. And the median top-quartile firm has an α_i of $q_\alpha(87.5\%) = 6.90\%$. Hence, the spread between the marginal top- and bottom-quartile firms, due to skill, is $q_\alpha(75\%) - q_\alpha(25\%) = 8.09\%$, annually. This calculation, however, assumes that the skill distribution is perfectly estimated. In Table IV, Specification I for Buyout funds shows a standard error of 0.008 for the σ_α estimate. Our Bayesian estimation procedure simulates the full posterior distribution of σ_α, and we can calculate the corresponding posterior distribution of $q_\alpha(75\%) - q_\alpha(25\%)$. The mean of this posterior distribution, which also accounts for the estimation error in the skill distribution, is 7.93%, as reported in Table IV. This estimate of 7.93% is close to the estimate of 8.09%, which was calculated using the point estimate of σ_α without adjusting for the estimation error, and the effect of this adjustment seems minor. Nevertheless, because the adjustment is simple to calculate, all the reported alpha spreads in Table IV adjusts for estimation error in the skill distribution.

---

[6]Testing for statistical significance of variance parameters is complicated by the asymmetric alternative hypothesis. We use a Bayes factor test to test $H_0 : \sigma_\alpha^2 = 0$ against $H_A : \sigma_\alpha^2 > 0$, as reported in Table VII and discussed in the Appendix.

We estimate two alpha spreads: The interquartile range is the difference between the marginal (worst) top-quartile and the marginal (best) bottom-quartile firms. It is denoted $q_\alpha(75\%) - q_\alpha(25\%)$, and our point estimate of this difference is 7.93%, annually. In other words, the marginal BO firm with top-quartile skills outperform the marginal firm with bottom-quartile skills by 7.93% annually. We also report the average difference in the performance between the median top- and bottom-quartile firms, denoted $q_\alpha(87.5\%) - q_\alpha(12.5\%)$, and our estimate of this difference is 13.63%, annually.

Note that these alpha spreads cannot be calculated as the empirical difference between the IRRs of top- and bottom-quartile funds, because top-quartile performance does not imply top-quartile skills, so this empirical difference confounds skill and luck. If σ_ε^2 is large, but σ_α^2 is zero, there is no long-term persistence, and $q_\alpha(75\%) - q_\alpha(25\%)$ is zero. But a large σ_ε^2 still implies a large difference in fund performance, albeit due to noise, so the empirical difference would still be large, and in this case it would overstate the performance that is due to heterogeneous skills. The empirical difference may also understate long-term persistence. In periods where a disproportionate number of high-quality (or low-quality) firms are active, the empirical difference may be too small, because it is calculated from funds in a narrow range of the α_i distribution. For this reason, it is important that our model includes PE firms that raise only a single fund. These firms are likely from the tail of the α_i distribution, so excluding them would lead to a downward bias in σ_α^2 and underestimate the long-term persistence.

Venture Capital For VC funds, the variance that is due to long-term persistence ($100 \times \sigma_\alpha^2$) is 0.243. This variance is similar to the one for BO funds, and therefore the alpha spreads are also similar. Specifically, for VC firms, $q_\alpha(87.5\%) - q_\alpha(12.5\%) = 11.17\%$ and $q_\alpha(75\%) - q_\alpha(25\%) = 6.50\%$, annually. The variance due to the overlap effect ($10 \times \sigma_\eta^2$) is 0.675, which is somewhat larger than for BO funds, but this difference disappears with year FEs. Importantly, there is a large difference between VC and BO funds in idiosyncratic variance, which is 3–4 times larger for VC funds. Hence, even though the difference in skills between good and bad firms is similar for VC and BO, the performance of VC firms is much more noisy. This noise may also explain the weaker persistence results for VC

funds using the AR(1) regression, because more noisy outcomes leads to weaker statistical power in this model.

Other   For Other funds, the overlap and long-term persistence effects are similar to those of BO and VC funds. In Table IV, Specification I shows that σ_α is very close for Other and VC funds, so their alpha spreads are almost identical: $q_\alpha(87.5\%) - q_\alpha(12.5\%) = 11.21\%$ and $q_\alpha(75\%) - q_\alpha(25\%) = 6.52\%$, annually. The idiosyncratic volatility, however, is lower for Other funds.

Overall, these estimates of the long-term persistence, as measured by the performance differences of high- and low-skilled funds, show that BO funds may have slightly more long-term persistence than VC and Other funds, which are very similar, but the differences in the economic magnitudes of these persistence estimates are modest. These alpha spreads are calculated from the underlying true distribution of the skills of PE firms. For an LP to earn the full spread, the LP must be able to perfectly assess these skills, and hence these spreads represent an upper bound on the value of the persistence to the LPs.

C   Overlap Effect

A partial overlap of funds induces correlation between subsequent funds. In Table IV, Specifications I and II for BO firms show overlap effects of 0.216 and 0.432. This overlap effect is easier to interpret with vintage-year FEs. Without time FEs, the $\eta_{i\tau}$ terms capture all correlations between contemporaneous funds, including correlations due to general market exposure. Funds that overlap are exposed to the same market returns during the overlap period, and this leads to a correlation in their performance, but this correlation is common across all PE firms, and it is not due to the actions of any particular firm. To control for these common exposures, Specification II includes year fixed effects. With year fixed effects, the overlap effect is isolated to funds that are managed by the same PE firm. Comparing these specifications across fund types, we see that BO funds have the largest, and Other funds have the smallest overlap effects.

This overlap effect is important, because the AR(1) regression of the performance of fund N on fund $N-1$ will find a positive coefficient due to this effect, but this positive coef-

ficient does not indicate persistence in the conventional sense. The descriptive statistics in Table I shows average overlaps in the range 5.8–6.8 years. With this amount of overlap, the variance estimates in Table IV imply that the covariance in the performance of funds with average overlaps are in the range 0.37–0.64. But 25.8%–61.8% of this covariance is due to the overlap, not long-term persistence, as measured by σ_α^2, suggesting that the magnitudes of the estimated coefficients in the AR(1) regression may be upward biased by 34%–168%. Moreover, this bias is smaller for Other firms, which have a smaller overlap effect, and this smaller upward bias in the coefficient from the AR(1) regression may partially explain the weaker persistence results for Other firms found using the AR(1) model.

IV   Learning and Investable Persistence

The previous section considered long-term persistence, defined as the performance of funds managed by PE firms with better skills, assuming that these skills are perfectly known. In practice, however, it is difficult for LPs to identify skilled firms. For example, Phalippou [2010] finds that the previous fund's interim performance, which is the only performance that is known when a subsequent fund is raised, is not statistically significant for predicting the performance of the subsequent fund. In other words, the interim performance of the previous fund, by itself, is insufficient for LPs to identify skilled PE firms. This finding points to a more general question: For an LP to evaluate an investment in a new fund, how much information does the LP need to determine the PE firm's skill? If LPs only need little information, such as just the interim performance of the previous fund, it would be easy for them to identify the skilled PE firms, and we would say that PE performance has a large amount of investable persistence.

We quantify the amount of investable persistence in two steps. First, we estimate the signal-to-noise ratio. This ratio is simple to calculate, it allows for a direct comparison of different types of firms, and it has a simple economic intuition based on the updating of the LPs' beliefs about a PE firm's skills. The disadvantage is that this ratio does not reflect all the features of the statistical model. Consequently, in the second step, we use the full model to estimate how many past funds an LP must observe to assess the firm's

skill with reasonable certainty. Overall, we find that the signal-to-noise ratio is low, and it is difficult for LPs to identify skilled PE firms based on their past performance alone. An LP would need to observe an excessive number of past funds to evaluate an PE firm's skill with reasonable confidence. Additional information is needed, such as details about individual deals, individual partners associated with these deals (see Ewens and Rhodes-Kropf [2013]), or other additional information.

A Signal-to-Noise

Our model has two types of shocks: The transitory shocks are drawn independently each period, as given by the $\eta_{i\tau}$ and ε_u terms. The persistent shocks, reflecting the heterogeneous skills of the PE firms, are given by α_i. As noted above, we find that VC, BO, and Other funds have similar amounts of long-term persistence, as measured by σ_α^2. But VC funds have much greater transitory shocks. VC performance is more noisy, and it is more difficult for LPs to identify skilled VC firms.

We define the signal-to-noise ratio, s_α, as the ratio of persistent to total variation (e.g., Cochrane (1988) uses a similar variance ratio to study the persistence of GDP shocks):

$$s_\alpha = \frac{100\sigma_\alpha^2}{\sigma_y^2}. \tag{6}$$

In our application, this ratio has a simple economic interpretation. In a Gaussian learning model, an LP would update its beliefs about the α_i of a given PE firm as follows. Let the LP's prior beliefs be $\alpha_i \sim \mathcal{N}\left(\alpha_0, \sigma_0^2\right)$, with $\alpha_0 = 0$, and the beliefs after observing N funds is $\mathcal{N}\left(\alpha_N, \sigma_N^2\right)$. After observing one additional fund, the mean and variance of the LP's updated beliefs are then:

$$\alpha_{N+1} = s_\alpha \times \frac{\left(y_{i,N+1} - X_u'\beta\right)}{10} + (1 - s_\alpha) \times \alpha_N, \tag{7}$$

and

$$\sigma_{N+1}^2 = (1 - s_\alpha) \times \sigma_N^2. \tag{8}$$

Hence, the signal-to-noise ratio reflects how much weight an LPs should place on new

information. When the ratio is low, new performance is largely uninformative about the firm's skills, and it is difficult for an LP to learn this skill. Conversely, the larger is s_α, the faster the LPs learns α_i, as measured by a lower σ^2_{N+1}.

Figure 3 plots the estimated posterior distribution of s_α for VC, BO, and Other firms. For VC firms, the large idiosyncratic variance means that relatively little variation in fund performance is due to long-term persistence, and VC performance is largely uninformative about the skills of VC firms. For BO firms, a greater fraction is due to long-term persistence, making it easier to assess their skill. Other firms have the least idiosyncratic variation, and the best signal-to-noise ratio, making it is easiest for LPs to identify skilled Other firms.

** FIGURE 3: ESTIMATES OF SIGNAL-TO-NOISE RATIO **

B Identifying Skilled PE Firms

Using the estimated parameters in the full model, Figure 4 shows how quickly an LP can identify the skills of a PE firm by observing the performance of the PE firm's past funds. To interpret Figure 4, consider first a limit case where there is no long-term persistence and σ^2_α is zero. Some PE firms show top-quartile performance, but this performance is entirely due to luck. Hence, without any persistence, the probability that a PE firm with top-quartile skills also has top-quartile performance is just 25%. With some persistence, the probability increases, and this probability is plotted in Figure 4.

With five past funds, a VC firm with top-quartile skills has a 37% probability of also having top-quartile performance, as illustrated in Figure 4. While this is better than the 25% probability in the uninformative case, without any persistence, it is not much better. For BO and Other firms, this probability increases to 47% and 51%. Figure 4 shows these probabilities for up to 50 observed past funds. Since no current PE firm has 50 fully liqui-dated independent funds, this case is an upper bound on the information that available from past fund performance. Even for this upper bound, a VC firm with top-quartile skills has just a 53% probability of showing top-quartile performance. In other words, if an LP that invests in all VC firms with top-quartile performance, even after observing 50 past funds

for each firm, then 47% of these VC firms do not actually have top-quartile skills. For BO and Other PE firms, the corresponding probabilities improve to 58% and 60%.

** FIGURE 4: LEARNING SPEED **

C   Investable Persistence

Combining the findings about long-term term persistence and the ability of LPs to identify skilled firms provides a measure of the investable persistence. Other firms have the best signal-to-noise ratio, and it is easiest for LPs to identify the skilled firms of this type. However, BO firms have greater long-term persistence than Other firms, so it is more valuable to identify skilled BO firms. To combine these two effects, Figure 5 reports the expected alpha that an LP earns when investing in the firms with top-quartile performance as a function of the number of observed funds. The more funds that the LP has observed, the better the LP can identify the skilled firms, particularly for Other firms. Figure 5 shows that BO has greater investable persistence than Other firms. The advantage of Other firms is that they have less idiosyncratic risk, so past performance is more informative and it is easier to identify the skilled firms. The advantage of BO firms is that they have greater long-term persistence, so the value of identifying their skills is larger, even if the skilled firms are harder to identify. VC firms, however, have weak investable persistence. They suffer both from a high idiosyncratic risk, so skilled VC firms are difficult to identify, and from low long-term persistence, making it less valuable to identify skilled VC firms.

** FIGURE 5: INVESTABLE PERSISTENCE **

V   Subsamples

Table VI shows estimates of our model for different subsamples. To illustrate how we calculate these estimates, when we compare persistence in large and small funds in the first panel in Table VI, we divide the funds into two separate samples, and estimate the model on each sample separately, so a PE firm that manages both small and large funds will be

represented in both samples, but only the small funds are included in the first estimation, and no information from this first estimation is used in the second one.

<center>** TABLE VI: SUB-SAMPLES **</center>

Fund Size   In Panel B of Table VI, the columns with $100\sigma_\alpha^2$, show that smaller funds have greater long-term persistence than larger funds, across fund types. For VC and Buyout funds, the small ones also have more volatile performance, as indicated by σ_y^2. But despite these more volatile performances, Panel A shows that the long-term performances are sufficiently large that the signal-to-noise ratios are better for smaller funds, across all types. For smaller funds, past performance is a better predictor of subsequent performance. In contrast, larger funds have less long-term persistence and worse signal-to-noise ratios. The worst ones are large VC funds, which have almost entirely uninformative performance.

GP Location   Across PE firm types, Panel B in Table VI shows that there is more long-term persistence for PE firms located in rest of world ("ROW"), with Europe second, and least persistence for US-based PE firms. However, the total performance volatility also follows this pattern, and ROW performance is substantially more volatile than performance in Europe or US. Comparing the informativeness of the performance, VC and Buyout performance is most informative in Europe, whereas the performance of Other PE firms is more informative in ROW. In fact, these Other PE firms in ROW have both the largest long-term persistence and the most informativeness performance.

Investment Style   We separate VC and Other funds into different investment styles, as classified by Preqin. VC firms with funds focusing on early-stage investments have the least long-term persistence and the least informative performance. Generalist VC firms have slightly more persistence, but slightly less informative performance than VC firms specialized in late-stage investments, which have much lower idiosyncratic volatility than other types of VC investments.

For other funds, we distinguish real-estate funds from fund-of-funds. These two types of funds are very different, but they have surprisingly similar persistence characteristics.

<center>21</center>

The differences are that fund-of-funds have slightly greater long-term persistence, and the performance of real-estate funds is slightly more informative.

Time Period We confirm the finding by Braun, Jenkinson, and Stoff [2013] and Harris, Jenkinson, Kaplan, and Stucke [2013], that persistence has declined over the sample period. Table VI reports persistence estimates in the earlier and later half of our sample. This decline is particularly pronounced for VC funds, were both the amount of long-term persistence and the performance informativeness have almost vanished. For buyout funds, there is still substantial persistence and informativeness. Recently, however, the performance of other funds is most informative, even if they show slightly lower long-term persistence than Buyout funds.

VI   Conclusion

We decompose private equity ("PE") performance into skill and luck. When performance is noisy, PE firms with top-quartile skills will not necessarily show top-quartile performance, and this distinction leads to two different notions of performance persistence. The first type, long-term persistence, reflects the performance differences between funds managed by skilled and unskilled PE firms. Across all types of PE firms, we find a large amount of long-term persistence, and skilled firms outperform by 7% to 8% annually. The second type of persistence is investable persistence. It reflects the ability of LPs to identify skilled PE firms using information about their past performance. We find that past performance is very noisy, it has a large component of luck and a poor signal-to-noise ratio, making it difficult for LPs to identify skilled firm. VC firms, in particular, have poor investable persistence. To identify skilled PE firms, LPs need information beyond what is contained in the performance of a PE firm's past funds.

Subsamples Comparing subsamples, we find that smaller funds have greater persistence than larger funds. In particular, large VC funds have poor long-term and investable persistence. Across geographical locations of PE firms, we find the least persistence for PE firms located in the US, followed by Europe, and the greatest persistence for PE firm located

22

in the rest of the world ("ROW"), although these PE firms also have more volatile performance. Finally, we confirm the finding by Harris, Jenkinson, Kaplan, and Stucke [2013] and Braun, Jenkinson, and Stoff [2013] that persistence has declined in the 2000s relative to the 1990s. This decline is largest for VC firms, though, and we find that Buyout and Other funds still show substantial long-term persistence, even post 2000.

Implications Our results have three implications for understanding PE performance: First, the low investable persistence explains LPs' increasing interest in collecting more detailed information about PE performance. For example, Ewens and Rhodes-Kropf [2013] study performance using deal- and partner-level information. We show that such detailed information is necessary for LPs to identify top PE firms. Second, our results provide an alternative explanation of why persistent outperformance is not competed away. When identifying skilled PE firms is difficult, LPs with this ability may also be scarce, and LPs with this ability should earn rents. Third, the large idiosyncratic risk and lower skill components of VC performance is consistent with the anecdotal saying among VCs that "I'd rather be lucky than smart."

Appendix: Estimation Procedure

We implement the model as a Bayesian multi-level hierarchical model, redefining the error terms to absorb the firm-specific random effects using hierarchical centering, as recommended by Gelfand, Sahu, and Carlin (1995). The performance of fund u of firm i is:

$$y_{iu} = X_{iu}\beta + \sum_{\tau=t_{iu}}^{t_{iu}+9} \eta_{i\tau} + \varepsilon_{iu}, \tag{9}$$

The conditional distributions of the random effects are given as:

$$\eta_{i\tau}|\alpha_i \;\sim\; \mathcal{N}\left(\alpha_i, \sigma_\eta^2\right), \tag{10}$$

$$\alpha_i \;\sim\; \mathcal{N}\left(0, \sigma_\alpha^2\right). \tag{11}$$

The fund-specific error term distribution is IID

$$\varepsilon_{iu} \sim \mathcal{N}\left(0, \sigma_\varepsilon^2\right). \tag{12}$$

We are interested in estimating the parameter vector $\theta \equiv \left(\beta, \sigma_\alpha^2, \sigma_\eta^2, \sigma_\varepsilon^2\right)$, given a dataset of fund returns, $\{y_{iu}\}$, the dates of inception and termination of each fund, and the set of observed fund-level covariates, X_{iu}. We augment the parameter vector with the latent α's and the $\eta's$, and use a Bayesian estimation algorithm that produces a set of draws from the posterior distribution, $f\left(\theta, \{\alpha_i\}, \{\eta_{it}\}|data\right)$, using a Gibbs sampler (Gelfand and Smith, 1990. See also Korteweg, 2013, for a detailed description). By the Hammersley-Clifford theorem, we can divide the posterior into five *complete conditionals* that are easy to sample from:

1. Latent firm-year random effects: $f\left(\{\eta_{it}\}|\{\alpha_i\}, \theta, data\right)$

2. Variance of fund-specific error term and β-coefficients: $f\left(\sigma_\varepsilon^2, \beta|\{\alpha_i\}, \{\eta_{it}\}, \sigma_\alpha^2, \sigma_\eta^2, data\right)$

3. Latent firm random effects: $f\left(\{\alpha_i\}|\{\eta_{it}\}, \theta, data\right)$

4. Variance of firm-year random effects: $f\left(\sigma_\eta^2|\{\alpha_i\}, \{\eta_{it}\}, \beta, \sigma_\alpha^2, \sigma_\varepsilon^2, data\right)$

5. Variance of firm random effects: $f\left(\sigma_\alpha^2|\{\alpha_i\}, \{\eta_{it}\}, \beta, \sigma_\eta^2, \sigma_\varepsilon^2, data\right)$

We sample from each distribution 1 through 5 in turn, after which we return back to step 1 and repeat. The resulting sequence of parameter draws forms a Markov chain, the stationary distribution of which is exactly the posterior distribution. Given a sample of draws of the posterior distribution, it is then straightforward to numerically integrate out the latent variables and obtain the marginal posterior of parameters, $f(\theta|data)$, or the distribution of the random effects, $f(\{\alpha_i\}|data)$ and $f(\{\eta_{it}\}|data)$, for example. We now discuss how to draw from each conditional distribution.

A1 Latent firm-year random effects

The firm-year random effects, η_{it}, are simulated using a Bayesian regression of the fund returns on a set of year indicator variables, with known variance. This is done on a firm-by-firm basis, as the random effects are assumed independent across firms (and time). For each firm, i, the regression model takes the form

$$y_i = X_i\beta + Z_i\eta_i + \varepsilon_i, \tag{13}$$

where y_i is a vector of stacked fund returns for the U_i funds of firm $\underline{i}$, and X_i is the sub-matrix of the covariates $\left[X'_{i1}\ldots X'_{iU_i}\right]'$ for which each row correspond to a fund of firm i. The vector η_i contains the firm-year random effects for the years in which firm i has at least one active fund. The length of the vector η_i is denoted T_i, and may vary by firm. The matrix Z_i is a $U_i \times T_i$ matrix of indicator variables. Each row represents a fund of firm i, and contains ones in the columns that correspond to the years that the fund is active, and zeros in all other columns.

Given the prior in equation (10), and using the standard Bayesian regression setup (e.g., Rossi, Allenby, and McCulloch, 2005), the posterior distribution is

$$\eta_i|\{\alpha_i\},\theta,data \quad \sim \quad \mathcal{N}\left(\mu_\eta,\sigma_\varepsilon^2\Omega^{-1}\right), \tag{14}$$

where

$$\Omega \quad = \quad \frac{\sigma_\varepsilon^2}{\sigma_\eta^2} \cdot \mathbb{I}_{T_i} + Z_i' Z_i \tag{15}$$

$$\mu_\eta \quad = \quad \Omega^{-1} \left(\alpha_i \cdot \frac{\sigma_\varepsilon^2}{\sigma_\eta^2} \cdot 1_{T_i} + Z_i'(y_i - X_i \beta) \right), \tag{16}$$

where $\mathbb{I}_{T_i}$ is the $T_i \times T_i$ identity matrix, and 1_{T_i} a $T_i \times 1$ vector of ones.

A2 Variance of fund-specific error term and β-coefficients

Given the conditioning on the random effects, η_{it}, this step is a standard Bayesian regression. With the conjugate prior

$$\sigma_\varepsilon^2 \quad \sim \quad IG(a_0, b_0) \tag{17}$$

$$\beta | \sigma_\varepsilon^2 \quad \sim \quad \mathcal{N} \left(\mu_0, \sigma_\varepsilon^2 \Sigma_0^{-1} \right), \tag{18}$$

the posterior distribution is

$$\sigma_\varepsilon^2 | \{\eta_i\}, data \quad \sim \quad IG(a, b) \tag{19}$$

$$\beta | \sigma_\varepsilon^2, \{\eta_i\}, data \quad \sim \quad \mathcal{N} \left(\mu, \sigma_\varepsilon^2 \Sigma^{-1} \right), \tag{20}$$

where

$$a \quad = \quad a_0 + \sum_{i=1}^{N} U_i \tag{21}$$

$$b \quad = \quad b_0 + e'e + (\mu - \mu_0) \Sigma_0 (\mu - \mu_0) \tag{22}$$

$$\Sigma \quad = \quad \Sigma_0 + X'X \tag{23}$$

$$\mu \quad = \quad \Sigma^{-1} \cdot \left(\Sigma_0 \mu_0 + X'(y - Z\eta) \right). \tag{24}$$

The vector $y = \left[y_i' \ldots y_N' \right]'$ contains the fund returns stacked across the N firms, X is the matrix of stacked X_i and Z the stacked Z_i. The vector $e = y - Z\eta - X\mu$ contains the stacked error terms .

A3 Latent firm random effects

Drawing the firm random effects, α_i, is similar in spirit to simulating the firm-year random effects in step 1. Write the estimation problem as a regression of the firm-year random effects on a set of indicator variables

$$\eta = W\alpha + \nu, \tag{25}$$

where $\eta = [\eta_1 \ldots \eta_N]'$, and $\alpha = [\alpha_1 \ldots \alpha_N]'$, and $\nu \sim \mathcal{N}\left(0, \sigma_\eta^2 \cdot \mathbb{I}_N\right)$. The matrix W is a $\sum_{i=1}^N T_i \times N$ matrix of indicator variables. Each row of W represents a firm-year, and contains a one in the column of the corresponding firm, and zeros in all other columns.

With the prior in equation (11), the posterior distribution is

$$\alpha | \{\eta_{it}\}, \theta, data \quad \sim \quad \mathcal{N}\left(\mu_\alpha, \sigma_\eta^2 A^{-1}\right), \tag{26}$$

where

$$A \quad = \quad \frac{\sigma_\eta^2}{\sigma_\alpha^2} \cdot \mathbb{I}_N + W'W \tag{27}$$

$$\mu_\alpha \quad = \quad A^{-1}\left(W'\eta\right). \tag{28}$$

A4 Variance of firm-year random effects

The variance of the firm-year random effects, σ_η^2, is the variance of the residuals $v = \eta - W\alpha$ from the regression in step 3. Using the inverse gamma prior

$$\sigma_\eta^2 \sim IG\left(c_0, d_0\right), \tag{29}$$

yields the posterior distribution

$$\sigma_\eta^2 | \{\alpha_i\}, \{\eta_i\}, data \sim IG\left(c, d\right), \tag{30}$$

where

$$c = c_0 + \sum_{i=1}^{N} T_i \tag{31}$$

$$d = d_0 + v'v. \tag{32}$$

A5 Variance of firm random effects

The variance of the firm random effects, σ_α^2, using the inverse gamma prior

$$\sigma_\alpha^2 \sim IG(f_0, g_0), \tag{33}$$

has posterior distribution

$$\sigma_\alpha^2 | \{\alpha_i\}, data \sim IG(f, g), \tag{34}$$

with parameters:

$$f = f_0 + N, \tag{35}$$

$$g = g_0 + \alpha'\alpha. \tag{36}$$

A6 Mixture of Normals Specification

For the mixtures of Normals specification we replace the distribution of the fund-specific error term in equation (12), with a mixture of K Normal distributions,

$$\varepsilon_{iu} \sim \sum_{k=1}^{K} p_k \cdot \mathcal{N}\left(\mu_k, \sigma_{\varepsilon,k}^2\right). \tag{37}$$

Setting K=1 reduces the model to the baseline Normal specification in (12). We drop the intercept in X_{iu} because it is absorbed by the error term, which has mean $E[\varepsilon_{iu}] = \sum_{k=1}^{K} p_k \mu_k$. This specification is equivalent to the specification with an intercept in X_{iu} and zero mean error ε_{iu}, but it is easier to implement because it avoids enforcing cross-parameter restric-

tions on the μ_k. To estimate the mixture model by Gibbs sampler, the procedure requires one more latent variable that indicates which of the K Normal distributions each observation is drawn from. Conditional on this indicator, the Gibbs steps described above remain largely unchanged. For details on estimation of the latent indicator and the parameters of the mixture components we refer to West (1992), Diebolt and Robert (1994), and Chen and Liu (2000).

A7 Priors and Starting Values

Our Gibbs sampler uses 10,000 iterations for the initial burn-in, followed by 100,000 iterations to simulate the posterior distribution. During the burn-in phase, the simulations converge quickly. We use diffuse prior distributions for the parameters, so that our results are driven by the data rather than prior assumptions. First, we set $a_0 = 2.1$, and $b_0 = 1$. This implies that our prior belief is that $E[\sigma_\varepsilon] = 0.854$, and that σ_ε is between 0.362 and 2.874 with 99% probability (note that this is for ten-year fund returns, so the annualized volatility is about a factor 3 lower). Second, we set $c_0 = f_0 = 2.1$, and $d_0 = g_0 = 0.15^2$. Since both the $\alpha's$ and η's are specified at the annual level, this implies that $E[\sigma_\alpha] = E[\sigma_\eta] = 0.128$ per year, and σ_α and σ_η are between 0.054 and 0.431 (annually) with 99% probability. Conditional on X, the prior ten-year fund return variance, $100\sigma_\alpha^2 + 10\sigma_\eta^2 + \sigma_\varepsilon^2$, has an expected value of 1.658, and is between 0.861 and 4.666 with 99% probability. Finally, we set the prior mean for β equal to zero ($\mu_0 = 0$), implying a prior mean fund return of zero. We set Σ_0 equal to the identity matrix, so that the prior β's are between -3.1 and +3.1 with 99% probability.

For the mixtures of Normals specifications we set the prior of each mixture component, $1 \ldots K$, equal to the prior of the error term ε in the Normal model, i.e., mean zero and Inverse Gamma prior parameters equal to a_0 and b_0. This ensures that the prior distribution of y is the same across all K, so that the Bayes Factor (see below) is a valid comparison across different mixtures. The prior distribution of the mixture probabilities, p, is the conjugate Dirichlet distribution, $Dir(K, \delta)$, with $\delta = 1_K \cdot 10$. This implies that all distributions in the mixture have equal prior mean probability, $1/K$.

We start the algorithm with all α's and β's equal to zero (their prior means). We initial-

ize all variances (σ_α^2, σ_η^2, and σ_ε^2) at their prior means. For the mixtures specification, we set the mixture probabilities to their prior mean, $1/K$. We do not need starting values for the η's, since they are the first variables we simulate.

References

Albert, James and Siddhartha Chib, 1993, Bayesian Analysis of Binary and Polychotomous Response Data, *Journal of the American Statistical Association* 88, 669–679.

Bailey, Martin J., Richard F. Muth, and Hugh O. Nourse, 1963, A Regression Method for Real Estate Price Index Construction, *Journal of the American Statistical Association* 58, 933–942.

Baquero, Guillermo, Jenke ter ter Horst, and Marno Verbeek, 2005, Survival, Look-Ahead Bias, and Persistence in Hedge Fund Performance, *Journal of Financial and Quantitative Analysis* 40, 493–517.

Bengtsson, Ola, 2013, Relational Venture Capital Financing of Serial Founders, *Journal of Financial Intermediation* 3, 308–344.

Billingsley, Patrick, 1961, Statistical Methods in Markov Chains, *The Annals of Mathematical Statistics* 32, 12–40.

Braun, Reiner, Tim Jenkinson, and Ingo Stoff, 2013, How Persistent is Private Equity Performance? Evidence from Deal-Level Data, *Working Paper*, Oxford University.

Burstein, L., K. H. Fischer, M. D. Miller, 1980, The Multilevel Effects of Background On Science Achievement: A Cross National Comparison, *Journal of Sociology* 53, 215–225.

Chen, Rong, and Jun S. Liu, 2000, Mixture Kalman Filters, *Journal of the Royal Statistical Society, Series B* 62, 493–508.

Chib, Siddhartha, 1995, Marginal Likelihood from the Gibbs Output, *Journal of the American Statistical Association* 90, 1313–1321.

Chung, Ji-Woong, 2010, Performance Persistence in Private Equity Funds, *Working Paper*, Chinese University of Hong Kong.

Cochrane, John, 1988, How Big is the Random Walk Component in GDP?, *Journal of Political Economy* 96, 893–920.

de Leeuw, Jan and Erik Meijer, 2008, Handbook of Multilevel Analysis, Springer, NY.

Diebolt, J., and Christian P. Robert, 1994, Estimation of Finite Mixture Distributions Through Bayesian Sampling, *Journal of the Royal Statistical Society, Series B,* 56, 363–375.

Ewens, Michael, and Matthew Rhodes-Kropf, 2013, Is a VC Partnership Greater Than the Sum of its Partners, *Working Paper*, Harvard Business School.

Gelfand, Alan E., Sujit K. Sahu, and Bradley P. Carlin, 1995, Efficient parametrisations for normal linear mixed models, *Biometrika* 82, 479–488.

Gompers, Paul, Anna Kovner, Josh Lerner, David Scharfstein, 2010, Performance Persistence in Entrepreneurship, *Journal of Financial Economics* 96, 18–32.

Harris, Robert, Tim Jenkinson, Steven Kaplan, 2013, Private Equity Performance: What Do We Know?, *Working Paper*, University of Chicago.

Harris, Robert, Tim Jenkinson, Steven Kaplan, and Ruediger Stucke, 2013, Has Persistence Persisted in Private Equity? Evidence From Buyout and Venture Capital Funds, *Working Paper*, University of Chicago.

Hochberg, Yael, Alexander Ljungqvist, and Anette Vissing-Jorgensen, 2010, Informational Hold-Up and Performance Persistence in Venture Capital, *Working Paper*, Kellogg.

Jagannathan, Ravi, Alexey Malakhov, and Dmitry Novikov, 2010, Do Hot Hands Exist among Hedge Fund Managers? An Empirical Evaluation, *Journal of Finance* 65, 217–255.

Korteweg, Arthur G., 2013, Markov Chain Monte Carlo Methods in Corporate Finance, in P. Damien, P. Dellaportas, N. Polson, and D. Stephens, eds.: Bayesian Theory and Applications.

Korteweg, Arthur G., and Stefan Nagel, 2013, Risk-adjusting the Returns to Venture Capital, *Working Paper*, NBER working paper #1934.

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

Kass, Robert E., and Adrian E. Raftery, 1995, Bayes Factors, *Journal of the American Statistical Association* 90, 773–795.

Lerner, Josh, Antoinette Schoar, and Wan Wongsunwai, 2007, Smart Institutions, Foolish Choices? The Limited Partner Performance Puzzle, *Journal of Finance* 62, 731–764.

Li, Yi, 2013, Reputation, Volatility and Performance Persistence of Private Equity, *Working Paper*, Princeton University.

Marin, Jean-Michel, and Christian P. Robert, 2008, Approximating the Marginal Likeli-

hood in Mixture Models, *Working Paper*, INRIA and Université Paris-Dauphine.

Marquez, Robert, Vikram Nanda, Deniz Yavuz, 2011, Performance Persistence in Private Equity Fund Returns, *Working Paper*, Purdue.

Phalippou, Ludovic, 2010, Venture capital funds: Flow-performance relationship and performance persistence, *Journal of Banking and Finance* 34, 568–577.

Phalippou, Ludovic and Oliver Gottschalg, 2009, The Performance of Private Equity Funds, *Review of Financial Studies* 22, 1747–1776.

Rossi, Peter E., Greg M. Allenby and Robert McCullough, 2005, Bayesian Statistics and Marketing, John Wiley and Sons, Chichester, UK.

Sensoy, Berk, Yingdi Wang, and Michael Weisbach, 2013, Limited Partner Performance and the Maturing of the Private Equity Industry, *Working Paper*, NBER.

West, M., 1992, Modeling With Mixtures, in Bayesian Statistics 4, eds. J. M. Bernardo, J. O. Berger, A. P. Dawid, and A. F. M. Smith, Oxford, U.K.: Oxford University Press, 503–524.

Figure 1: Vintage Year Fixed Effects. Plots of the posterior mean of the vintage years fixed effects from Specification II in Table IV. The vintage year effects are for transformed to IRR equivalents (annualized, in percent). The shaded bands represents the (1%, 99%) Bayesian credible interval (confidence bounds). The broken line shows the average IRR calculated annually.

Panel A: VC
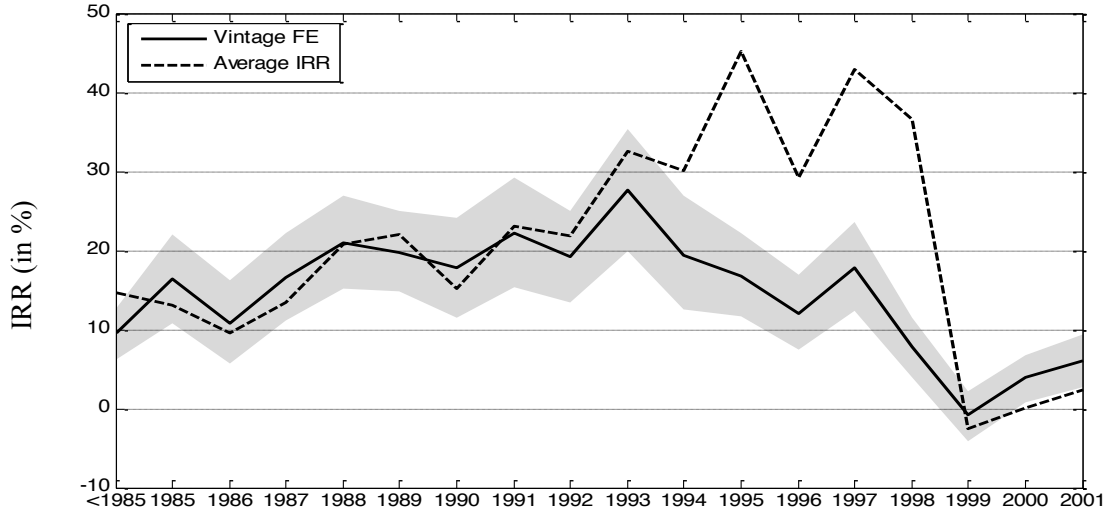
Panel B: Buyout
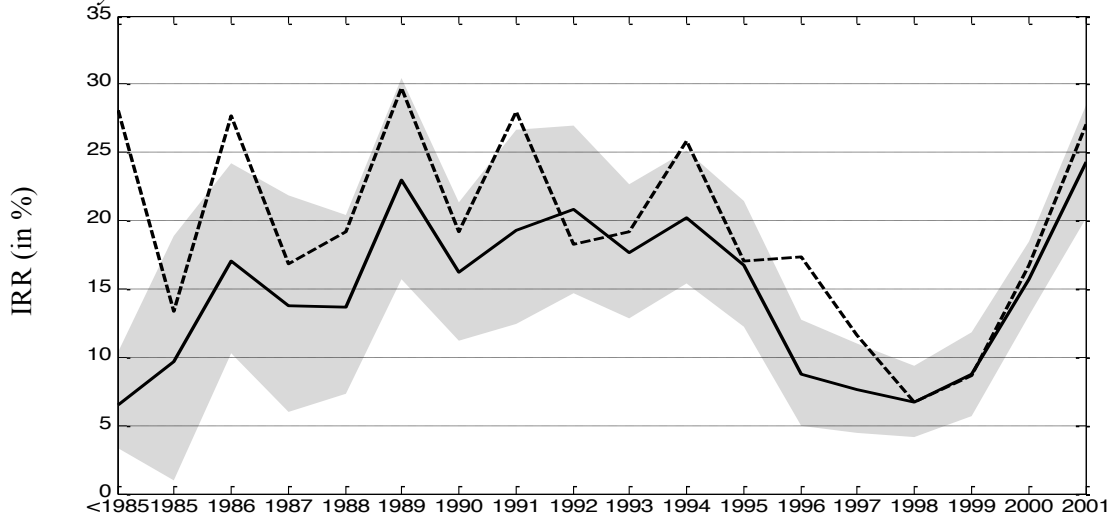
Panel C: Other
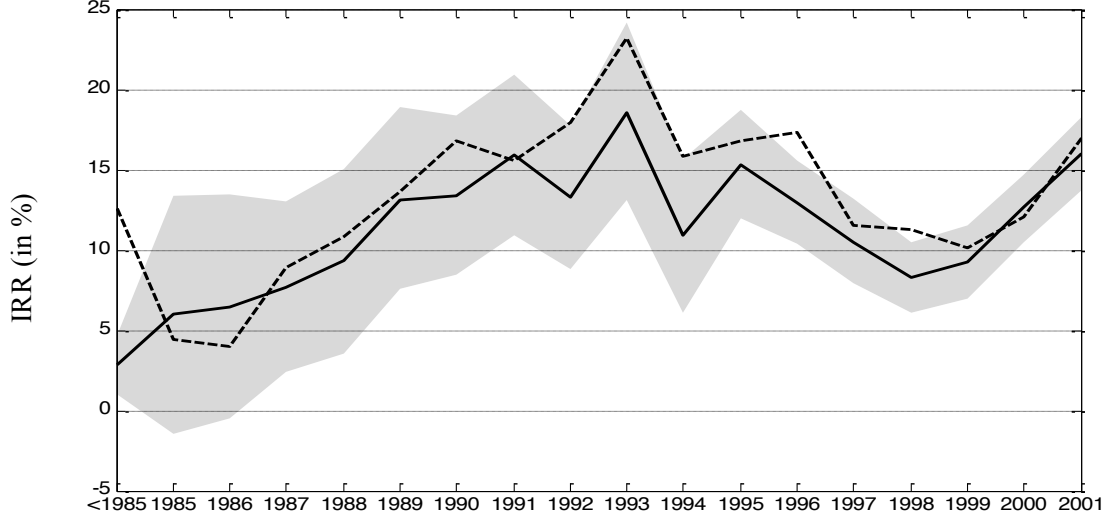
Figure 2: Fund Overlap and Covariance. The figure shows the covariance between total fund returns as a function of the overlap (in years) between two funds managed by the same firm, using the variance estimates for Specifications II in Table IV.
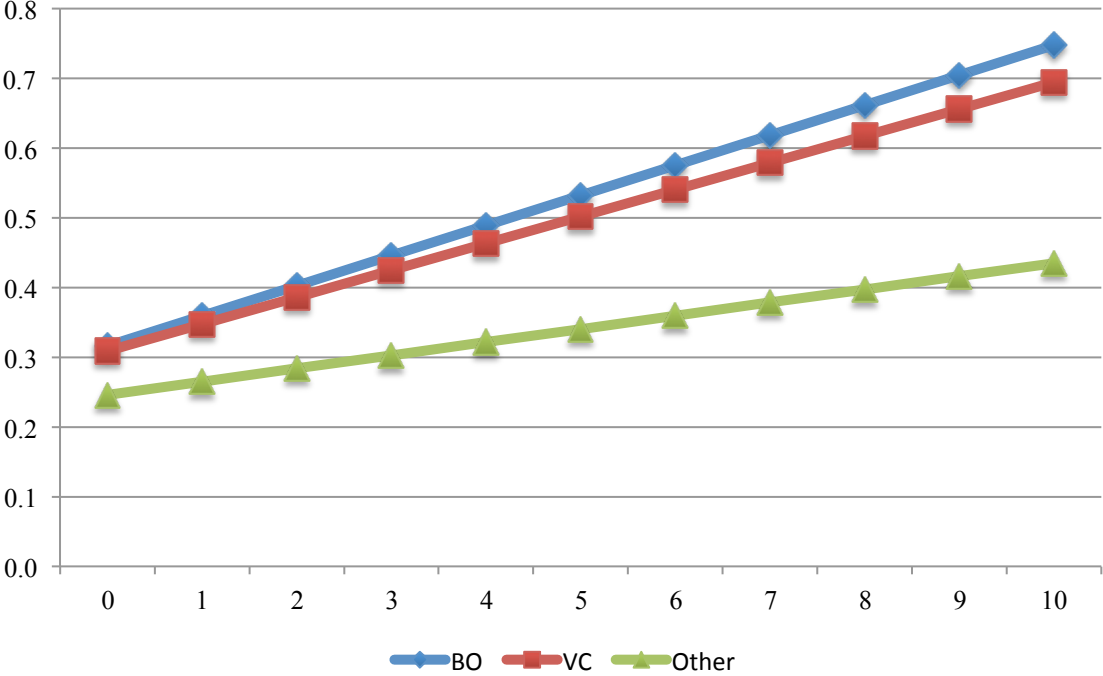
Figure 3: Estimates of Signal-to-Noise Ratio. Histograms of the estimated posterior distribution of the signal-to-noise ratio, s_α, by fund type, from the specifications reported in Table IV. The solid line is the kernel plot for specification I (without year fixed effects), and the broken line is the kernel plot for specification II (with year fixed effects).

Panel A: VC

Panel B: Buyout

Panel C: Other
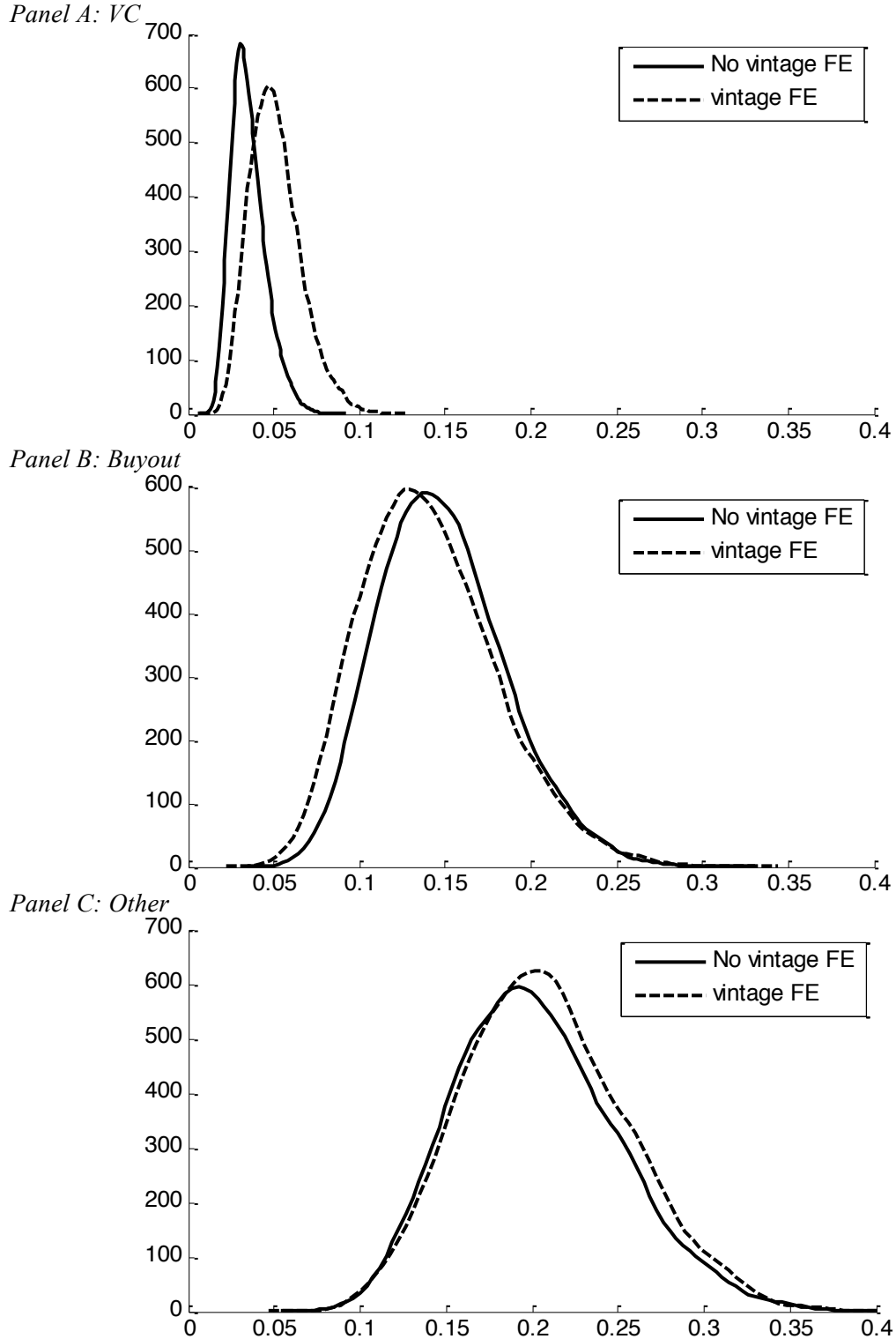
Figure 4: Speed of Learning. Graph of the posterior probability that a fund is in the top quartile of funds. Probabilities are calculated from 100,000 simulated firms, each with a different alpha that is drawn from the top 25% of the distribution. For each firm the simulation draws a sequence of 50 non-overlapping fund returns. Reported probabilities are averages of the posterior mean probability across the simulated firms after observing a given number of realized fund returns for each firm (*Fund history*, on the horizontal axis). The figure uses the parameter estimates from Table IV specification II, with vintage year fixed effects.
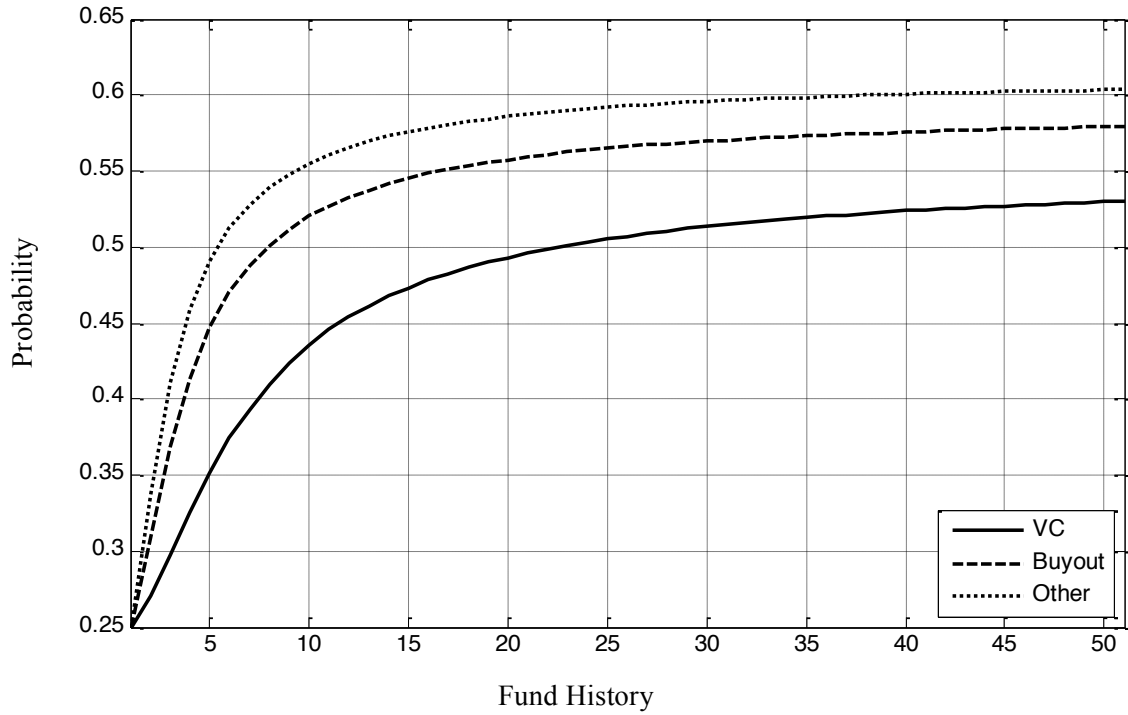
Figure 5: Investable Persistence. This figure shows the expected alpha of investing in funds raised by PE firms with top-quartile performance depending on the number of observed fund returns for the firm (*Fund history*). The expected alpha is calculated from 100,000 simulations of fund histories, using the parameter estimates from Table IV specification II, with year fixed effects.
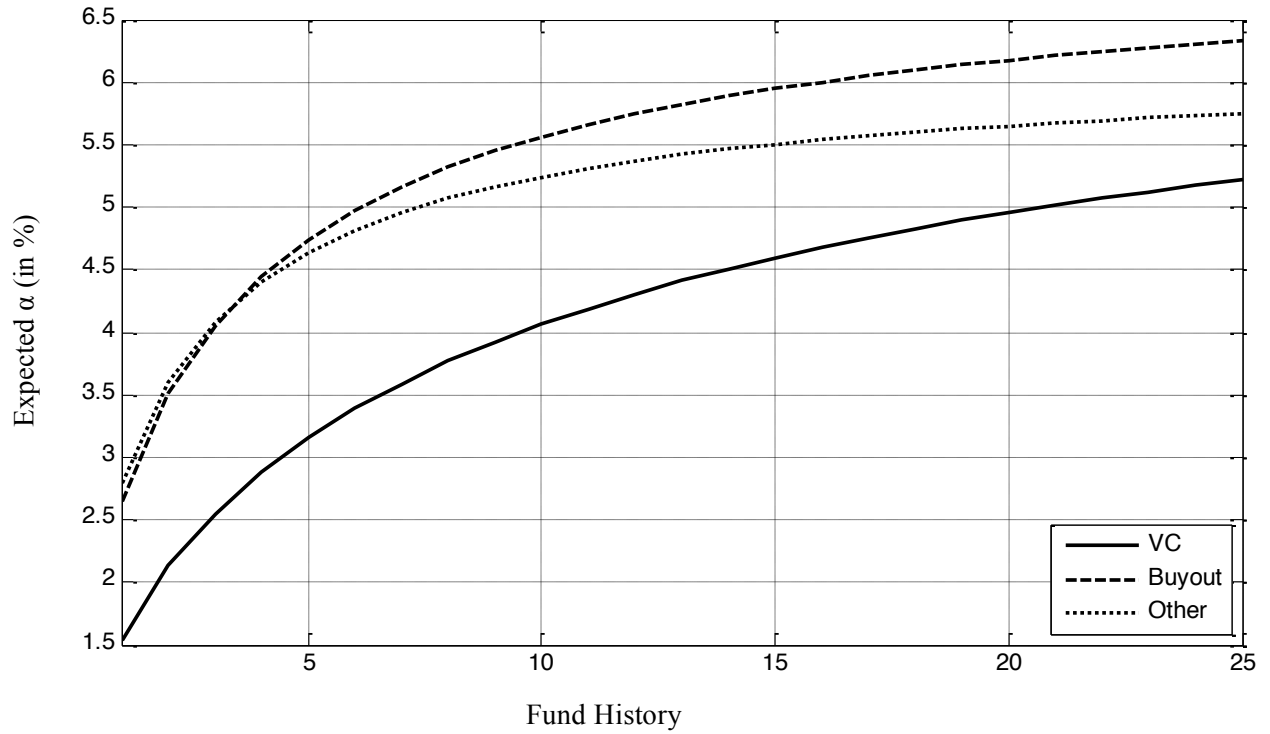
Figure 6: Marginal Log-likelihood. Plots of the marginal log-likelihood as a function of the number of Normal mixtures in the error term distribution, by fund type. The solid line represents specification I of Table IV (which has no vintage year fixed effects), and the striped line represents specification II (with vintage year fixed effects).

Panel A: VC
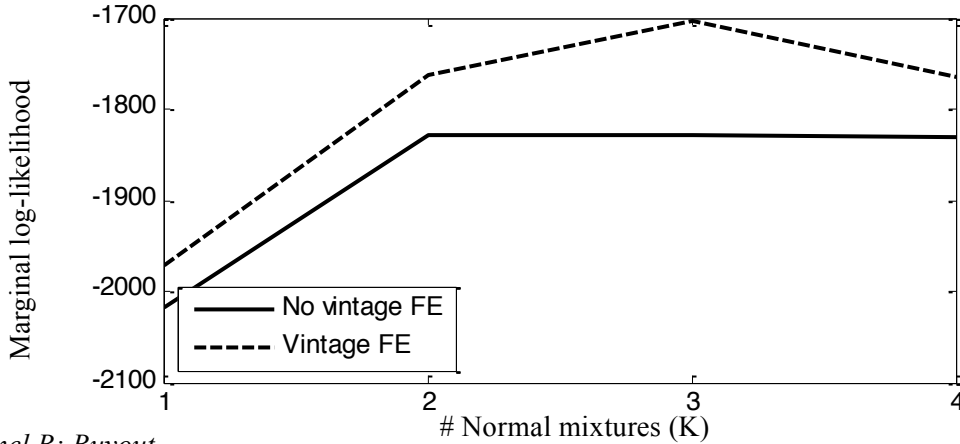
Panel B: Buyout
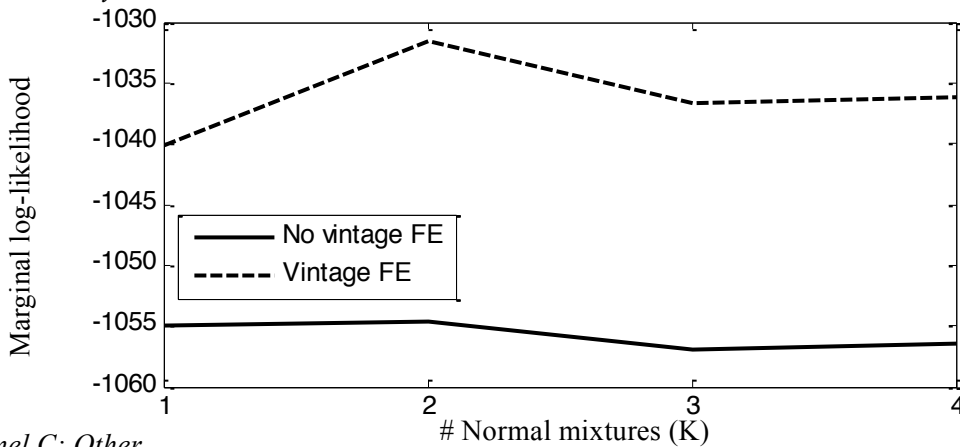
Panel C: Other
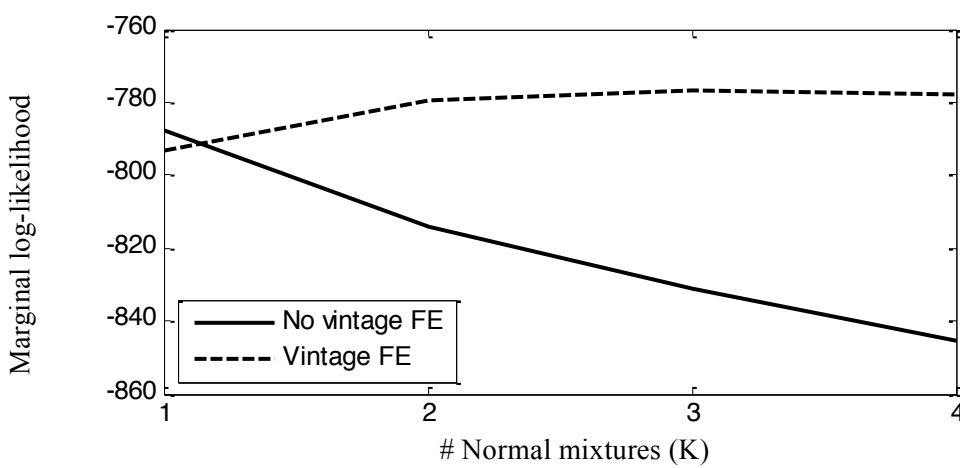
Table I: Summary Statistics. Descriptive statistics of the sample of private equity funds, by fund type (VC, Buyout, or Other). The sample contains 1,924 fully-liquidated funds raised between 1969 and 2001, with at least $5 million in committed capital (in 1990 dollars) and with non-missing returns data. The funds are raised by 831 individual PE firms (some firms manage funds of more than one fund type). Fund size is the committed capital in millions of dollars. *IRR* is the fund's internal rate of return, net of fees (where 0.1 represents a 10% return). The *Ten-year log return* is computed as $10 \cdot \ln(1 + IRR)$. *Overlap* is the number of years of overlap for funds of the same firm and type that overlap. Source: Preqin.

Panel A: Broad fund categories

|	VC	Buyout	Other
# Funds	842	562	518
# Firms	409	285	197
# Funds / firm			
Mean	2.059	1.972	2.629
Median	1	1	2
Std. dev.	1.941	1.563	2.575
10th percentile	1	1	1
90th percentile	4	4	5
Fund size ($m)			
Mean	206.900	694.073	373.269
Median	110.000	300.000	206.750
Std. dev	276.130	1,035.595	517.109
10th percentile	27.008	52.576	32.965
90th percentile	500.000	1,823.556	863.000
IRR			
Mean	0.177	0.169	0.139
Median	0.086	0.149	0.119
Std. dev.	0.548	0.186	0.129
10th percentile	-0.104	-0.017	0.004
90th percentile	0.460	0.379	0.289
Ten-year log return			
Mean	1.173	1.438	1.245
Median	0.825	1.385	1.124
Std. dev.	2.623	1.552	1.075
10th percentile	-1.101	-0.170	0.040
90th percentile	3.786	3.216	2.542
Overlap (years)			
# fund pairs	891	512	968
Mean	5.819	5.795	6.769
Median	6	6	7
Std. dev.	2.454	2.305	2.338
10th percentile	2	2	3
90th percentile	9	9	9

Panel B: Fund sub-categories

		# funds	# firms	Fund size		IRR		
				mean	median	mean	median	std. dev.
VC								
	Early-stage	177	105	186.151	92.000	0.237	0.069	0.908
	Late-stage	153	89	289.026	154.000	0.123	0.105	0.170
	Generalist	512	239	187.470	104.000	0.172	0.085	0.447
Other								
	Real estate	202	86	389.251	263.000	0.138	0.129	0.097
	Fund-of-funds	144	48	349.316	138.340	0.112	0.083	0.137
	Distressed Debt	36	14	662.069	438.000	0.146	0.140	0.108
	Natural Resources	58	25	265.372	154.750	0.170	0.152	0.157
	Secondaries	34	12	336.842	263.500	0.168	0.150	0.122
	Infrastructure, Turnaround, Special Situations, Co-Investments, Venture Debt	44	26	292.975	150.000	0.169	0.130	0.186

Table II: Fund Internal Rates of Return by Vintage Year. This table shows the number of private equity funds in the sample, and their internal rates of return, by vintage year and fund type (VC, Buyout, and Other). The sample comprises 1,924 fully-liquidated funds over the period 1969 to 2001, with at least $5 million in committed capital (in 1990 dollars) and non-missing returns data. Weighted average IRRs are based on funds for which size data is available. Source: Preqin.

	VC				Buyout				Other			
		IRR (in %)				IRR (in %)				IRR (in %)		
Vintage year	Funds	Avg.	Median	Weighted Avg.	Funds	Avg.	Median	Weighted Avg.	Funds	Avg.	Median	Weighted Avg.
1969	1	8.7	8.7	8.7	0	-	-	-	0	-	-	-
1970	0	-	-	-	0	-	-	-	0	-	-	-
1971	0	-	-	-	0	-	-	-	0	-	-	-
1972	1	21.5	21.5	21.5	0	-	-	-	0	-	-	-
1973	0	-	-	-	0	-	-	-	0	-	-	-
1974	0	-	-	-	0	-	-	-	0	-	-	-
1975	0	-	-	-	0	-	-	-	0	-	-	-
1976	0	-	-	-	0	-	-	-	0	-	-	-
1977	0	-	-	-	1	35.5	35.5	35.5	0	-	-	-
1978	2	48.6	48.6	51.0	0	-	-	-	0	-	-	-
1979	1	18.5	18.5	18.5	1	19.4	19.4	19.4	1	18.0	18.0	18.0
1980	6	16.5	14.0	25.9	3	23.7	25.8	25.7	1	12.0	12.0	12.0
1981	7	18.9	11.3	19.1	0	-	-	-	2	11.1	11.1	-
1982	11	13.8	9.3	19.3	1	39.2	39.2	39.2	1	10.0	10.0	10.0
1983	12	9.8	9.3	9.4	2	21.9	21.9	21.4	3	14.7	5.9	36.5
1984	20	12.3	12.0	12.0	7	30.4	18.4	30.4	1	7.1	7.1	7.1
1985	21	13.1	13.0	14.9	4	13.4	10.7	17.9	4	4.4	2.7	15.3
1986	20	9.5	8.5	9.3	12	27.6	18.8	25.6	3	4.0	4.0	8.5
1987	22	13.4	14.8	12.6	9	16.8	18.9	9.7	8	8.9	8.3	10.7
1988	26	20.8	21.4	27.0	12	19.2	14.2	14.6	6	10.9	11.6	9.0
1989	35	22.0	16.4	30.1	12	29.8	27.5	28.6	8	13.6	11.1	14.7
1990	24	15.1	16.5	17.6	23	19.2	15.4	15.5	10	16.8	16.0	27.6

1991	18	23.1	21.1	27.2	10	28.0	25.2	28.4	10	15.5	12.2	18.3
1992	30	21.8	16.0	24.3	21	18.3	21.2	31.5	15	18.0	16.3	21.1
1993	36	32.6	29.5	35.9	23	19.2	16.9	19.7	16	23.2	19.8	24.6
1994	30	30.0	25.5	38.2	36	25.9	21.5	35.5	18	15.9	14.0	10.1
1995	41	45.2	17.5	42.5	33	17.1	17.6	15.8	30	16.8	17.0	17.6
1996	43	29.3	10.3	21.3	35	17.4	10.4	11.4	45	17.3	12.7	14.9
1997	62	42.8	20.6	37.6	54	11.6	9.0	9.3	45	11.5	8.4	10.8
1998	73	36.6	7.0	25.0	73	6.7	8.3	5.7	65	11.2	8.3	11.7
1999	95	-2.5	-3.5	-3.4	57	8.6	8.5	7.1	58	10.1	9.3	8.1
2000	120	0.1	-0.9	-0.2	90	16.8	17.5	16.3	80	12.1	12.4	13.4
2001	85	2.4	1.0	2.9	43	27.1	28.0	28.0	88	17.0	15.2	22.3
Average	842	20.2	14.9	21.1	562	21.4	19.6	21.4	518	13.0	11.4	16.4

Table III: IRR Regressions. The table shows AR(1) regressions using fund IRRs, by fund type. In panel A, $IRR_{i,N-1}$ is the IRR of the most recent fund of the same type (VC, Buyout, Other) and the same firm, and $IRR_{i,N-2}$ is the IRR of the second previous fund. In Panel B, $IRR_{i,N-1}$ is the net IRR on the most recent, non-overlapping fund of the same type and the same firm. All regressions include vintage year fixed effects. Standard errors are clustered by firm, and are shown in brackets. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A: IRR Regressions

	VC		Buyout		Other	
	I	II	I	II	I	II
$IRR_{i,N-1}$	0.125	0.263	0.315	0.304	0.275	0.149
	(0.075)*	(0.065)***	(0.069)***	(0.075)***	(0.096)***	(0.107)
$IRR_{i,N-2}$		-0.059		0.156		0.108
		(0.066)		(0.103)		(0.113)
Log(Fund size)	1.492	2.988	-0.189	-0.237	0.439	0.128
	(1.775)	(2.862)	(0.786)	(1.014)	(0.694)	(0.955)
Log(Fund sequence)	7.510	1.741	1.872	3.020	-1.476	5.945
	(3.616)**	(6.678)	(3.188)	(3.966)	(1.773)	(2.947)**
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
N	411	232	255	133	243	139
R^2	0.250	0.282	0.366	0.488	0.200	0.195

Panel B: Non-overlapping funds

	VC	Buyout	Other
IRR$_{i,N-1}$	0.044	0.190	-0.488
	(0.058)	(0.194)	(0.219)*
Log(Fund size)	3.603	0.316	0.545
	(3.265)	(4.078)	(2.886)
Log (Fund sequence)	-1.684	-3.694	-9.273
	(11.747)	(11.228)	(9.713)
Year fixed effects	Yes	Yes	Yes
N	146	59	31
R^2	0.211	0.182	0.518

Table IV: Parameter Estimates. This table reports parameter estimates of the model described in the text. The error term ε_{iu} is a mixture of K normal distributions, chosen as the best fit according to the marginal log-likelihood. The vector X includes either a single intercept (specification I) or vintage year fixed effects, grouping the pre-1985 vintages into one bucket (specification II). The model is estimated separately for each fund type (Venture Capital, Buyout, Other), by Markov chain Monte Carlo (MCMC) using 10,000 burn-in cycles followed by 100,000 samples, saving every 10th draw. Panel A shows the parameter estimates, and panel B shows the variance decomposition estimates for the same parameters. Posterior standard deviations (Bayesian standard errors) are in brackets. Panel C shows the alpha spreads across percentiles of the posterior distribution.

	VC		Buyout		Other	
	I	II	I	II	I	II
Panel A: Parameter estimates						
σ_α	0.049	0.055	0.060	0.056	0.049	0.049
	(0.007)	(0.008)	(0.008)	(0.008)	(0.006)	(0.006)
σ_η	0.258	0.193	0.142	0.203	0.202	0.135
	(0.031)	(0.037)	(0.039)	(0.043)	(0.028)	(0.028)
σ_ε	2.449	2.326	1.359	1.225	0.807	0.865
	(0.123)	(0.121)	(0.058)	(0.073)	(0.039)	(0.050)
Panel B: Variance decomposition:						
$100 \cdot \sigma_\alpha^2$	0.243	0.309	0.361	0.316	0.244	0.246
	(0.067)	(0.087)	(0.094)	(0.089)	(0.065)	(0.061)
$10 \cdot \sigma_\eta^2$	0.675	0.386	0.216	0.432	0.416	0.189
	(0.158)	(0.141)	(0.113)	(0.160)	(0.111)	(0.076)
σ_ε^2	6.015	5.426	1.852	1.505	0.654	0.751
	(0.604)	(0.567)	(0.159)	(0.180)	(0.064)	(0.087)
σ_y^2	6.933	6.120	2.428	2.253	1.314	1.186
	(0.596)	(0.561)	(0.152)	(0.168)	(0.084)	(0.090)
Panel C: Spread in α:						
$q_\alpha(75\%)-q_\alpha(25\%)$	6.50%	7.31%	7.93%	7.40%	6.52%	6.57%
$q_\alpha(87.5\%)-q_\alpha(12.5\%)$	11.17%	12.58%	13.63%	12.73%	11.21%	11.29%
Vintage FE	N	Y	N	Y	N	Y
K	3	3	2	2	1	3
N	842	842	562	562	518	518

Table V Speed of Learning. This table shows the expected alpha of investing in the top quartile of a universe of 100 PE firms, after observing a history of non-overlapping funds of a given length (*Fund history*). The expected alphas are calculated from 100,000 simulations of 100 firms, and reported after observing a sequence of 1 through 25 funds for each firm. The three left-most columns use the parameter estimates from Table IV specification I (which has no vintage year fixed effects), and the three right-most columns show specification II (with vintage year fixed effects). For each model, the table shows the expected true α of investing in the top quartile of firms, the probability of achieving a true alpha higher than zero, the probability of attaining a higher true alpha compared to a strategy of random allocations to funds, and expected true α from a strategy of going long the top quartile and short the bottom quartile of funds.

Fund history	Without vintage year FE				With vintage year FE			
	E[α] top quartile	Prob(α>0)	Prob(α > random)	E[α] top -bottom quartile	E[α] top quartile	Prob(α>0)	Prob(α > random)	E[α] top -bottom quartile
Panel A: VC								
1	1.15%	88.1%	82.0%	2.30%	1.54%	92.1%	86.4%	3.08%
2	1.60%	95.1%	90.1%	3.20%	2.13%	97.6%	93.8%	4.26%
3	1.93%	97.8%	92.7%	3.86%	2.55%	99.2%	95.8%	5.10%
4	2.19%	99.0%	96.3%	4.38%	2.88%	99.7%	98.3%	5.77%
5	2.41%	99.5%	97.4%	4.83%	3.16%	99.9%	98.9%	6.31%
10	3.18%	100.0%	100.0%	6.36%	4.06%	100.0%	100.0%	8.12%
25	4.25%	100.0%	100.0%	8.50%	5.22%	100.0%	100.0%	10.43%
Panel B: Buyout								
1	2.92%	99.5%	97.5%	5.84%	2.65%	99.3%	97.1%	5.29%
2	3.85%	100.0%	99.6%	7.71%	3.50%	100.0%	99.5%	7.01%
3	4.44%	100.0%	99.8%	8.88%	4.05%	100.0%	99.8%	8.10%
4	4.86%	100.0%	100.0%	9.72%	4.44%	100.0%	100.0%	8.88%
5	5.17%	100.0%	100.0%	10.34%	4.74%	100.0%	100.0%	9.47%
10	6.04%	100.0%	100.0%	12.07%	5.56%	100.0%	100.0%	11.12%
25	6.82%	100.0%	100.0%	13.64%	6.33%	100.0%	100.0%	12.65%
Panel C: Other								
1	2.88%	99.9%	99.0%	5.76%	2.80%	99.9%	99.0%	5.59%
2	3.72%	100.0%	99.9%	7.44%	3.60%	100.0%	99.9%	7.20%
3	4.21%	100.0%	100.0%	8.43%	4.08%	100.0%	100.0%	8.16%
4	4.55%	100.0%	100.0%	9.10%	4.40%	100.0%	100.0%	8.80%
5	4.80%	100.0%	100.0%	9.59%	4.64%	100.0%	100.0%	9.27%
10	5.44%	100.0%	100.0%	10.87%	5.24%	100.0%	100.0%	10.48%
25	5.97%	100.0%	100.0%	11.94%	5.74%	100.0%	100.0%	11.49%

Table VI: Sub-samples. This table shows estimates of signal-to-noise ratio, s_α, for different subsamples. The columns labeled *mean* and *std. dev.* report the posterior mean and standard deviation of the variance decomposition, respectively. *IQR* is the interquartile range of the posterior distribution of s_α, i.e., the spread in between the 75th and 25th percentile of the posterior distribution. *N* is the number of funds in the subsample. The estimates are based on model specification II in Table IV (which includes vintage year fixed effects). The top row corresponds to the signal-to-noise ratio of the full sample, as depicted in Figure 4. The number of funds in the fund size sub-samples do not add up to the number of funds in the full sample, because size is not observed for some funds.

Panel A: Signal-to-noise

	VC				Buyout				Other			
	mean	std. dev.	IQR	N	mean	std. dev.	IQR	N	mean	std. dev.	IQR	N
Full sample	0.051	0.015	0.020	842	0.141	0.040	0.053	562	0.208	0.048	0.065	518
Fund size												
Small (< median)	0.111	0.043	0.057	373	0.183	0.068	0.096	261	0.291	0.103	0.146	210
Large (>= median)	0.051	0.022	0.028	374	0.123	0.049	0.064	270	0.168	0.056	0.077	210
GP location												
US	0.051	0.020	0.027	675	0.140	0.065	0.092	416	0.079	0.042	0.050	439
Europe (incl. UK)	0.180	0.079	0.110	93	0.269	0.081	0.112	113	0.310	0.116	0.159	59
ROW	0.057	0.036	0.042	74	0.083	0.058	0.061	33	0.584	0.187	0.292	20
Style												
Early-stage	0.023	0.011	0.013	177	-	-	-	-	-	-	-	-
Late-stage	0.141	0.063	0.080	153	-	-	-	-	-	-	-	-
Generalist	0.067	0.029	0.040	512	-	-	-	-	-	-	-	-
Real Estate	-	-	-	-	-	-	-	-	0.222	0.061	0.084	202
Fund-of-Funds	-	-	-	-	-	-	-	-	0.173	0.061	0.080	144
Sample period												
Early (< 1997)	0.108	0.035	0.047	407	0.211	0.091	0.129	245	0.271	0.088	0.125	227
Late (>= 1997)	0.040	0.016	0.021	435	0.178	0.058	0.080	317	0.243	0.066	0.091	291

Panel B: Long-term Persistence

	VC				Buyout				Other			
	$100\sigma_\alpha^2$	$10\sigma_\eta^2$	σ_y^2	$q_\alpha(75\%)-$ $q_\alpha(25\%)$	$100\sigma_\alpha^2$	$10\sigma_\eta^2$	σ_y^2	$q_\alpha(75\%)-$ $q_\alpha(25\%)$	$100\sigma_\alpha^2$	$10\sigma_\eta^2$	σ_y^2	$q_\alpha(75\%)-$ $q_\alpha(25\%)$
Full sample	0.309	0.386	6.120	7.31%	0.316	0.432	2.253	7.40%	0.246	0.189	1.186	6.57%
Fund size												
Small (< median)	0.823	0.522	7.537	11.62%	0.489	0.207	2.667	8.96%	0.326	0.245	1.110	7.31%
Large (>= median)	0.239	0.328	4.808	6.26%	0.234	0.452	1.107	6.21%	0.203	0.086	1.217	5.85%
GP location												
US	0.300	0.339	5.897	7.06%	0.286	0.438	2.046	6.63%	0.089	0.319	1.130	3.69%
Europe (incl. UK)	0.491	0.151	2.743	8.79%	0.427	0.146	1.578	8.47%	0.327	0.121	1.031	7.18%
ROW	0.562	0.258	10.465	9.15%	0.444	0.152	5.643	7.98%	1.434	0.180	2.306	14.42%
Style												
Early-stage	0.217	0.106	9.866	5.93%	-	-	-	-	-	-	-	-
Late-stage	0.272	0.272	1.949	6.63%	-	-	-	-	-	-	-	-
Generalist	0.377	0.386	5.674	7.82%	-	-	-	-	-	-	-	-
Real Estate	-	-	-	-	-	-	-	-	0.149	0.079	0.674	5.07%
Fund-of-Funds	-	-	-	-	-	-	-	-	0.182	0.074	1.056	5.51%
Sample period												
Early (< 1997)	0.529	0.429	4.947	9.49%	0.611	0.372	2.879	9.83%	0.338	0.123	1.239	7.50%
Late (>= 1997)	0.288	0.117	7.383	6.90%	0.303	0.156	1.709	7.16%	0.279	0.097	1.143	6.92%

Table VII: Model Specification Tests. This table shows tests of the model specification. Column I reproduces specification I of Table IV. Column II drops the transient firm effect, η, from the model, and Column III drops the long-run firm-specific effect, α. Columns IV to VI show the same for specification II of Table IV, which includes vintage year fixed effects. The *Bayes factor* represents the ratio of marginal likelihoods, indicating the weight of evidence of each model relative to the full model specification in column IV, where a Bayes Factor of one indicates that the two models have equal support in the data. For each model the number of distributions in the error term (K) is chosen to find the best model fit by marginal log-likelihood. Posterior standard deviations (Bayesian standard errors) are in brackets.

Panel A: VC

	I	II	III	IV	V	VI
σ_α	0.049	0.042		0.055	0.040	
	(0.007)	(0.004)		(0.008)	(0.004)	
σ_η	0.258		0.313	0.193		0.268
	(0.031)		(0.023)	(0.037)		(0.025)
σ_ε	2.449	2.604	2.431	2.326	2.465	2.316
	(0.123)	(0.107)	(0.114)	(0.121)	(0.115)	(0.123)
σ_y^2	6.933	6.971	6.906	6.120	6.255	6.106
	(0.596)	(0.562)	(0.553)	(0.561)	(0.572)	(0.567)
Vintage FE	N	N	N	Y	Y	Y
K	3	3	2	3	3	3
N	842	842	842	842	842	842
Marginal log-L	-1,829.5	-1,755.4	-1,823.6	-1,703.8	-1,669.2	-1,727.4
Bayes factor	0.000	0.000	0.000	N/A	1.0E+15	0.000

Panel B: Buyout

	I	II	III	IV	V	VI
σ_α	0.060	0.043		0.056	0.044	
	(0.008)	(0.005)		(0.008)	(0.005)	
σ_η	0.142		0.277	0.203		0.279
	(0.039)		(0.034)	(0.043)		(0.020)
σ_ε	1.359	1.542	1.261	1.225	1.469	1.196
	(0.058)	(0.051)	(0.065)	(0.073)	(0.054)	(0.071)
σ_y^2	2.428	2.567	2.373	2.253	2.359	2.217
	(0.152)	(0.163)	(0.121)	(0.168)	(0.163)	(0.167)
Vintage FE	N	N	N	Y	Y	Y
K	2	2	1	2	2	2
N	562	562	562	562	562	562
Marginal log-L	-1,054.6	-1,039.7	-1,054.6	-1,031.6	-1,019.7	-1,035.0
Bayes factor	0.000	0.000	0.000	N/A	1.4E+05	0.033

Panel C: Other

	I	II	III	IV	V	VI
σ_α	0.049	0.039		0.049	0.037	
	(0.006)	(0.004)		(0.006)	(0.004)	
σ_η	0.202		0.245	0.135		0.199
	(0.028)		(0.022)	(0.028)		(0.017)
σ_ε	0.807	1.059	0.793	0.865	1.038	0.848
	(0.039)	(0.033)	(0.041)	(0.050)	(0.037)	(0.050)
σ_y^2	1.314	1.272	1.234	1.186	1.217	1.119
	(0.084)	(0.075)	(0.079)	(0.090)	(0.081)	(0.087)
Vintage FE	N	N	N	Y	Y	Y
K	1	2	1	3	3	3
N	518	518	518	518	518	518
Marginal log-L	-787.8	-777.6	-781.7	-776.9	-771.3	-774.5
Bayes factor	0.000	0.491	0.008	N/A	256.185	10.677