Risk-Adjusted Returns of Private Equity Funds: A New Approach

ARTHUR KORTEWEG and STEFAN NAGEL*

ABSTRACT

This paper introduces a new approach to benchmarking the performance of individual private equity funds. The new metric, α, is substantially less sensitive to noise in individual fund cash flows compared to the popular public market equivalent (PME) measure and its generalized version (GPME), while having the same aggregate pricing implications for private equity as GPME. We illustrate in simulated data that α is an overall more accurate measure of risk-adjusted fund returns than both PME and GPME. For a large data set of private equity fund cash flows, the standard deviation of α across venture capital funds is 20% lower than PME and GPME. For buyout funds, PME and α are quite close, but the metrics deviate in subsamples where beta is different from one. Our more accurate performance metric increases the statistical power in regressions of performance on fund size, and improves performance predictability of managers' future funds.

*Arthur Korteweg is from the University of Southern California Marshall School of Business, and Stefan Nagel is from the University of Chicago, NBER, and CEPR. We thank Burgiss for providing the data, the Private Equity Research Consortium and the Institute for Private Capital for providing nonfinancial research support, and Wendy Hu for excellent research assistance. We thank Elise Gourier (discussant), Chris Schwarz (discussant), Juhani Linnainmaa, Luke Taylor and seminar participants at the 2021 SFS Cavalcade, the 2016 Southern California Law and Finance Conference, Boston College, Carnegie-Mellon, Dartmouth, UC San Diego, and the University of Melbourne for comments.

Allocations to private equity (PE) by U.S. pension funds and endowments have been rising in recent years.[1] For example, the 2020 NACUBO-TIAA Study of Endowments reports that university endowments invested 23% of total assets in venture capital and buyout funds, and another 6% in private equity real estate. The top 200 largest U.S. defined benefit public (corporate) pension funds in 2016 allocated an average of 9.0% (5.8%) of their assets to venture capital and leveraged buyout, and 8.3% (5%) to real estate.[2] The increased prominence of PE brings the problem of performance assessment to the forefront. Despite recent progress in assessing risk-adjusted returns at the asset class level, little is known about the properties of performance metrics when applied to individual PE funds. Assessing an individual fund's performance is, however, the typical problem in practice, as investors have to decide how to allocate money to managers. It is also an important issue for academic research on topics such as the degree of performance persistence and the relation between fund performance and characteristics of funds or managers.[3] This paper shows that the popular public market equivalent (PME) metric, and its generalized version, are very noisy measures of the risk-adjusted return to individual PE funds, and proposes a new measure that yields more accurate estimates.

Performance evaluation in PE is complicated because regularly observed returns are not available. A PE fund is organized as a limited partnership that formally lasts ten years. Lim-

[1] For the purposes of this paper, the term "private equity" encompasses all types of private equity, including but not limited to investments in leveraged buyout, venture capital, real estate, and natural resources.

[2] Numbers are from the Pensions and Investments annual survey of pension funds:
`http://www.pionline.com/article/20170206/PRINT/302069976/assets-of-top-funds-up-62-to-94-trillion`

[3] Papers that study performance persistence in private equity include Kaplan and Schoar (2005), Phalippou and Gottschalg (2009), Phalippou (2010), Chung (2012), Harris, Jenkinson, Kaplan, and Stucke (2014), Hochberg, Ljungqvist, and Vissing-Jorgensen (2014), Robinson and Sensoy (2016), and Braun, Jenkinson, and Stoff (2017), Korteweg and Sorensen (2017), Nanda, Samila, and Sorenson (2020). Examples of studies that relate PE fund performance to fund or manager characteristics include papers on returns to scale (e.g., Kaplan and Schoar, 2005, Metrick and Yasuda, 2010, Harris, Jenkinson, and Kaplan, 2014, Lopez-de-Silanes, Phalippou, and Gottschalg, 2015, Rossi, 2019), pay for performance (e.g., Chung, Sensoy, Stern, and Weisbach, 2012, Hüther, Robinson, Sievers, and Hartmann-Wendels, 2020), the relation between VC networks and performance (e.g., Hochberg, Ljungqvist, and Lu, 2007, Abell and Nisar, 2007, Liu and Chen, 2014), the relation between performance and LP identities (e.g., Lerner, Schoar, and Wongsunwai, 2007, Hochberg and Rauh, 2013, Sensoy, Wang, and Weisbach, 2014, Andonov, Hochberg, and Rauh, 2018, Cavagnaro, Sensoy, Wang, and Weisbach, 2019), the potential manipulation of interm performance during fundraising (e.g., Jenkinson, Sousa, and Stucke, 2013, Barber and Yasuda, 2017, Chakraborty and Ewens, 2018, Brown, Gredil, and Kaplan, 2019, Jenkinson, Landsman, Rountree, and Soonawalla, 2020, Hüther, 2022), and the relation between performance and idiosyncratic volatility (e.g., Ewens, Jones, and Rhodes-Kropf, 2013, Peters, 2018, Opp, 2019), amongst others.

ited partners (LPs) commit capital to the fund, which is called by the general partner (GP) when suitable investment opportunities are identified, and to pay management fees. When portfolio companies are sold, distributions are paid to the limited partners, after performance fees. The standard approach in the mutual and hedge fund industries that uses factor models to benchmark returns and estimate fund-level alphas, cannot be readily applied to a such a sequence of cash flows occurring at random times.

The generalized public market equivalent (GPME) approach of Korteweg and Nagel (2016) values the PE cash flow stream from an LP's perspective using a stochastic discount factor (SDF) that is calibrated to exactly price a set of benchmark assets (for example, public equities and risk-free bonds). If the PE cash flows can be replicated using a (possibly levered) investment in the benchmark assets, then the GPME is zero in expectation. This is a robust approach that accommodates the skewness and irregular horizon of PE cash flows, and it is well-suited to address the question of aggregate outperformance of the PE industry. As such, the measure works well in large data sets, taking averages over a large number of funds of varying vintages. However, for an *individual* fund, even if its expected GPME is zero, the realized value of the GPME may be far from zero. Typically, GPME realizations for individual funds are far too noisy to draw meaningful inferences about individual fund performance.

The root of the problem can be understood by analogy with factor model regressions for public market funds that have regularly observed realized fund returns. With observable returns, the typical exercise is to compare the realized fund return, R_t, to a benchmark portfolio return, R_t^b, which is a combination of factor portfolio returns scaled by their corresponding factor loadings. Subtracting R_t^b from R_t not only generates an excess return that is zero in expectation under the null of zero outperformance, but it also absorbs much of the period-by-period random variation in R_t that originates from unexpected shocks to the risk factors in R_t^b. As a consequence, $R_t - R_t^b$ delivers a relatively precise estimate of abnormal performance even for an individual fund. In contrast, discounting PE cash flows with the SDF in the GPME approach does not remove such common factor shocks to the same degree, which results in a much noisier estimate of abnormal performance.

In this paper we build on the GPME approach to develop a new benchmarking method that removes common factor shocks and is suitable for individual fund performance evaluation. To achieve this, the method requires two additional assumptions over and above the assumptions underlying the GPME method. First, we assume that fund cash flows and benchmark portfolio payoffs are jointly lognormal with constant covariance matrix and constant means. This is restrictive, but so are typical factor model regression approaches for public market funds which are essentially based on a mean-variance optimization framework. Second, we assume that betas (i.e., factor loadings) are the same across funds within an asset class (e.g., VC or buyout) and constant over time. This assumption is necessary to infer betas from data on aggregate performance. It could be relaxed by conditioning on fund characteristics and state variables, although this may not be feasible in practice given the small sizes of available data sets.

We construct a benchmark portfolio that matches the systematic risk of the PE cash flows. It takes continuously rebalanced levered positions in the SDF risk factors with weights given by factor betas. We show that if the benchmark portfolio uses the true betas as weights, then deflating PE cash flows with the return on this benchmark portfolio delivers exactly the same expected abnormal payoff as the GPME approach. In addition, deflating by the benchmark portfolio removes the common risk factor shocks from the PE cash flows, just as factor model regression in public market funds returns data do. As a consequence, we obtain a much less noisy measure of individual fund abnormal performance than the GPME. We use α as the label for the abnormal performance relative to the benchmark portfolio.

In practice we don't know the true betas. To estimate betas, we exploit the fact that deflating PE cash flows with the return on this benchmark portfolio must deliver the same expected abnormal payoff as the GPME approach. As a first step, we therefore estimate the GPME for a large group of funds and vintages within an asset class. We then apply the benchmark portfolio approach and search for the betas that deliver average abnormal payoffs that are exactly equal to the GPME and that minimize, at the same time, the cross-sectional variance of abnormal performance across funds. Finally, we evaluate individual fund performance with a benchmark portfolio that uses these estimated asset-class specific betas as factor weights.

Simulations confirm that α is overall a more accurate measure of fund-level abnormal PE return than GPME or the more restrictive public market equivalent (PME) of Kaplan and Schoar (2005) that deflates PE cash flows with the market return. There are only two cases where α is not strictly preferred. First, when beta is truly equal to one, then α and the abnormal performance measured by PME coincide, and the PME may perform slightly better in practice as it avoids the estimation of the parameters needed to compute the benchmark portfolio. Of course, in practice one would not know if beta is truly equal to one. Second, when betas are equal to the SDF loading on the corresponding factors (i.e., the factors' prices of risk), GPME and α coincide. We also show that these results are robust to reasonable violations of the main assumptions of constant betas and lognormality.

We use the Burgiss Manager Universe, one of the largest samples of PE fund cash flow histories available, to take the benchmark portfolio approach to the data. Our sample includes a total of 1,219 VC funds and 879 leveraged buyout funds. The data are sourced exclusively from a large number of LPs, avoiding the natural biases introduced by sourcing data from GPs. The data are of very high quality because they are used for control purposes (audit and performance measurement) and cross-checked when multiple LPs invest in the same fund. In our baseline analysis, we use a CAPM SDF and, correspondingly, a benchmark portfolio with the market portfolio return as the only factor.

For venture capital funds, the differences between α and GPME at the individual fund level turn out to be modest. The across-fund standard deviation of α is 20% lower compared with the GPME. The differences are not big because the estimated VC beta of 2.43 is quite close to the estimate of the price of risk of the market factor in the SDF (3.46). As a consequence, discounting with the SDF or deflating with the benchmark portfolio produces relatively similar results. Put differently, with a market price of risk for the market factor in the vicinity of beta, discounting by the SDF removes common factor shocks almost as well as the deflating with the benchmark portfolio does.

In contrast, for the sample of leveraged buyout funds, we find much bigger differences between α and GPME. Here, the across-fund standard deviation of α is 66.7% lower compared with the GPME. The reason is that the difference between beta (0.79) and market factor price of risk in the SDF (3.59) is much larger. This means that discounting by the SDF yields realized pricing errors that have large common factor components greatly magnify the noise and cross-sectional dispersion in fund-level GPME realizations compared with α.

We find different estimates of betas across early and late time periods, across fund size splits of the data, and for subclasses of venture capital. Therefore, the degree to which PME and GPME deviate from each other, and from α, varies over time and across fund characteristics. Although this may raise some concern regarding violations of the assumption of constant betas in the benchmark method, the simulations show that for the range of betas estimated in the data, the alpha metric is robust and more accurate than both PME and GPME.

When we repeat the analysis using two-factor models where we augment the market factor with an additional factor from the portfolios that underlie the Fama-French size and value factors, we find results that are similar. We again find much bigger differences between α and GPME at the individual fund level for buyout funds than for VC funds. Interestingly, VC funds load most strongly on the big-growth portfolio rather than on the small-growth portfolio that is often thought to be a good approximation of firms in VC's portfolios. Buyout seems to be most closely related to big-value.

Finally, we demonstrate the usefulness of the α performance measure in two applications that study the properties of PE funds' abnormal performance. First, we examine the relation between fund performance and fund size. The additional noise in the PME and GPME performance metrics substantially reduces statistical power to reject the null hypotheses in regressions of fund performance on fund size. Coefficient standard errors are as much as one-third to two-thirds lower when using α instead of GPME as the dependent variable, depending on the regression specification and asset class. Second, we look at performance persistence. We find that α performs better than (G)PME in predicting future fund performance, especially at longer horizons (when predicting performance of future funds that do not overlap with the present fund). This is consistent with less noise contamination in α compared with (G)PME, and hence smaller errors-in-variables bias. Compared to PME, predicting future performance based on α works best for top quartile VCs. Relative to GPME, α works well across the board in buyout, and for predicting poorly performing VCs.

Our work relates to several other recent papers that propose advances in PE performance evaluation methodology. Several papers develop alternative approaches for individual fund performance evaluation. Gupta and Van Nieuwerburgh (2021) pursue an approach that is closely related to ours. Based on a multi-factor SDF and VAR dynamics for state variables that include valuation ratios and dividends of equity factors, they construct model-implied prices and cash flows of bond and equity strips (i.e., claims to single future cash flows at various horizons). Linear combinations of these strips then serve as benchmark portfolio prices and payoffs. A main difference to our approach is complexity. Their approach requires estimation of hundreds of parameters (the VAR companion matrix alone has more than 300, given there 18 state variables) which is only feasible with ad-hoc parameter restrictions. In our case, with the CAPM SDF, we have four parameters: 2 for the SDF and 2 benchmark portfolio parameters. Brown, Ghysels, and Gredil (2022) develop a state-space model to filter unobserved fund NAVs and estimate fund-level alphas and betas. Like Gupta and Van Nieuwerburgh, this approach also involves estimating many parameters and latent states over time (they filter weekly NAVs for each fund). They estimate an average buyout fund beta of 1.0 to 1.3, depending on the model specification, which is slightly higher than our estimate of 0.8. Similar to us, they estimate a positive risk-adjusted return to the average buyout fund. Their average VC fund has an estimated market beta of 1.4 to 1.8, somewhat lower than the 2.4 beta that we estimate. Like us, they estimate a negative risk-adjusted VC return in three out of their four specifications. Driessen, Lin, and Phalippou (2012) propose a heuristic approach that uses, like the PME method, a realized benchmark return to discount PE cash flows, but in their case the benchmark return is a beta-adjusted market return and includes an abnormal return parameter. Robinson and Sensoy (2013) use this approach, too, but they exogenously choose a beta instead of estimating it. In contrast, our approach is rigorously justified based on asset-pricing theory and our statistical assumptions. That said, the beta of 2.7 for VC funds that Driessen et al. estimate is very close to our estimate of 2.4.

Other recent papers propose other innovations in performance evaluation methodology, but without tackling the problem of individual fund performance evaluation. Gredil, Sorensen, and Waller (2020) use SDFs of habit-formation and long-run risks consumption-based asset pricing models. Using these SDFs, they find that venture funds have better risk-adjusted performance than with a CAPM SDF. Like the GPME method, their approach uses realized cross-products of SDFs and cash flows (albeit with an adjustment aimed at improving finite-sample performance of the estimator) and is hence not well-suited for evaluating individual fund performance. Ang, Chen, Goetzmann and Phalippou (2018) use Bayesian methods to extract from cash flow data and public markets factor returns an implied series of returns for groups of PE funds. This method is not applicable to individual funds, however. Boyer, Nadauld, Vorkink, and Weisbach (2021) construct an aggregate PE index from secondary market transactions of PE fund shares. Stafford (2022) construct a public markets replicating portfolio for PE buyout funds in aggregate.

The paper is structured as follows. Section I introduces theoretical background and our proposed method for estimating individual fund alphas. Section II discusses estimation. Section III shows simulation results and robustness. Section IV reports results from PE fund data, and Section V concludes.

I. Risk-adjusting Returns of Private Equity Funds

Our model setup is as follows. We take the perspective of the limited partners in a private equity fund. Let $\{C\} = \{C_1, C_2, .., C_K\}$ be the stream of (nonnegative) capital calls by the fund, and $\{X\} = \{X_1, X_2, .., X_J\}$ the (nonnegative) distributions to the LPs. For example, $C = \$1$ means the LP sends one dollar to the GP, and $X = 2$ means the GP distributes \$2 back to the LP. See Figure 1 for an example of a fund with $K = 3$ capital calls and $J = 4$ distributions. To keep notation simple, we suppress fund-identifier subscripts in this section. Throughout the paper we use uppercase R for arithmetic returns, and lowercase r for log returns, that is, $r \equiv \log R$.

A. Generalized Public Market Equivalent

Korteweg and Nagel (2016) (KN) introduce a stochastic discount factor (SDF) approach to evaluating private equity fund outperformance. If an SDF M perfectly prices fund payoffs, then

$$E\sum_{j=1}^{J} M_{h(j)}X_j = E\sum_{k=1}^{K} M_{h(k)}C_k. \tag{1}$$

where $h(k)$ is the time since fund inception until the date of capital call k, and $h(j)$ accordingly for distributions. The SDF M_h is the one that prices payoffs over the relevant time horizon, h. KN assume an exponentially-affine CAPM SDF,

$$M_h = \exp(\delta h - \gamma r_h^m), \tag{2}$$

where r_h^m is the log return on the stock market portfolio from fund inception to h periods later. In analogy to the CAPM in its linear factor model version, this SDF evaluates payoffs from the viewpoint of an investor who is invested in the stock market portfolio and the risk-free asset. If an investment opportunity offers abnormal payoffs based on this SDF, this means that this investor would benefit at the margin from adding this asset to the portfolio in addition to the risk-free asset and the stock market portfolio. Unlike the linear factor model version of the CAPM, the SDF in (2) can be applied to multi-period valuation problems. The multi-period SDF for a payoff at horizon h is simply the single-period SDF compounded over that horizon. Additional risk factors, such as those in Fama and French (1993), can be added easily, but to keep notation simple we focus on the single-factor case in laying out our approach.

KN estimate the SDF parameters δ and γ by requiring the SDF to perfectly price pseudo-funds that invest in T-bills and the stock market index. Then, if the private equity cash flow stream can be replicated by a (possibly levered) portfolio of the market index and the risk-free asset (or, more generally, the portfolios included in M_h), then (1) must hold with the SDF (2). Based on this, KN define the generalized public market equivalent (GPME) for a given fund as

$$GPME = \sum_{j=1}^{J} M_{h(j)} X_j - \sum_{k=1}^{K} M_{h(k)} C_k, \tag{3}$$

where the X_j and C_k are distributions and capital calls scaled by the sum of capital calls discounted with the risk-free rate. These scaled distribution cash flows, $\{X\}$, can then be interpreted as the dividends from an investment of approximately \$1 in the fund, and the expected GPME can be interpreted as the NPV of that investment. If the capital call cash flows are deterministic, this interpretation becomes exact. If there are no abnormal returns to investing in PE, then the *expected* GPME equals zero.

The GPME generalizes the well-known public market equivalent (PME) metric of Kaplan and Schoar (2005), which is a special case of equation (2) with $\delta = 0$ and $\gamma = 1$ (see Sorensen and Jagannathan, 2015, and Korteweg and Nagel, 2016). We use a slightly different definition of PME than the original Kaplan and Schoar paper, which uses the ratio of discounted distributions and capital calls. Our definition uses the difference, which avoids some Jensen inequality issues with the ratio, as explained in Korteweg and Nagel (2016). In the Kaplan-Schoar paper, a fund that can be replicated by the market has a PME of one, while according to our definition that fund's PME is zero.

B. Benchmark portfolio for individual fund returns

GPME works well in expectation, that is, when estimated across a large number of funds that cover different time periods. However, a typical problem both in academia and in practice is to assess the performance of an individual fund. In simulations below we show that the GPME, though statistically consistent, can be a very noisy measure of a single fund's outperformance. In this section we develop a benchmark portfolio comparison method that works equally well in expectation, but yields more accurate estimates of fund-level abnormal returns.

In order to achieve a similar result for PE funds, we need to impose a little more structure on the problem than the GPME framework does. This gives up some of the generality of the GPME approach, but it enhances statistical power. Specifically, we add the assumption that market factor returns and fund payoffs are IID over time and distributed jointly lognormal. This means that the joint distribution that the log market return $r^m_{h(j)}$ and log payoff x_j realized between time t and $t + h(j)$ are drawn from the same bivariate normal distribution for all t. We further assume a constant risk-free rate r_f. Finally, we assume that the market beta of log payoffs, $\beta = cov(x_j, r^m_h)/var(r^m_h)$, is horizon-independent and does not vary across funds within an asset class (henceforth we drop the j index in $h(j)$ to reduce notational clutter). Defining $\mu = \log E[R^m_1] - r_f$ and $\sigma^2 = var(r^m_1)$, these assumptions imply

$$\begin{pmatrix} r^m_h \\ x_j \end{pmatrix} \sim \mathcal{N}\left[\begin{pmatrix} \mu h + r_f h - \frac{h}{2}\sigma^2 \\ E[x_j] \end{pmatrix}, \begin{pmatrix} \sigma^2 h & \beta\sigma^2 h \\ \beta\sigma^2 h & var(x_j) \end{pmatrix}\right]. \tag{4}$$

Since h-period log market returns are the sum of single-period market returns, expected return and variance scale linearly with h. In contrast, the x_j payoff may not necessarily have a representation as a sum of $h(j)$ single-period IID random variable realizations. Therefore, $E[x_j]$ and $var(x_j)$ do not necessarily scale with $h(j)$ (but the component of x_j spanned by r^m_h scales with $h(j)$ and the definition of β implies that $var(x_j) \geq \beta^2\sigma^2 h$ holds.)

Under these assumptions, the GPME method with SDF (2) corresponds to benchmarking PE returns with a log-linear version of the CAPM,

$$\log E[R_h] - r^f_h = \beta\left(\log E\left[R^m_h\right] - r^f_h\right). \tag{5}$$

To see this, note that the SDF parameters are pinned down by the requirement that the SDF correctly prices the risk-free asset and the market portfolio of public equities. The pricing equations for these two assets can be solved for

$$\begin{aligned} \delta &= -r_f - \frac{1}{2}\gamma^2\sigma^2 + \gamma(r_f + \mu - \frac{1}{2}\sigma^2), \\ \gamma &= \frac{\mu}{\sigma^2}. \end{aligned} \tag{6}$$

Evaluating the log of the pricing equation $E[M_h R_h] = 1$ with these parameter values then yields (5).

Even though PE fund returns are not observable, we *can* benchmark fund cash flows under the statistical assumptions we made above. Consider the valuation of a single cash flow, X_j, in (3). Using the SDF (2) with the parameter values in (6), and taking expectations, we obtain

$$E[M_h X_j] = E\left[\exp\left\{-r_h^f - \beta(r_h^m - r_h^f) + \frac{h}{2}\beta(\beta-1)\sigma^2\right\} X_j\right]. \tag{7}$$

(PROOF: See appendix.)

Since β is horizon-independent, equation (7) applies to the valuation of every other cash flow in the summations in (3), using the same β. Therefore, we get

$$E[GPME] = \left(E\sum_{j=1}^{J}\frac{X_j}{R^b_{h(j)}}\right) - \left(E\sum_{k=1}^{K}\frac{C_k}{R^b_{h(k)}}\right) \tag{8}$$

with the benchmark portfolio return

$$R_h^b = \exp\left\{r_h^f + \beta(r_h^m - r_h^f) - \frac{h}{2}\beta(\beta-1)\sigma^2\right\} \tag{9}$$

Thus, if we deflate each cash flow X_j by the realized gross return that the benchmark portfolio would have earned over the horizon $h(j)$ from initial investment into the fund until realization of X_j, we get the same abnormal payoff in expectation, $E[GPME]$, as by discounting with the SDF in (3).

The return R_h^b can be interpreted as the return on a levered portfolio that matches the systematic risk of the fund cash flows. More precisely, if time is continuous, R_h^b is the h-period return on a continuously rebalanced portfolio that, at every instant, invests a share β in the market portfolio and $1-\beta$ in the risk-free asset (alternatively, one can view R_h^b as the exponential of a second-order Taylor approximation of the log return of a portfolio that invests a share β in the market portfolio and $1-\beta$ at the beginning of the return measurement interval; see Campbell and Viceira (2002), Chapter 2).

While deflating cash flows with R_h^b delivers the same expected abnormal payoff as discounting with the SDF does, it is important to note that $1/R_h^b$ is not an SDF, as it depends on an asset-specific parameter (β).[4] Moreover, while expected values are the same, discounting with the SDF and deflating with the benchmark portfolio produces a different realized abnormal payoffs for an individual fund. Define the fund-level abnormal performance metric

$$\alpha = \sum_{j=1}^{J} \frac{X_j}{R^b_{h(j)}} - \sum_{k=1}^{K} \frac{C_k}{R^b_{h(k)}}. \tag{10}$$

Compared to (3), α is a more accurate measure of abnormal fund-level performance, because the benchmark portfolio eliminates variation in payoffs due to the factors in the benchmark portfolio. To see this, consider again a single cash flow X_j. In logs, the cash flow deflated by the benchmark portfolio return is $x_j - r^b_h$ while the realized product of the cash flow with the SDF is $x_j + m_h$. Now note that

$$x_j - E[x_j] = \beta(r^m_h - E[r^m_h]) + \epsilon_j \tag{11}$$

$$r^b_h - E[r^b_h] = \beta(r^m_h - E[r^m_h]) \tag{12}$$

$$m_h - E[m_h] = -\gamma(r^m_h - E[r^m_h]) \tag{13}$$

where ϵ_j is an idiosyncratic shock. Therefore, after deflating by R^b_h, the unexpected log payoff is

$$x_j - E[x_j] - (r^b_h - E[r^b_h]) = \epsilon_j. \tag{14}$$

By virtue of β being a regression coefficient in a regression of x_j on r^m_h, deflation with R^b completely removes the component of x_j related to r^m_h and only the idiosyncratic part ϵ_j remains. In contrast, the unexpected log product with the SDF is

$$x_j - E[x_j] + (m_h - E[m_h]) = (\beta - \gamma)(r^m_h - E[r^m_h]) + \epsilon_j. \tag{15}$$

Unless it happens to be the case that $\beta = \gamma$, the SDF based measure in (3) is therefore subject to an additional source of noise coming from the $r^m_h - E[r^m_h]$ component. Especially when

[4] In contrast, the reciprocal of the SDF *is* a proper portfolio return: the return on the growth-optimal portfolio, as shown in Long (1990). This is the return on a continuously rebalanced portfolio that, at every instant, invests a share γ in the market portfolio and $1-\gamma$ in the risk-free asset. In the special case of log-utility investors, $\gamma = 1$, the growth-optimal portfolio is the market portfolio, and we obtain the PME metric.

$|\beta - \gamma|$ is big or when the realized average market portfolio return during a fund's lifetime was far from its expected value, the SDF-based GPME can deviate far from $E[GPME]$.

Thus, while the SDF approach provides a valid measure of abnormal performance in the sense that it is correct on average in a large-T sample, the benchmark portfolio approach, while also correct on average, should minimize the noise in performance assessment and therefore provide a more accurate measure of abnormal performance suitable for benchmarking individual funds.

The above analysis ignores estimation of β and the SDF parameters, δ and γ. Estimation is not a trivial task because PE fund returns are unobserved in practice makes, which materially complicates the estimation of β compared to, say, mutual funds applications. To make estimation feasible, we impose the assumption that β is identical within a group of PE funds. As we explain in the next section, this allows us to use information from the cross-section of fund pricing errors to pin down β. The performance of the GPME and alpha measures, both in absolute and relative terms, will depend on the degree of estimation error in these parameters. The performance of alpha will also depend on how reasonable the additional assumptions of homogeneous betas and lognormality are in the data, and the robustness of the approach to violations of these assumptions. We analyze these issues in the simulation section below. The next section describes estimation.

II. Estimation

As a first step, we estimate the parameters δ and γ in the SDF (2) as discussed in KN, by fitting the SDF to artificial PE funds that invest only in Treasury Bills and the stock market portfolio. We then apply this SDF to calculate the expected GPME (a single number for the whole sample of funds), again exactly as in KN. We label this asset-class level abnormal performance as $\bar{\alpha}$.

The next step is to estimate β. To incorporate a degree of robustness to the lognormality assumption, we use σ_h^2, the empirical variance of factor returns over horizon h, instead of $h\sigma^2$

in the Jensen inequality adjustment.

$$R_h^b = \exp\left\{ r_h^f + \beta(r_h^m - r_h^f) - \frac{1}{2}\beta(\beta-1)\sigma_h^2 \right\}. \tag{16}$$

We then estimate β by exploiting the fact that deflating with R_h^b should produce the same pricing error $\bar{\alpha}$ that we get from discounting with the SDF. One might therefore think that to estimate β, one could remove $\bar{\alpha}$ from the right and left-hand side of (10), averaged over all a funds within an asset class, and then simply solve this equation for β. Intuitively, this approach looks for the degree of leverage that the benchmark portfolio needs to have to produce a pricing error equal to $E[GPME]$. However, this equation has two solutions. Note that $1/R_h^b$ and M_h both have the functional form $\exp(ah - br_h^m)$ with $b = \beta$ in the case of R_h^b and $b = \gamma$ for the SDF. Since both deflating with R_h^b and discounting with M_h produces the same pricing error $\bar{\alpha}$, both yield an expected value of (10) of zero after $\bar{\alpha}$ has been removed from both sides of this equation. In other words, one of the solutions just reproduces the SDF that was used to obtain $\bar{\alpha}$.

To find the solution with $b = \beta$ rather than $b = \gamma$, we exploit the fact that the solution with $b = \beta$ should produce smaller cross-sectional dispersion in abnormal performance across funds than the solution with $b = \gamma$. The reason is that abnormal performance based on the SDF method with $b = \gamma$ includes the additional component $(\beta - \gamma)(r_h^m - E[r_h^m])$ in (15). Funds that do not have completely identical lifetimes and cash-flow timing will have different realizations for this component in addition to the differences that arise from the idiosyncratic component ϵ_j. In contrast, the benchmark portfolio method abnormal performance measure in (14) has only the ϵ_j term, and hence smaller cross-sectional dispersion.

Using a data set of $i = 1 \ldots N$ funds, we therefore estimate

$$\widehat{\beta} = \arg\min_{\beta} Q(\beta) \;\; \text{with} \;\; Q(\beta) \equiv \frac{1}{N}\sum_{i=1}^{N}(\alpha_i - \bar{\alpha})^2$$
$$\text{subject to } \frac{1}{N}\sum_{i=1}^{N}(\alpha_i - \bar{\alpha}) = 0 \tag{17}$$

where α_i is a function of β and $\bar{\alpha}$ is given from the SDF-based $E[GPME]$ estimation. Thus, among the two solutions that satisfy the pricing constraint, we look for the one that minimizes the cross-sectional dispersion in abnormal performance.

III. Simulations

In this section we illustrate the performance of the algorithm using simulations. We simulate data sets of N funds spanning T years, with an equal number of funds starting each vintage year. The yearly log market return is drawn from a normal distribution with $\mu = 11\%$, and $\sigma = 15\%$. The log risk-free rate is 2% per year. Given these parameters, an SDF with $\delta = 0.15$ and $\gamma = 4$ exactly prices the risk-free rate and public stocks. These SDF parameters are close to the empirical point estimates in the next section. Funds call \$1 in capital in their vintage year, and they distribute $J = 25$ cash flows at random times, drawn from a uniform distribution over funds' ten year lifetimes. Cash flows are generated as

$$X_j = \frac{1}{J} \exp\left(r^f_{h(j)} + \beta(r^m_{h(j)} - r^f_{h(j)}) - \frac{h(j)}{2}\beta(\beta - 1)\sigma^2 + \eta_{h(j)} \right). \tag{18}$$

This specification implies that the true expected alpha and GPME are zero for all funds. The idiosyncratic log return, η, is iid normally distributed with standard deviation, ω, of 25% per year, and mean $-\frac{h}{2}\omega^2$. We allow the fund-level r_h to be cross-sectionally correlated with a correlation coefficient of 0.1. Using these assumptions, the simulated funds have realistic return properties. For example, for a PE beta of two, the simulated mean internal rate of return (IRR) is 16.3%, which is close to the average IRR for both venture capital and buyout funds in the data (shown below). The standard deviation of simulated IRRs is 16.4%, a few percentage points higher than the observed standard deviation of buyout funds, but lower than the 41.0% standard deviation for venture capital funds in our data.

Table I reports the simulation results. Panel A uses one very large data set that spans 1,000,000 years and has one fund per vintage year, for a total of one million funds. Although this is implausibly large, it allows us to approximate the asymptotic properties of the α and (G)PME performance metrics. Each row shows results for a different value of the true common fund beta, ranging from 0.5 to 3. For each fund we compute its true realized alpha as $\sum_{j=1}^{J} z_j \exp\left(\alpha_{h(j)} + \eta_{h(j)}\right)$. These realized alphas are the benchmark for estimated performance metrics, since they are the best estimates of fund-level outperformance one can get if the true beta were known. For all values of beta we use the same draws for the market and idiosyncratic returns, so realized alphas are by design the same across betas, with a mean of -0.001 (very close to the true alpha of zero) and a standard deviation across funds of 0.571. Comparing our estimates of α and GPME, a number of results jump out. First, $\bar{\alpha}$ (i.e., average GPME and, mechanically, average α) is a very good estimate of the true mean fund alpha, ranging between -0.006 to +0.012 across the different values of beta. Second, the cross-sectional variation in α is close to the variation in realized alphas for all values of beta, ranging between 0.567 and 0.581. In contrast, the cross-sectional variation in GPMEs is much higher for most betas, and never lower than the variation in α. Third, the root mean squared error (RMSE) of α, measured relative to realized alphas, is small, ranging from 0.002 to 0.035. In comparison, the RMSE of GPME is very high, with values between 0.374 and 2.409. Fourth, the correlation of α with realized alphas is essentially one. This is much higher than the correlation between GPME and realized alphas, which is as low as 0.229 and no higher than 0.841. These results reveal that α is a more accurate measure of fund-level realized alphas than GPME. In fact, the α metric is different from realized alphas only because the estimated beta and σ^2 differ from their true values.

A case of special interest occurs when $\beta = 1$, so that $R_h^b = R_h^m$. Because the underlying assumption of $\beta = 1$ is true, the PME measure exactly equals the realized alpha for all funds. Fund alphas are very close, but differ slightly because the estimated beta is 1.006, not exactly one. For $\beta \neq 1$, PME performs poorly both as a measure of aggregate outperformance and as a measure of individual fund outperformance. For example, if the true β equals two, the average PME is 0.534, its standard deviation is 1.171, its RMSE is 0.942, and its correlation with realized alphas is 0.819. Generally, for $\beta > 1$, PME overestimates excess returns because it assumes a discount rate that is too low, and vice versa for $\beta < 1$.[5]

Another case of special interest is $\beta = \gamma$, because in this case GPME and α are the same. In our simulations we have $\gamma = 4$ and $\beta = 4$ is not included in the table, but it is clearly

[5] Another case of interest is $\beta = \gamma$, which equals 4 in this example. In this case, the GPME and alpha measures coincide, because $1/R_h^b$ and M_h are identical.

apparent that the cross-sectional standard deviation of GPME approaches the cross-sectional standard deviation of α.

Panel B shows results from simulated data sets of 1,200 funds over 30 years, which is a representative sample size for data sets observed in practice. To assess the small sample properties of the performance metrics, we repeat the simulation 50,000 times and report averages of the statistics across data sets. The mean realized alpha is -0.0003, with a standard deviation of 0.551, on average across data sets. Because parameter estimates are noisier in smaller data sets, α and GPME are less accurate compared to the large data set. Still, for betas of two or lower, $\bar{\alpha}$ continues to work quite well as a measure of aggregate outperformance, ranging between -0.042 and +0.006. For these values of beta, the estimated α's are close to the true realized fund alphas, with variances ranging from 0.523 to 0.552, RMSEs between 0.090 and 0.126, and correlations of 0.983 and higher. When beta reaches three, $\bar{\alpha}$ clearly diverges from the mean realized alpha, and estimated fund alphas substantially understate the cross-sectional variation in realized alphas, though their correlation is still close to one.

The PME metric does not suffer much performance loss compared to the large data set (and does better in many cases), as no parameters need to be estimated. However, PME performed poorly with large amounts of data, and continues to do so in smaller data sets. As before, only when beta is truly equal to one does PME perform well, as it coincides with realized alphas in this special case. However, even in this smaller data set the difference with the α metric remains relatively small. For all other values of beta, the α metric clearly captures both aggregate outperformance and realized fund betas better than PME.

A. Robustness

The simulations demonstrate that the alpha method works better than (G)PME when the additional assumptions of common fund betas and lognormality are satisfied. In this section we show that the α method is robust to modest amounts of cross-sectional variation in fund betas, and to reasonably large departures from lognormality.

A.1. Cross-sectional variation in fund betas

We simulate data in which fund betas are drawn randomly from a normal distribution with various combinations of means and variances. We use the same set of mean betas as the common fund betas in Table I, and two scenarios for standard deviations. In the first scenario the standard deviation is 0.25, so that 95% of fund betas are within 0.5 of the mean. We calculate results for both the large and the realistic size data sets using the same draws for market and idiosyncratic returns, so that true realized alphas are the same as in Table I. The results in Table II show that in the first scenario, $\bar{\alpha}$ is not affected by the cross-sectional variation in betas. This is not surprising for the large data set, since SDF pricing does not impose any assumption on betas, and the expected GPME (and hence expected α) should remain unchanged. However, the same is not true for the individual fund α and GPME estimates. These estimates are noisier, which shows up as higher cross-sectional variation, higher RMSEs, and a lower correlation with true realized alphas compared to the case of a common fund beta. Still, the differences with Table I are small, and we conclude that a modest amount of cross-sectional variation in betas is harmless in both large and realistic size data sets. One important difference to note is that while PME performed extremely well when the true beta equals one for all funds, when the average beta is one but there is moderate cross-sectional variation in fund betas, PME performs comparably to α in terms of RMSE and correlation with true realized alphas. When average beta is not one, PME performs much worse than α, as before.

The second scenario raises the standard deviation of beta to one, a rather extreme case in which 95% of betas are within 2 above or below the mean. For the large data set, Table II reports only small changes when increasing the dispersion in betas. For the realistic size data set, the standard deviation of α increases by about 20% to 50%, the RMSEs double to triple, while correlations remain at 0.65 or above. Despite the deteriorated performance, α still beats both GPME and PME as a performance measure even with the extreme variation in fund betas. In particular, even when the average beta is one, large heterogeneity in betas across funds cause PME to be a significantly worse measure of fund performance than α.

A.2. Deviations from lognormality

The alpha metric derivation assumes joint lognormality of market and PE returns. Although it is plausible that market returns deviate from lognormality, more serious violations of this assumption are likely to come from the idiosyncratic PE returns. For example, it is well known that in venture capital, a large fraction of a fund's investments fail, and most of the fund's return is generated by one or a few "home run" winners. As a result, idiosyncratic fund returns are right-skewed, resembling returns to a portfolio of long call options. The lognormal distribution, though right-skewed by nature, may not fully capture the thick right tail of the empirical distribution.

We consider two scenarios that depart from lognormality by simulating the idiosyncratic log fund return, η, as a mixture of two components. In the first scenario, there is a 10% chance that the fund's cash flow in a given year is magnified by exactly 50%. In the second scenario, the fund's payout in any year is 100% higher with 1% probability. When there is no such "home run" payout, η is drawn from a normal distribution with mean and variance chosen such that the *unconditional* mean and variance of the idiosyncratic return, $\exp(\eta)$, are equal to those of the lognormal distribution in the base case scenario of Table I. Relative to the base case, these two scenarios increase the skewness in fund IRRs by about 20% and 30%, respectively, and excess kurtosis by 35% and 100%, while preserving the first two unconditional moments.[6]

The results in Table III show that deviation of lognormality have little impact on the properties of α and GPME. In the large data set, the means and variances of α and (G)PME, as well as their RMSEs and correlations with true realized alphas, are very close to those reported in Table I for all values of beta and for both scenarios. For the realistic data set, the properties are also very similar, and if anything, α is closer to true realized alphas compared to the lognormal simulations. However, the true realized alphas are no longer the same as in the previous simulations, due to the new sampling procedure for idiosyncratic returns. Nevertheless, the main takeaway from Table III is clear: departures from lognormality do not

[6]Note that the simulations with cross-sectional heterogeneity in betas in Table II also induce skewness and kurtosis in PE returns, but only in the cross-section, not within fund returns.

appear to have a detrimental effect on the properties of α as a measure of fund performance.

IV. Data

We use private equity fund cash flow data provided by Burgiss, a global provider of investment decision support tools for investors in private capital markets. The data are sourced exclusively from hundreds of LPs, including public and private pension funds, endowments and foundations. We have the complete cash flow (net of management fees and carried interest (profit shares) paid to the GPs) and valuation history of all funds that the participating LPs invested in. The data are accurate and up-to-date, because they are used for audit, performance measurement and reporting purposes, and they are cross-checked when different LPs invest in the same fund. Market return and risk-free rate data are from Ken French's data library.[7]

A. Descriptive statistics

We separate our sample into 1,219 VC funds and 879 buyout funds, representing all U.S. funds raised before 2013 that have at least $5 million in committed capital in 1990 dollars (following Kaplan and Schoar, 2005). The cash flow data run through the end of 2017, so we have at least five years of cash flow data for all funds in our sample.

Table IV Panel A reports that the average (median) VC fund has $249 million ($160 million) in committed capital, and involves 38.5 (35.0) net cash flows spanning 13.4 years (13.4 years). There are 449 unique VC firms in the sample, with an average (median) of 2.7 (2.0) funds per GP.

The average (median) buyout fund has committed capital of $1,053 million ($450 million), runs for 12.1 (11.9) years, comprising 57.5 (52.0) cash flows. The sample represents 384 unique GPs, with 2.3 (2.0) funds per GP.

As of the end of 2017, 45% of VC funds and 34% of buyout funds were fully liquidated. These are funds that are at least 10 years old and have zero remaining net asset value (NAV). Some funds run longer than 10 years but have small net asset balances, so we also consider

[7] We thank Ken French for providing the data on his website.

a less strict definition that considers funds to be liquidated if they are over 10 years old and their most recent reported NAV is less than 5% of committed capital. By this "95% liquidated" measure, 54% of VC funds and 49% of buyout funds are liquidated.

Performance of private equity funds is typically measured by their internal rate of return (IRR), and their total value to paid-in capital (TVPI) multiple. The TVPI is defined as the total distributions plus the NAV of unrealized investments, divided by total capital contributions to the fund by LPs. Both measures are net of fees and carried interest paid to the GP. The IRR of the average (median) VC fund is 14.1% (7.5%) per year, compared to 13.9% (13.1%) for buyout. Weighted by fund size, the average IRRs are 11.0% and 13.6% for VC and buyout, respectively. The average (median) TVPI is 2.0 (1.4) for VC funds, and 1.7 (1.6) for buyout funds. The size-weighted TVPIs are 1.8 and 1.7, respectively.

Panel B of Table IV reports the distribution of the number of funds and their performance metrics by vintage year (defined as the year of the fund's first capital call). IRRs and TVPIs appear cyclical for both venture capital and buyout. Venture capital experienced high returns in the early to mid-1990s vintages followed by a prolonged period of poor returns until the end of the first decade of the new millenniums. Buyout returns have been relatively more stable, but showed some slumps in in the late 1990s and around 2005. Note that funds raised after 2008 have not yet reached their ten-year lifetimes by the end of 2017, which is the end of our sample period, and performance statistics depend increasingly on unrealized NAV toward the end of the sample. IRR and TVPI for fully liquidated funds tend to be higher, because many cash distributions occur towards the end of a fund's life, and in many cases GPs are conservative in updating NAVs (Jenkinson, Sousa, and Stucke (2013) and Brown, Gredil, and Kaplan (2019)).[8]

Harris, Jenkinson, and Kaplan (2014) perform a careful comparison of Burgiss with commonly used private equity data sets (including Preqin, Venture Economics, and Cambridge Associates). They find that compared to Preqin, Burgiss contains considerably more funds with cash flow histories, which are necessary to compute the performance metrics in this

[8] Since 2008, PE firms (both venture capital and buyout) are required by the Financial Accounting Standards Board under ASC Topic 820 (formerly known as FAS 157) to value their assets at fair value each quarter. While valuations have become more accurate since (Jenkinson et al., 2020, Easton et al., 2021), they remain conservative (Brown et al., 2019).

study. In their sample, Burgiss' coverage of VC funds is less extensive than Venture Economics and Cambridge Associates in the early years, but increases significantly over the sample period. For buyout, they find that Burgiss offers excellent coverage since 2000, but contains relatively few funds before 1993.

In terms of performance, Harris et al. report that Burgiss yields qualitatively and quantitatively similar results to the Cambridge Associates and Preqin data, whereas Venture Economics data are downward biased (see also Stucke, 2011). Crucially, performance does not appear to be biased by the voluntary approval of LPs to allow their data to be included by Burgiss.

Compared to Harris et al.'s sample, our Burgiss data includes four more years of data (their sample ends with the 2008 vintage funds). More importantly, many more LPs have allowed their data to be included since the conclusion of their study, resulting in additional fund data going back all the way to the earliest vintages. Consequently, our sample of VC and buyout funds is about 50% larger than the Burgiss sample in Harris et al. (2014).

B. Fund-level abnormal returns

Table V reports key statistics of the risk-adjusted fund performance metrics (α, GPME, and PME) as well as the parameter estimates of the SDF and the common fund beta, for venture capital and buyout separately. The distribution of fund performance differs substantially across the three metrics. For venture capital, the mean GPME (and hence, mean α) is negative but insignificant at -0.207. While similarly statistically insignificant, this result is lower than Korteweg and Nagel (2016), who estimate a mean GPME of -0.103 in a shorter sample of Preqin VC funds that includes vintages up to 2008. In contrast, the mean PME in Table V is 0.129, which is positive and significant (Korteweg and Nagel (2016) estimate a lower PME of 0.048). PME overstates the risk-adjusted return because the beta of VC is estimated at 2.4, far above the PME assumption that beta equals one and resulting in a PME-implied expected return that is too low, as in the simulations of Table I.[9] The litera-

[9] A different but equivalent perspective of the difference between PME and GPME is that PME assumes that the SDF is the reciprocal of the return on wealth, such that the SDF parameters δ and γ are zero and one, respectively. However, as Table V shows, the estimate of γ in the VC sample is also far from one.

ture reports VC fund beta estimates between 1.0 and 2.8 (see see Korteweg, 2019, 2022, for reviews of the literature). The wide range of estimated betas is due to differences in data sources, sample periods, and methodologies, but it's reassuring that our estimate falls in the middle of that range.

Despite the difference in central location of the GPME and PME distributions, the level of dispersion around the mean is very similar. Both measures have standard deviations around 1.4. In contrast, the standard deviation of fund-level α's is moderately lower, at 1.1. This is in line with the simulation results of Table I, which showed that both GPME and PME overstate the dispersion in true realized alphas when beta is not equal to the γ parameter in the SDF. Here beta is neither equal to the value of $\gamma = 1$ assumed in the PME nor the estimate of $\gamma = 3.5$ used in the GPME calculation. As a consequence, GPME and PME are contaminated with additional noise at the individual fund level that is eliminated in the α calculation. To illustrate the comparison between PME, GPME and α more clearly, Figure 2 plots the histograms of these three performance metrics. Many funds have GPMEs below -2 or above 2, whereas almost all fund α's are confined with this range. This implies that there are far fewer funds with extreme outcomes than the (G)PME metrics suggest. The PME distributions are visually closer to α, but are also more dispersed, as is most clearly seen from the lower peaks of the distributions at the mode and the higher incidence of outliers. Below we will analyze the implications of these differences for the evidence on performance persistence, and the relation between performance and fund characteristics.

Turning to the buyout results, we find that the mean GPME of 0.207 is above the mean PME of 0.148. The means of the two metrics are considerably closer than in VC, because the estimated buyout beta is 0.8, not so far from the PME assumption that beta equals one. This estimate is at the lower end of the 0.7 to 2.7 range of buyout fund betas reported in the literature, based on a variety of methodologies, data sources, and sample periods (see Korteweg, 2019, 2022). Since the beta estimate is below one, the mean GPME is above the mean PME, unlike in venture capital where a beta above one resulted in a lower mean GPME.

It might appear puzzling that despite being higher, the mean GPME is not statistically different from zero whereas the mean PME is highly significant. The reason is that the GPME requires estimation of the SDF parameters, whereas PME assumes they are fixed and given. Relaxing the PME restrictions raises the standard error of estimated GPME.

Comparing α and PME, the difference between the two distributions is considerably smaller in buyout than in VC. Though α has a higher mean and median, the standard deviation, skewness, and kurtosis of the two distributions are close. The reason is again that the estimated beta is not far from $\gamma = 1$ (or, equivalently, $\beta = 1$) assumed in the PME calculation. In the full sample of buyout funds considered here, PME and α thus look similar. However, the simulations suggest that α yields more accurate results when beta is not exactly one, and we show below that the differences between α and PME become more stark in certain subperiods and subsamples where beta is further away from one.

Comparing α and GPME, we find a huge difference in cross-sectional standard deviation. For α the cross-sectional standard deviation is only one third of the cross-sectional standard deviation of GPME. The reason is that the buy out fund beta of 0.8 is very far from the estimated $\gamma = 3.6$. Figure 2 further illustrates that the distribution of GPME looks very different from the distribution of α for buyout funds, and even more diffuse than in VC. As illustrated by the simulations, GPME includes a significant amount of noise at the fund level when β is far from γ, as is the case here. Thus, although mean GPME is a good metric for industry-level risk-adjusted performance, individual fund GPMEs are very noisy estimates of fund-level performance.

Overall, the results demonstrate clearly that the benchmark portfolio approach typically delivers a less noisy measure of individual fund abnormal performance than the (G)PME.

C. Two-factor Models

Table VI applies the benchmark portfolio approach in a two-factor setting. In this analysis, we the return of one of the four corner portfolios (small/big and low/high book-to market, i.e., SL, SH, BL, BH) of the six size-value portfolios underlying the Fama-French factors as a second factor in addition to the market factor in the SDF. More precisely, to obtain easily interpretable betas, we express the second factor as a log excess return relative to the market factor.[10]

Several points in Table VI are noteworthy. First, across all different two-factor specifications, we see the same pattern in terms of cross-sectional standard deviations of α and GPME as in the one-factor case. Both in Panel A, for VC, and Panel B, for buyout funds, the dispersion of α is always lower than the dispersion of GPME, consistent with more precisely estimated abnormal performance. Also, as in the one-factor case, the reduction in dispersion with α relative to GPME is much stronger for buyout funds than for VC funds.

Second, we see interesting patterns in factor betas and cross-sectional standard deviation of α across factor models. For VC funds, the factor model that produces the smallest cross-sectional standard deviation of α and the largest beta on the second factor is the one with the BL-M factor.[11] Somewhat surprisingly perhaps, a benchmark portfolio tilted toward large growth stocks (rather than small growth stocks) provides the closest approximation of VC fund payoffs. For buyout funds in Panel B, either a tilt away from small growth stocks (indicated by the negative factor beta on SL-M) or a tilt toward big value stocks (indicated by the positive factor beta on BH-M) produce the lowest cross-sectional standard deviation of α. These results show that the method recovers economically plausible factor exposures.

D. Subsamples

Table VII reports estimates when the performance metrics are estimated on the subsamples of fully liquidated funds, funds raised before and after the year 2000, and funds below and above median size (computed relative to the median fund size in a fund's vintage year).[12] For brevity, the table does not show the SDF parameter estimates or the higher order moments of the performance distributions.

---

[10]With a second factor, x, the SDF becomes $M_h = \exp(\delta h - \gamma_m r_h^m - \gamma_x r_h^x)$. The benchmark portfolio from equation (16) becomes

$$R_h^b = \exp\left\{r_h^f + \beta' r_h + \frac{1}{2}\beta' \mathrm{diag}(\Sigma_h) - \frac{1}{2}\beta' \Sigma_h \beta\right\}$$

where $r = \left[r_h^m - r_h^f; r_h^x\right]$, $\beta = [\beta_m; \beta_x]$, and Σ_h is the covariance matrix of r_h.

[11]The standard deviation of α is the square root of the objective function Q in equation (17), so the model with the lowest standard deviation is the best fitting model.

[12]A number of funds are exactly at median size (especially in venture capital) so the numbers of funds in the subsamples of small and large funds are not equal.

Table VII shows that the estimated VC betas are fairly constant across the subsamples, ranging from 2.2 to 2.5. The differences between the three performance metrics are therefore comparable to those of the full sample results. Mean α and GPME are consistently about 0.35 below the mean PME, and the standard deviations of GPME and PME are substantially higher than the standard deviation of α in most subsamples. The means vary quite substantially across subsamples. Liquidated VC funds have an insignificant mean alpha, which is higher than the full sample mean at least in part because many of these funds were raised earlier in the sample, and Panel B shows that VC funds raised before the year 2000 have positive, but insignificant, mean alpha. Funds raised in 2000 or later, and smaller funds, have significantly negative mean alpha. The variation in alpha across funds also differs by subsample. The standard deviation of α for liquidated funds and those raised before 2000 is about twice as large as that for post-2000 funds. Small funds have somewhat lower variation in alphas than large funds.

In contrast to VC, buyout betas vary quite dramatically across subsamples. The estimated beta of 0.7 for funds raised after the year 2000 is near the lowest across all subsamples, whereas pre-2000 funds have the highest beta at 1.1. This difference may be due to the fact that leveraged buyout deals were more highly levered in the 1980s and 1990s, differences in the types of industries that underwent buyouts, and differences in buyout strategies. The Burgiss coverage is also not very extensive before 1993, so some degree of sample selection is possible. Another potential explanation is that the pre-2000 period contains more liquidated funds. Non-liquidated fund performance is calculated using the most recent NAV, which may not have been updated by GPs (especially prior to the ASC 820/FAS 157 Fair Value accounting standard of 2007) and therefore understate the extent to which performance fluctuates with the benchmark factors. However, this is not likely to be dominant reason. Although the beta for the subsample of liquidated funds is higher than the full-sample beta, at 0.9 versus 0.8, we do not see much difference in these beta estimates for VC funds, which are likely more subject to the NAV updating issue than buyout funds. Finally, when we split the sample by fund size, we find that small funds have a beta of 0.6, compared to 1.0 for large funds.

Since the estimated beta for liquidated and large funds are close to one, α and PME are also close. This is not the case for the post-2000 and small funds, where the mean α is almost double the mean PME, due to their low beta. Conversely, the mean alpha of 0.142 for pre-2000 funds is slightly below the mean PME of 0.157, as the estimated beta is above one. For all subsamples, the standard deviation of alpha and PME are close, and fairly constant across subsamples, being somewhat higher for smaller and liquidated funds, and somewhat lower for large funds. As in the full sample results, the variance in GPME is substantially higher, and the difference with the variance of alpha is larger for subsamples with lower betas.

Table VIII considers three subclasses of venture capital, namely, generalist VCs who invest in companies of all stages, and VCs who invest in early-stage or late-stage startups.[13] Unlike the VC subsamples in Table VII, beta estimates show more variation across the subclasses. Generalist VCs are similar to post-2000 funds, with a beta of 2.2, a mean α of -0.269, which is significantly different from zero at the 5% level. Unlike post-2000 funds, the mean PME is near zero at -0.003, and insignificant. Early-stage VCs have a mean beta of 2.7, while the late-stage beta is 1.5. This drop in beta from early-stage to late-stage funds is consistent with surviving startups turning risky growth opportunities into assets in place as they mature (Berk, Green, and Naik, 1999, and Fisher, Carlson, and Giammarino, 2004, but see Zhang (2005) for a different perspective). As Berk, Green, and Naik (2004) point out, the systematic risk of the early-stage company is higher even if the early-stage risk is purely idiosyncratic (for example, due to technological risk). The mean α of early-stage VCs is -0.216, while late-stage VCs have a positive mean α of 0.062, though neither are significantly different from zero. The mean PME of these two subclasses is nearly identical, 0.206 for early-stage and 0.182 for late-stage VCs. Finally, the dispersion in alphas is higher for early-stage VCs compared to late-stage and generalist VCs. This could be due to larger heterogeneity in skill in early-stage investing, but is probably mostly due to a higher degree of idiosyncratic risk resulting in more extreme realized outcomes.

We recognize that the variation in estimated betas across subsamples and subclasses is a potential source of concern given the assumption of common fund betas underlying the alpha method. However, the overall range of betas, from 1.5 to 2.7 in VC, and from 0.6 to 1.1 in

[13] The data do not contain subclass designations for buyout.

buyout, is within the range where the simulations suggest the alpha method works well.

E. Performance regressions

We introduce two exercises to illustrate the importance of accurate fund-level performance estimates. The first exercise is to replicate common regressions in the literature that have fund performance as the dependent variable. Specifically, we consider regressions that relate performance to fund size. The second exercise considers the persistence of PE performance. We focus on comparing the α and GPME metrics, which are less restrictive than PME. All results in this section are based on the full-sample performance estimates reported in table V.

E.1. Performance and fund size

Table IX reports regression results of fund performance on size quartiles, with the smallest quartile being the omitted category. All regressions include vintage year fixed effects. The regressions in columns 1 and 2 use α as the dependent variable, columns 3 and 4 use GPME, and the final two columns use PME.

The results for venture capital are in Panel A. Column 1 shows that a fund alphas are monotonically increasing in size quartiles. The performance of the fourth (largest) quartile is significantly higher than the first (smallest) quartile, at the 5% level. For GPME in column 3, the relation is not quite monotonic, but the result for the largest quartile remains. The standard errors in column 1 are about a third lower than those for GPME in column 3, due to lower degree of noise when using alpha rather than GPME as the dependent variable. All else equal, this increases t-statistics by about 50%. The PME standard errors in column 5 are roughly the same order of magnitude as their GPME counterparts, and therefore also considerably larger than in the alpha regressions. The increased power of test statistics is an important advantage to using alpha in regressions that include a fund-level performance metric.

Columns 2, 4, and 6 include the natural logarithm of the fund's sequence number as an additional explanatory variable.[14] The coefficient on sequence number is strongly positive and significant. Including fund sequence changes the sign of the size quartile coefficients, rendering all of them insignificant. Put differently, the effect of size in the earlier regressions appears appears to be a survivorship effect. Successful VCs survive and get to raise larger funds, and this is captured by the fund's sequence number.

Turning to buyout funds in Panel B, we do not find any significant coefficients on size quartiles or the fund sequence number for any performance metric or any specification. Still, the choice of performance metric has a large effects on standard errors. The standard errors for the alpha regressions are as low as one third the size of those for GPME. Perhaps surprisingly, the adjusted R-squared is higher for the GPME regressions. This is due to the vintage fixed effects soaking up more of the variance in buyout GPMEs compared to alphas. In contrast, the difference between the α and PME regressions is marginal, since the buyout beta is close to one.

Several papers consider the relation between fund size and PME. Kaplan and Schoar (2005) and Harris, Jenkinson, and Kaplan (2014) find similar results for the relation between size and VC performance, though the former paper finds an insignificant coefficient on the fund sequence number, and the latter does not control for sequence. Like us, the literature finds no significant relation between buyout fund size and performance (for example, Kaplan and Schoar, 2005, Harris, Jenkinson, and Kaplan, 2014, and Lopez-de-Silanes, Phalippou, and Gottschalg, 2015, Rossi, 2019).

E.2. Performance persistence

We assess the degree of performance persistence in VC and buyout by considering how well a current fund's performance predicts the risk-adjusted returns of future funds by the same manager (GP). We adopt the following procedure. First, for each vintage year between 1985 and 2010 we rank funds by their α metric (we start in 1985 to have a reasonable number of observations for both VC and buyout funds, and we stop in 2010 to allow GPs some time to raise a next fund). For each fund in the top and bottom quartile of its vintage year, we

[14] Burgiss did not provide fund sequence numbers. We constructed sequence numbers based on GP identifiers and vintage years.

then compute the risk-adjusted performance (measured by α, GPME, and PME) of future funds of the same type (VC or buyout) that were raised by the same GP. The first three columns of Table X show the results for VC (in panel A) and buyout (panel B). The first four rows of each panel show the future performance of GPs with a fund in the top quartile, and the second four rows show the performance of bottom quartile GPs. Within each block of four rows, we report the average performance of the next fund raised (both the very next fund, which may overlap with the present fund, and the next nonoverlapping fund[15]), as well as of all future funds (again either considering all future funds regardless of overlap, or only nonoverlapping future funds). As a comparison, the middle and rightmost sets of three columns of Table X repeat the same exercise but using GPME and PME as alternative performance measures to rank funds in the first step of the procedure.

The key takeaway from Table X is that using the α metric to sort GPs into top and bottom quartiles gives better predictive results in most cases, compared to using (G)PME. For example, GPs with top quartile VC funds by α, generate an average α of 0.57 for their next nonoverlapping fund. Using PME to sort GPs instead, yields an average α of 0.41. This difference is economically meaningful: on a dollar invested, the difference in net present value is \$0.16 over the life of the fund.

Moreover, even if (G)PME is used as the performance evaluation metric of choice, using α to rank GPs results in better sorting in most cases. To continue the prior example, the PME of the next nonoverlapping fund of a top quartile VC GP as ranked by α is 1.01, compared to only 0.81 for a top-quartile GP based on sorting the current funds by PME itself.

The differences between α and PME for top quartile GPs are similar if we consider all future nonoverlapping funds instead of only the next one. In contrast, for overlapping funds, α and PME sorts yield results that are roughly comparable to each other.

Results for the bottom quartile of VCs are similar: α does a better job in identifying poorly performing GPs for nonoverlapping funds, as both the α and PME are lower (more negative) compared to sorting on PME. For overlapping funds, sorting on PME gives slightly

[15] Using the next nonoverlapping fund eliminates concerns of mechanical correlation (see, for example, Korteweg and Sorensen, 2017, for a discussion), although overlap should be less of a concern for risk-adjusted (as opposed to raw) performance.

better results than sorting on α.

Compared to sorting on GPME, classifying VC funds by their α does a better job identifying GPs with poor future performance, both in terms of α and GPME and irrespective of whether one considers only the next fund or all future funds, or overlapping or nonoverlapping funds. For top quartile VCs, α sorts result in slightly better average performance of overlapping funds, but somewhat worse performance in nonoverlapping funds.

For buyout, sorting on α better predicts future fund performance than sorting on GPME. both for the top and bottom quartiles, and irrespective of whether funds overlap or whether one considers the next fund or all future funds. The results for poorly performing managers are especially strong. Not surprisingly, given their similarity for buyout, α and PME yield similar predictive results.

V. Conclusion

We introduce a new method to benchmark payoffs of individual private equity funds when only cash flow data are available, but not returns. The method has a clear theoretical foundation in asset pricing theory and, with a one-factor benchmark, it requires estimation of only four parameters. Risk-adjusted payoffs based on this new measure are identical to the GPME measure of Korteweg and Nagel (2016) at the asset-class level, but they are less noisy than the GPME at the individual fund level. For this reason, the new measure provides a more accurate assessment of individual fund performance than the PME or its generalized version, although the measures are identical in in the special case where beta happens to be equal to the market price of risk parameter in the SDF. The method assumes log-normality of payoffs and within-asset-class homogeneity of betas, but simulations show that it is robust to reasonably large deviations from these assumptions. The method works well even in smaller data sets.

The empirical results from a large data set of private equity fund cash flows reveal that the variation in estimated risk-adjusted performance across venture capital funds is considerably smaller with our new method than with the PME and GPME. For example, for buyout funds, the cross-sectional standard deviation of abnormal payoffs according to our new measure is only one third of the cross-sectional standard deviation of the GPME. This illustrates clearly the much lower level of noise in the new measure.

We also extend the new measure to a two-factor setting, where the benchmark portfolio combines the market factor with one of the four corner size/value portfolios underling the Fama-French factors. We obtain economically plausible factor exposures. For venture capital funds, we find that a benchmark portfolio tilted toward big growth stocks best explains individual fund payoffs and hence produces the smallest cross-sectional dispersion abnormal payoffs. For buyout funds, we find the that a tilt away from small growth or toward big value works best. The method can be readily extended to include additional benchmark assets. It also has applications to other asset classes such as real estate funds.

Our new measure enhances the statistical power of tests of the determinants of fund-level performance. For example, the lower level of noise in our new measure results in substantially lower standard errors in regressions of fund performance on the size of the fund, and improvements in fund performance predictability. In some cases the results are very different depending on the performance metric used. For example, we can much better predict the future performance of buyout managers, especially poorly performing GPs, when ranking their current fund's performance on α rather than on GPME.

REFERENCES

Abell, Peter, and Tahir M. Nisar, 2007, Performance effects of venture capital firm networks, *Management Decision* 45, 923–936.

Andonov, Aleksandar, Yael V. Hochberg, and Joshua D. Rauh (2018), Political representation and governance: Evidence from the investment decisions of public pension funds, *Journal of Finance* 73, 2041–2086.

Ang, Andrew, Bingxu Chen, William N. Goetzmann, and Ludovic Phalippou, 2018, Estimating private equity returns from limited partner cash flows, *Journal of Finance* 73, 1751–1783.

Barber, Brad M., and Ayako Yasuda, 2017, The impact of interim performance on private equity fundraising, *Journal of Financial Economics* 124, 172–194.

Berk, Jonathan B., Richard C. Green, and Vasant Naik, 1999, Optimal investment, growth options, and security returns, *Journal of Finance* 54, 1553–1607.

Berk, Jonathan B., Richard C. Green, and Vasant Naik, 2004, Valuation and return dynamics of new ventures, *Review of Financial Studies* 17, 1–35.

Boyer, Brian, Taylor D. Nadauld, Keith P. Vorkink, and Michael S. Weisbach, 2021, Discount rate risk in private equity: Evidence from secondary market transactions, *Journal of Finance*, forthcoming.

Braun, Reiner, Tim Jenkinson, and Ingo Stoff, 2017, How persistent is private equity performance? Evidence from deal-level data, *Journal of Financial Economics* 123, 273–291.

Brown, Gregory W., Eric Ghysels, and Oleg Gredil, 2022, Nowcasting net asset values: The case of private equity, *Review of Financial Studies*, forthcoming.

Brown, Gregory W., Oleg Gredil, and Steven N. Kaplan, 2019, Do private equity funds manipulate reported returns? *Journal of Financial Economics* 132, 267–297.

Campbell, John Y., and Luis M. Viceira, 2002. *Strategic asset allocation: Portfolio choice for long-term investors.* Oxford University Press, Oxford, UK.

Cao, Jerry, and Josh Lerner, 2009, The performance of reverse leveraged buyouts, *Journal of Financial Economics* 91, 139–157.

Carlson, Murray, Adlai Fisher, and Ron Giammarino, 2004, Corporate investment and asset price dynamics: Implications for the cross-section of returns, *Journal of Finance* 59, 2577–2603.

Cavagnaro, Daniel R., Berk A. Sensoy, Yingdi Wang, and Michael S. Weisbach, 2019, Measuring institutional investors' skill at making private equity investments, *Journal of Finance* 74, 3089–3134.

Chakraborty, Indraneel and Michael Ewens, 2018, Managing performance signals through delay: Evidence from venture capital, *Management Science* 64, 2875–2900.

Chung, Ji-Woong, 2012, Performance persistence in private equity funds, Working paper, Chinese University of Hong Kong.

Chung, Ji-Woong, Berk A. Sensoy, Lea Stern, and Michael S. Weisbach, 2012, Pay for performance from future fund flows: The case of private equity, *Review of Financial Studies* 25, 3259–3304.

Cochrane, John H., 2005, The risk and return of venture capital, *Journal of Financial Economics* 75, 3–52.

Driessen, Joost, Tse-Chun Lin, and Ludovic Phalippou, 2012, A new method to estimate risk and return of nontraded assets from cash flows: the case of private equity funds, *Journal of Financial and Quantitative Analysis* 47, 511–535.

Gupta, Arpit, and Stijn Van Nieuwerburgh, 2021, Valuing private equity investments strip by strip, *Journal of Finance* 76, 3255–3307.

Gredil, Oleg, Morten Sorensen, and William Waller, 2020, Evaluating private equity performance using stochastic discount factors, Working paper, Tuck School of Business.

Driessen, Joost, Tse-Chun Lin, and Ludovic Phalippou, 2012, A new method to estimate risk and return of nontraded assets from cash flows: The case of private equity funds, *Journal of Financial and Quantitative Analysis* 47, 511–535.

Easton, Peter D., Stephannie Larocque, and Jennifer Sustersic Stevens, 2021. Private equity valuation before and after ASC 820, *Journal of Investment Management* 19, 105–135.

Ewens, Michael, Charles M. Jones, and Matthew Rhodes-Kropf, 2013, The price of diversifiable risk in venture capital and private equity, *Review of Financial Studies* 26, 1854–1889.

Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3-56.

Franzoni, Francesco, Eric Nowak, and Ludovic Phalippou, 2012, Private equity performance and liquidity risk, *Journal of Finance* 67, 2341–2373.

Gompers, Paul A., and Josh Lerner, 1997, Risk and reward in private equity investments: The challenge of performance assessment, *Journal of Private Equity* 1, 5–12.

Harris, Robert S., Tim Jenkinson, and Steven N. Kaplan, 2014, Private equity performance: What do we know? *Journal of Finance* 69, 1851–1882.

Harris, Robert S., Tim Jenkinson, Steven N. Kaplan, and Rüdiger Stucke, 2014, Has persistence persisted in private equity? Evidence from buyout and venture capital funds, Working paper, University of Virginia, University of Chicago, and University of Oxford.

Hochberg, Yael V., Alexander Ljungqvist, and Yang Lu, 2007, Whom you know matters: Venture capital networks and investment performance, *Journal of Finance* 62, 251–301.

Hochberg, Yael V., Alexander Ljungqvist, and Annette Vissing-Jorgensen, 2014, Informational holdup and performance persistence in venture capital, *Review of Financial Studies* 27, 102–152.

Hochberg, Yael V., and Joshua D. Rauh, 2013, Local overweighting and underperformance: Evidence from limited partner private equity investments, *Review of Financial Studies* 26, 403–451.

Hüther, Niklas, 2022, Do private equity managers raise funds on (sur)real returns? Evidence from deal-level data, *Journal of Financial and Quantitative Analysis*, forthcoming.

Hüther, Niklas, David T. Robinson, Sönke Sievers, and Thomas Hartmann-Wendels, 2020, Paying for performance in private equity: Evidence from venture capital partnerships, *Management Science* 66, 1756–1782.

Jegadeesh, Narasimhan, Roman Kraussl, and Joshua M. Pollet, 2015, Risk and expected returns of private equity investments: Evidence based on market prices, *Review of Financial Studies* 28, 3269–3302.

Jenkinson, Tim, Wayne R. Landsman, Brian R. Rountree, and Kazbi Soonawalla, 2020. Private equity net asset values and future cash flows, *The Accounting Review* 95, 191–210.

Jenkinson, Tim, Miguel Sousa, and Rüdiger Stucke, 2013, How fair are the valuations of private equity funds? Working paper, Oxford University.

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

Korteweg, Arthur, 2019, Risk adjustment in private equity returns, *Annual Review of Financial Economics* 11, 131–152.

Korteweg, Arthur, 2022, Risk and return in private equity, *Handbook in Economics: Corpo-*

rate Finance I, forthcoming.

Korteweg, Arthur, and Stefan Nagel, 2016, Risk-adjusting the returns to venture capital, *Journal of Finance* 71, 1437–1470.

Korteweg, Arthur, and Morten Sorensen, 2010, Risk and return characteristics of venture capital-backed entrepreneurial companies, *Review of Financial Studies* 23, 3738–3772.

Korteweg, Arthur, and Morten Sorensen, 2017, Skill and luck in private equity performance, *Journal of Financial Economics* 124, 535–562.

Lerner, Josh, Antoinette Schoar, and Wan Wongsunwai, 2007, Smart institutions, foolish choices: The limited partner performance puzzle, *Journal of Finance* 62, 731–764.

Liu, Zhiyang, and Zhiqi Chen, 2014, Venture capital networks and investment performance in China, *Australian Economic Papers* 53, 97–111.

Long, John B., 1990, The numeraire portfolio, *Journal of Financial Economics* 26, 29–69.

Lopez-de-Silanes, Florencio, Ludovic Phalippou, and Oliver Gottschalg, 2015, Giants at the gate: Investment returns and diseconomies of scale in private equity, *Journal of Financial and Quantitative Analysis* 50, 377–411.

Metrick, Andrew, and Ayako Yasuda, 2010, The economics of private equity funds, *Review of Financial Studies* 23, 2303–2341.

Nanda, Ramana, Sampsa Samila, and Olav Sorenson, 2020, The persistent effect of initial success: Evidence from venture capital, *Journal of Financial Economics* 137, 231–248.

Opp, Christian C., 2019, Venture capital and the macroeconomy, *Review of Financial Studies* 32, 4387–4446.

Peng, Liang, 2001, Building a venture capital index, Working paper, University of Colorado, Boulder.

Peters, Ryan, 2018, Volatility and venture capital, Working paper, Tulane University.

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.

Robinson, David T., and Berk A. Sensoy, 2013, Do private equity fund managers earn their fees? Compensation, ownership, and cash flow performance, *Review of Financial Studies* 26, 2760–2797.

Robinson, David T., and Berk A. Sensoy, 2016. Cyclicality, performance measurement, and cash flow liquidity in private equity, *Journal of Financial Economics* 122, 521–543.

Rossi, Andrea, 2019, Decreasing returns or mean-reversion of luck? The case of private equity fund growth, Working paper, The Ohio State University.

Sensoy, Berk A., Yingdi Wang, and Michael S. Weisbach, 2014, Limited partner performance and the maturing of the private equity industry, *Journal of Financial Economics* 112, 320–343.

Sorensen, Morten, and Ravi Jagannathan, 2015, The public market equivalent and private equity performance, *Financial Analysts Journal* 71, 43–50.

Stafford, Erik, 2022, Replicating private equity with value investing, homemade leverage, and hold-to-maturity accounting, *Review of Financial Studies* 35, 299–342.

Stucke, Rüdiger, 2011, Updating history, Working paper, University of Oxford.

Woodward, Susan, 2009, Measuring risk for venture capital and private equity portfolios, Working paper, Sand Hill Econometrics.

Zhang, Lu, 2005, The value premium, *Journal of Finance* 60, 67–103.

Appendix: Proofs

Proof of equation (7)**.** The left-hand side of equation (7), with the solutions for δ and γ substituted in, equals

$$E[\exp(\delta h - \gamma r_h^m) X_j] = -r^f h + E[x_j] + \frac{1}{2}\sigma_x^2 - \frac{\mu}{\sigma^2} cov(x_j, r_h^m), \tag{A.1}$$

$$= -r^f h + E[x_j] + \frac{1}{2}\sigma_x^2 - \beta\mu h. \tag{A.2}$$

The right-hand side is

$$E\left[\exp\left\{-r^f h + \frac{h}{2}\beta(\beta-1)\sigma^2 - \beta(r_h^m - r^f h)\right\} X_j\right]$$

$$= -r^f h - \frac{h}{2}\beta\sigma^2 - \beta E[r_h^m - r^f h] + E[x_j] + \frac{1}{2}\sigma_x^2, \tag{A.3}$$

$$= -r^f h + E[x_j] + \frac{1}{2}\sigma_x^2 - \beta\mu h, \tag{A.4}$$

where, for the first equality we used $\beta cov(x_j, r_h^m) = \beta^2 h \sigma^2$. Hence, the left and right-hand sides of equation (7) are equal. ■

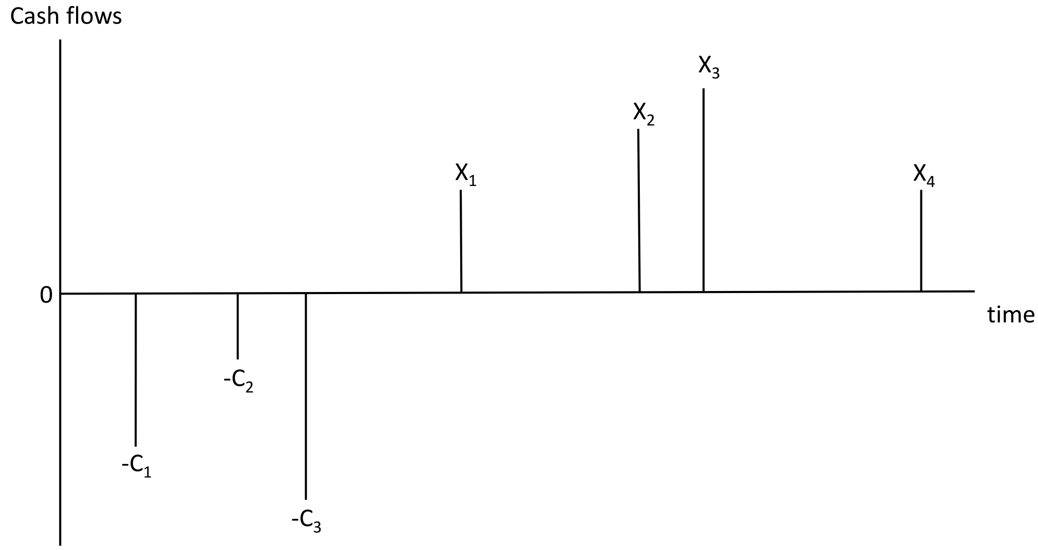

Figure 1. Example of the timing of cash flows to limited partners of a fictitious private equity fund with three capital calls $(C_1, \ldots, C_3)$ and four distributions $(X_1, \ldots, X_4)$.

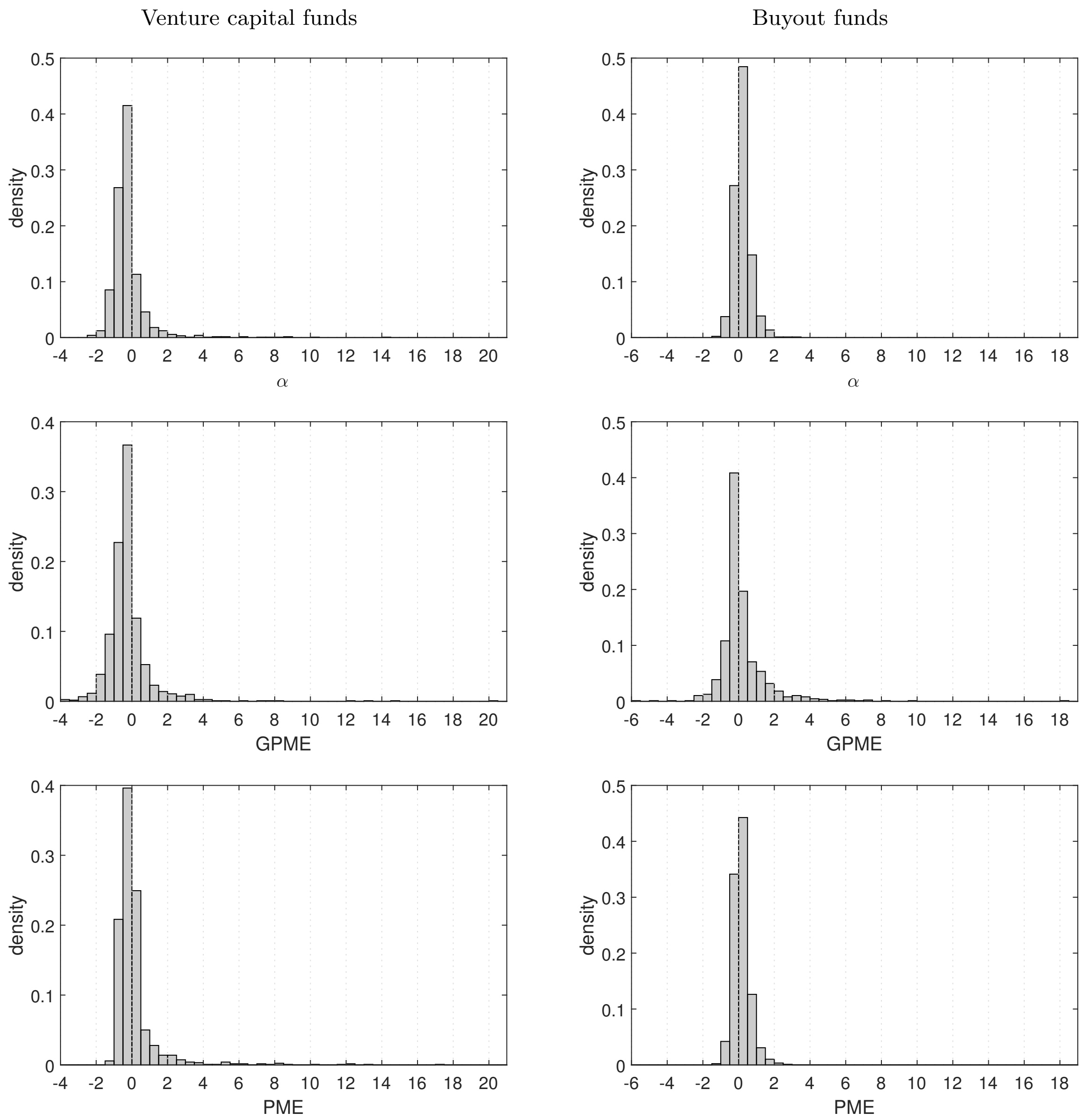

Figure 2. This figure shows histograms of the fund performance metrics estimated on the Burgiss data. The top, middle, and bottom rows show the distribution of fund alphas, GPMEs, and PMEs, respectively. The performance metrics are described in Table V, using the market return as the sole risk factor in the SDF. The figures in the left column show the performance distribution for venture capital, and the right column shows leveraged buyout fund performance.

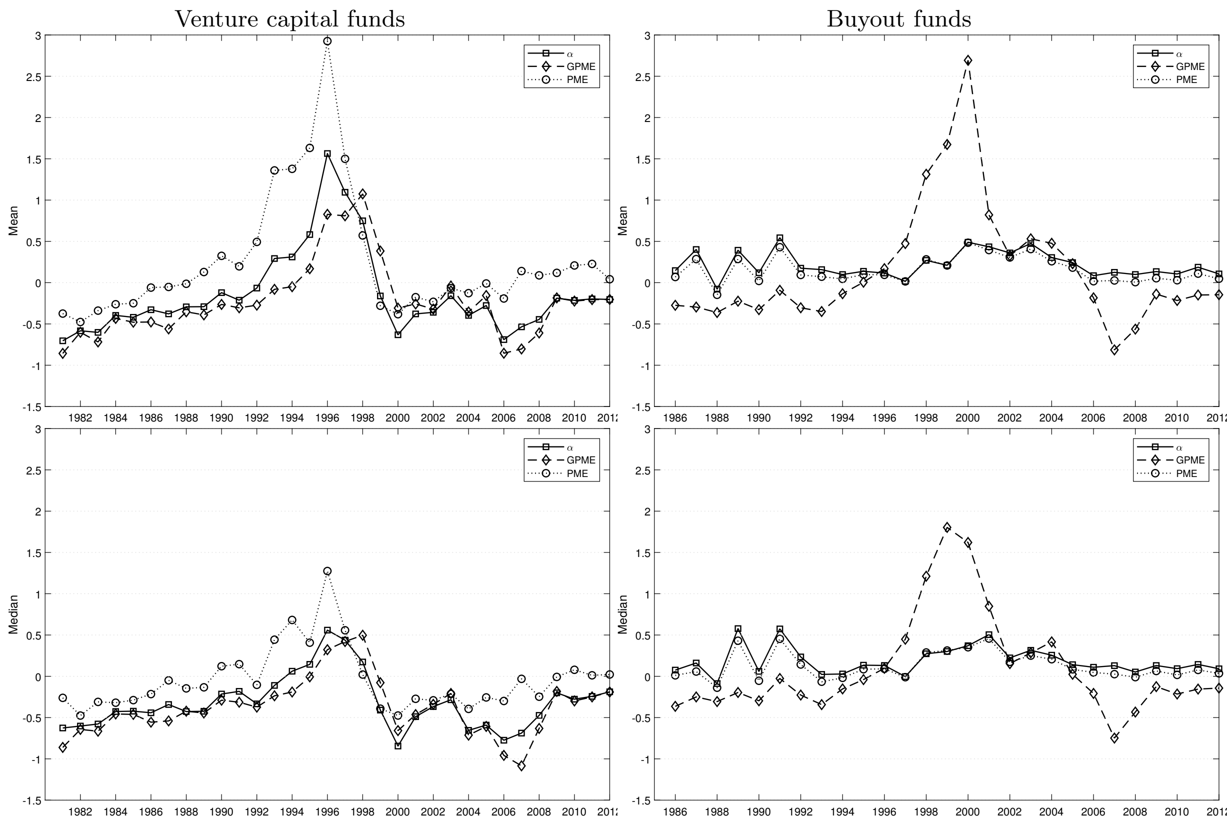

Figure 3. This figure shows time series of the fund performance metrics estimated on the Burgiss data. The top and bottom rows show the mean and median, respectively. The performance metrics are described in Table V, using the market return as the sole risk factor in the SDF. The figures in the left column show the performance for venture capital, and the right column shows leveraged buyout fund performance. Years on the horizontal axis are fund vintage years.

Table I

Simulation Results

This table reports summary statistics of the fund-level abnormal return measures α, GPME and PME, based on simulated data sets of N funds that are active over T years. Panel A uses one large data set of 1,000,000 vintages with one fund per vintage (yielding one million funds in total) to approximate asymptotic results, and Panel B uses 50,000 data sets of similar size as observed private equity fund data sets (30 years with 40 funds per vintage, for a total of 1,200 funds). Funds call \$1 in capital in their vintage year, and distribute 25 cash flows at random times, distributed uniformly over their ten year lifetimes. The underlying fund returns, on which cash flows are based, are generated using the true market β shown in the first column and with idiosyncratic volatility of 25% per year. The log risk-free rate is 2% per year, and the annual log market return is drawn from a normal distribution with $\mu = 11\%$ and $\sigma = 15\%$. The benchmark portfolio used to estimate α is

$$R_h^b = \exp(r_h^f + \beta(r_h^m - r_h^f) - \frac{h}{2}\beta(\beta - 1)\sigma^2),$$

for a period from fund inception to h years later. Panel A reports the estimates for the common fund beta, $\widehat{\beta}$, and Panel B reports its mean and standard deviation (in parentheses) across the simulated data sets. The estimated variance of market returns, $\widehat{\sigma^2}$, is not shown. The GPME performance measure is computed using an exponential affine stochastic discount factor (SDF)

$$M_h = \exp(\delta h - \gamma r_h^m).$$

PME is a restricted version of the alpha metric that assumes $\beta = 1$ regardless of the true beta, or equivalently, a restricted version of GPME with $\delta = 0$ and $\gamma = 1$. For each fund performance metric, Panel A reports its mean and standard deviation (in parentheses) across funds, its root mean squared error (*RMSE*) relative to true realized fund alphas, and the correlation (*Corr.*) between the metric and true realized alphas (in square brackets). Panel B reports the mean of these metrics across the 50,000 simulated data sets.

β	$\widehat{\beta}$	α		GPME		PME	
		Mean	RMSE	Mean	RMSE	Mean	RMSE
		(St.Dev)	[Corr.]	(St.Dev)	[Corr.]	(St.Dev)	[Corr.]
Panel A: Large Data Set (N = 1,000,000; T = 1,000,000)							
0.5	0.511	-0.006	0.008	-0.006	2.409	-0.180	0.265
		(0.567)	[1.000]	(2.475)	[0.229]	(0.471)	[0.948]
1	1.006	-0.003	0.004	-0.003	1.645	-0.001	0.000
		(0.569)	[1.000]	(1.741)	[0.327]	(0.571)	[1.000]
1.5	1.497	0.000	0.002	0.000	1.183	0.231	0.379
		(0.571)	[1.000]	(1.314)	[0.436]	(0.778)	[0.947]
2	1.983	0.004	0.009	0.004	0.854	0.534	0.942
		(0.574)	[1.000]	(1.030)	[0.558]	(1.171)	[0.819]
2.5	2.460	0.008	0.019	0.008	0.595	0.933	1.819
		(0.577)	[1.000]	(0.829)	[0.696]	(1.884)	[0.668]
3	2.917	0.012	0.035	0.012	0.374	1.460	3.251
		(0.581)	[0.999]	(0.690)	[0.841]	(3.163)	[0.526]

Table I - Continued

Panel B: 50,000 Realistic Data Sets (N = 1,200; T = 30)							
0.5	0.499	0.006	0.092	0.006	1.644	-0.179	0.251
	(0.191)	(0.552)	[0.995]	(1.717)	[0.313]	(0.447)	[0.966]
1	1.001	0.000	0.090	0.000	1.538	-0.000	0.000
	(0.207)	(0.548)	[0.994]	(1.623)	[0.344]	(0.551)	[1.000]
1.5	1.514	-0.017	0.097	-0.018	1.413	0.230	0.343
	(0.245)	(0.536)	[0.990]	(1.510)	[0.376]	(0.732)	[0.967]
2	2.033	-0.042	0.126	-0.048	1.281	0.531	0.818
	(0.317)	(0.523)	[0.983]	(1.384)	[0.407]	(1.035)	[0.891]
2.5	2.535	-0.068	0.178	-0.087	1.153	0.927	1.486
	(0.407)	(0.511)	[0.974]	(1.253)	[0.436]	(1.518)	[0.804]
3	3.011	-0.093	0.248	-0.133	1.042	1.451	2.438
	(0.508)	(0.502)	[0.962]	(1.125)	[0.462]	(2.264)	[0.722]

Table II

Simulations with Heterogeneous Betas

This table reports summary statistics of the fund-level abnormal return measures α, GPME, and PME, based on simulated data sets of N funds over T years, with heterogeneity in fund factor loadings. Panel A uses one large data set of 1,000,000 vintages with one fund per vintage (yielding one million funds in total) to approximate asymptotic results, and Panel B uses 50,000 data sets of similar size as observed private equity fund data sets (30 years with 40 funds per vintage, for a total of 1,200 funds). Simulations are as described in Table I, except that fund betas are generated from a normal distribution with mean μ_β and standard deviation σ_β. For each fund performance metric, Panel A reports its mean and standard deviation (in parentheses) across funds, its root mean squared error (*RMSE*) relative to true realized fund alphas, and the correlation (*Corr.*) between the metric and true realized alphas (in square brackets). Panel B reports the mean of these metrics across the 50,000 simulated data sets.

μ_β	$\widehat{\beta}$	α		GPME		PME	
		Mean (St.Dev)	RMSE [Corr.]	Mean (St.Dev)	RMSE [Corr.]	Mean (St.Dev)	RMSE [Corr.]
Panel A: Large Data Set (N = 1,000,000; T = 1,000,000)							
Scenario 1: Low dispersion in β ($\sigma_\beta = 0.25$)							
0.5	0.528	-0.006	0.008	-0.006	2.374	-0.174	0.257
		(0.567)	[1.000]	(2.441)	[0.232]	(0.473)	[0.951]
1	1.023	-0.003	0.004	-0.003	1.626	0.006	0.011
		(0.569)	[1.000]	(1.723)	[0.331]	(0.576)	[1.000]
1.5	1.514	0.001	0.002	0.001	1.170	0.240	0.394
		(0.572)	[1.000]	(1.302)	[0.439]	(0.787)	[0.944]
2	1.999	0.004	0.009	0.004	0.844	0.546	0.966
		(0.574)	[1.000]	(1.021)	[0.563]	(1.189)	[0.814]
2.5	2.476	0.008	0.019	0.008	0.587	0.949	1.857
		(0.577)	[1.000]	(0.823)	[0.701]	(1.917)	[0.663]
3	2.932	0.012	0.036	0.012	0.367	1.481	3.314
		(0.581)	[0.998]	(0.686)	[0.845]	(3.222)	[0.521]
Scenario 2: High dispersion in β ($\sigma_\beta = 1$)							
0.5	0.579	-0.005	0.007	-0.005	2.275	-0.158	0.233
		(0.567)	[1.000]	(2.344)	[0.242]	(0.480)	[0.961]
1	1.074	-0.003	0.003	-0.003	1.569	0.028	0.044
		(0.569)	[1.000]	(1.670)	[0.342]	(0.591)	[0.999]
1.5	1.564	0.001	0.003	0.001	1.132	0.268	0.443
		(0.572)	[1.000]	(1.268)	[0.451]	(0.818)	[0.933]
2	2.049	0.005	0.010	0.005	0.816	0.583	1.040
		(0.575)	[1.000]	(0.998)	[0.576]	(1.246)	[0.799]
2.5	2.524	0.009	0.021	0.009	0.563	0.997	1.975
		(0.578)	[1.000]	(0.806)	[0.716]	(2.018)	[0.647]
3	2.977	0.013	0.039	0.013	0.345	1.545	3.512
		(0.581)	[0.998]	(0.675)	[0.859]	(3.405)	[0.508]

Table II - Continued

Panel B: 50,000 Realistic Data Sets (N = 1,200; T = 30)							
Scenario 1: Low dispersion in β ($\sigma_\beta = 0.25$)							
0.5	0.518	0.005	0.195	0.005	1.642	-0.174	0.266
	(0.195)	(0.577)	[0.952]	(1.715)	[0.313]	(0.463)	[0.938]
1	1.017	-0.001	0.179	-0.001	1.538	0.006	0.148
	(0.211)	(0.567)	[0.958]	(1.622)	[0.343]	(0.577)	[0.964]
1.5	1.527	-0.019	0.168	-0.020	1.415	0.239	0.412
	(0.249)	(0.551)	[0.962]	(1.510)	[0.374]	(0.772)	[0.930]
2	2.042	-0.043	0.176	-0.049	1.285	0.543	0.888
	(0.321)	(0.535)	[0.962]	(1.386)	[0.405]	(1.094)	[0.860]
2.5	2.539	-0.068	0.212	-0.088	1.159	0.943	1.574
	(0.410)	(0.520)	[0.957]	(1.256)	[0.433]	(1.603)	[0.780]
3	3.011	-0.093	0.271	-0.133	1.050	1.472	2.558
	(0.510)	(0.511)	[0.948]	(1.130)	[0.458]	(2.387)	[0.705]
Scenario 2: High dispersion in β ($\sigma_\beta = 1$)							
0.5	0.909	-0.017	0.654	-0.017	1.624	-0.093	0.597
	(0.977)	(0.850)	[0.655]	(1.689)	[0.304]	(0.775)	[0.681]
1	1.320	-0.022	0.573	-0.023	1.533	0.113	0.827
	(0.780)	(0.785)	[0.699]	(1.607)	[0.328]	(1.054)	[0.675]
1.5	1.745	-0.036	0.502	-0.039	1.433	0.383	1.289
	(0.626)	(0.723)	[0.740]	(1.511)	[0.352]	(1.486)	[0.655]
2	2.179	-0.054	0.455	-0.065	1.328	0.737	2.031
	(0.536)	(0.674)	[0.774]	(1.408)	[0.374]	(2.142)	[0.628]
2.5	2.608	-0.074	0.443	-0.099	1.229	1.205	3.137
	(0.524)	(0.642)	[0.798]	(1.302)	[0.394]	(3.125)	[0.597]
3	3.024	-0.093	0.468	-0.139	1.143	1.828	4.752
	(0.575)	(0.627)	[0.810]	(1.200)	[0.411]	(4.583)	[0.566]

Table III

Simulations with Non-Lognormal Returns

This table reports summary statistics of the fund-level abnormal return measures α, GPME, and PME, based on simulated data sets of N funds over T years, with heterogeneity in fund factor loadings. Panel A uses one large data set of 1,000,000 vintages with one fund per vintage (yielding one million total funds) to approximate asymptotic results, and Panel B uses 50,000 data sets of similar size as observed private equity fund data sets (30 years with 40 funds per vintage, for a total of 1,200 funds). Simulations are as described in Table I, except that there is some probability that idiosyncratic PE returns are substantially higher in any given year. In Scenario 1, there is a 10% probability that the fund's cash flow for the year is 50% higher. In Scenario 2, there is a 1% chance of a 100% higher payout. The mean and variance of the idiosyncratic return are chosen such that the idiosyncratic mixture distribution has the same mean and standard deviation as the lognormal distribution in Table I. All funds have same the beta, shown in the column labeled β. For each fund performance metric, Panel A reports its mean and standard deviation (in parentheses) across funds, its root mean squared error (*RMSE*) relative to true realized fund alphas, and the correlation (*Corr.*) between the metric and true realized alphas (in square brackets). Panel B reports the mean of these metrics across the 50,000 simulated data sets.

β	$\widehat{\beta}$	α		GPME		PME	
		Mean (St.Dev)	RMSE [Corr.]	Mean (St.Dev)	RMSE [Corr.]	Mean (St.Dev)	RMSE [Corr.]
Panel A: Large Data Set (N = 1,000,000; T = 1,000,000)							
Scenario 1: 10% probability of 50% higher cash flow							
0.5	0.503	-0.000	0.002	-0.000	2.865	-0.178	0.264
		(0.569)	[1.000]	(2.922)	[0.197]	(0.471)	[0.948]
1	0.998	0.002	0.001	0.002	1.732	0.001	0.000
		(0.570)	[1.000]	(1.826)	[0.316]	(0.570)	[1.000]
1.5	1.488	0.005	0.007	0.005	1.196	0.233	0.376
		(0.573)	[1.000]	(1.328)	[0.435]	(0.774)	[0.948]
2	1.972	0.009	0.014	0.009	0.856	0.535	0.932
		(0.575)	[1.000]	(1.033)	[0.560]	(1.158)	[0.820]
2.5	2.446	0.013	0.025	0.013	0.594	0.933	1.788
		(0.579)	[0.999]	(0.831)	[0.699]	(1.849)	[0.670]
3	2.897	0.017	0.044	0.017	0.372	1.458	3.172
		(0.582)	[0.998]	(0.691)	[0.844]	(3.076)	[0.528]
Scenario 2: 1% probability of 100% higher cash flow							
0.5	0.501	0.001	0.001	0.001	3.030	-0.178	0.267
		(0.584)	[1.000]	(3.083)	[0.186]	(0.480)	[0.949]
1	0.998	0.002	0.002	0.002	1.781	0.002	0.000
		(0.585)	[1.000]	(1.872)	[0.308]	(0.584)	[1.000]
1.5	1.489	0.006	0.007	0.006	1.207	0.234	0.382
		(0.587)	[1.000]	(1.340)	[0.435]	(0.795)	[0.948]
2	1.973	0.009	0.014	0.009	0.858	0.537	0.954
		(0.590)	[1.000]	(1.040)	[0.565]	(1.195)	[0.821]
2.5	2.447	0.013	0.025	0.013	0.595	0.936	1.849
		(0.593)	[0.999]	(0.839)	[0.705]	(1.926)	[0.669]
3	2.898	0.018	0.044	0.018	0.373	1.465	3.328
		(0.597)	[0.998]	(0.703)	[0.848]	(3.254)	[0.524]

Table III - Continued

Panel B: 50,000 Realistic Data Sets (N = 1,200; T = 30)							
Scenario 1: 10% probability of 50% higher cash flow							
0.5	0.501	0.006	0.058	0.006	1.651	-0.178	0.253
	(0.122)	(0.563)	[0.998]	(1.730)	[0.319]	(0.455)	[0.966]
1	1.002	0.000	0.056	0.000	1.544	0.000	0.000
	(0.134)	(0.559)	[0.998]	(1.636)	[0.349]	(0.561)	[1.000]
1.5	1.516	-0.018	0.068	-0.018	1.420	0.231	0.346
	(0.173)	(0.547)	[0.995]	(1.522)	[0.381]	(0.746)	[0.967]
2	2.036	-0.042	0.109	-0.048	1.287	0.533	0.826
	(0.263)	(0.534)	[0.988]	(1.396)	[0.412]	(1.055)	[0.892]
2.5	2.540	-0.068	0.171	-0.087	1.158	0.930	1.503
	(0.369)	(0.522)	[0.978]	(1.263)	[0.441]	(1.548)	[0.806]
3	3.014	-0.093	0.246	-0.132	1.047	1.455	2.469
	(0.479)	(0.513)	[0.966]	(1.135)	[0.466]	(2.310)	[0.726]
Scenario 2: 1% probability of 100% higher cash flow							
0.5	0.503	0.006	0.074	0.006	1.646	-0.178	0.252
	(0.175)	(0.553)	[0.997]	(1.720)	[0.313]	(0.448)	[0.965]
1	1.004	0.000	0.072	0.000	1.539	0.000	0.000
	(0.186)	(0.549)	[0.996]	(1.627)	[0.343]	(0.552)	[1.000]
1.5	1.519	-0.018	0.082	-0.018	1.415	0.231	0.344
	(0.227)	(0.538)	[0.992]	(1.513)	[0.375]	(0.735)	[0.967]
2	2.039	-0.042	0.117	-0.048	1.283	0.533	0.821
	(0.300)	(0.525)	[0.985]	(1.387)	[0.407]	(1.039)	[0.890]
2.5	2.541	-0.068	0.174	-0.087	1.154	0.930	1.492
	(0.396)	(0.513)	[0.975]	(1.256)	[0.436]	(1.524)	[0.802]
3	3.017	-0.093	0.247	-0.132	1.044	1.454	2.448
	(0.498)	(0.504)	[0.964]	(1.128)	[0.461]	(2.273)	[0.720]

Table IV

Descriptive Statistics

This table reports descriptive statistics of U.S. venture capital and buyout funds with vintages between 1978 and 2012 that have committed capital of at least \$5 million in 1990 dollars. Panel A shows statistics across all funds of a given asset class. *Funds* is the number of funds. *Percentage of funds liquidated* is the percentage of funds that are over 10 years old with a final NAV of zero (*100% liquidated*), or the percentage of funds that are over 10 years old with a most recently reported NAV of less than 5% of committed capital (*95% liquidated*), or less than 10% of committed capital (*90% liquidated*). *Firms* is the number of GPs in the sample, and *Funds per firm* is the number of funds raised by a GP. *Fund size* is the total commitment to a fund, in millions of dollars. *Fund effective years* is the time between the first and last observed cash flows of a fund, and *Cash flows per fund* is the number of cash flows observed for a fund (counting multiple cash flows on the same date as one cash flow). Cash flows data runs until the end of 2017. A fund's *IRR* is computed using the final observed NAV of the fund. *TVPI* is total value to paid-in capital, computed as the sum of cash distributions to LPs plus the final NAV divided by the sum of cash takedowns by the fund from LPs. The table also reports fund size weighted averages of the *IRR* and *TVPI*. Panel B reports *IRR* and *TVPI* by vintage year. *N* is the number of funds, and *Wtd* is the size weighted average of the performance measures. Source: Burgiss.

Panel A: Descriptive Statistics

	Venture capital funds			Buyout funds		
	Mean	Median	St.Dev.	Mean	Median	St.Dev.
Funds	1,219			879		
Percentage of funds liquidated:						
100% liquidated	45			34		
95% liquidated	54			49		
90% liquidated	58			55		
Firms	449			384		
Funds per firm	2.71	2.00	2.48	2.29	2.00	1.76
Fund size (\$m)	248.79	160.00	294.25	1,053.20	450.00	1,928.25
Fund effective years	13.40	13.37	4.22	12.10	11.90	3.90
Cash flows / fund	38.51	35.00	18.69	57.54	52.00	33.88
IRR (%)	14.10	7.47	40.64	13.88	13.11	15.28
Size-weighted	11.01		31.09	13.60		11.55
TVPI	2.03	1.42	2.79	1.74	1.60	0.87
Size-weighted	1.81		2.23	1.67		0.58

Table IV - Continued

Panel B: Performance Statistics by Vintage Year

	Venture capital funds							Buyout funds						
	N	IRR			TVPI			N	IRR			TVPI		
		Mean	Median	Wtd	Mean	Median	Wtd		Mean	Median	Wtd	Mean	Median	Wtd
≤ 1980	11	28.72	29.30	23.57	2.93	2.67	2.68							
≤ 1985								12	25.91	20.34	28.93	3.73	2.61	3.63
1981	10	8.84	10.44	8.77	1.75	1.78	1.66							
1982	12	5.13	6.39	5.15	1.46	1.57	1.47							
1983	20	8.44	8.96	11.14	1.75	1.69	2.00							
1984	26	7.75	7.73	8.19	1.67	1.57	1.67							
1985	24	6.51	8.73	8.85	1.99	1.73	2.06							
1986	22	11.71	11.36	15.84	2.19	1.94	3.19	6	14.90	13.00	17.98	3.19	1.83	4.53
1987	29	12.11	14.37	17.45	2.22	2.00	2.46	10	19.53	14.73	12.69	3.03	2.39	2.23
1988	27	12.48	11.75	21.32	2.12	1.92	2.70	9	11.88	11.17	15.87	1.85	1.66	2.10
1989	27	16.83	15.13	18.17	2.50	2.11	2.58	11	23.23	28.61	30.71	2.55	3.05	3.03
1990	16	22.24	20.88	28.03	2.80	2.23	3.29	7	18.26	16.26	16.08	2.37	2.36	2.10
1991	7	24.79	23.62	28.12	2.78	2.43	2.83	5	34.53	34.52	30.88	2.73	2.77	2.75
1992	19	26.19	15.15	28.20	3.18	1.79	2.85	10	25.00	23.19	35.89	1.97	1.91	2.16
1993	18	50.54	41.15	52.86	5.70	3.20	5.77	9	22.25	18.18	21.32	2.11	1.74	2.08
1994	18	44.37	36.39	60.87	5.77	4.07	9.37	19	17.15	15.44	22.59	1.74	1.47	1.92
1995	26	61.69	32.64	60.07	5.50	2.70	5.38	28	14.90	12.12	14.07	1.67	1.70	1.61
1996	16	106.65	64.05	95.67	7.17	3.68	7.17	17	11.07	9.32	15.59	1.53	1.36	1.69
1997	43	79.24	32.44	80.89	3.63	2.04	3.55	29	3.56	3.15	8.52	1.23	1.17	1.49
1998	57	26.98	3.07	26.47	1.89	1.10	1.94	48	5.79	8.27	3.52	1.45	1.42	1.31
1999	88	-4.61	-5.21	-5.56	0.87	0.76	0.90	34	6.26	9.05	7.16	1.43	1.52	1.46
2000	112	-4.34	-4.24	-2.02	0.93	0.81	1.00	48	14.03	13.45	16.87	1.76	1.60	1.82
2001	65	0.36	2.23	1.99	1.26	1.13	1.25	32	21.86	21.12	24.10	1.84	1.87	1.96
2002	24	0.58	1.88	2.03	1.13	1.11	1.13	24	19.09	18.44	19.73	1.87	1.77	1.93
2003	22	0.20	2.22	2.91	1.36	1.10	1.60	24	19.90	15.11	23.81	2.11	1.91	2.05
2004	44	1.57	-0.38	3.37	1.44	0.95	1.44	45	15.09	12.94	16.08	1.72	1.60	1.82
2005	62	4.41	4.43	5.58	1.57	1.28	1.75	66	11.04	10.37	10.82	1.70	1.59	1.69
2006	79	2.51	4.85	4.99	1.45	1.33	1.53	78	7.92	8.34	7.00	1.52	1.54	1.48
2007	74	12.09	11.39	13.51	1.97	1.66	2.03	76	11.55	12.04	11.56	1.66	1.63	1.62
2008	62	10.00	5.86	12.07	2.06	1.32	2.07	71	13.66	13.43	12.88	1.68	1.61	1.59
2009	28	16.62	14.95	17.34	2.21	1.88	2.09	28	16.22	18.90	18.86	1.80	1.85	1.92
2010	35	19.37	17.74	18.16	2.25	1.83	2.12	34	17.59	14.86	15.66	1.63	1.60	1.57
2011	37	16.73	14.91	18.26	2.20	1.58	2.27	47	19.14	17.98	19.18	1.77	1.61	1.82
2012	59	14.19	15.10	17.01	1.65	1.57	1.81	52	16.01	15.29	20.26	1.51	1.46	1.58
<1990s	208	11.50	10.17	14.96	2.07	1.77	2.41	48	19.96	16.59	18.49	2.89	2.29	2.70
1990s	308	34.47	10.76	22.82	3.02	1.46	2.38	206	11.05	9.35	10.16	1.60	1.47	1.54
2000s	703	6.00	5.50	7.84	1.58	1.31	1.64	625	14.34	13.45	14.12	1.70	1.61	1.67

Table V

Performance Estimates

This table reports statistics of the abnormal return measures α, GPME, and the PME across venture capital funds (in the first three columns) and across buyout funds (in the last three columns). The α estimates use a benchmark portfolio, whose return for a period from fund inception to h years later is equation (16) in the main text, reproduced here for reference:

$$R_h^b = \exp\left\{ r_h^f + \beta(r_h^m - r_h^f) - \frac{1}{2}\beta(\beta-1)\sigma_h^2 \right\}.$$

The table reports the estimates of the common fund beta, β.
The GPME performance measure is computed using an exponential affine stochastic discount factor (SDF)

$$M_h = \exp(\delta h - \gamma r_h^m).$$

The table reports the SDF parameter estimates of δ and γ.
PME is a restricted version of the alpha metric that assumes $\beta = 1$ regardless of the true beta, or equivalently, a restricted version of GPME with $\delta = 0$ and $\gamma = 1$. The table shows the p-value for the test that the mean (G)PME is equal to zero. Standard errors are in parentheses.

	Venture capital (N=1,219)			Buyout (N=879)		
	α	GPME	PME	α	GPME	PME
Mean	-0.207	-0.207	0.129	0.207	0.207	0.148
Percentiles:						
10th	-1.003	-1.223	-0.692	-0.296	-0.816	-0.339
25th	-0.686	-0.751	-0.437	-0.056	-0.311	-0.116
Median	-0.351	-0.340	-0.158	0.152	-0.080	0.087
75th	-0.066	-0.003	0.208	0.400	0.416	0.349
90th	0.508	0.748	0.865	0.786	1.465	0.689
St.Dev.	1.103	1.387	1.387	0.468	1.401	0.457
Skewness	5.473	5.818	5.796	0.981	4.099	1.026
Kurtosis	51.059	65.946	48.880	6.233	41.362	6.300
p-value		0.132	0.000		0.205	0.000
β	2.434			0.790		
SDF Parameters						
δ		0.162	0		0.185	0
		(0.055)			(0.062)	
γ		3.456	1		3.586	1
		(0.641)			(0.617)	

Table VI

Performance across Factor Models

This table reports statistics of the abormal return measures α and GPME across factor models specifications. Panel A reports results for venture capital funds, and Panel B for buyout funds. Model (1) reproduces the one-factor results from Table V, for reference. Models (2) through (5) report results for various two-factor models. In addition to the excess log market return, these models include the log return of one of the corner portfolios of the Fama-French size and book-to-market portfolios in excess of the log market return: small-growth (SL-M), small-value (SH-M), large-growth (BL-M) and large-value (BH-M). The GPME performance measure is thus computed using an exponential affine stochastic discount factor (SDF)

$$M_h = \exp(\delta h - \gamma_m r_h^m - \gamma_x r_h^x),$$

where x is the second factor (SL-M, SH-M, BL-M, or BH-M). The benchmark portfolio has two beta loadings, on the excess market return (β_m) and on the return to the factor x (β_x). For the α metrics, the table also reports the value of the objective function Q from eqn (17) at estimated parameters, where lower values indicate a better model fit. For GPME, the table show the p-value for the test that the mean GPME is equal to zero. Standard errors are in parentheses.

	α					GPME				
	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)
x =	No x	SL-M	SH-M	BL-M	BH-M	No x	SL-M	SH-M	BL-M	BH-M
Panel A: Venture Capital Funds										
Mean	-0.207	-0.210	0.385	-0.190	-0.206	-0.207	-0.210	0.385	-0.190	-0.206
Percentiles:										
10th	-1.003	-1.003	-0.685	-0.962	-0.995	-1.223	-1.229	-0.657	-1.278	-1.233
Median	-0.351	-0.353	0.031	-0.338	-0.347	-0.340	-0.341	-0.084	-0.327	-0.340
90th	0.508	0.505	1.466	0.530	0.511	0.748	0.737	1.277	0.809	0.752
St.Dev.	1.103	1.104	1.700	1.085	1.100	1.387	1.383	2.182	1.466	1.396
p-value						0.132	0.118	0.000	0.206	0.140
β_m	2.434	2.425	0.753	2.358	2.391					
β_x		0.081	-0.554	1.010	-0.247					
δ						0.162	0.162	0.061	0.172	0.163
						(0.055)	(0.055)	(0.347)	(0.046)	(0.044)
γ_m						3.456	3.470	1.088	3.539	3.475
						(0.641)	(0.624)	(2.923)	(0.595)	(0.541)
γ_x							-0.108	1.863	-0.839	0.064
							(1.327)	(2.294)	(2.761)	(1.157)

Table VI - Continued

	α					GPME				
	(1)	(2)	(3)	(5)	(6)	(1)	(2)	(3)	(4)	(5)
x =	No x	SL-M	SH-M	BL-M	BH-M	No x	SL-M	SH-M	BL-M	BH-M
					Panel B: Buyout Funds					
Mean	0.207	0.213	0.345	0.229	0.221	0.207	0.213	0.345	0.229	0.221
Percentiles:										
10th	-0.296	-0.295	-0.243	-0.292	-0.297	-0.816	-0.758	-1.160	-0.716	-0.649
Median	0.152	0.162	0.292	0.188	0.181	-0.080	-0.056	-0.037	-0.127	-0.089
90th	0.786	0.792	0.944	0.826	0.807	1.465	1.385	2.867	1.523	1.299
St.Dev.	0.468	0.467	0.515	0.471	0.467	1.401	1.346	3.022	1.487	1.418
p-value						0.205	0.192	0.070	0.217	0.200
β_m	0.790	0.753	0.351	0.753	0.753					
β_x		-0.267	0.097	-0.622	0.382					
δ						0.185	0.176	0.450	0.160	0.149
						(0.062)	(0.058)	(0.085)	(0.044)	(0.042)
γ_m						3.586	3.433	5.047	3.404	3.145
						(0.617)	(0.560)	(0.666)	(0.522)	(0.479)
γ_x							0.759	3.102	2.887	-1.660
							(0.983)	(0.534)	(3.078)	(1.314)

Table VII

Performance Estimates for Subsamples

This table reports statistics of the fund-level abnormal return measures α, GPME, and the PME for subsamples of the Burgiss dataset, using the market return as the sole risk factor. Panel A shows performance estimates for the subsample of funds that were fully liquidated by the end of the sample period, in 2015. These are funds that are over 10 years old with a final reported net asset value equal to zero. Panel B uses funds raised before the year 2000, while Panel C uses the funds raised in 2000 or later. Panels D and E consider funds with committed capital below or above the median size of funds raised in the same vintage year. The performance measures and reported results are as defined in Table V.

	Venture capital			Buyout		
	α	GPME	PME	α	GPME	PME
Panel A: Fully Liquidated Funds						
Mean	0.041	-0.048	0.312	0.210	0.210	0.194
Percentiles:						
10th	-0.811	-0.827	-0.694	-0.373	-0.461	-0.377
Median	-0.247	-0.296	-0.159	0.144	0.024	0.128
90th	1.095	0.884	1.511	0.928	1.129	0.897
St.Dev.	1.350	1.245	1.831	0.528	0.821	0.525
No. of funds	547	547	547	299	299	299
p-value		0.798	0.000		0.087	0.000
β	2.520			0.894		
Panel B: All Funds of Pre-2000 Vintage						
Mean	0.135	0.069	0.393	0.142	0.142	0.157
Percentiles:						
10th	-0.722	-0.733	-0.649	-0.371	-0.417	-0.368
Median	-0.206	-0.238	-0.105	0.070	0.014	0.084
90th	1.104	0.966	1.660	0.746	1.013	0.804
St.Dev.	1.391	1.333	1.880	0.462	0.593	0.469
No. of funds	516	516	516	254	254	254
p-value		0.736	0.000		0.160	0.000
β	2.398			1.131		

Table VII - Continued

	Venture capital			Buyout		
	α	GPME	PME	α	GPME	PME
Panel C: All Funds of Post-2000 Vintage						
Mean	-0.421	-0.421	-0.064	0.260	0.260	0.145
Percentiles:						
10th	-1.035	-1.848	-0.719	-0.215	-1.087	-0.304
Median	-0.422	-0.420	-0.186	0.198	-0.125	0.088
90th	0.018	0.596	0.445	0.796	1.582	0.644
St.Dev.	0.700	1.648	0.812	0.465	2.151	0.452
No. of funds	703	703	703	625	625	625
p-value		0.008	0.000		0.237	0.000
β	2.180			0.662		
Panel D: Funds below Median Size						
Mean	-0.311	-0.311	0.045	0.251	0.251	0.132
Percentiles:						
10th	-1.171	-1.469	-0.728	-0.379	-0.929	-0.422
Median	-0.415	-0.395	-0.236	0.214	-0.098	0.081
90th	0.546	0.760	0.843	0.935	1.761	0.755
St.Dev.	1.023	1.273	1.266	0.539	1.753	0.513
No. of funds	589	589	589	432	432	432
p-value		0.023	0.000		0.211	0.000
β	2.686			0.598		
Panel E: Funds above Median Size						
Mean	-0.118	-0.118	0.208	0.168	0.168	0.164
Percentiles:						
10th	-0.901	-1.101	-0.634	-0.214	-0.695	-0.215
Median	-0.290	-0.299	-0.082	0.096	-0.074	0.091
90th	0.510	0.748	0.931	0.645	1.186	0.642
St.Dev.	1.179	1.461	1.489	0.396	1.025	0.395
No. of funds	630	630	630	447	447	447
p-value		0.407	0.000		0.216	0.000
β	2.268			0.986		

Table VIII

Performance Estimates for Venture Capital Subclasses

This table reports statistics of the fund-level abnormal return measures α, GPME, and the PME for subclasses of venture capital funds in the Burgiss dataset. The performance measures and reported results are as defined in Table V.

	Generalist VCs (N = 302)			Early-Stage VCs (N = 745)			Late-Stage VCs (N = 120)		
	α	GPME	PME	α	GPME	PME	α	GPME	PME
Mean	-0.269	-0.269	-0.003	-0.216	-0.216	0.206	0.062	0.062	0.182
Percentiles:									
10th	-0.875	-1.252	-0.653	-1.156	-1.339	-0.729	-0.547	-0.746	-0.412
Median	-0.331	-0.339	-0.130	-0.390	-0.354	-0.204	-0.055	-0.189	0.071
90th	0.369	0.761	0.586	0.691	0.662	1.197	0.760	1.258	0.863
St.Dev.	0.530	1.006	0.624	1.362	1.546	1.682	0.833	1.037	0.893
p-value		0.038	0.610		0.141	0.000		0.662	0.000
β	2.159			2.706			1.467		

Table IX

Performance and Fund Characteristics

This table reports regression results using the α, GPME, or PME performance metric as the dependent variable.The market return is the sole risk factor in the SDF. Panel A reports results for venture capital funds, and Panel B for buyout funds. Size quartiles are defined by fund type (venture capital or buyout) and decade (pre-1990, 1990s, and 2000 onwards) based on fund sizes measured in 1990 dollars. The lowest size quartile is the omitted category. Some specifications include the natural logarithm of a fund's sequence number, *log(Sequence)*. All regressions include vintage year fixed effects. Heteroscedasticity-consistent standard errors are in parentheses and p-values are in square brackets.

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable	α	α	GPME	GPME	PME	PME
Panel A: Venture Capital Funds						
Size quartile 2	0.008	-0.030	-0.037	-0.081	0.108	0.058
	(0.073)	(0.075)	(0.090)	(0.092)	(0.098)	(0.101)
	[0.908]	[0.690]	[0.678]	[0.379]	[0.272]	[0.565]
Size quartile 3	0.031	-0.064	-0.004	-0.113	0.122	-0.002
	(0.071)	(0.077)	(0.087)	(0.096)	(0.088)	(0.093)
	[0.665]	[0.405]	[0.960]	[0.243]	[0.165]	[0.982]
Size quartile 4	0.198	0.048	0.286	0.116	0.229	0.035
(largest)	(0.085)	(0.085)	(0.122)	(0.117)	(0.093)	(0.102)
	[0.020]	[0.569]	[0.019]	[0.319]	[0.014]	[0.734]
log(Sequence)		0.165		0.188		0.215
		(0.048)		(0.064)		(0.059)
		[0.001]		[0.004]		[0.000]
Vintage fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R^2	0.161	0.170	0.107	0.113	0.161	0.170
Number of funds	1,219	1,219	1,219	1,219	1,219	1,219
Panel B: Buyout Funds						
Size quartile 2	0.070	0.068	0.067	0.056	0.070	0.069
	(0.048)	(0.049)	(0.128)	(0.134)	(0.047)	(0.047)
	[0.148]	[0.164]	[0.599]	[0.678]	[0.131]	[0.140]
Size quartile 3	0.050	0.045	-0.005	-0.031	0.051	0.048
	(0.046)	(0.047)	(0.121)	(0.137)	(0.044)	(0.046)
	[0.272]	[0.338]	[0.969]	[0.819]	[0.253]	[0.297]
Size quartile 4	0.028	0.018	-0.040	-0.089	0.030	0.025
(largest)	(0.045)	(0.051)	(0.139)	(0.175)	(0.044)	(0.050)
	[0.538]	[0.719]	[0.772]	[0.610]	[0.498]	[0.623]
log(Sequence)		0.011		0.058		0.007
		(0.026)		(0.076)		(0.025)
		[0.668]		[0.443]		[0.797]
Vintage fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R^2	0.059	0.058	0.342	0.342	0.079	0.078
Number of funds	879	879	879	879	879	879

Table X

Performance Persistence

This table reports the performance of future funds of managers (i.e., General Partners) with a fund in the top or bottom quartile in a given vintage year. The first 3 columns report results from ranking funds in each vintage from 1985 until 2010 by the alpha metric (with the market return as the sole risk factor in the SDF), selecting the top and bottom quartile funds in each vintage, and computing the average performance (alpha, GPME, and PME) of future funds of the same managers. The first row only considers the performance of the next fund by the same manager that is raised at least 10 years after the current fund (such that it does not overlap in time with the current fund). The second row allows for funds to overlap. The third and fourth rows report performance of all future funds by the manager, either nonoverlapping (third row) or overlapping (fourth row). The middle and right sets of three columns repeat the exercise but using GPME and PME as the metric to sort funds. Panel A reports results for venture capital funds, and Panel B for buyout funds.

Current funds sorted on:	α			GPME			PME		
Future fund(s) metric:	α	GPME	PME	α	GPME	PME	α	GPME	PME
Panel A: Venture Capital									
Managers with Top Quartile Funds:									
Next fund, nonoverlapping	0.57	0.86	1.01	0.66	0.91	1.17	0.41	0.68	0.81
Next fund, overlapping	0.66	0.64	1.23	0.64	0.57	1.30	0.69	0.67	1.27
All future funds, nonoverlapping	0.35	0.57	0.82	0.41	0.61	0.91	0.21	0.42	0.64
All future funds, overlapping	0.18	0.30	0.62	0.18	0.29	0.63	0.16	0.28	0.59
Managers with Bottom Quartile Funds:									
Next fund, nonoverlapping	-0.50	-0.66	-0.07	-0.35	-0.43	0.01	-0.44	-0.61	-0.03
Next fund, overlapping	-0.38	-0.38	-0.13	-0.33	-0.33	-0.06	-0.45	-0.43	-0.28
All future funds, nonoverlapping	-0.50	-0.64	-0.07	-0.36	-0.43	-0.00	-0.47	-0.62	-0.06
All future funds, overlapping	-0.41	-0.43	-0.13	-0.37	-0.38	-0.08	-0.45	-0.46	-0.21
Panel B: Buyout									
Managers with Top Quartile Funds:									
Next fund, nonoverlapping	0.23	0.58	0.18	0.17	0.20	0.11	0.24	0.59	0.18
Next fund, overlapping	0.28	0.26	0.22	0.25	0.15	0.19	0.28	0.26	0.22
All future funds, nonoverlapping	0.23	0.41	0.17	0.18	0.10	0.12	0.24	0.41	0.18
All future funds, overlapping	0.22	0.22	0.15	0.22	0.15	0.15	0.22	0.22	0.16
Managers with Bottom Quartile Funds:									
Next fund, nonoverlapping	0.23	0.05	0.16	0.44	0.68	0.38	0.20	0.05	0.14
Next fund, overlapping	0.07	0.15	0.03	0.09	0.20	0.05	0.06	0.15	0.03
All future funds, nonoverlapping	0.25	0.02	0.18	0.37	0.39	0.30	0.22	0.02	0.16
All future funds, overlapping	0.02	-0.07	-0.03	0.09	0.03	0.04	0.02	-0.07	-0.03