Risk-Adjusting the Returns to Venture Capital*

Arthur Korteweg[†]
University of Southern California

Stefan Nagel[‡]
University of Michigan, NBER and CEPR

February 2015

ABSTRACT

We adapt stochastic discount factor (SDF) valuation methods for venture capital (VC) performance evaluation. Our approach generalizes the popular Public Market Equivalent (PME) method and it allows statistical inference in the presence of cross-sectionally dependent, skewed VC payoffs. We relax SDF restrictions implicit in the PME so that the SDF can accurately reflect risk-free rates and returns of public equity markets during the sample period. This generalized PME yields substantially different abnormal performance estimates for VC funds and start-up investments, especially in times of strongly rising public equity markets and for investments with betas far from one.

*We thank Susan Woodward at Sand Hill Econometrics, and Mark Broadie, Michael Chernov, and Michael Johannes for providing access to data, and John Cochrane, Michael Ewens, Bob Hodrick, Steve Kaplan, Ludovic Phalippou, Berk Sensoy, Morten Sørensen, Michael Roberts (the editor), an associate editor, three anonymous referees, and participants of the 2013 NBER Entrepeneurship meeting, Carnegie-Mellon University, Stanford finance brown bag, and the 2014 Western Finance Association Meetings for valuable comments and suggestions.

†University of Southern California, Marshall School of Business, 3670 Trousdale Pkwy., Los Angeles, CA 90089, email: korteweg@marshall.usc.edu

‡Ross School of Business and Department of Economics, University of Michigan, 701 Tappan St., Ann Arbor, MI 48019, e-mail: stenagel@umich.edu

Venture capital (VC) investments play an important role in supporting the formation of young enterprises in the economy. Assessing the risk and return from this type of investments has thus far proven difficult for a number of reasons: payoffs are infrequent, realized over multiple periods with varying time horizons, highly skewed, and cross-sectionally dependent. Linear factor model techniques that are common in the mutual funds and hedge funds literature are not easily applicable unless one is willing to impose strong assumptions on the return-generating process as, for example, in Cochrane (2005) and Korteweg and Sorensen (2010). Much of the literature therefore relies on heuristic performance metrics such as the Public Markets Equivalent (PME) measure of Kaplan and Schoar (2005).

While the PME method has many desirable features that make it suitable for irregularly-spaced and skewed private-equity cash flow data, the method still suffers from two deficiencies compared with standard methods in mutual fund and hedge fund performance evaluation. The first is that the PME does not correctly adjust for the extent to which high-beta assets mechanically outperform a public market index in times of rising public equity markets. The second is that so far no methods exist to conduct statistical inference in a PME-style performance evaluation. Our objective in this paper is to address both challenges and to analyze empirically how much these methodological improvements alter performance evaluation results.

To address the first challenge, we propose a generalization of the PME that we label the *Generalized Public Market Equivalent* (GPME). The GPME can account for the mechanical relationships between VC payoffs and contemporaneous risk factor realizations—just like a linear factor model would. At the same time, it retains the desirable feature of the PME method that it can accommodate irregularly-spaced, skewed payoffs without relying on strong distributional assumptions. In our baseline approach we value VC payoffs with a stochastic discount factor (SDF) of the form

$$M_{t+1} = \exp(a - b r_{m,t+1}), \tag{1}$$

where $r_{m,t+1}$ is the log return on the public equity market portfolio and a and b are parameters. We also consider extensions to multiple factors.

To fix ideas, consider the simplified case of a one-dollar investment in start-up company i at time t that will lead to a single uncertain payoff, or gross return, R_{t+1} one period later. Under the null hypothesis of no abnormal performance, the present value of a one-dollar investment in any asset should equal one dollar, which means

$$1 = E[M_{t+1} \cdot R_{t+1}]. \tag{2}$$

We choose a and b in (1) so that (2) holds for the public equity market portfolio return and the risk-free asset return. We then apply the SDF to VC payoffs with (2) as the null hypothesis. Our approach nests the PME of Kaplan and Schoar (2005) as the special case of (1) with the parameter restrictions $a = 0$ and $b = 1$ so that $M_{t+1} = 1/R_{m,t+1}$. The problem with imposing this restriction is that it prevents the SDF from properly adjusting for market risk exposure.

To see how the GPME method works and how it relates to the PME, it is useful to consider the special case in which R_{t+1} and $R_{m,t+1} = \exp(r_{m,t+1})$ are jointly log-normal. In this case, the GPME method is equivalent to performance evaluation with the CAPM in a conventional (log-)linear beta-formulation. Taking logs of (2), with a and b chosen to match the public equity market return and risk-free asset returns, and using the properties of the log-normal distribution,[1] we get

$$\log E[R_{t+1}] = r_f + \beta(\log E[R_{m,t+1}] - r_f), \tag{3}$$

where r_f is the log risk-free rate and $\beta \equiv \mathrm{Cov}\,(r_{t+1}, r_{m,t+1})/\sigma_m^2$, and $\sigma_m^2 \equiv \mathrm{Var}\,(r_{m,t+1})$. Although it is not directly apparent from (2), the expression for the log-normal case in (3)

[1]See Online Appendix A for a detailed derivation. While the log-normality assumption is useful for illustration, there is no need to rely on this assumption when applying the valuation model (1) and (2). With irregularly-spaced and skewed cash flow data from VC investments, it is not straightforward to apply the log-normal model (3) and the strong distributional assumptions are questionable.

makes it transparent that our method accounts for an asset's beta exposure to the market factor just like standard factor models in the mutual funds and hedge funds literatures do.

To illustrate the assumptions implicit in the PME, we follow the same steps as above, but now imposing the restrictions $a = 0$ and $b = 1$. We get

$$\log E[R_{t+1}] = \log E[R_{m,t+1}] - \sigma_m^2 + \beta\sigma_m^2. \quad (4)$$

Thus, comparing (4) and (3), we see that the PME implicitly restricts the equity premium to σ_m^2 and the log risk-free rate to $\log E[R_{m,t+1}] - \sigma_m^2$.[2] These restrictions are inconsequential for assets with $\beta = 1$ because the errors from imposing the wrong risk-free rate and market risk premium cancel out in this case: the right-hand sides of both (4) and (3) become $\log E[R_{m,t+1}]$ in this special case. But these restrictions do matter if $\beta \neq 1$. The magnitude of the bias depends on how much the PME's implied restrictions are violated in the data that a researcher analyzes. For example, since VC investments typically have $\beta > 1$, they mechanically outperform the market index in times of a rising equity market.[3] By fixing the equity premium at $\log E[R_{m,t+1}] - r_f = \sigma_m^2$, the PME does not take this into account, which results in an overstatement of abnormal performance in times of strongly rising public equity markets. Furthermore, if payoffs are not jointly log-normal with $R_{m,t+1}$, then even $\beta = 1$ no longer guarantees that the performance assessment with the PME is accurate.

These violations can be substantial. For example, many VC studies use data sets concentrated in the 1990s. From the beginning of 1990 to the end of 1999, the annualized log return from rolling over one-month T-bills was 4.81%, the log average return on the CRSP value-

[2]The fact that the PME restricts the equity premium and risk-free rate can also be seen by recognizing that (2) is equivalent to valuation with the SDF from the log-utility CAPM, where $R_{m,t+1}$ is the return on aggregate wealth. In the log-utility model, relative risk aversion is restricted to a value of one, which restricts the equity premium. Under joint log-normality, this is the restriction $\log E[R_{m,t+1}] - r_f = \sigma_m^2$ that is built into (4). See also independent and contemporaneous work by Sorensen and Jagannathan (2013) that points out the log-utility assumption behind the PME.

[3]Recent β estimates in the literature include Gompers and Lerner (1997) who find values from 1.1 to 1.4, while Peng (2001) finds 1.3 to 2.4, Cochrane (2005) finds 1.9, Woodward (2009) finds 2.2, Korteweg and Sorensen (2010) find 2.8, Driessen, Lin, and Phalippou (2012) find 2.7, and Ewens, Jones, and Rhodes-Kropf (2013) find 1.2.

weighted index was 15.71%, while its annualized variance was 1.86%. As a consequence, an asset with $\beta = 2$ would have a benchmarked return of $4.81\% + 2 \times (15.71\% - 4.81\%) = 26.61\%$ according to the standard factor model in (3), while the PME-implied benchmarked return following (4) is $15.71\% + (2-1) \times 1.86\% = 17.57\%$, a difference of almost 10 percentage points. In such samples, the PME could therefore lead to the conclusion that VC payoffs were abnormally high when, in fact, the performance was just in line with properly risk-matched public markets benchmarks.

To implement the GPME, we need to estimate the SDF parameters. To ensure that the SDF valuation properly benchmarks against contemporaneous stock market performance, a and b should be chosen to correctly price stock market index and T-bill returns during the sample period from which VC payoffs are drawn. Under log-normality, this is equivalent to evaluating performance with the sample version of (3). Thus, this approach is analogous to the standard method of estimating factor risk premia in linear factor models with the average excess return of the factor during the performance evaluation period.

The second challenge that we aim to address in this paper is statistical inference. Reported PMEs in the literature are typically simple point estimates without standard errors. Statistical inference is difficult, because changes in start-up company valuations between funding rounds, or the cash flows of VC funds over the fund life, are overlapping in time to varying degrees and are likely subject to similar common factor shocks that induce positive cross-sectional correlation. To conduct statistical inference, we implement the GPME estimation within a GMM framework. To derive test statistics that are robust to cross-correlation between VC payoffs, we borrow methods from the spatial GMM literature. In applying these methods, we treat the degree of time-overlap between a pair of payoff series as analogous to a spatial distance that determines the level of correlation.

We apply our method to a data set of VC funds from 1979 to 2012, and a data set of start-up company financing rounds from 1987 to 2005. The VC fund payoffs are measured net of fees and carried interest, whereas the payoffs in the start-up company data reflect the returns from VC investments before fees and carried interest are paid to the fund managers. In our baseline estimation, we use the public equity market return as the risk factor in (1). In line with our discussion above, the GPME differs substantially from the PME during sample periods such as the 1980s and 1990s, in which the public equity markets earned high realized returns. For example, for VC funds of pre-1998 vintages, the PME overstates the fund NPV by roughly 21 percentage points. For investments in pre-1998 start-up funding rounds, the PME overstates the NPV by about 42 percentage points. In contrast, in sample periods in which the parameter restrictions implicit in the PME method are, coincidentally, more or less in line with the data on realized equity market returns and risk-free rates, the discrepancies between PME and GPME are much smaller.

We further illustrate this crucial difference between the GPME and PME by analyzing artificially levered investments in VC funds and start-up rounds. Artificial leverage creates high-beta assets. With proper risk-adjustment, levering up an investment by a factor k should raise the NPV by the same factor k. We show that the GPME satisfies this requirement. The PME, however, does not. The NPV estimates delivered by the PME method increase by more than k. This is a reflection of the fact that the PME does not correctly price the risk-free asset during the sample period. Levering up the VC payoffs builds this pricing error into the abnormal performance measured by the PME.

Thus, depending on the public markets returns during the sample period and the degree of systematic risk of the assets, performance metrics based on the PME can be severely distorted. However, in the full samples, the differences between GPME and PME are relatively small, because the assumptions underlying the PME are not grossly violated by the data. Our full-sample results are broadly in line with earlier literature: Based on the GPME we find slightly negative (equally-weighted) abnormal returns of VC funds and strongly positive abnormal returns of VC investments in start-up companies.[4]

[4]For comparison, Ljungqvist and Richardson (2003), Kaplan and Schoar (2005), Woodward (2009), Robinson and Sensoy (2011), Stucke (2011), and Ewens, Jones, and Rhodes-Kropf (2013) report VC fund returns that are equal to or slightly above the return on the market portfolio. Conversely, Phalippou and Zollo (2005), Phalippou and Gottschalg (2009), and Driessen, Lin, and Phalippou (2012) find below-market returns. Har-

In interpreting our results, one must keep in mind that venture capital data is usually subject to some degree of selection bias, and this paper is no exception. Our method can address some of the existing empirical challenges—such as adjustment for systematic risk and tools for statistical inference—but it does not remove the residual selection biases in the data set.

Regarding the economic interpretation of an SDF like (1), the standard asset-pricing perspective would be to view this SDF as a CAPM that arises within a representative-agent IID economy under the assumption of constant relative risk aversion (Giovannini and Weil 1989). However, in a performance measurement application like ours, the objective is not to test an equilibrium asset pricing model but to ask whether an available investment opportunity would, at the margin, improve the risk-return tradeoff of a particular investor who is not necessarily a representative agent. Following this interpretation, the SDF valuation equation (2) represents the first-order condition of an investor whose wealth is invested in the public equity market. This interpretation is analogous to the typical interpretation of performance measures in the mutual funds and hedge funds literature.

One could augment the SDF with additional risk factors to capture other assets that the investor may have access to and state-variable risks that he or she may care about. We explore this route by adding the return on a portfolio of small growth stocks to the SDF in (1) to allow for a comparison of VC payoffs with the returns of publicly traded companies that share some of the characteristics of start-up companies. Augmenting the SDF with this factor raises the abnormal returns of VC funds and start-up company investments in some samples in which small growth stocks in public markets performed relatively poorly. One could add additional factors such as those in Longstaff (2009), Franzoni, Nowak, and Phalippou (2012), and Sorensen, Wang, and Yang (2014).

Our method can be applied to other infrequently traded asset classes with highly levered

ris, Jenkinson, and Kaplan (2014), using what is probably the most comprehensive data set on VC funds currently available, report that VC funds outperformed public equities in the 1990s but underperformed in the 2000s. Cochrane (2005), Hall and Woodward (2007), and Korteweg and Sorensen (2010) find positive abnormal round-to-round returns (gross of fees) from VC investments in start-up companies.

or option-like returns, such as leveraged buyouts or real estate. Our approach also has implications for the risk and reward to entrepreneurs (Moskowitz and Vissing-Jorgensen 2002), and the performance of different types of limited partners (Lerner, Schoar, and Wongsunwai 2007) or general partners (Ewens and Rhodes-Kropf 2013) in private equity investments.

The paper is organized as follows. Section I outlines our SDF pricing approach to risk-adjusting VC returns. Section II presents our VC fund results. Section III presents our start-up company results, and Section IV concludes.

I. Risk-Adjusting Venture Capital Returns

We start by describing our approach in more detail. Throughout the paper, we use lower case letters for logs. For example R_t is an arithmetic return, and r_t is a log return (i.e., $r_t = \log(R_t)$). The time-t value, V_t, of an asset that pays a single cash flow at time $t+1$ is the expected value of the cash flow, C_{t+1}, discounted with an SDF, M_{t+1},

$$V_t = E_t[M_{t+1} \cdot C_{t+1}]. \tag{5}$$

If the asset has a continuation value, V_{t+1}, at time $t+1$, then

$$V_t = E_t\left[M_{t+1} \cdot (C_{t+1} + V_{t+1})\right]. \tag{6}$$

Dividing both sides of the above equation by V_t yields the pricing relation

$$1 = E_t[M_{t+1} \cdot R_{t+1}], \tag{7}$$

where $R_{t+1} \equiv \left(C_{t+1} + V_{t+1}\right)/V_t$ is a gross return.

Our specification of the single-period SDF is exponentially-affine,

$$M_{t+1} = \exp(a - bf_{t+1}). \tag{8}$$

As we discussed in the introduction, in the log-normal special case this specification implies the standard CAPM beta-pricing relationship that is commonly used as a basic performance benchmark in the mutual funds and hedge funds literature. However, the CAPM specification (8) is also valid more generally when asset payoffs are not log-normal. Additional risk factors, such as those of Fama and French (1993), can be added easily, but to keep notation simple, we focus on the single-factor case to lay out our approach. In a performance-evaluation application like ours, the relevant interpretation of an SDF like (8) is not as an equilibrium asset pricing model, but rather as the reflection of the return on wealth and the state-variable risks of an investor who is evaluating whether the performance of the assets in question would add value to his or her wealth portfolio.

One particularly pressing problem with payoffs on non-public equity—and illiquid asset classes more generally—is the multi-period nature of the payoffs. Cash flows of VC funds occur at irregularly spaced points in time throughout the life of the fund. The timing of funding rounds and exit events of VC-backed start-up companies (times at which valuation estimates are available) is similarly irregular. The exponentially-affine SDF setup is particularly well-suited for multi-period payoffs measured over varying time horizons, because the pricing relation holds for longer horizons

$$V_t = E_t[M^h_{t+h} \cdot C_{t+h}], \tag{9}$$

where M^h_{t+h}, the multi-period SDF from t to $t+h$, simply compounds the single-period discount factors

$$\begin{aligned} M^h_{t+h} &\equiv \prod_{i=1}^{h} M_{t+i} \qquad (10) \\ &= \exp(ah - bf^h_{t+h}), \qquad (11) \end{aligned}$$

where

$$f^h_{t+h} \equiv \sum_{i=1}^{h} f_{t+i}. \tag{12}$$

A common approach in the VC and private equity literature is to work with linearized versions of the SDF, which imply a linear beta-pricing relationship (see, for example, Ljungqvist and Richardson (2003), Hall and Woodward (2007), Driessen, Lin, and Phalippou (2012), Ewens, Jones, and Rhodes-Kropf (2013)). Our approach avoids this linearization. In Online Appendix B we show that linearizing the model leads to problems with multi-period discounting. Moreover, it can lead to specification errors when applied to highly non-linear, option-like payoffs.

Applying the SDF valuation to a whole series of cash flows from a fund or company i, including all cash distributions to investors and capital takedowns into the fund, we obtain the realized GPME as

$$GPME_i \equiv \sum_{j=1}^{J} M^{h(j)}_{t+h(j)} \cdot C_{i,t+h(j)}, \tag{13}$$

where t is the date of the first cash flow, and $C_{i,t+h(j)}$ is the net cash flow (distributions minus takedowns) for fund i at date $t+h(j)$. The number of cash flows, J, and the initial cash flow date, t, vary by fund, but we suppress dependence on i for notational simplicity. Since the cash flow series includes the initial investment, the valuation equation (9) implies, under the null hypothesis, that $E[GPME_i] = 0$, i.e., the NPV is zero.

Our approach nests the PME of Kaplan and Schoar Kaplan and Schoar (2005) as the special case with $a = 0$ and $b = 1$ in (8). As discussed in the introduction, fixing a and b in this way implicitly restricts the equity premium and risk-free rate, which leads to the undesirable feature that the performance benchmark does not properly reflect the mechanical dependence of asset payoffs on the public market return. In our generalized method, a and b are chosen so that the SDF perfectly prices the stock market and risk-free asset payoffs during the sample period in direct analogy to linear factor models where the average realized excess return on a factor during the sample period pins down the estimate of the factor's risk premium. In this way, the GPME allows the SDF to properly account for the returns available in public markets during the sample period. At the same time, because a and b are pinned down by moments of public market returns and are not chosen to fit VC payoffs, the GPME does not offer more degrees of freedom than the PME in explaining VC payoffs.

The PME measure is traditionally defined as the sum of cash distributions discounted at the realized public market return, divided by the sum of capital takedowns, while we work with the *difference* between discounted values of inflows and outflows for each fund, rather than their *ratio*. Asset pricing theory implies that the expected difference (the NPV) is zero, but when inflows are stochastic there are no clear predictions about the expected ratio because of a Jensen's inequality effect.[5] Working with the difference also simplifies the econometrics. From here on, when we use the label PME, we refer to this redefined PME.

If we find abnormal performance with the GPME method, this means that the VC investments under evaluation would, at the margin, add value to the investor's portfolio that is not attainable from the public market factors in the SDF. Whether it would be a good idea for a specific investor to invest in VC assets of course depends on the particular situation of this investor. An economic reason why abnormal performance could exists and persist is that the typical VC investor's wealth portfolio return and state-variable risks are not spanned by the factors in the SDF or that VC investors require compensation for the illiquidity of private-equity investments.[6] Our goal in developing the GPME method is to devise a properly risk-matched public markets benchmark for VC payoffs so that we can gauge to what extent payoffs from VC investments just replicate what's available in public markets and to what extent they expand the investment opportunity set. The GPME estimates then

[5] If there is only one initial investment that occurs at the time of the valuation (as in the case of round-to-round returns of start-up companies) the initial inflow is non-stochastic, and hence the Jensen's inequality problem does not arise. In the case of VC funds, however, investors typically make multiple capital contributions during the life of the fund. The amounts and timing of these contributions are initially uncertain, which implies that the inflows are stochastic.

[6] While VC investments are clearly illiquid, it does not automatically follow that there must be a large illiquidity discount. For example, investors focusing on VC may be patient long-horizon investors with little need to trade, requiring only small compensation for illiquidity, as in Constantinides (1986). We thank an anonymous referee for pointing this out.

allow researchers to assess and further investigate the potential role of illiquidity effects and other channels that may give rise to abnormal performance, relative to public markets, in equilibrium.

Thus far we have treated the payoff horizon, h, as given. In reality, start-ups endogenously decide when to raise new financing, and venture funds have the option to extend their life for several years beyond the typical ten-year limit, making h an endogenous variable. For example, projects that are more successful may come back to investors for a new financing round sooner, in order to scale up the business model. The return horizon is thus endogenous in the sense that it is correlated with the unexpected degree of success of the project or the fund. For valuation purposes within the SDF framework, this endogeneity does not present a problem as long as the realized payoffs are ultimately observed. Even if there is an endogenous state-dependence cash flows, the appropriate valuation a payoff in a certain state is still the product of the state's probability and the SDF in that state. The endogeneity concerns that remain relevant in our approach are those that may cause a right-censoring problem and a sample selection problem. We discuss in the data sections below how we deal with these two remaining problems.

A. Setting up public markets benchmarks

Application of our method requires estimates of the SDF parameters a and b that reflect the risk-free rate and market risk premium. For the method to work well, we want these estimates to reflect the risk-free rate and market risk premium during the investment period. This requires a careful structuring of the data.

For the start-up companies data, this is relatively simple, because we can easily match each individual round-to-round return with public markets returns. We pool the data across firms and funding rounds. For each round-to-round observation $i = 1, ..., N$ we observe returns, $R^h_{i,t+h}$, between t and $t + h$, where t and h can be different for each i. We match each return with returns on the CRSP value-weighted market index, F^h_{t+h}, and the return of one-month Treasury Bills, $R^h_{f,t+h}$, over the same horizon from t to $t+h$.[7] Thus, we work with N observations of $Y_i = (R^h_{i,t+h}, F^h_{t+h}, R^h_{f,t+h})'$.[8] Let θ denote the parameters of the SDF, and define the vector

$$u_i(\theta) \equiv M^h_{t+h}(\theta) \cdot Y_i - 1. \tag{14}$$

The first element of u_i is the $GPME_i$ of equation (13) for funding round i. In other words, its expected value is the NPV of investing \$1 in round i.

For the VC funds data, this is more complicated, because each fund has a long series of cash in- and outflows at various points in time. There is no simple definition of an investment period, because the amount invested varies depending on the timing and magnitude of in- and out-flows. For example, if a fund has paid out almost all of its capital towards the end of its life, the total NPV of the fund will no longer be sensitive to risk factor realizations after this point. More generally, the sensitivity to risk-factor realizations depends on how much capital the fund has at a given point in time. Thus, the realized market risk premium and risk-free rate that we should benchmark a fund against must be weighted according to funds' typical pattern of capital accumulation and payout. To accomplish this, we match each fund with artificial benchmark funds that invest in T-bills and the public equity market index (and benchmark funds for additional risk factors in expanded versions of the SDF). We then estimate the SDF parameters a and b to price these benchmark funds.

More precisely, in the VC funds data, we observe cash flows realized at $j = 1, ..., J$ to and from each fund $i = 1, ..., N$, and possibly a final net asset value (NAV) if the fund is not yet liquidated. The first cash flow of the fund occurs at date t. As before, we suppress the dependence of t and J on i for notational simplicity. As a first step, we normalize each fund's cash flows by fund size so that estimation gives each fund equal weight and the normalized cash flows, $C_{i,t+h(j)}$, can be interpreted as resulting from an investment with a

[7] We use subscript $t + h$ for the Treasury Bill return, not t, as one typically would with a conditionally risk-free asset, because it is the return on rolling over short-term Treasury Bills from t to $t+h$, and hence the return is not known until $t + h$.

[8] Note that in Y we use gross arithmetic (not log) returns to conform to the pricing equation (7).

total commitment of \$1. Since t is defined as the time at which the first cash flows into the fund, we have $h(1) = 0$.

We then set up two benchmark funds matched to each VC fund i. One invests in T-bills and the other in the CRSP value-weighted index. We take the inflows into these funds to be exactly identical in magnitude and timing to those of the VC fund that we match. If fund i makes a payout at $t + h(j)$, then we assume that the benchmark funds also make a payout equal to the sum of two components. The first component is equal to the return accumulated since the last cash flow date $t + h(j-1)$. The second component pays out a fraction π_j of the capital that was in the benchmark fund after the last cash flow at $t + h(j-1)$ occurred. The payout ratio is determined as

$$\pi_j = \min\left(\frac{h(j) - p}{10 - p}, 1\right), \tag{15}$$

where p is the time (measured relative to fund inception) of the most recent payout prior to time $t + h(j)$. If there was no prior payout yet, then $p = 0$. Time periods are measured in years. For example, if one year has passed since the last payout, and at the date of this last payout the fund had been in existence for 2 years, then $\pi_j = 1/8$. This assumption effectively sets the fund life-time to roughly 10 years (positive returns towards the end of the benchmark fund's life can lengthen the life-time to some extent because the second payout component is based on the lagged, not current level of capital in the fund). With this assumption, the pattern of payouts over the life-cycle of the average benchmark fund resembles the pattern of payouts of the average VC fund quite well.

With the benchmark fund cash flow streams set up, we now match each VC fund cash flow, $C_{i,t+h(j)}$, with the market return benchmark fund cash flow, $C_{M,t+h(j)}$, and the Treasury

Bill benchmark fund cash flow, $C_{f,t+h(j)}$, collected in the vector

$$Y_{i,t+h(j)} \equiv \begin{pmatrix} C_{i,t+h(j)} \\ C_{if,t+h(j)} \\ C_{iM,t+h(j)} \end{pmatrix}, \tag{16}$$

We define

$$u_i(\theta) \equiv \sum_{j=1}^{J} M_{t+h(j)}^{h(j)}(\theta) \cdot Y_{i,t+h(j)}. \tag{17}$$

Unlike in (14), the price vector here is a vector of zeros and hence it does not appear in (17). Prices are zero because each cash flow series already includes the initial investment. Hence, the expected discounted sum of these payoffs represents the NPV, which is zero if there is no abnormal performance.

B. GMM estimation

For both the VC fund and start-up company data, we employ the GMM estimator

$$\hat{\theta} = \arg\min_{\theta} \left(\frac{1}{N} \sum_i u_i(\theta) \right)' W \left(\frac{1}{N} \sum_i u_i(\theta) \right). \tag{18}$$

Our objective is to evaluate how venture investment payoffs compare, on a risk-adjusted basis, to investments in publicly traded securities. Consistent with this objective, we estimate the SDF parameters, θ, from public capital market data alone. More precisely, we choose as a weighting matrix a diagonal matrix with entries of zero for the element corresponding to VC payoffs (the first element in u_i) and ones for the remaining elements. This leads to exact identification and parameter estimates that allow the SDF to price the public equity markets and Treasury Bill benchmarks with zero pricing errors.

For the estimation of GPME to work well, the data set needs to have sufficiently many non-overlapping time-windows over which the payoffs of the cross-sectional units, i, are realized. More formally, application of the usual asymptotic results for GMM estimation requires large

T; a large number of cross-sectional units is not sufficient. This can be seen most clearly in the extreme hypothetical case in which all VC payoffs are observed over the same time-window. In this case, the realization of the SDF is exactly the same for each cross-sectional unit. Asset-pricing theory does not predict that the cross-sectional average of returns multiplied by this single realization of the SDF value equals one. Instead, it predicts that the *expected* cross-product of the SDF and the return equals one. Precise estimation of this expected value requires sufficiently large T. Online Appendix C conducts Monte Carlo experiments that assess the performance of our estimator for various values of T.

To estimate the SDF parameters from public markets data, one could potentially choose a sample that is longer in the time dimension than the typical data set on VC investments. In a performance evaluation application like ours, there is, however, no benefit from doing so. A longer sample of public markets data would allow the SDF parameters to more precisely reflect the *ex-ante* equity premium, but for the purpose of performance evaluation of funds or fund investments the relevant benchmark is the *ex-post* equity premium during the performance evaluation period.[9] Consider, for example, the evaluation of an investment fund's performance. For simplicity, assume that the risk-free rate is zero and that the fund has a beta of one so that its returns follow $r_t = \alpha + f_t + \varepsilon_t$, where α is the fund's expected abnormal performance, f_t is the market factor return, with $\mu \equiv E[f_t]$, and ε_t is a mean-zero residual uncorrelated with f_t. Comparing this fund's sample mean return, $\bar{r}_t$, to the mean market return, $\bar{f}_t$, over the same period, i.e., $\bar{r}_t - \bar{f}_t = \alpha + \bar{\varepsilon}_t$, yields a less noisy estimate of α than a comparison with the ex-ante equity premium μ, i.e., $\bar{r}_t - \mu = \alpha + \bar{f}_t - \mu + \bar{\varepsilon}_t$. The latter is more noisy because it does not remove the uncertainty stemming from differences between $\bar{f}_t$ and μ. Following this reasoning, tests that use the ex-post equity premium as a performance benchmark should be more powerful than tests that use the ex-ante equity premium. In the Online Appendix, Section D, we present Monte Carlo evidence that confirms this point:

[9] This point has been noted, for example, by Shanken (1992), p. 11. Linear factor model regression estimates of abnormal performance in the mutual funds and hedge funds literature also use the ex-post equity premium as a benchmark.

tests in which the SDF parameters are estimated using data from the performance evaluation period have greater power to detect abnormal performance than tests in which the SDF parameters are set to their true population values.

To conduct statistical inference, we need a metric to assess the magnitudes of the pricing errors, i.e., the GPME, associated with the venture investment payoffs. A key ingredient of the standard GMM formulas is the spectral density matrix

$$S = \sum_{k=-\infty}^{+\infty} E[u_i u_k']. \tag{19}$$

The complication here is that there is likely to be substantial correlation between u_i and u_k if they are measured over fully or partly overlapping time periods. Earlier work in the existing literature either does not report PME standard errors, or assumes uncorrelatedness across funds or start-up firms. In our framework, GMM techniques allow us to avoid this implausible assumption.

The typical application of our estimator will feature a short sample with small T. To ensure reasonable small-sample properties, it is useful to impose some economically motivated restrictions on S. Specifically, we exploit the implications of the null hypothesis that the SDF correctly prices the assets. Under this assumption, the realized pricing errors, u_i, are increments of a martingale difference sequence, i.e., they are mutually uncorrelated unless they are measured over overlapping time periods. Under this assumption, the spectral density, or long-run covariance matrix, in (19) simplifies. Decompose S as

$$S = \Lambda^{\frac{1}{2}} \Gamma \Lambda^{\frac{1}{2}}, \tag{20}$$

into a 3×3 matrix of correlations, Γ, and a diagonal matrix of variances, Λ. For serially uncorrelated variables x and y, the long-run correlation of multiperiod sums of x and y is equal to the contemporaneous correlation. Thus, Γ can be estimated consistently as the correlation matrix component of $E[u_i u_i']$, without adding the additional summation terms where $k \neq i$ in (19).[10] Accordingly, our estimator for Γ is

$$\hat{\Gamma} = \left[\frac{1}{N}\sum_{i=1}^{N} \operatorname{diag}(u_i \circ u_i)\right]^{-\frac{1}{2}} \left(\frac{1}{N}\sum_{i=1}^{N} u_i u_i'\right) \left[\frac{1}{N}\sum_{i=1}^{N} \operatorname{diag}(u_i \circ u_i)\right]^{-\frac{1}{2}}, \tag{21}$$

where $\circ$ denotes an elementwise (Hadamard) product.

For variances, however, the cross-products of u_i and u_k with $i \neq k$ are important, even under the null hypothesis. Our estimator for Λ takes these into account. We assume that the correlation of payoffs i and k depends on the degree of overlap of the time-windows over which these payoffs are measured. In analogy to spatial GMM methods (Conley (1999)), we treat the degree of overlap as a measure of distance, assuming that correlation declines with distance. Let $t(i)$ and $t(k)$ be the start, and $t(i)+h(i)$ and $t(k)+h(k)$ be the end of the time windows for observation i and k, respectively. Define the distance as

$$d(i,k) \equiv 1 - \frac{\min[t(i)+h(i), t(k)+h(k)] - \max[t(i), t(k)]}{\max[t(i)+h(i), t(k)+h(k)] - \min[t(i), t(k)]}. \tag{22}$$

If the return measurement windows exactly overlap, then $d(i,k) = 0$, if they are adjacent, but non-overlapping, then $d(i,k) = 1$, and $d(i,k) > 1$ if there is a gap between them. To ensure that S is positive semidefinite, we apply Bartlett-type weights that decline with greater distance,

$$\hat{\Lambda} = \frac{1}{N}\sum_{k=1}^{N}\sum_{i=1}^{N} \max(1 - d(i,k)/\bar{d}, 0) \operatorname{diag}(u_i \circ u_k), \tag{23}$$

where $\bar{d}$ is a constant. In our empirical work, we set $\bar{d} = 1.5$, which means that observation-pairs with a gap between return measurement windows still get some weight in the calculation

[10] For example, consider two serially uncorrelated random variables x and y with pairwise contemporaneous correlation ρ and variances σ_x^2 and σ_y^2. If data is observed as overlapping sums $u_t = (x_t + x_{t-1}, y_t + y_{t-1})'$, then the off-diagonal elements of the spectral density (19) are $S_{xy} = \text{Cov}(x_t + x_{t-1}, y_t + y_{t-1}) + 2\text{Cov}(x_t, y_t) = 4\sigma_x\sigma_y\rho$. The diagonal elements are $S_{xx} = 4\sigma_x^2$ and $S_{yy} = 4\sigma_y^2$. Thus, we can write $S = \Lambda^{\frac{1}{2}}\Gamma\Lambda^{\frac{1}{2}}$ as in (20), where Λ has $4\sigma_x^2$ and $4\sigma_y^2$ on the diagonal and Γ has ones on the diagonal and ρ, i.e., the simple correlation of x_t and y_t as off-diagonal elements.

of $\hat{\Lambda}$. The estimator of S then follows as

$$\hat{S} = \hat{\Lambda}^{\frac{1}{2}} \hat{\Gamma} \hat{\Lambda}^{\frac{1}{2}}. \tag{24}$$

In small-sample Monte Carlo simulations, we found that imposing the restrictions on Γ that lead to the specification of $\hat{\Gamma}$ in (21) helps to substantially reduce size distortions of the test. If one does not impose this assumption, the addition of cross-products of u_i and u_k with $i \neq k$ in the correlation estimator often results in extreme off-diagonal entries in S that appear to distort small-sample inferences. As we show in the Online Appendix C, the S matrix estimator (24) performs reasonably well even with very small T.

Finally, we follow standard practice and use a centered version of S where we demean the u_i before computing $\hat{\Gamma}$ and $\hat{\Lambda}$. While not necessary under the null hypothesis, doing so tends to lead to greater power of the test under misspecification (Hall (2000)).

With $\hat{S}$ in hand, we can use standard GMM formulas to compute a J-statistic for a test that the pricing error for venture investment payoffs (the first element of u_i) is zero, which is equivalent to a test that $GPME = 0$, taking into account the estimation error in the SDF parameters. Since the SDF parameters are estimated using only information on pricing factor and T-Bill returns, but not venture investment payoffs, this J-statistic has an asymptotic $\chi^2(1)$ distribution.

II. VC Funds

We now apply the GPME method to data on VC investments. We start with VC funds.

A. Data

We use a large sample of VC fund cash flows between 1979 and 2012, obtained from Preqin. The data contain capital calls by the fund from LPs (also known as capital takedowns, these are cash flows into the private equity partnership), cash distributions from the fund to LPs, and quarterly Net Asset Values (NAVs). For comparison to the prior literature, we follow Kaplan and Schoar (2005) and restrict the sample to U.S. funds with committed capital of at least $5 million in 1990 dollars. We also drop funds that are raised after 2008, as very few of their investments will have realized by the end of the sample period, and their performance is therefore likely to be misleading. For funds that are not yet liquidated by the end of the sample period, performance metrics are calculated using self-reported Net Asset Value (NAV) at the end of the sample period proxy for a final cash flow. These NAVs are known to have issues.[11] In Online Appendix E we describe a subsample of funds that were close to fully liquidated by the end of our sample period, and Appendix F shows robustness of our main results in this subsample.

Panel A of Table I shows descriptive statistics for the sample of funds. The sample comprises 545 funds, raised by 278 VC firms. The median (average) VC firm in our sample raised one (two) fund(s), with a total commitment of $226 ($358) million. There is a long right tail in the distribution of fund sizes, mostly due to older VC firms that have raised more than one fund, as these firms tend to raise funds that are increasingly larger. The median (average) fund has 26 (30.28) unique cash flows (including both capital calls and distributions) during the sample period. The time between the first and last observed cash flow for the median (average) fund is 11.09 (10.40) years. The median (average) fund has an internal rate of return (IRR) of 4.37% (8.84%) and a total value to paid-in capital (TVPI, which is defined as the sum of cash distributions to the fund's limited partners (LPs) plus any net asset value remaining in the fund divided by the sum of capital takedowns, that is, cash transfers from LPs to the fund) of 1.16 (1.57). Panel B shows the distribution of funds by vintage year. Most of the funds are of post-1990 vintages. The highest returns are realized in the 1990s vintages before 1999. Performance from 1999 until the end of the sample period is poorer, with average IRRs in the single digits and even sometimes negative. We will discuss

[11] Woodward (2009) and Phalippou and Gottschalg (2009) suggest that NAVs may be stale. Jenkinson, Sousa, and Stucke (2013) and Brown, Gredil, and Kaplan (2013) show that most GPs tend to be conservative in reporting NAVs, while bottom GPs tend to be aggressive during fundraising periods.

Table I
Summary Statistics: VC Fund Data

This table reports descriptive statistics of the sample of U.S. VC funds with cash flow data from Preqin, for vintages between 1979 and 2008, eliminating funds with committed capital below $5 million in 1990 dollars. *Fund size* is the total commitment to the fund, in millions of dollars. *Fund vintage year* is the calendar year in which the fund is raised. *Fund effective years* is the time between the first and the last observed cash flow of a fund. The fund *IRR* is computed using the final observed NAV of the fund. *TVPI* stands for total value to paid-in capital, and is computed as the sum of cash distributions to Limited Partners plus final NAV divided by the sum of cash takedowns by the fund from LPs.

Panel A: Descriptive statistics			
	mean	median	st.dev.
# Funds	545		
# VC firms	278		
# Funds / VC firm	1.96	1.00	1.42
Fund size ($m)	358.46	226.00	450.35
Fund effective years	10.40	11.09	4.32
# Cash flows / fund	30.28	26.00	16.96
IRR (%)	8.84	4.37	37.11
TVPI	1.57	1.16	2.26

Table I - Continued

Panel B: Performance statistics by vintage year					
	# funds	IRR		TVPI	
		mean	median	mean	median
1979	1	18.49	18.49	2.23	2.23
1980	1	15.26	15.26	2.22	2.22
1981	1	27.04	27.04	2.63	2.63
1982	3	8.83	8.50	1.73	1.74
1983	2	9.04	9.04	1.74	1.74
1984	3	11.49	11.82	1.91	1.99
1985	4	13.17	14.31	2.05	1.95
1986	7	5.89	6.33	1.50	1.44
1987	5	14.69	14.84	2.36	2.32
1988	5	17.84	15.72	2.43	2.33
1989	5	19.61	15.44	2.55	1.66
1990	7	17.67	14.16	2.17	2.19
1991	2	21.04	21.04	2.46	2.46
1992	12	27.87	25.32	3.32	2.64
1993	9	38.98	32.13	3.88	3.40
1994	13	34.78	32.12	5.16	2.74
1995	16	46.27	19.95	3.56	1.72
1996	16	33.33	12.95	3.24	1.71
1997	22	28.93	11.32	1.92	1.35
1998	29	19.94	0.74	1.66	1.03
1999	40	-3.24	-5.72	0.82	0.74
2000	67	-3.10	-2.56	0.91	0.88
2001	43	1.04	2.85	1.20	1.13
2002	25	-1.68	-1.84	0.98	0.91
2003	19	2.00	3.08	1.07	1.13
2004	26	2.65	0.74	1.36	1.03
2005	28	2.09	4.81	1.18	1.21
2006	52	1.12	2.31	1.07	1.08
2007	46	8.06	6.26	1.29	1.18
2008	36	7.53	4.77	1.19	1.09

this in more detail below.[12]

B. GMM Results

Table II shows the GPME estimates for various specifications of the SDF. Column (i) shows the PME measure, expressed as a difference between in- and outflows rather than the traditional ratio. Panel A shows that the PME estimate across all funds is 0.048. This means that on a \$1 fund commitment, the abnormal profit is 4.8 cents, or 4.8%. Note that this is the present value of the abnormal cash flow for the fund, i.e., the net present value, and not an annualized number. This PME is not statistically significantly different from zero, though, as indicated by the p-value of the J-test. As with IRR and TVPI, the literature reports a wide range of average PMEs (bearing in mind that a PME of zero in our method corresponds to a traditional PME of one): Kaplan and Schoar (2005) report 0.96, Robinson and Sensoy (2011) find 1.06 (1.03 for fully liquidated funds), and Harris, Jenkinson, and Kaplan (2014) 1.36. The latter number is from the most comprehensive and, likely, most representative sample of VC funds in the literature. This is comparable to our PME estimate of 0.276 in a sample of nearly liquidated funds, as reported in Online Appendix E, which is statistically significantly different from zero, though we caution against broad generalizations based on our smaller sample. Rather, we focus on our main empirical contribution, analyzing the impact of changing the SDF on abnormal performance estimates.[13]

The second column in Table II shows the estimates for the more general CAPM SDF in equation (8). This specification relaxes the equity premium and risk-free rate restrictions implicit in the PME. In contrast to column (i), the GPME estimate of -0.103 in column

[12] Preqin also provides a larger data set of VC funds that report multiples and IRRs, but not individual cash flows between limited partners and the fund. Our sample funds tend to be larger, and the median IRR is 3.5 percentage points lower. Performance is comparable across the subsamples of liquidated funds. Online Appendix E compares our sample to this larger data set in more detail, as well as to other popular data sets in the literature. Given the variation in reported performance between the various data sets, one should be cautious in generalizing our findings.

[13] Online Appendix E also shows that, though the PME and GPME estimates are higher in the liquidated funds subsample than in the full sample reported here, the relative differences between PME and GPME—which is our main focus—are robust.

(ii) is negative and we can reject the null of zero pricing errors. The difference between the estimates in column (i) and (ii) is sizable. It amounts to about 15 cents per dollar of fund commitment. The origin of this discrepancy can be seen by inspecting the SDF parameters. The point estimate of the coefficient b_1 in column (ii) is 2.65, substantially above the value of 1.0 imposed by the log-utility model in column (i) that the PME method is based on. This means that in this sample period the PME implicitly understates the public markets equity premium, which, together with the fact that VC funds have $\beta > 1$, overstates the abnormal return.

The intuition for this can be obtained from the log-normal special case discussed in the introduction. Taking the difference of the log expected returns implied by the GPME and PME, respectively, in equations (3) and (4), we get

$$(\beta - 1)(\log E[R_{m,t+1}] - r_f - \sigma_m^2) \quad (25)$$

Thus, if the equity premium is high so that $\log E[R_{m,t+1}] - r_f > \sigma_m^2$ (which is true in our sample), then $\beta > 1$ implies that the PME overstates the abnormal return. Without log-normality, this does not hold exactly, but the intuition should still apply approximately.

In column (iii), the SDF includes the log return on a small-growth factor in excess of the log market return. The small-growth factor portfolio is one of the six portfolios underlying the Fama and French (1993) factors.[14] The estimated price of risk coefficient, b_2, is close to zero for this factor. As a consequence, the addition of this factor has little effect on the GPME compared with the CAPM in column (ii).

To illustrate the source of the differences between the PME and the GPME, we split the sample into one sub-sample during which the public equity market had high returns and one in which the market returns were more moderate. Since funds have long and varying lifetimes, an exact split by time periods is not possible. However, splitting it by fund vintage into pre-1998 vintages (Panel B) and vintages 1998 and later (Panel C) achieves some separation

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

Table II
Generalized Public Market Equivalents for VC Funds

We estimate the Generalized Public Market Equivalent (GPME) by discounting VC fund cash flows with the stochastic discount factor

$$M_{t+h}^{h} = \exp(ah - b_1 r_{m,t+h}^{h} - b_2 r_{x,t+h}^{h}),$$

summing each fund's discounted cash flows, and averaging across all funds. Fund cash flows are normalized by fund size to a total commitment of \$1. The log-utility CAPM special case in column (i) with $a = 0$, $b_1 = 1$, and $b_2 = 0$ corresponds to the Public Market Equivalent of Kaplan and Schoar (2005). In column (iii), the second factor, $r_{x,t+h}^{h}$, is the log return of the small-growth portfolio in excess of the log market return. The SDF parameters are chosen, with exact identification, to correctly price benchmark funds that receive the same inflows as the VC funds, but which invest in the CRSP value-weighted index, Treasury Bills, and, in column (iii), the small-growth portfolio. Standard errors of the SDF parameter estimates and the GPME are in parentheses, and p-values of the J-test of $GPME = 0$ are in square brackets. The spectral density matrix used in the computation of the J-statistic takes into account error dependence arising from overlapping fund life times as described in the text.

	(i) Log-utility CAPM	(ii) CAPM	(iii) CAPM w/ small growth
Panel A: All funds ($N = 545$)			
GPME	0.048	-0.103	-0.106
	(0.061)	(0.042)	(0.039)
	[0.430]	[0.015]	[0.007]
SDF parameters			
a	0	0.088	0.090
		(0.037)	(0.036)
b_1	1	2.650	2.680
		(0.560)	(0.526)
b_2			-0.154
			(0.913)
Panel B: All funds of vintage < 1998 ($N = 134$)			
GPME	0.423	0.111	0.148
	(0.109)	(0.142)	(0.604)
	[0.000]	[0.434]	[0.626]
SDF parameters			
a	0	0.309	0.596
		(0.094)	(0.227)
b_1	1	4.361	5.655
		(0.992)	(1.866)
b_2			-8.068
			(6.446)

Table II - Continued

	(i) Log-utility CAPM	(ii) CAPM	(iii) CAPM w/ small growth
Panel C: All funds of vintage $\geq$ 1998 ($N = 411$)			
GPME	-0.074	-0.144	-0.105
	(0.030)	(0.040)	(0.045)
	[0.015]	[0.000]	[0.019]
SDF parameters			
a	0	0.062	0.024
		(0.044)	(0.033)
b_1	1	2.361	1.199
		(0.768)	(0.689)
b_2			3.773
			(1.555)
Panel D: GPME of funds by stage			
Generalist	0.075	-0.142	-0.157
(N = 301)	(0.090)	(0.062)	(0.050)
	[0.408]	[0.022]	[0.002]
Early	-0.069	-0.180	-0.160
(N = 145)	(0.046)	(0.045)	(0.049)
	[0.129]	[0.000]	[0.001]
Late	0.205	0.107	0.133
(N = 41)	(0.071)	(0.109)	(0.119)
	[0.004]	[0.323]	[0.263]
Mezzanine	0.095	0.175	0.175
(N = 58)	(0.036)	(0.038)	(0.038)
	[0.009]	[0.000]	[0.000]
Panel E: GPME of funds by size			
Small	0.165	-0.049	-0.060
(< median)	(0.090)	(0.064)	(0.062)
(N = 268)	[0.067]	[0.449]	[0.332]
Large	-0.066	-0.152	-0.126
($\geq$ median)	(0.035)	(0.040)	(0.049)
(N = 256)	[0.055]	[0.000]	[0.010]

of funds that had their investments concentrated in the two "bull market" decades leading up to the year 2000 from those that were exposed to the declines in equity markets in 2000 and 2007 to 2009. Accordingly, the SDF parameter estimates in column (ii) of Panel B feature a much higher price of risk coefficient b_1 at almost twice the magnitude of the estimate in Panel A. As a consequence, the difference between the PME in column (i) and the GPME in column (ii) is about twice as high as in Panel A. Moreover, the GPME in column (ii) of Panel B is not statistically significant, while the PME in column (i) is.

These facts are consistent with our earlier discussion that one would expect the PME to overstate abnormal performance of VC investments during times of strongly rising public equity markets. The estimation of the SDF parameters in column (ii) from public markets data lets the SDF adapt to the high realized market returns during the sample period. The PME in column (i) with fixed $a = 0$ and $b_1 = 1$ does not adjust the performance benchmark in this manner. These restrictions do not allow the model to properly adjust for the fact that VC fund returns for pre-1998 vintages are quite high simply because public equity markets were buoyant during much of the lifetime of these funds.

The results in Panel C are quite similar to those on Panel A because most of the observations in the full sample in Panel A are from funds with vintages 1998 and later. The difference between the PME in column (i) and the CAPM GPME in column (ii) is slightly smaller than in Panel A, consistent with the slightly smaller estimate of the price of risk coefficient b_1 in the ≥ 1998 sub-sample.

Panels D and E explore cross-sectional cuts of the data by the type of fund and the size of the fund. With regards to differences between the PME and the GPME, which is our focus here, the picture is quite homogeneous. One exception are Mezzanine funds in Panel D. They are the only category for which the GPME is higher than the PME. This is exactly what one would expect. Many of these funds' investments are structured as debt claims. The debt usually comes with "equity-kickers" in the form of warrants or other equity-like features, but its beta should still be quite low and likely below one. Evidence that mezzanine investments

Table III
Generalized Public Market Equivalents for Artificially Levered VC Funds

We use the VC funds, with cash flows $C_{i,t+h(j)}$, and the matched Treasury Bill benchmark funds, with cash flows $C_{if,t+h(j)}$ to create artificially levered VC funds with cash flows

$$L_{i,t+h(j)} = C_{i,t+h(j)} + k\left(C_{i,t+h(j)} - C_{if,t+h(j)}\right).$$

We apply the SDFs from Table II, Panel A, to these artificially levered funds for a range of different leverage ratios k and we report the resulting GPME estimates. Standard errors of the GPME are in parentheses, and p-values of the J-test of $GPME = 0$ are in square brackets. The spectral density matrix used in the computation of the J-statistic takes into account error dependence arising from overlapping fund life times as described in the text.

k	(i) Log-utility CAPM	(ii) CAPM
-0.8	-0.053	-0.021
	(0.016)	(0.026)
	[0.001]	[0.428]
-0.6	-0.027	-0.041
	(0.023)	(0.041)
	[0.234]	[0.320]
0	0.048	-0.103
	(0.061)	(0.042)
	[0.430]	[0.015]
1	0.174	-0.205
	(0.131)	(0.121)
	[0.182]	[0.089]
2	0.301	-0.308
	(0.201)	(0.209)
	[0.135]	[0.141]

have lower risk than other VC investments also point in this direction (see, e.g., Schmidt (2006)). Following equation (25), one should therefore expect the PME to understate rather than overstate the abnormal performance of Mezzanine funds in samples like ours in which the estimated b_1 is greater than one due to a relatively high realized equity premium.

The Mezzanine funds results are suggestive, but the sample is relatively small. Furthermore, based on fund cash flow data it is difficult to ascertain the beta, or, more generally the covariance between risk factors and fund returns. For this reason, we explore an alternative route to show that it is important to relax the parameter restrictions implicit in the PME to properly adjust for systematic risk exposure. We create artificially levered funds. With leverage, we can achieve any desired level of systematic risk and we can explore the effect on PME and GPME. We use the VC funds, with cash flows $C_{i,t+h(j)}$, and the matched Treasury Bill benchmark funds, with cash flows $C_{if,t+h(j)}$, to create artificially levered VC funds with cash flows

$$L_{i,t+h(j)} = C_{i,t+h(j)} + k\left(C_{i,t+h(j)} - C_{if,t+h(j)}\right). \tag{26}$$

The parameter k controls leverage. With $k = 0$ there is no leverage, $k = 1$ implies that the VC fund assets are levered two-to-one. A negative k generates a portfolio that is partly invested in the VC fund and partly in a long position in Treasury Bills.

Table III presents estimates of the PME and our GPME when we apply the SDFs from Table II, Panel A, to these artificially levered funds for a number of different values of k. A necessary condition for a risk-adjustment to be accurate is that levering up, say, two-to-one should raise the abnormal performance by the same factor of two. As Table III shows, the GPME satisfies this requirement: Going from no leverage ($k = 0$, same as in Table II) to two-to-one leverage ($k = 1$) exactly doubles the (negative) GPME. In contrast, the PME in column (i) rises more than three-fold.

By choosing k to be far below zero, we can also reproduce the Mezzanine funds result from Table II where the PME was lower than the GPME. For example, with $k = -0.8$, the beta of the resulting VC fund/T-bill portfolio should be lower than one. In this case, following equation (25), the PME should understate rather than overstate the abnormal performance. As the table shows, this is what we find.

Overall, the findings in Tables II and III illustrate the benefits of our proposed performance evaluation approach. The estimation of SDF parameters with public markets data—rather than fixing them at $a = 0$ and $b_1 = 1$ as in the PME method—allows the GPME performance benchmark to reflect public market returns during the sample period in a risk-adjusted manner. Our evidence indicates that the PME overstates VC fund performance in the past few decades, except for funds focusing on low-beta mezzanine investments.

III. VC-backed Start-up Companies

A. Data

We use data of financing rounds for VC-backed start-up companies, provided to us by Sand Hill Econometrics (SHE). SHE has combined two commercially available databases, VentureXpert (from Thomson Venture Economics) and VentureSource (formerly Venture One), and invested substantial time and effort to track down exits for companies not previously shown as exited, which are mostly shutdowns or disappointing acquisitions, and to fill in missing private financing rounds.[15] To ensure accuracy of the data, SHE has removed duplicate investment rounds and consolidated rounds, such that each round corresponds to a single investment by one or more VCs.

The full SHE dataset contains 61,356 financing rounds for 18,237 unique start-ups between 1987 and 2005. Of these start-ups, 1,891 (10.4%) ultimately went public in an IPO, 4,271 (23.4%) were acquired, and 2,892 (15.9%) were liquidated. The ultimate outcome for the remaining 9,183 firms (50.4%) was unknown by the end of the sample. Some of these firms were still operating as private firms, but many of them have likely been liquidated (the so-called "zombie" firms). For these firms, the average (median) time since the last financing round was 57 (41) months by the end of the sample. Of the 52,302 rounds in which new venture capital was raised (i.e., non-IPO, non-acquisition, non-liquidation rounds), 1,393 (2.7%) were seed rounds, 34,066 (65.1%) were early rounds, 16,466 (31.5%) were late rounds, and 377 (0.7%) were designated as mezzanine rounds.[16]

We construct round-to-round returns as the change in valuation of the start-up from the

[15] Gompers and Lerner (1999) and Kaplan, Sensoy, and Strömberg (2002) find that missing investments in VentureXpert are predominantly smaller and more idiosyncratic ones.

[16] The label "early" versus "late" is somewhat subjective, and the mezzanine round designation is even more fuzzy. Typically a mezzanine round is the financing round that bridges the 6 to 12 month gap to a liquidity event—IPO, or sometimes acquisition—but sometimes refers to the round between the early and late stage.

post-money valuation in a given round to the pre-money valuation in the subsequent round. Post-money is a term used in the VC industry to denote the value of the start-up including the new investment. Pre-money is defined as the post-money valuation minus the amount invested in that round. With consecutive financing rounds at time t and $t+h$, the return is,

$$R^h_{t+h} = \frac{V^{PRE}_{t+h}}{V^{POST}_t}, \tag{27}$$

where V^{PRE}_t is the pre-money valuation for a financing round that takes place at time t, and V^{POST}_t is the post-money valuation of that same round. The return to a buy-and-hold investor who holds on to her initial investment, and does not invest any additional money in future rounds, is simply the compounded return across rounds.[17]

Not all rounds have valuations filled in, so we can compute round-to-round returns only for a subset of the data. We observe 6,861 round-to-round returns (after dropping one return that was less than -100%) for 3,497 unique firms. Table IV Panel A reports summary statistics for this sample.

This sample suffers from two potential forms of selection bias, which are both ultimately related to the endogeneity of financing rounds. First, there could be a right-censoring problem: we observe financing rounds between 1987 and 2005, so for firms that were started towards the mid-2000s we only observe financing rounds for a relatively short time period. Since the most successful firms are likely those that proceed most quickly through the VC funding process towards exit, we are more likely to observe the final outcomes for these successful firms. Less successful firms might not have a final valuation available, and the values

[17] Consider a simple example of a buy-and-hold investor. Suppose an investor invests $40 into a series A round, for 100 shares of stock, representing 40% of the equity of the firm, implying a post-money valuation of $100. In the next, series B round, another investor purchases 150 shares of new stock (i.e., a stake of 150/(250+150) = 37.5%) at $1 per share, for a total investment of $150. The post-money B-round valuation is $400 (400 shares at $1 each), and the pre-money round B valuation is $400-$150=$250. The return to our buy-and-hold investor from round A to B is $250/$100 = 2.5 (of course we could more easily compute the return from the share prices, but these are usually not reported in the SHE data). Now suppose the start-up is acquired a few months later, at a share price of $2. The return for round B to IPO is $2 · 400 / $400 = 2. Our buy-and-hold investor initially bought her shares at $0.40/share, and she realized a total return on her investment of $2/$0.40 = 5. This exactly equals the compounded return over the rounds, 2.5 times 2.

Table IV
Summary Statistics: VC Rounds Data

Descriptive statistics for the sample of VC financing rounds of start-up companies from Sand Hill Econometrics over the period 1987 to 2005. Panel A reports statistics for the sample of round-to-round returns. The top row of Panel B shows the sampling frequencies by funding stage for the sample with observed end-of-round events. The second row shows the frequencies for the sample if we assume that companies with unobserved exits are assumed to be liquidated, while the third row uses only observations that have data on round-to-round returns available or which are liquidations (for which we assume a liquidation return). Panel C shows the mean gross round-to-round returns in the resampled sample that matches the frequencies in the second row of Panel B (see the text for details), where the unobserved liquidation returns are assumed to be -90% (first row), -70% (second row), and -50% (third row). Panel D reports the mean gross return on investments in the CRSP value-weighted index that are matched in time to the resampled round-to-round returns, and Panel E shows the mean return horizons (in years), by funding stage.

			quantile				
	mean	st.dev	10	50	90		
Panel A: Raw VC financing rounds data (rounds with a return)							
# Rounds	6,861						
# Start-up companies	3,497						
# Rounds / company	2.96	1.30	2.00	2.00	5.00		
Time between rounds (yrs)	1.01	0.83	0.25	0.83	1.92		
Panel B: Sampling frequencies							
	Seed	Early	Late	Mezz	Acq	IPO	Liq
Observed end-of-round event	0.005	0.414	0.363	0.008	0.099	0.044	0.067
Unobs. end-of-round event treated as liq.	0.004	0.341	0.299	0.007	0.082	0.036	0.231
Observed return or assumed liq. return	0.001	0.325	0.217	0.017	0.048	0.094	0.297
Panel C: Mean gross round-to-round returns (in resampled sample)							
Ass. liq.ret. = -90%	1.256	2.389	1.710	2.410	3.190	3.331	0.101
Ass. liq.ret. = -70%	1.256	2.389	1.710	2.410	3.190	3.331	0.301
Ass. liq.ret. = -50%	1.256	2.389	1.710	2.410	3.190	3.331	0.501
Panel D: Time-matched payoffs on CRSP value-weighted index (in resampled sample)							
Mean gross return	1.110	1.138	1.125	1.203	1.221	1.189	1.191
Panel E: Round-to-round time horizon (in resampled sample)							
Mean (in years)	0.589	0.955	0.916	1.076	1.675	1.056	2.166

Table V
Summary Statistics: VC Rounds Data by End-of-Round Year

This table shows the distribution of observations across end-of-round years for the sample of round-to-round returns in the Sand Hill Econometrics data, by end-of-round event (Seed, Early, Late, Mezzanine, Acquisition, IPO, or Liquidation). The sample corresponds to the sample used in the first line in Panel B of Table IV.

Year	Seed	Early	Late	Mezz	Acq	IPO	Liq
1987	0.0003	0.0036	0.0050	0.0000	0.0004	0.0002	0.0000
1988	0.0003	0.0109	0.0111	0.0000	0.0011	0.0005	0.0000
1989	0.0003	0.0131	0.0146	0.0002	0.0017	0.0008	0.0000
1990	0.0003	0.0126	0.0148	0.0002	0.0017	0.0009	0.0001
1991	0.0003	0.0107	0.0156	0.0005	0.0017	0.0023	0.0002
1992	0.0003	0.0121	0.0160	0.0006	0.0021	0.0033	0.0003
1993	0.0003	0.0096	0.0141	0.0009	0.0026	0.0033	0.0001
1994	0.0001	0.0112	0.0130	0.0007	0.0028	0.0027	0.0003
1995	0.0003	0.0123	0.0143	0.0010	0.0035	0.0035	0.0012
1996	0.0004	0.0215	0.0159	0.0008	0.0046	0.0056	0.0003
1997	0.0006	0.0285	0.0169	0.0004	0.0050	0.0031	0.0053
1998	0.0003	0.0322	0.0217	0.0002	0.0061	0.0019	0.0038
1999	0.0003	0.0442	0.0297	0.0012	0.0083	0.0064	0.0028
2000	0.0000	0.0668	0.0331	0.0011	0.0123	0.0053	0.0062
2001	0.0000	0.0443	0.0299	0.0001	0.0117	0.0007	0.0216
2002	0.0007	0.0214	0.0273	0.0000	0.0103	0.0005	0.0121
2003	0.0001	0.0214	0.0265	0.0000	0.0086	0.0005	0.0076
2004	0.0001	0.0215	0.0253	0.0001	0.0094	0.0018	0.0050
2005	0.0000	0.0163	0.0179	0.0000	0.0054	0.0008	0.0003
Total	0.0049	0.4143	0.3626	0.0080	0.0992	0.0438	0.0672

from the last available financing round before the end of the sample are likely out of date. To mitigate the censoring problem, we treat firms with unobserved valuations at the end of the sample as liquidations, and we explore how our results depend on the assumed liquidation return of -90%, -70%, or -50%. As this assumption may understate VC payoffs, we also perform robustness checks where we only include companies with initial financing rounds before January 2000, such that it is more likely that we have seen most of the final outcomes by the end of the sample (see Online Appendix F).

Second, throughout the whole sample, data is missing for reasons that are likely to be related to performance. Two types of missing data concerns are relevant here: missing information on the end-of-round event and missing returns. Concerning the nature of the end-of-round event, the first row of Panel B in Table IV shows observed frequencies of different investment and outcome stages (Seed, Early, Late, Mezzanine, Acquisition, IPO, or Liquidation) without correcting for the missing outcomes, and Table V shows the distribution of these outcome stages across sample years. However, the ultimate fate of a substantial number of firms is unknown (these firms are also known as "zombie" firms). These tend to be companies with bad outcomes: unsuccessful start-ups do not return for new financing rounds, and it is difficult to find data on liquidations of such small firms. As a first step to dealing with the missing data problem, we treat zombie rounds (where the final observed round for a firm is not an exit, i.e., an IPO, acquisition, or known liquidation) as liquidations. The observed frequencies for the resulting sample are shown in the second row of Panel B in Table IV.

Concerning the availability of round-to-round returns, the problem is that data is more likely to be missing for unsuccessful outcomes. This is due to the fact that IPO and acquisitions data tend to be more widely publicized and thus easier to find and backfill (for example using S-1 registration documents) than information about unsuccessful firms. To illustrate the success bias, we have the full return history (i.e., where we observe valuations for every single round) for only 742 firms. Of these firms, 524 ultimately went public, and 218 were ultimately acquired, and none were liquidated. After assuming a liquidation return for known liquidations and "zombies," the success bias is not as strong anymore, but, as shown in the third row of Panel B of Table IV, the observed frequencies are still not quite the same as those in the second row.

We correct for this remaining selection bias by resampling round-to-round returns with replacement (including their time-matched risk-factor payoffs) to match the distribution across investment stages of the full sample that includes the missing liquidation returns in the second row of Panel B of Table IV. More precisely, we take 9,651 draws from the 9,651 observations for which we either observe a round-to-round return or which are liquidations (with an assumed liquidation return), and for which we have put factor returns.[18] The sample size of 9,651 after resampling is arbitrary. Our cross-correlation robust spectral density matrix in the GMM estimation ensures that the choice of this sample size does not affect the results. Drawing a greater number of observations would make it more likely that the same observation is drawn more often into the sample, raising the cross-sectional correlation of pricing errors, but the cross-correlation robust estimator adjusts for this.

This procedure corrects for the oversampling of stages for which returns are more likely to be observed. Still, within each stage of investment, there may remain some residual selection bias, for example, if the acquisitions with observed returns are more successful than unobserved acquisitions. Our resampling procedure assumes that the missing returns of each stage come from the same distribution as the observed returns. This is a particular concern for acquisitions. Unlike IPO returns, which are quite comprehensively observed, many acquisition returns are not known. We therefore check the robustness of our results under the assumption that unknown acquisition returns are either 0% or liquidated (see Online Appendix F).

The mean gross returns per funding round category are shown in Panel C of Table IV.

[18]Due to the later start date here compared with the VC funds data, we are only missing put returns for around 1% of observations. Considering this small number, we conduct the estimation of all models, not only the one with put factor returns, based on the sample with available put factor data.

For robustness, we estimate the GPME in our analysis below assuming that liquidations with unknown returns (including the zombie rounds) can have a -90%, -70% or -50% return. The final column of Panel C in Table IV shows how this affects the average return in the liquidation category.

B. GMM Results

Table VI reports the GPME estimates for a variety of models. Panel A shows results for the full sample. The PME estimate is 0.544 under the assumption that liquidation returns are -90%. For higher liquidation returns of -70% and -50%, we find PME estimates of 0.589 and 0.633, respectively. Relaxing the SDF restrictions to the more general CAPM specification (8) in column (ii) lowers the GPME by between 0.019 and 0.028, depending on the liquidation return assumption. Relative to the magnitude of the GPME point estimate, this is a small decrease. However, in terms of absolute economic magnitudes this is actually a significant reduction. For example, a reduction of the pricing error by 0.02 for investments that have a roughly one-year horizon, on average, corresponds roughly to a reduction of the annual abnormal return by about 2 percentage points. The reason why the GPME and PME results do not differ to a greater extent is that the SDF parameter estimates obtained from public market data, as shown in the lower part of column (ii), are very close to $a = 0$ and $b = 1$. In other words, the performance of public markets during the full sample periods is, coincidentally, almost consistent with the assumptions underlying the PME.

The magnitudes of the GPME estimates are big, but they are roughly consistent with the arithmetic alphas that Cochrane (2005) (32% annualized) and Korteweg and Sorensen (2010) (3.5% monthly) obtain by combining a log-normal model with a selection model for endogenous funding and exit events. The approach that we propose here delivers these results in a much simpler and more robust fashion without the need for the specific distributional assumptions and the rather cumbersome estimation of a selection model. Without correcting for selection, Cochrane (2005) and Korteweg and Sorensen (2010) obtain much higher

Table VI
Generalized Public Market Equivalents for VC Round-to-Round Returns

We match each start-up company round-to-round gross return, measured over horizon h from time t to $t+h$, with the return of the CRSP value-weighted index ($r^h_{m,t+h}$ in logs) from t to $t+h$, the return from rolling over 1-month T-bills, and the return on a small-growth portfolio, $r^h_{x,t+h}$. We estimate the Generalized Public Market Equivalent (GPME) by discounting the round-to-round gross returns with the stochastic discount factor

$$M^h_{t+h} = \exp(ah - b_1 r^h_{m,t+h} - b_2 r^h_{x,t+h})$$

and averaging across all observations. The log-utility CAPM special case in column (i) with $a = 0$, $b_1 = 1$, and $b_2 = 0$ corresponds to the Public Market Equivalent of Kaplan and Schoar (2005). In column (iii), the second factor, $r^h_{x,t+h}$, is the log return of the small-growth portfolio in excess of the log market return. The SDF parameters are chosen, with exact identification, to correctly price benchmark funds that receive the same inflows as the VC funds, but which invest in the CRSP value-weighted index, Treasury Bills, and, in column (iii), the small-growth portfolio. Standard errors of the SDF parameter estimates and the GPME are in parentheses, and p-values of the J-test of $GPME = 0$ are in square brackets. The spectral density matrix used in the computation of the J-statistic takes into account error dependence arising from overlapping fund life times as described in the text.

	(i) Log-utility CAPM	(ii) CAPM	(iii) CAPM w/ small growth
Panel A: Full sample ($N = 9,651$)			
Ass.: Unobserved liquidation returns = -90%			
GPME	0.544	0.516	0.598
	(0.195)	(0.180)	(0.283)
	[0.000]	[0.004]	[0.034]
Ass.: Unobserved liquidation returns = -70%			
GPME	0.589	0.565	0.648
	(0.183)	(0.167)	(0.277)
	[0.000]	[0.001]	[0.019]
Ass.: Unobserved liquidation returns = -50%			
GPME	0.633	0.614	0.698
	(0.172)	(0.155)	(0.272)
	[0.000]	[0.000]	[0.010]
SDF parameters			
a	0	0.033	0.096
		(0.101)	(0.132)
b_1	1	1.444	1.484
		(1.204)	(1.402)
b_2			-2.057
			(1.357)

Table VI - Continued

	(i) Log-utility CAPM	(ii) CAPM	(iii) CAPM w/ small growth
Panel B: Round start date < 1998 ($N = 3,896$)			
Ass.: Unobserved liquidation returns = -90%			
GPME	0.672	0.250	0.179
	(0.085)	(0.213)	(0.308)
	[0.000]	[0.240]	[0.563]
		SDF parameters	
a	0	1.054	2.454
		(0.147)	(0.322)
b_1	1	10.345	21.225
		(1.252)	(4.089)
b_2			-30.498
			(6.411)
Panel C: Round start date ≥ 1998 and < 2001 ($N = 4,268$)			
Ass.: Unobserved liquidation returns = -90%			
GPME	0.907	1.173	1.092
	(0.368)	(0.569)	(0.523)
	[0.000]	[0.039]	[0.037]
		SDF parameters	
a	0	-0.053	-0.076
		(0.019)	(0.035)
b_1	1	-0.918	-1.305
		(1.792)	(2.137)
b_2			1.284
			(2.460)

Table VI - Continued

	(i) Log-utility CAPM	(ii) CAPM	(iii) CAPM w/ small growth
Panel D: Round start date $\geq$ 2001 ($N = 1,487$)			
Ass.: Unobserved liquidation returns = -90%			
GPME	-0.323	-0.326	-0.248
	(0.045)	(0.084)	(0.154)
	[0.000]	[0.000]	[0.108]
	SDF parameters		
a	0	-0.016	-0.052
		(0.299)	(0.060)
b_1	1	-0.299	-5.001
		(1.748)	(3.089)
b_2			10.178
			(4.963)
Panel E: GPME by stage			
Ass.: Unobserved liquidation returns = -90%			
Seed	1.353	1.246	0.909
(N = 398)	(0.167)	(0.344)	(0.407)
	[0.000]	[0.000]	[0.026]
Early	0.543	0.553	0.625
(N = 6,230)	(0.238)	(0.256)	(0.319)
	[0.000]	[0.031]	[0.050]
Late	0.283	0.255	0.310
(N = 2,804)	(0.170)	(0.154)	(0.221)
	[0.000]	[0.097]	[0.161]
Mezzanine	0.312	0.299	0.447
(N = 219)	(0.158)	(0.151)	(0.169)
	[0.000]	[0.048]	[0.008]

estimates of arithmetic alphas.

Adding a small-growth factor to the SDF in column (iii) further raises the GPME. As in the VC funds data, this reflects the relatively poor performance of small growth stocks over the sample period and the positive exposure of start-up companies to the small-growth factor.[19] Across all specifications, the J-test rejects the null hypothesis that the GPME is zero in all cases at conventional significance levels. As Panel A further shows, making a less conservative assumption about the unobserved liquidation returns raises the GPME estimates.

One reason why these estimates are higher than the GPMEs for VC funds is that we are looking at the returns gross of fees to General Partners (before management fees and carried interest), whereas fund-level GPMEs are net of fees. Another reason is that the start-up company sample is weighed more towards earlier years, when VC performed better, compared to the funds data, which has a higher proportion of observations after the year 2000 when VC returns were poor (see for example, Korteweg and Sorensen (2010), Harris, Jenkinson, and Kaplan (2014)).

Although some discrepancy is to be expected since carried interest fees are higher when performance is good, a simple back-of-the-envelope calculation suggests that fees are not the full explanation. Assume a fund with a standard fee structure of 20% carried interest and 2% per year management fees (Metrick and Yasuda (2010)). Suppose all of the fund's investments have a GPME of 0.516 (column (ii) in Table VI, Panel A), then the GPME to LPs is roughly 80% of $0.516 = 0.413$, minus the present value of the management fees of 0.154 (based on a 5% discount rate over 10 years), resulting in a pseudo fund GPME of 0.259. This gets closer, but does not fully match the GPME of 0.111 from pre-1998 VC funds from Table II column (ii).

In Online Appendix F we look at a sample of start-up companies that had their first round

[19]With more than one risk factor in the SDF, it is no longer possible, even in the log-normal special case, to read off the sign of risk premia directly from the factors' price of risk coefficients in the SDF (unless the factors are uncorrelated), because translating one into the other involves multiplication with the inverse of the factors' covariance matrix.

before 1998, but where we include all subsequent funding rounds for these firms. This way of splitting the sample should generate an even closer match to the pre-1998 vintage funds sample in Table II. We find that the GPME estimates are close to those in Panel A of Table VI and, hence, still very different from the pre-1998 vintage funds results. This suggests that there are still some residual sampling differences in the pre-1998 period between the funds and round-to-round data sets. In Online Appendix F we also show that alternative assumptions about the unobserved acquisition returns help to close the gap to the fund results.

Finally, the additional tests in Online Appendix F with firms that had their first funding round before 1998 are also useful to evaluate performance in a way that reduces the impact of the right-censoring problem. For these firms, it is considerably more likely that the eventual final outcome is recorded in the data set. The right-censoring issue should therefore be much less of a concern for this sample. The fact that we find GPME estimates that are close to those in Panel A of Table VI suggests that the high GPME estimates in the pre-1998 period are, at least for the most part, not a consequence of the right-censoring problem.

Panels B, C, and D of Table VI show that the sample period has a large effect on the results. Since the time span between funding rounds is much shorter than the typical life-time of the VC funds that we examined in Section II, the rounds data is better suited for the construction of sub-samples with a big dispersion in public equity market performance. Panel B limits the sample to funding rounds with start dates before 1998, while Panel C looks at returns of rounds started after 1997 and before 2001, while Panel D looks at those started in 2001 and later. The first sub-sample in Panel B was one with extraordinarily high public equity market returns, which is reflected in the high price of risk coefficient $b_1 = 10.345$. This leads to an enormous difference of 0.422 between the GPME in column (ii) and the PME in column (i). The latter is based on the assumption of $b_1 = 1$, which is grossly at odds with the data in this period and which leads the PME to strongly overstate abnormal performance of VC investments. In contrast, public equity markets performed poorly in the sub-samples in Panels C and D, as reflected in the negative b_1 estimates. In these cases, the PME in column

(i) understates the abnormal round-to-round returns.

Panel E examines round-to-round returns from various stages in the life-cycle of start-up companies. With regards to differences between PME and GPME, as in the full sample, these cross-sectional splits produce only small differences between PME and GPME.

Table VII considers artificially levered round-to-round returns. With round-to-round returns, $R^h_{i,t+h}$, and matched Treasury Bill returns, $R^h_{f,t+h}$, the artificially levered round-to-round returns are given by

$$L^h_{i,t+h} = R^h_{i,t+h} + k\left(R^h_{i,t+h} - R^h_{f,t+h}\right). \tag{28}$$

We apply the SDFs from Table VI Panel A, to these artificially levered funds for a range of different leverage ratios k.

Unlike in the VC funds sample in Section II, we would expect only small discrepancies here between the PME and the GPME because the parameter restrictions of the PME, $a = 0$ and $b_1 = 1$, are close to being true in the full round-to-round returns sample (see Panel A of Table VI). Consistent with this logic, Table VII shows that the discrepancies are small indeed. For example, going from no leverage ($k = 0$) to three-to-one leverage ($k = 2$) raises the GPME exactly three-fold, while the PME rises just a little more than that.

Overall, our results demonstrate that the risk-adjustment implicit in the PME is well-behaved—for example, in the sense that abnormal returns scale correctly with the degree of leverage—only if the assumptions of the PME are coincidentally satisfied in a sample of data. If these assumptions are at odds with the data, as in the VC funds sample in Table III and the pre-1998 rounds data in Table VI, the PME delivers a distorted estimate of abnormal returns.

In addition, the analyses with artificially levered funds or round-to-round returns also underscores that it is useful to estimate the SDF parameters in the GPME method with data drawn from the same sample period as the VC payoffs that are being valued. The abnormal return properly scales with leverage only if the SDF perfectly prices the Treasury Bill payoffs

Table VII
Generalized Public Market Equivalents for Artificially Levered VC Round Investments

We use the round-to-round returns, $R^h_{i,t+h}$, and the matched Treasury Bill returns, $R^h_{f,t+h}$ to create artificially levered round-to-round returns

$$L^h_{i,t+h} = R^h_{i,t+h} + k\left(R^h_{i,t+h} - R^h_{f,t+h}\right).$$

We apply the SDFs from Table VI Panel A, to these artificially levered funds for a range of different leverage ratios k and we report the resulting GPME estimates. Standard errors of the GPME are in parentheses, and p-values of the J-test of $GPME = 0$ are in square brackets. The spectral density matrix used in the computation of the J-statistic takes into account error dependence arising from overlapping fund life times as described in the text.

k	(i) Log-utility CAPM	(ii) CAPM
-0.8	0.089	0.103
	(0.040)	(0.070)
	[0.000]	[0.142]
-0.6	0.203	0.206
	(0.062)	(0.069)
	[0.000]	[0.003]
0	0.544	0.516
	(0.195)	(0.180)
	[0.000]	[0.004]
1	1.114	1.032
	(0.431)	(0.432)
	[0.000]	[0.017]
2	1.683	1.547
	(0.670)	(0.688)
	[0.000]	[0.025]

during the VC data sample period. If one estimated the parameters from a longer sample (or, hypothetically, used the population values of the parameters), this would not be true.

IV. Conclusions

We propose a stochastic discount factor approach to valuation of private equity payoffs. Our approach shares the advantages of the Public Market Equivalent (PME) method in that it avoids strong distributional assumptions and is well-suited for irregularly-spaced, skewed, and endogenously timed payoffs. Our method generalizes the PME by allowing the SDF to reflect the realized equity premium and risk-free rate in public markets during the sample period. In this way, the performance benchmark for VC payoffs properly adjusts for systematic risk exposure. We also devise a procedure to conduct statistical inference that is robust to cross-sectional dependence in payoffs.

We show that that there can be substantial differences in abnormal performance estimates between the PME and our generalized PME (GPME). Unless the assumptions about equity premium and risk-free rate implicit in the PME are coincidentally satisfied in a particular sample, the abnormal performance estimates obtained from the PME can be severely distorted. These distortions are particularly large in samples in which public equity markets earned high rates of return such as the 1980s and 1990s. We construct artificially levered funds to show that these distortions are especially large for assets with high-beta payoffs.

The GPME fills a methodological gap between the private equity literature and performance evaluation methods in mutual funds and hedge funds research. With the linear factor models that are common in the mutual funds and hedge funds literature, it is straightforward to adjust a fund's performance by contemporaneously realized risk factor returns multiplied with the risk exposures to each factor. With irregularly-spaced and skewed VC cash flow data, linear factor models are not easily applicable without strong distributional assumptions. The GPME method avoids these strong assumptions, but is effective in performing risk-adjustments in a similar way as linear factor models do.

In this paper we have focused on the risk and return to venture capital. Our methodology can also be applied to other infrequently traded assets, in particular those with highly levered or option-like returns, such as real estate or leveraged buyouts.

References

Brown, Gregory W., Oleg Gredil, and Steven N. Kaplan, 2013, Do private equity funds game returns?, Working paper, University of Chicago.

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

Conley, T.G., 1999, GMM estimation with cross sectional dependence, *Journal of Econometrics* 92, 1 – 45.

Constantinides, George M, 1986, Capital market equilibrium with transaction costs, *Journal of Political Economy*, 842–862.

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.

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.

Ewens, Michael, and Matthew Rhodes-Kropf, 2013, Is a VC partnership greater than the sum of its partners?, Working paper, Carnegie Mellon University and Harvard University.

Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 23–49.

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

Giovannini, Alberto, and Philippe Weil, 1989, Risk aversion and intertemporal substitution in the capital asset pricing model, Working paper, NBER.

Gompers, Paul, and Josh Lerner, 1999, An analysis of compensation in the us venture capital partnership, *Journal of Financial Economics* 51, 3–44.

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

Hall, Alastair R., 2000, Covariance matrix estimation and the power of the overidentifying restrictions test, *Econometrica* 68, 1517–1527.

Hall, Robert E., and Susan E. Woodward, 2007, The incentives to start new companies: Evidence from venture capital, Working paper, NBER.

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

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, *The Journal of Finance* 60, 1791–1823.

Kaplan, Steven N., Berk A. Sensoy, and Per Strömberg, 2002, How well do venture capital databases reflect actual investments, Working paper, University of Chicago.

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

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

Ljungqvist, Alexander, and Matthew Richardson, 2003, The cash flow, return and risk characteristics of private equity, Working paper, NYU Stern.

Longstaff, Francis A., 2009, September, Portfolio claustrophobia: Asset pricing in markets with illiquid assets, *American Economic Review* 99, 1119–44.

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

Moskowitz, Tobias J, and Annette Vissing-Jorgensen, 2002, The returns to entrepreneurial investment: A private equity premium puzzle?, *American Economic Review* 92, 745–778.

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

Phalippou, Ludovic, and Oliver Gottschalg, 2009, The performance of private equity funds, *Review of Financial Studies* 22, 1747–1776.

Phalippou, Ludovic, and Maurizio Zollo, 2005, What drives private equity fund performance?, Working paper, Wharton Working Paper 05-42.

Robinson, David T., and Berk A. Sensoy, 2011, Cyclicality, performance measurement, and cash flow liquidity in private equity, Working paper, Ohio State and Duke University.

Schmidt, Daniel M, 2006, Private equity versus stocks: Do the alternative asset's risk and return characteristics add value to the portfolio?, *The Journal of Alternative Investments* 9, 28–47.

Shanken, Jay, 1992, On the estimation of beta-pricing models, *Review of Financial Studies* 5, 1–33.

Sorensen, Morten, and Ravi Jagannathan, 2013, The public market equivalent and private equity performance, Working paper, Columbia University and Northwestern University.

Sorensen, Morten, Neng Wang, and Jinqiang Yang, 2014, Valuing private equity, *Review of Financial Studies*, forthcoming.

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.