Online Appendix
for
"Risk-Adjusting the Returns to Venture Capital"

Arthur Korteweg*
University of Southern California

Stefan Nagel†
University of Michigan, NBER and CEPR

February 2015

This appendix presents a number of supporting results, extensions, and robustness checks to supplement the analysis in the main paper. Section A provides a detailed derivation of the results for the log-normal model referred to in the introduction of the main paper. Section B discusses the relationship between our method and tests with linearized versions of SDFs that are common in the literature. Section C conducts Monte Carlo experiments to study the size of the tests we use to evaluate VC fund and start-up company payoffs. Section D looks at how the estimation of SDF parameters affects the size and power of tests. Section E provides additional detail on the data sets. Section F presents robustness checks where we consider further subsample results, including a sample of liquidated funds, and we explore the sensitivity of our findings to the assumptions about the unobserved acquisition returns of start-up company investments. In Section G, we compare our approach to earlier methods in the literature that are based on a log-normal model of returns.

*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

A Expected Returns in the Jointly Log-Normal Case

We show results for expected returns for the special case of jointly log-normal market returns and asset returns $R_{i,t+1}$ for the general SDF specification

$$M_{t+1} = \exp(a - br_{m,t+1}). \tag{OA.1}$$

The PME method discussed at the beginning of the introduction of the main paper is a special case with $a = 0$ and $b = 1$. Substituting the SDF into the valuation equation

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

and taking logs, we get, using the properties of the log-normal distribution,

$$\begin{aligned}
0 &= E[r_{i,t+1} + m_{t+1}] + \frac{1}{2}\text{Var}\left(r_{i,t+1} + m_{t+1}\right) \\
&= E[r_{i,t+1}] + a - bE[r_{m,t+1}] + (1/2)\text{Var}\left(r_{i,t+1}\right) + (1/2)b^2\text{Var}\left(r_{m,t+1}\right) - b\text{Cov}\left(r_{i,t+1}, r_{m,t+1}\right) \\
&= \log E[R_{i,t+1}] + a - b\log E[R_{m,t+1}] + b^2\text{Var}\left(r_{m,t+1}\right) - b\text{Cov}\left(r_{i,t+1}, r_{m,t+1}\right) \qquad \text{(OA.3)}
\end{aligned}$$

In the PME special case with $a = 0$ and $b = 1$, this yields the result stated in the introduction

$$\log E[R_{i,t+1}] = \log E[R_{m,t+1}] - \sigma_m^2 + \beta_i\sigma_m^2, \tag{OA.4}$$

where $\sigma_m^2 \equiv \text{Var}\left(r_{m,t+1}\right)$ and $\beta_i \equiv \text{Cov}\left(r_{i,t+1}, r_{m,t+1}\right)/\sigma_m^2$. To obtain the CAPM without restricting the risk-free rate and the equity premium, we can use the general case of (OA.3) applied to the risk-free asset (with $\beta_i = 0$) and the market portfolio (with $\beta_i = 1$), and we obtain two equations that we can solve for

$$a = -r_f + \frac{\log E[R_{m,t+1}] - r_f}{\mathrm{Var}\,(r_{m,t+1})} \log E[r_{m,t+1}] - \left(\frac{\log E[R_{m,t+1}] - r_f}{\mathrm{Var}\,(r_{m,t+1})} \right)^2 \mathrm{Var}\,(r_{m,t+1}) \quad \text{(OA.5)}$$

$$b = \frac{\log E[R_{m,t+1}] - r_f}{\mathrm{Var}\,(r_{m,t+1})} \quad \text{(OA.6)}$$

Using this solution in (OA.3), yields

$$\log E[R_{i,t+1}] - r_f = \beta_i \left(\log E[R_{m,t+1}] - r_f \right), \quad \text{(OA.7)}$$

the result stated in the introduction.

B Non-linear Payoffs and Linearized Factor Models

A common approach in the VC and Private Equity area 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)). Performing a first-order Taylor approximation of

$$M_{t+h}^h = \exp(ah - bf_{t+h}^h) \quad \text{(OA.8)}$$

around $F_{t+h}^h = 1$ and $h = 1$ yields

$$\widetilde{M}_{t+h}^h \approx \exp(a) \cdot [1 - b(F_{t+h}^h - 1) + a(h-1)]. \quad \text{(OA.9)}$$

Redefining parameters,

$$\widetilde{M}_{t+h}^h \approx c + \tilde{a} \cdot h - \tilde{b}(F_{t+h}^h - 1), \quad \text{(OA.10)}$$

which implies a linear beta-pricing specification for expected returns,

$$E[R^h_{i,t+h}] - R^h_F = \beta^h \left[\tilde{b}\mathrm{Var}\,(F^h_{t+h})R^h_F\right], \tag{OA.11}$$

where $\beta^h \equiv \frac{\mathrm{Cov}\,(R^h_{i,t+h}, F^h_{t+h})}{\mathrm{Var}\,(F^h_{t+h})}$.

As we point out in the main paper, there is not really any need to linearize the model, and the linearized model is cumbersome for multi-period payoffs, as it loses the nice compounding properties of the exponential-affine model. Linearized approximations like (OA.10) (and hence linear beta-pricing formulations for expected returns) could also lead to specification errors because it is possible to have negative realizations of the SDF, implying that some states of nature have negative state prices, which is inconsistent with the absence of arbitrage opportunities. This is especially problematic for assets with highly non-linear payoffs such as options.

The highly nonlinear nature of venture investments therefore suggests that linear approximations of the SDF are best avoided. The SDF approach underlying our calculation of the GPME can handle arbitrary payoff non-linearities, and the exponentially-affine specifications of SDFs that we examine in the main paper are strictly positive and hence consistent with the absence of arbitrage opportunities. Thus, these SDFs can, in principle, price any payoff, including option payoffs.

C Monte Carlo evidence on test size

We simulate the finite-sample distribution of the J-statistic that we use in the main analysis to evaluate the performance of VC funds and VC investments in start-ups. Our emphasis in these simulations is on finite-sample distortions in the size of the J-test that may arise in short samples with small T that are typical for data sets on private equity and venture capital, including the data sets that we use in our empirical analysis. We also explore the role of cross-sectional correlation of shocks. The estimator of the spectral density matrix that we propose in the paper is designed to take such correlation into account, but it is an open question to what extent this improves the performance of the J-test in small-T samples.

We assume a constant log risk-free rate r_f, and we simulate single period log returns (for $t = 1, 2, ..., T$) on the market factor as

$$f_t = r_f + \gamma\sigma^2 - \frac{1}{2}\sigma^2 + \sigma\varepsilon_t \tag{OA.12}$$

where $\varepsilon_t \sim \mathcal{N}(0, 1)$. This assumption implies that the log SDF, $m_t = a - \gamma f_t$, prices the risk-free rate and market factor perfectly in population, as one can verify by plugging this expression for f_t into the population pricing equation $E[\exp(f_t)\exp(m_t)] = 1$.[1] We simulate N vectors of (unobserved) single-period fund log returns (for $t = 1, 2, ..., T$) as

$$r_t = f_t + \eta_t - \frac{1}{2}\omega^2 \tag{OA.13}$$

where $\eta_t \sim \mathcal{MVN}(0, \Gamma)$ and $\Gamma = [\rho + I_N(1 - \rho)]\omega^2$, with I_N an identity matrix. Thus, the vector η_t is drawn from a multivariate normal distribution with identical variances of ω^2 for each element and pairwise correlation of ρ. A correlation of $\rho > 0$ is plausible: it reflects the fact that market-adjusted VC investment payoffs are positively correlated. For example, payoffs may be correlated because many VC investments are in the same industries. Substituting (OA.13) into the pricing equation, one can verify that the same log SDF also prices these returns correctly.

The observed fund payoffs are generated from the underlying r_t process. We assume that each fund has a start date, s, at which an initial investment of \$1 is made, followed by a series of J positive cash flows that are realized within a time span of h^{max}, i.e., until $s + h^{max}$.[2] For each of the N funds, we draw s, an integer, randomly from a discrete uniform distribution on the integer interval $[1, T - h^{max}]$. We further draw the J cash-flow realization dates $h(j)$, for

[1] With $a = (\gamma - 1)r_f + \frac{1}{2}\gamma(\gamma - 1)\sigma^2$

[2] The definition of the number of cash flows, J, here is slightly different from the main text where it also includes the initial investment.

$j = 1, 2, ..., J$, from a discrete uniform distribution on the integer interval $[1, h^{max}]$. With the initial investment and cash-flow realization dates drawn for each of the $i = 1, 2, .., N$ funds, we construct the cash flows, in levels, by cumulating the single period returns from (OA.13)

$$C_{i,s+h(j)} = \frac{1}{J} \exp\left(\sum_{\tau=1}^{h(j)} r_{i,s+\tau}\right). \tag{OA.14}$$

In similar fashion, we construct time-period matched payoffs on the market factor and the risk-free asset as

$$F^{h(j)}_{s+h(j)} = \frac{1}{J} \exp\left(\sum_{\tau=1}^{h(j)} f_{s+\tau}\right) \tag{OA.15}$$

$$R^{h(j)}_{f,s+h(j)} = \frac{1}{J} \exp\left(\sum_{\tau=1}^{h(j)} r_{f,s+\tau}\right). \tag{OA.16}$$

The division by J in all three cases ensures that the (stochastically) discounted sum of each cash flow stream has a price of one.

We then use the simulated data to estimate the CAPM SDF from the time-period matched payoffs on the market factor and the risk-free asset and we evaluate the pricing errors for the simulated VC fund payoffs in the same way as in the empirical analysis in the main text. We repeat this 10,000 times and record the J-statistic in each simulation run.

We simulate the model at annual frequency, with $\sigma = 0.15$ and $\gamma = 2$, which implies an equity premium of $\gamma\sigma^2 = 0.045$. The log risk-free rate is $r_f = 0.02$. The parameters of the idiosyncratic shocks are $\omega = 0.25$ and $\rho = 0.1$. We set $N = 700$.

To generate simulated data that resemble the VC funds data, we set $J = 25$ (together with the initial cash outflow this matches the median number of 26 cash flows in Table I in the main paper), and $h^{max} = 11$, which leads to the median fund having cash flows realized over an 11 year period (equal to the median in Table I).

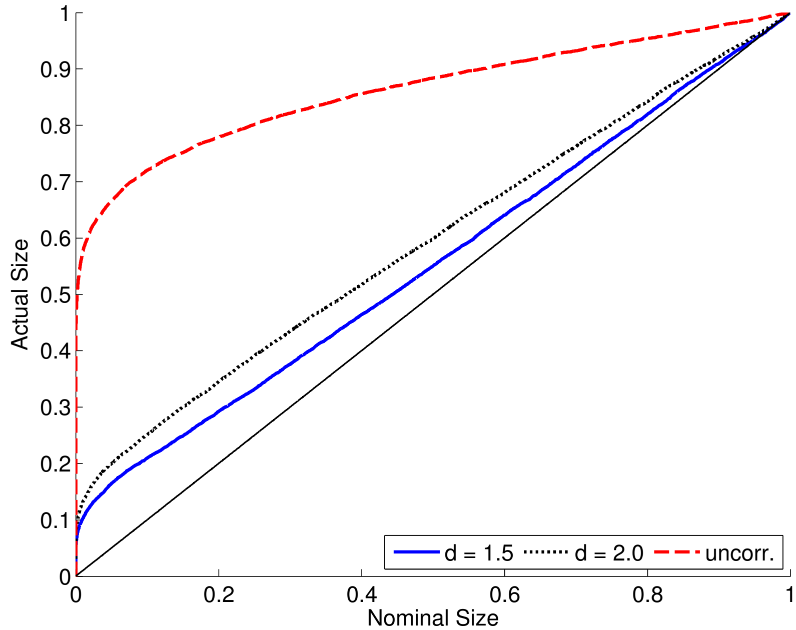
To generate simulated data that are similar to the start-up company data in our empirical analysis, we set $J = 1$ and $h^{max} = 1$, reflecting the fact that a typical round-to-round return is measured over a one-year period. The other parameters are set to the same values as in the VC fund data simulation.

For the sample size in the time dimension, T, we choose two values: $T = 25$, which is between the the actual sample sizes of the VC funds and start-up company samples in our analysis in the main paper, and $T = 200$ to check how the tests behave in a much larger sample. We further check the finite-sample properties of the J-test under three specifications of the spectral density matrix (see equation (26) in the main text): (i) $d = 1.5$, (ii) $d = 2$, (iii) assuming u_i and u_k are uncorrelated if $i \neq k$.

Figure OA.1 shows the results for the simulated VC fund data. The plots show the empirical cumulative distribution of the p-values of the J-test based on the simulated data. Under the asymptotic distribution of the test statistic under the null hypothesis, the p-values should have a uniform distribution and hence the empirical cumulative distribution should line up with the 45 degree line shown in the plot, i.e., nominal size should equal actual size. For example, focusing on Panel (a) where $T = 25$, if we pick the point 0.1 on the horizontal axis, and we look at the solid line corresponding to $d = 1.5$, we are asking how often the J-test rejects (about 20% of the time) if we set the nominal size of the test to 10%. Thus, the plot shows that the test is oversized in this case of a short sample, i.e., it rejects too often. The tendency to overreject is somewhat stronger for $d = 2$. If one uses a spectral density estimator that assumes uncorrelatedness, as shown by the dashed line, the overrejection problem is much worse. The reason is that this estimator severely understates the estimation uncertainty by ignoring the cross-correlations in pricing errors induced by the overlapping nature of the data.

Panel (b) shows that the J-test with $d = 1.5$ or $d = 2$ has a weak tendency to underreject in a larger sample with $T = 200$. This sample size is much larger than the sample sizes available in typical empirical analyses. However, the version of the J-test that uses the spectral density estimator with uncorrelatedness remains substantially oversized.

(a) $T = 25$



(b) $T = 200$

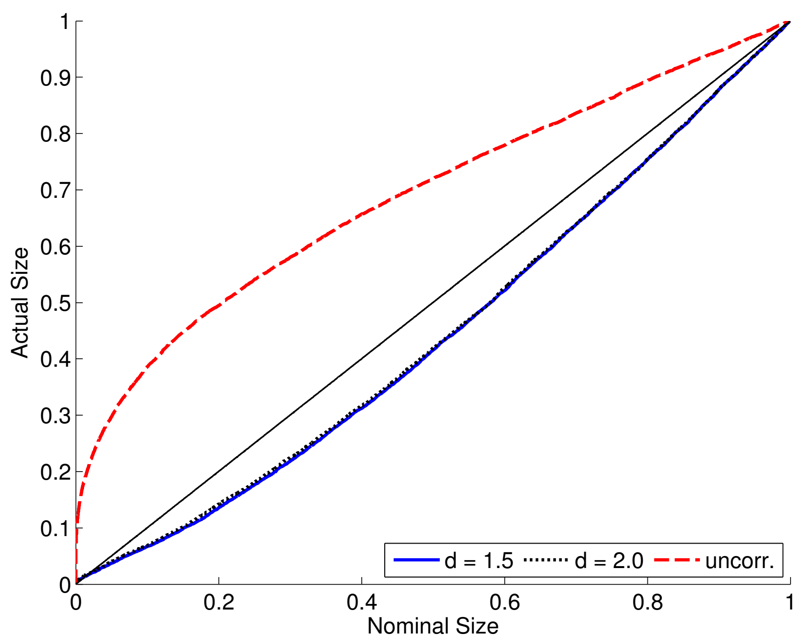


Figure OA.1: Actual and nominal size of the J-test in simulated VC funds data ($J = 25$ and $h^{max} = 11$)

Figure OA.2 presents the results for the simulated start-up company data. In these data, the degree of overlap in time of the payoff observations is much smaller. As a consequence, the tests behave quite differently. For $d = 1.5$ or $d = 2$ we get undersized tests even with the short $T = 25$ sample in Panel (a). Only the test that assumes uncorrelatedness in the construction of the spectral density matrix overrejects, and does so to a substantial extent. With a large sample in Panel (b), the actual sizes converge closer to the nominal size.

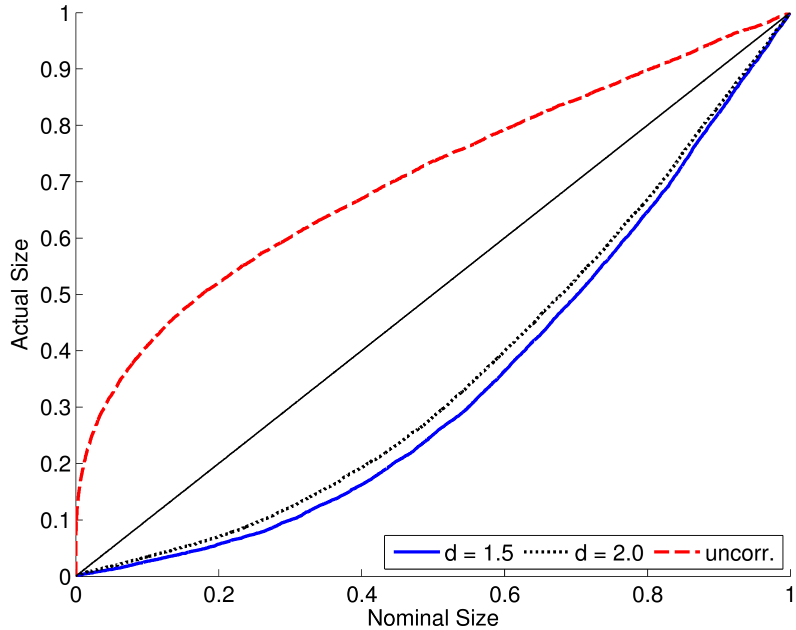
Overall, we draw the following conclusions from these Monte Carlo experiments. First, assuming uncorrelatedness in the construction of the spectral density matrix leads to severe overrejection. Second, the cross-correlation robust estimator that we propose in the main paper reduces this tendency to overreject. Even so, in short samples with a very high degree of time overlap in return observations, the J-test still has size distortions. Small-sample problems of this sort are common with GMM estimators (see, e.g., Hansen, Heaton, and Yaron (1996)) and our proposed estimator is no exception.

D Monte Carlo evidence on the effect of SDF parameter estimation

This section compares the size and power of tests when the SDF parameters are estimated from sample data with the size and power when the SDF parameters are set to their true values. The former case corresponds to using the ex-post equity premium during the performance evaluation period as a performance benchmark. The latter case corresponds to using the ex-ante equity premium or the asymptotic limit in a large sample. The case in which the SDF parameters are estimated from a sample of public markets data that is larger in the time dimension than the VC investments data sample, but not necessarily large enough to get close to the asymptotic limit, would fall between those two cases.

The simulations are similar to those reported in Section C of this appendix, only now with one set of simulations in which the J-statistic is calculated based on estimated SDF

(a) $T = 25$



(b) $T = 200$

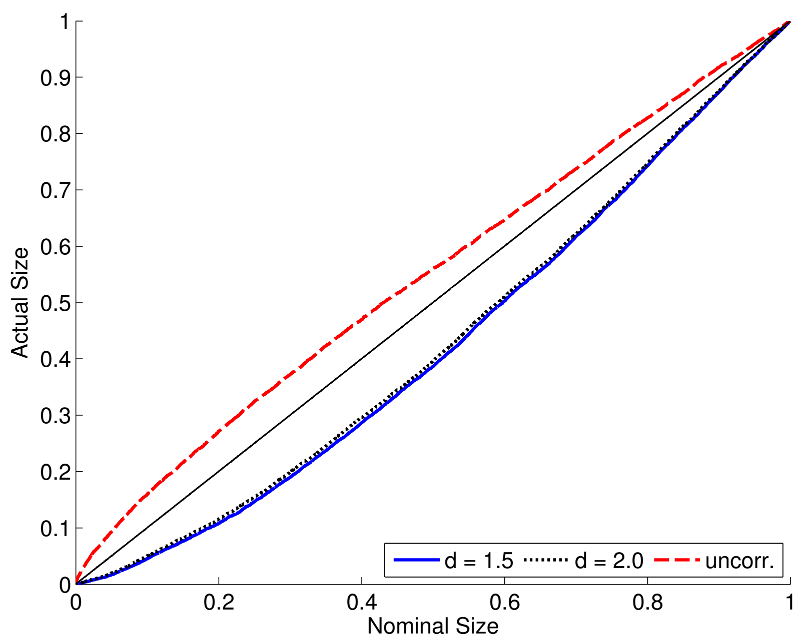


Figure OA.2: Actual and nominal size of the J-test in simulated start-up company data ($J = 1$ and $h^{max} = 1$)

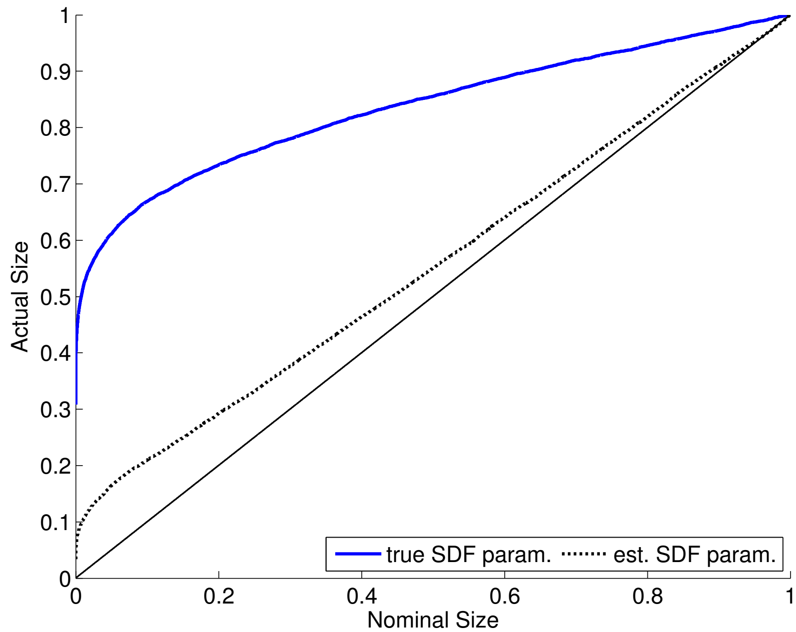
parameters (as in Section C), and another set in which the J-statistic is calculated based on true values of the SDF parameters. All simulations in this section use the cross-correlation robust spectral density matrix with $d = 1.5$.

Panel (a) of Figure OA.3 shows how SDF parameter estimation affects the size of the test with simulated VC funds data in the case of $T = 25$. As the figure shows, both versions of the tests have a tendency to overreject, but this size distortion is much worse for the test based on true SDF parameters. The additional noise from using the ex-ante rather than ex-post equity premium as a benchmark for performance comparison amplifies the size distortions.

To investigate power, we simulate data with fixed misspecification, where the SDF produces a pricing error of $0.30 (for a total investment of $1). Panel (b) plots the rejection probability against the actual size of the test. Plotting power against the actual rather than nominal size corrects for the size distortions that are evident in Panel (a). For example, actual size of 10% means that we set the critical value for the J-test to a value that leads it to reject under the null hypothesis—in terms of actual small-sample rejection probabilities—10% of the time. As Panel (b) shows, with this size-distortion-adjusted critical value, the test based on estimated SDF parameters rejects about 55% of the time under our misspecification alternative when the actual size of the test is 10%. Thus, the test has some power, even in this very short sample. In contrast, for the test based on true SDF parameters we cannot even evaluate power at actual size of 10%, because the the size distortions are so severe that there is no critical value for the J-statistic at which the test rejects less than 35% of the time under the null hypothesis in our simulated sample of 10,000 draws. This can be seen in Panel (a) where the actual size hits the vertical axes at a value above 35%. For an actual size above 35%, however, is clear from Panel (b) that the test based on estimated SDF parameters is much more powerful.

Figure OA.4 repeats this analysis with simulated start-up company data, again with $T = 25$, but now with misspecification of $0.10 (for a total investment of $1, which pays off over a shorter horizon than in the VC fund data simulations). The results are similar

(a) Size



(b) Power

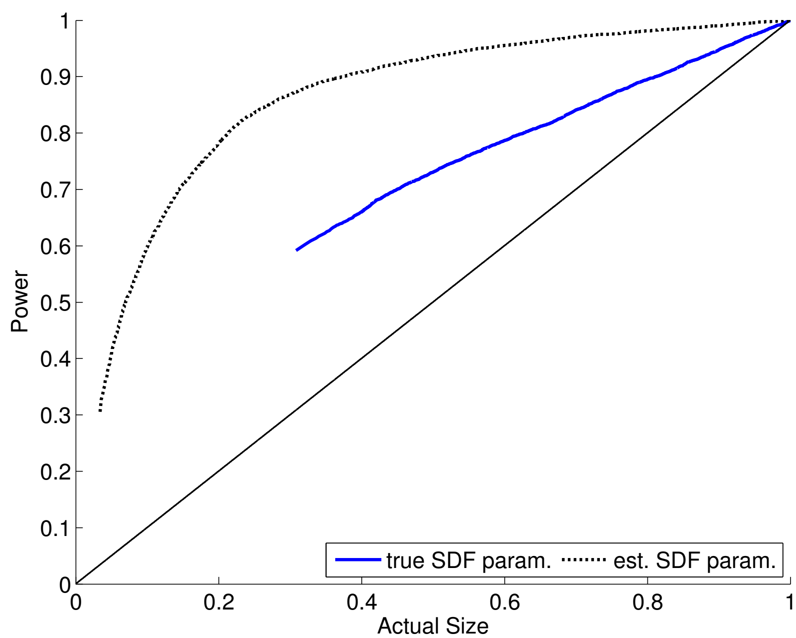


Figure OA.3: Comparison of size and power of the J-test in simulated VC funds data ($J = 7$, $h^{max} = 13$, $d = 1.5$, and $T = 25$) when SDF parameters are estimated or set to true values. Power in Panel (b) is evaluated with misspecification fixed at a pricing error of \$0.30 for a total investment of \$1.

to those for the simulated VC funds data. The only difference is that the size distortions in Panel (a) are now less severe. The test with true SDF parameters now exhibits almost no size distortions, while the test based on estimated SDF parameters has a tendency to underreject. Because size distortions are less severe, we can now also evaluate power over the whole range of actual size all the way to zero. As Panel (b) shows, the test based on estimated SDF parameters is again more powerful than the test based on true SDF parameters. For example, at an actual size of 10%, the test based on estimated SDF parameters rejects more close to 100% of the time under our misspecification alternative, while the test based on the true SDF parameters rejects only about 85% of the time.

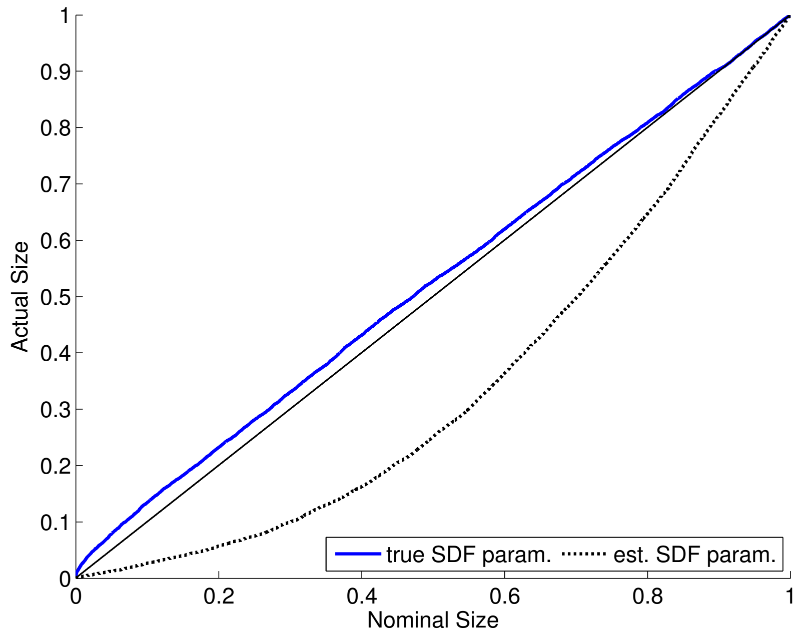
We have also investigated size and power with larger sample sizes such as $T = 200$. The result that the J-test is more powerful when SDF parameters are estimated over the performance evaluation period holds in this case, too.

These results confirm the intuition that we discussed in the main paper according to which performance evaluation metrics are less noisy when computed relative to the ex-post equity premium during the performance evaluation period (which is captured by the estimated SDF parameters) rather than the ex-ante equity premium (which is reflected in the true parameter values or in estimates from a long time series of factor returns).

E Data

For robustness, we construct a sample of funds that are nearly fully liquidated. This sample includes those funds with NAVs that are less than 5% of the fund's size by the end of the sample period, but that have called at least 50% of their total commitment at some point in the past (to exclude funds that report low NAVs because they have only recently started investing). The columns labeled "cash flow sample" in Table OA.1 compare the liquidated funds sample with the sample used in the paper. The table shows that the 181 nearly liquidated funds are smaller and show better performance than the sample of funds from the paper, with median (average) IRR of 10.04% (20.50%) and TVPI of 1.46 (2.38). This

(a) Size



(b) Power

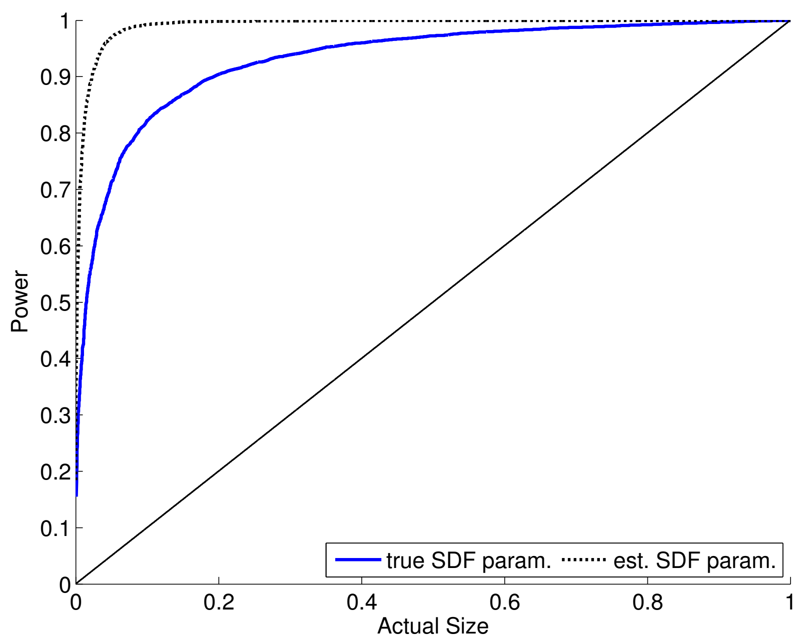


Figure OA.4: Comparison of size and power of the J-test in simulated start-up company data ($J = 1$, $h^{max} = 1$, $d = 1.5$, and $T = 25$ when SDF parameters are estimated or set to true values. Power in Panel (b) is evaluated with misspecification fixed at a pricing error of \$0.10 for a total investment of \$1.

better performance is driven not only by the fact that most distributions arrive at the end of a fund's life (see, for example, Figure 3 in Hochberg, Ljungqvist, and Vissing-Jorgensen (2014)), but also because most of the liquidated funds have been raised well before the turn of the new millennium when times were relatively good for VC.

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.[3] Although we cannot use this broad sample in our empirical work as it lacks cash flows, we can assess potential selection biases in the cash flow sample by comparing the two data sets. Table OA.1 shows that the broad sample contains 1,074 funds, which is nearly twice the size of the cash flow sample. Funds in the broad sample tend to be smaller (the median fund is $150 million compared to $226 million in the cash flow sample), and there are slightly more funds per VC firm in the broad sample. The median (average) IRR in the broad sample is 7.80% (12.89%), three to four percentage points higher than in the cash flow sample. TVPI is also higher, median 1.37 versus 1.16 in the cash flow sample (average 1.83 versus 1.57). Panel B shows that there are a number of years with large differences in performance between the two samples, such as 1988 and 1997, whereas in many other years the differences are small.[4] The performance difference between the cash flow and broad samples is slightly smaller for the liquidated funds compared to the full sample, reinforcing the importance of assessing the robustness of our results in the subsample of liquidated funds.

Table OA.2 shows the IRR and TVPI for venture capital funds reported in the recent literature. Though not comprehensive, the table illustrates the diversity in data sets and observed performance across studies.[5] The literature focuses almost exclusively on U.S. funds.

[3]Harris, Jenkinson, and Kaplan (2014) suggest that that the performance of all of the funds in Preqin is qualitatively and quantitatively similar to the performance in Burgiss and Cambridge Associates. In particular, the Preqin data appears to be more representative than VentureXpert (from Thomson Venture Economics), which has important updating problems in its post-2001 funds data (see also Stucke (2011)). However, we should be careful to point out that not all funds in Preqin report performance.

[4]The differences-in-means t-statistics in Table OA.1 provide an upper bound on the statistical significance of differences between the two samples as they presume IID sampling (which ignores the likely positive cross-sectional correlation between funds).

[5]We do not show performance results for Ljungqvist and Richardson (2003), whose sample of private data from a large institutional LP contains only 19 venture funds. Phalippou and Gottschalg (2009) and Driessen,

Table OA.1
Summary Statistics: VC Fund Data

This table reports descriptive statistics for the sample of U.S. VC funds from Preqin of vintages between 1979 and 2008, eliminating funds with committed capital below $5 million in 1990 dollars. The "cash flow sample" is the sample of Preqin funds with cash flows between the fund and its limited partners ending December 31, 2012. The "broad sample" is a broader sample of funds that report multiples and IRRs but no cash flows. The left three columns show summary statistics for all liquidated and non-liquidated funds, and the right three columns show the statistics for the sub-sample of liquidated funds. Funds in the cash flow sample are considered liquidated if the net asset value (NAV) is below 5% of the fund size, and the fund was at least 50% invested at some point during its lifetime. *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. The table reports averages, with medians in brackets. The t-statistic for the differences in means test between the cash flow and broad samples is shown in the columns labeled *t-stat*.

Panel A: Descriptive statistics						
	All funds			Liquidated funds		
	Cash flow sample	Broad sample	t-stat	Cash flow sample	Broad sample	t-stat
# Funds	545	1,074		181	395	
# VC firms	278	471		113	221	
# Funds / VC firm	1.96	2.28	-2.47	1.60	1.79	-1.36
	(1.00)	(2.00)		(1.00)	(1.00)	
Fund size ($m)	358.46	269.70	3.21	176.89	110.45	5.01
	(226.00)	(150.00)		(130.50)	(70.50)	
Fund effective years	10.40	-	-	14.09	-	-
	(11.09)	-		(14.08)	-	
# Cash flows / fund	30.28	-	-	33.49	-	-
	(26.00)	-		(29.00)	-	
IRR (%)	8.84	12.89	-2.09	20.50	23.49	-0.63
	(4.37)	(7.80)		(10.04)	(13.40)	
TVPI	1.57	1.83	-2.29	2.38	2.60	-0.76
	(1.16)	(1.37)		(1.46)	(1.86)	

Table OA.1 - Continued

Panel B: Performance statistics by vintage year										
	All funds					Liquidated funds				
	Cash flow sample		Broad sample		t-stat	Cash flow sample		Broad sample		t-stat
	# funds	IRR	#funds	IRR		#funds	IRR	#funds	IRR	
1979	1	18.50	1	18.50	-	1	18.50	1	18.50	-
1980	1	15.26	6	16.47	-0.06	1	15.26	6	16.47	-0.06
1981	1	27.04	7	19.94	0.30	1	27.04	7	19.94	0.30
1982	3	8.83	9	13.90	-0.44	3	8.83	9	13.90	-0.44
1983	2	9.04	10	9.27	-0.04	2	9.04	10	9.27	-0.04
1984	3	11.49	16	13.32	-0.33	3	11.49	16	13.32	-0.33
1985	4	13.17	17	12.76	0.14	4	13.17	16	12.10	0.39
1986	7	5.89	22	9.67	-1.13	7	5.89	22	9.67	-1.13
1987	5	14.69	21	13.11	0.31	5	14.69	19	13.75	0.18
1988	5	17.84	23	23.70	-0.88	4	18.37	22	24.30	-0.79
1989	5	19.61	36	20.88	-0.08	5	19.61	36	20.88	-0.08
1990	7	17.67	21	16.49	0.13	7	17.67	19	15.72	0.21
1991	2	21.04	12	28.63	-0.57	2	21.04	11	25.51	-0.40
1992	12	27.87	22	22.81	0.50	12	27.87	21	22.68	0.50
1993	9	38.98	30	31.11	0.66	9	38.98	29	29.78	0.78
1994	13	34.78	24	29.03	0.48	12	38.53	16	26.72	0.87
1995	16	46.27	29	55.17	-0.33	16	46.27	23	59.36	-0.43
1996	16	33.33	36	31.52	0.12	13	26.32	26	34.51	-0.44
1997	22	28.93	49	46.70	-1.20	18	29.53	28	54.11	-1.28
1998	29	19.94	57	22.97	-0.14	17	33.79	24	25.75	0.21
1999	40	-3.24	79	-0.73	-0.45	22	-3.04	20	1.22	-0.30
2000	67	-3.10	106	0.12	-1.53	8	-13.24	10	-8.31	-0.29
2001	43	1.04	63	4.54	-1.33	7	-17.10	2	3.20	-0.98
2002	25	-1.68	34	-0.70	-0.32	2	6.92	2	-25.60	1.50
2003	19	2.00	36	2.74	-0.23	0	-	0	-	-
2004	26	2.65	46	2.67	-0.01	0	-	0	-	-
2005	28	2.09	60	2.12	-0.01	0	-	0	-	-
2006	52	1.12	74	3.80	-1.56	0	-	0	-	-
2007	46	8.06	66	9.87	-0.55	0	-	0	-	-
2008	36	7.53	62	8.05	-0.17	0	-	0	-	-

One exception is Robinson and Sensoy (2011) who have data on 295 worldwide funds. They do not separately report U.S. fund returns, but their reported statistics are U.S. dominated since 260 of the 295 funds are U.S. funds. Many papers report performance for samples based on selection criteria similar to our "all funds" sample, although there is some variation in the criteria used. Most importantly, like our sample, these samples are restricted to funds that are at least 4 to 5 years old but not necessarily liquidated, and thus have to rely, to some extent, on NAVs to compute performance. Our "all funds" performance measures are similar to Robinson and Sensoy (2011), but lower than Ewens, Jones, and Rhodes-Kropf (2013) and Harris, Jenkinson, and Kaplan (2014). The latter two papers use larger, more comprehensive samples of venture funds. When compared to samples of liquidated funds in the literature, our performance metrics are more in line, albeit slightly higher, than the larger samples of Kaplan and Schoar (2005) and Hochberg, Ljungqvist, and Vissing-Jorgensen (2014).[6] The latter study matches ours most closely in terms of the time period studied. Given the variation in reported performance between our sample and the various data sets in the literature, one should be cautious in generalizing from our findings. Still, our data set should be useful to illustrate the GMM-SDF approach and to examine whether different specifications of the SDF produce different performance assessment.

F Robustness Checks

Table OA.3 shows the GPME estimates for the sample of funds that are liquidated by the end of the sample period. The GPMEs for liquidated funds are higher than for the sample in the paper that includes non-liquidated funds, consistent with the higher IRRs and TVPIs for this

Lin, and Phalippou (2012) use Venture Economics data but their performance statistics are difficult to compare to most of the literature as they do not report equal-weighted results. Jenkinson, Sousa, and Stucke (2013) use data from Calpers but do not show IRRs or TVPIs. Finally, we do not report performance statistics from Ang, Chen, Goetzmann, and Phalippou (2014) as they use the same Preqin cash flow data set as this paper, with minor differences because their sample ends in June 2011 whereas ours ends in December 2012.

[6] Hochberg, Ljungqvist, and Vissing-Jorgensen (2014) do not explicitly limit the sample to liquidated funds but since each fund has at least 10 years of data, most, if not all, of these funds will be either liquidated or have very low residual NAVs.

Table OA.2

VC Fund Performance in the Literature

This table shows the performance of VC funds reported in the literature. KS05 is Kaplan and Schoar (2005), RS11 is Robinson and Sensoy (2011), EJR13 is Ewens, Jones, and Rhodes-Kropf (2013), HLV14 is Hochberg, Ljungqvist, and Vissing-Jorgensen (2014), and HJK14 is Harris, Jenkinson, and Kaplan (2014). The performance measures are equal-weighted across funds. VE stands for Thomson Venture Economics, B is Burgiss, and P is Preqin. Geography refers to the investment focus of funds. The table reports average *IRR* and *TVPI*, with medians in brackets.

	KS05	RS11	EJR13	HLV14	HJK14	this paper
Data source(s)	VE	Large LP	VE, P & LP Source	VE, P	B	P
Sample period	1980–2001	1984–2010	1980–2011	1980–2012	1984–2011	1979–2012
Geography	U.S.	World	U.S.	U.S.	North America	U.S.
			Panel A: All funds			
# Funds		295	1,040		775	545
Vintages		1984–2009	1980–2007		1984–2008	1979–2008
IRR		8%	15.27%		16.8%	8.84%
		(1%)	(6.43%)		(11.1%)	(4.37%)
TVPI		1.38			2.34	1.57
		(1.03)			(1.73)	(1.16)
Panel B:Liquidated funds						
# Funds	577	192		1,052		181
Vintages	1980–1995	1984–2005		1980–2002		1979–2002
IRR	17%	9%		15.17%		20.50%
	(11%)	(2%)		(5.6%)		(10.04%)
TVPI		1.44				2.38
		(1.05)				(1.46)

subsample reported in Table OA.1. Most importantly, the results on the relative differences between the PME and GPME are robust: The PME overstates the return relative to the GPME, and the difference is greater in the pre-1998 period. The exceptions are mezzanine funds (as in the full sample results in the main paper) because they have a beta below one, and pre-1998 vintage funds because the realized market premium was negative for this particular subsample which – following the intuition from equation (25) in the main paper – causes the PME to understate the abnormal return for a beta ¿ 1 security.

Table OA.4 Panel A, reports additional robustness checks for the start-up company round-to-round data. Panel A limits the sample to firms that had their first funding round before the year 1998. This sample is quite similar to the sample in Panel B of Table VI in the main paper that uses data from funding rounds that were started before 1998 only. The one difference is that in Table OA.4 Panel A we also include funding rounds started in 1998 or later, as long as the company's first funding round took place before 1998. As we discuss in Section III.A of the main paper, cutting the sample off after early funding rounds could result in a selection bias. Table OA.4 provides an alternative perspective on the earlier part of the start-up company sample that is robust to this selection problem.

Focusing on the GPME estimates in the first row of Panel A of Table OA.4, the results are sensible. For example, for the CAPM model in column (ii), The GPME estimate of 0.520 is close to the GPME over the full sample period of 0.516 reported in Panel A of Table VI of the paper, suggesting that the results in the paper do not suffer from strong selection bias. Note that the estimate is higher than the pre-1998 estimate of 0.250 in Panel B of Table VI in the paper, due to the inclusion of subsequent funding rounds in the post-1998 period, when GPMEs were higher. The results in columns (i) and (iii) for the PME and the SDF that includes a small-growth factor, respectively, are also close to the full-sample estimates of Table VI Panel A.

Table OA.3

Generalized Public Market Equivalents for VC Funds: Liquidated Sample

This table replicates Table II in the paper for the subsample of VC funds that are liquidated by the end of the sample period. 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 paper.

	(i) Log-utility CAPM	(ii) CAPM	(iii) CAPM w/ small growth
Panel A: All liquidated funds ($N = 181$)			
GPME	0.276	0.019	-0.074
	(0.089)	(0.089)	(0.102)
	[0.002]	[0.834]	[0.470]
SDF parameters			
a	0	0.080	0.109
		(0.044)	(0.050)
b_1	1	2.272	2.255
		(0.541)	(0.639)
b_2			-2.685
			(1.299)
Panel B: All funds of vintage < 1998 ($N = 125$)			
GPME	0.414	0.067	0.074
	(0.111)	(0.129)	(0.277)
	[0.000]	[0.604]	[0.789]
SDF parameters			
a	0	0.319	0.641
		(0.094)	(0.258)
b_1	1	4.400	5.701
		(0.983)	(2.155)
b_2			-9.754
			(8.118)

Table OA.3 - Continued

	(i) Log-utility CAPM	(ii) CAPM	(iii) CAPM w/ small growth
Panel C: All funds of vintage $\geq$ 1998 ($N = 56$)			
GPME	-0.032	0.086	0.135
	(0.132)	(0.165)	(0.192)
	[0.808]	[0.602]	[0.484]
	SDF parameters		
a	0	-0.034	-0.046
		(0.007)	(0.008)
b_1	1	-0.372	-0.833
		(0.409)	(0.453)
b_2			1.452
			(0.454)
Panel D: GPME of funds by stage			
Generalist	0.396	0.028	-0.077
(N = 116)	(0.110)	(0.114)	(0.123)
	[0.000]	[0.805]	[0.532]
Early	0.016	-0.080	-0.196
(N = 31)	(0.130)	(0.118)	(0.106)
	[0.904]	[0.499]	[0.063]
Late	0.329	0.183	0.093
(N = 11)	(0.128)	(0.138)	(0.127)
	[0.010]	[0.185]	[0.468]
Mezzanine	-0.006	0.056	0.061
(N = 23)	(0.042)	(0.056)	(0.054)
	[0.893]	[0.318]	[0.255]
Panel E: GPME of funds by size			
Small	0.156	-0.143	-0.177
(< median)	(0.087)	(0.072)	(0.075)
(N = 81)	[0.074]	[0.048]	[0.018]
Large	0.434	0.240	0.087
($\geq$ median)	(0.131)	(0.136)	(0.113)
(N = 83)	[0.001]	[0.078]	[0.439]

Table OA.4

Generalized Public Market Equivalents for VC Round-to-Round Returns: Robustness Checks

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. 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. In Panel A, we restrict the sample to start-up companies that had their first funding round before 1998. We include all subsequent rounds of those firms. Panel B uses the full sample, but with different assumptions about acquisition returns.

	(i) Log-utility CAPM	(ii) CAPM	(iii) CAPM augm. w/ small growth
Panel A: First round $<$ 1998 ($N = 5,394$)			
Ass.: Unobserved liquidation returns = -90%			
GPME	0.632	0.520	0.485
	(0.066)	(0.130)	(0.239)
	[0.000]	[0.000]	[0.042]
Ass.: Unobserved liquidation returns = -70%			
GPME	0.649	0.553	0.520
	(0.064)	(0.142)	(0.246)
	[0.000]	[0.000]	[0.035]
Ass.: Unobserved liquidation returns = -50%			
GPME	0.667	0.586	0.556
	(0.061)	(0.154)	(0.254)
	[0.000]	[0.000]	[0.029]
SDF parameters			
a	0	0.330	0.451
		(0.176)	(0.207)
b_1	1	3.862	3.916
		(1.255)	(1.379)
b_2			-3.380
			(1.900)

Table OA.4 - Continued

	(i) Log-utility CAPM	(ii) CAPM	(iii) CAPM augm. w/ small growth
Panel B: acquisition return assumptions ($N = 13,345$)			
Ass.: Unobserved acquisition returns = 0%, unobserved liquidation returns = -90%.			
GPME	0.383	0.368	0.407
	(0.173)	(0.163)	(0.212)
	[0.000]	[0.024]	[0.055]
Ass.: Unobserved acquisition returns = 0%, unobserved liquidation returns = -70%.			
GPME	0.431	0.419	0.458
	(0.158)	(0.147)	(0.202)
	[0.000]	[0.004]	[0.023]
Ass.: Unobserved acquisition returns = 0%, unobserved liquidation returns = -50%.			
GPME	0.479	0.470	0.509
	(0.144)	(0.131)	(0.194)
	[0.000]	[0.000]	[0.009]
Ass.: Unobserved acquisition returns = -90%, unobserved liquidation returns = -90%.			
GPME	0.323	0.307	0.344
	(0.181)	(0.173)	(0.218)
	[0.000]	[0.076]	[0.114]
Ass.: Unobserved acquisition returns = -70%, unobserved liquidation returns = -70%.			
GPME	0.385	0.372	0.409
	(0.164)	(0.154)	(0.206)
	[0.000]	[0.016]	[0.047]
Ass.: Unobserved acquisition returns = -50%, unobserved liquidation returns = -50%.			
GPME	0.446	0.436	0.474
	(0.148)	(0.136)	(0.196)
	[0.000]	[0.001]	[0.015]
SDF parameters			
a	0	0.019	0.076
		(0.090)	(0.120)
b_1	1	1.273	1.306
		(1.186)	(1.350)
b_2			-1.987
			(1.280)

Panel B of Table OA.4 checks robustness regarding acquisition returns. Our resampling procedure assumes that the missing acquisition returns are drawn from the same distribution as the observed acquisition returns. Here we check how the results change if we make the (possibly conservative) assumption that unobserved acquisition returns are zero (first three blocks of rows) or equal to the return that we assume for unobserved liquidations (last three blocks of rows). Comparing the results to those in Table VI Panel A, we find that the acquisition return assumption is a significant one. The magnitudes of our GPME point estimates are substantially lower with the more conservative acquisition return assumptions in Table OA.4. For example, for the CAPM in column (ii), assuming an acquisition return of zero rather than sampling the return from the distribution of observed acquisition returns leads to a drop in the GPME from 0.516 to 0.368. However, even the extreme assumption that missing acquisition returns are liquidations does not change the general conclusion that the GPMEs are substantially greater than zero.

The SDF parameters in Panel B of Table OA.4 are slightly different from those in Table VI Panel A of the main paper because the factor returns are different. In the main paper, if a round-to-round return is an unobserved acquisition return, our re-sampling method draws both the acquisition return and the time-matched factor returns from the observed returns sample. Thus, the factor returns associated with the rounds that end with an unobserved acquisition are not included in the analysis. They are replaced by the factor return that is paired with the randomly drawn acquisition return. However, in Panel B of Table OA.4 we replace unobserved acquisition returns with a fixed number, and so in this case the factor returns associated with the rounds ending in an acquisition event are used and not replaced by resampled factor observations.

G Comparison with Log-Normal Model

In order to deal with both the multi-period nature of payoffs and the sample selection problem in the round-to-round returns, an alternative approach in the earlier literature (e.g., Cochrane

(2005) and Korteweg and Sorensen (2010)) makes strong distributional assumptions that lead to a linear factor model in logs that can be paired with a selection model, and estimated with maximum likelihood (ML) methods. We now turn to a brief comparison of our approach with this alternative method.

If asset returns, R_t^h, and the risk factor, F_t, are IID jointly log-normal, the pricing restriction associated with SDF (OA.8), under the null hypothesis $GPME = 0$, and applied to multi-period returns, yields

$$E[r_{t+h}^h] - r_f^h + \frac{h}{2}\sigma^2 = \beta\left(E[f_{t+h}^h] - r_f^h + \frac{h}{2}\sigma_f^2\right), \tag{OA.17}$$

where $\beta \equiv \mathrm{Cov}\,(r_{t+1}, f_{t+1})/\sigma_f^2$ and $\sigma^2 \equiv \mathrm{Var}\,(r_{t+1})$. Allowing for a pricing error, β can be estimated with the regression

$$r_{t+h}^h - r_f^h = g(h) + \beta(f_{t+h}^h - r_f^h) + \varepsilon_{t+h}^h, \tag{OA.18}$$

where comparison with (OA.17) shows that the abnormal return due to the pricing error is

$$\alpha^h = g(h) + \frac{h}{2}(\sigma^2 - \beta\sigma_f^2) \tag{OA.19}$$

$$= g(h) + \frac{h}{2}(\sigma_\varepsilon^2 + \beta(\beta - 1)\sigma_f^2). \tag{OA.20}$$

However, the null hypothesis gives no guidance about the specification of the intercept term $g(h)$ in (OA.18). Cochrane (2005) and Korteweg and Sorensen (2010) assume $g(h) = \gamma h$, which accumulates nicely over time as longer horizon returns are simply summations of shorter-horizon returns, but it is not obvious that this is the correct specification.[7] A perhaps equally plausible alternative specification of randomness in start-up company investments is that nature initially draws whether a project is a success or not. This initial draw fixes the (initially unobserved) abnormal return of the project. Projects differ in the amount of time

[7] Driessen, Lin, and Phalippou (2012) do not assume log-normality, but use a similar compounding of the intercept in a linear factor model.

it takes for the project value to be revealed, but a longer horizon does not imply that more abnormal return will be accumulated, nor that idiosyncratic variance is higher, let alone that it grows linearly in the horizon. Instead, $g(h) = \gamma$ and the variance of ε_{t+h} is constant and equal to σ_ε^2. We get

$$r_{t+h}^h - r_f^h = \gamma + \beta(f_{t+h}^h - r_f^h) + \varepsilon_{t+h}, \tag{OA.21}$$

and

$$\alpha^h = \gamma + \frac{1}{2}(\sigma_\varepsilon^2 + \beta(\beta - 1)h\sigma_f^2). \tag{OA.22}$$

The ad-hoc assumptions about $g(h)$ and σ_ε^h can have a large impact on the empirical estimates. For example, if the process for a start-up's valuation is closer to (OA.21) than to (OA.18) with $g(h) = \gamma h$, this could lead to a large upward bias in the estimate of the arithmetic alpha in a sample with heterogeneous payoff horizons. To see this, suppose the alternative log-normal model (OA.21) is the true model. For a simple illustration, consider the special case where $\beta = 0$ and $r_f = 0$ so that $r_{t+h}^h = \gamma + \varepsilon_t$. We are interested in the arithmetic alpha for an investment with horizon $h = 1$, which, in this $\beta = 0$ case, is simply

$$\alpha^1 = \gamma + \frac{1}{2}\sigma_\varepsilon^2. \tag{OA.23}$$

If the log-normal model (OA.18) with the (false) assumption $g(h) = \gamma h$ is employed, the ML estimator of γ in a large sample is $\hat{\gamma} = \gamma / E[h]$,[8] which implies an arithmetic annualized alpha of

$$\hat{\alpha}^1 = \hat{\gamma} + \frac{1}{2}E\left[\left(\frac{r_{t+h}^h}{\sqrt{h}} - \hat{\gamma}\sqrt{h}\right)^2\right] \tag{OA.24}$$

$$= \frac{\gamma}{E[h]} + \frac{1}{2}\sigma_\varepsilon^2 E\left[\frac{1}{h}\right] + \gamma^2\left(E\left[\frac{1}{h}\right] - \frac{1}{E[h]}\right), \tag{OA.25}$$

which is not generally equal to α^1. The multiplication of γ and σ_ε^2 in the first two terms by

[8] In this case, the ML estimator is equivalent to the OLS slope estimator in a regression of $r_{t+h}^h/\sqrt{h}$ on $\sqrt{h}$ without intercept.

$1/E[h]$ and $E[1/h]$ arises, respectively, because application of the false model (OA.18) with $g(h) = \gamma h$ scales the alpha by horizon, which is not consistent with the true model (OA.21). For example, if all payoffs have $h > 1$, the application of the false model scales the arithmetic alpha towards zero, because the model assumes that $g(h)$ and volatility would be smaller in magnitude at the shorter horizon of $h = 1$. Thus, the effect of the first two terms in (OA.25) depends on the characteristics of the data. In our data we have $1/E[h] < 1$, $\gamma < 0$, and $E[1/h] > 1$ in terms of sample equivalents, which suggests a positive inconsistency. The third term is always positive, which adds to the positive inconsistency, and it is bigger the greater the dispersion of h in the data. Taken together, if (OA.21) is true, but the econometrician applies (OA.18) with the assumption $g(h) = \gamma h$ in estimation, the arithmetic alpha estimates could be substantially inconsistent.

In the log-normal model, the payoff endogeneity discussed above leads to an additional inconsistency because of the additional restrictive assumptions about the process of a venture's value. The endogeneity implies that ε^h_{t+h} in (OA.18) is negatively correlated with h, as more successful firms are more likely to proceed quickly to the next funding round or exit. As a consequence, if estimation proceeds under the assumption $g(h) = \gamma h$, the estimate of γ will be inconsistent (but not in the alternative log-normal model (OA.21) where the intercept term does not depend on h). The selection models in Cochrane (2005) and Korteweg and Sorensen (2010) help ameliorate this potential shortcoming of the log-normal model by specifying the probability of a new funding round or an exit event as a function of the value of the start-up company. As the firm value rises, the probability of obtaining a new funding round or a successful exit rises as well. This limits the extent to which the firm value within a funding round can grow with h, which gets the model closer to (OA.21). These selection models, however, require rather strong assumptions.

In comparison, our SDF approach circumvents the need to make strong distributional assumptions. This avoids the inconsistency that can arise from misspecification of these assumptions. Nevertheless, to facilitate comparison with the prior literature, we estimate the

Table OA.5
Results from a Log-Normal Model for VC Round-to-Round Returns

We match each start-up company round-to-round gross return, measured over horizon h from t to $t+h$, with the return of the CRSP value-weighted index ($r^h_{m,t+h}$ in logs) from t to $t+h$ and the return from rolling over 1-month Treasury Bills (r^h_{ft+h} in logs). We estimate a linear regression

$$r^h_{it+h} - r^h_{ft+h} = g(h) + \beta(r^h_{m,t+h} - r^h_{f,t+h}) + \varepsilon^h_{t+h},$$

with OLS, where $g(h) = \gamma h$ in columns (i) to (iii) and $g(h) = \gamma$ in columns (iv) to (vi). The standard errors in parentheses are calculated assuming independent errors, in line with common practice based on this model in the existing literature. The annualized arithmetic alpha is calculated from γ with a Jensen's inequality adjustment. The calculation of variance for this adjustment assumes that the variance of ε^h_{t+h} is proportional to h in columns (i) to (iii), and constant in columns (iv) to (vi). The number reported in the table is the average of these arithmetic alphas across the whole sample. The implied GPME is calculated for each observation as the arithmetic alpha for the observation's horizon, α^h, as $GPME = \exp(\alpha^h) - 1$, and then averaged across the whole sample.

		$g(h) = \gamma h$			$g(h) = \gamma$	
	(i)	(ii)	(iii)	(iv)	(v)	(vi)
	liq. ret. = -90%	liq. ret. = -70%	liq. ret. = -50%	liq. ret. = -90%	liq. ret. = -70%	liq. ret. = -50%
γ	-0.346	-0.123	-0.019	-0.356	-0.098	0.022
	(0.011)	(0.008)	(0.008)	(0.014)	(0.011)	(0.010)
β	2.924	2.077	1.683	1.284	0.962	0.812
	(0.078)	(0.059)	(0.053)	(0.086)	(0.060)	(0.050)
Annualized arithmetic alpha	0.702	0.599	0.617	0.595	0.464	0.473
Implied GPME	1.458	1.153	1.202	0.817	0.590	0.603

log-normal model (OA.18) on our data. We estimate both the specification with $g(h) = \gamma h$, and the alternative model (OA.21) where $g(h) = \gamma$. Comparing these specifications helps us assess which stochastic process best describes start-up company project values and how to think about the economics of entrepreneurial ventures. In each case, we convert the alpha to a GPME measure as $GPME = \exp(\alpha^h) - 1$, averaged across all round-to-round observations (which differ in h).

Table OA.5 shows the results for the log-normal model. The first three specifications use the assumption $g(h) = \gamma h$ from Cochrane (2005) and Korteweg and Sorensen (2010), while the last three specifications use the alternative assumption, $g(h) = \gamma$.

Across in columns (i) to (iii) in all panels, the β estimates are between 1.8 and 3.5, and this range partly overlaps with recent estimates in the literature (Gompers and Lerner (1997) estimate betas from 1.1 to 1.4, Peng (2001) finds 1.3 to 2.4, 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). The implied GPMEs are considerably higher than what we find with the SDF approach that does not make strong distributional assumptions. Evidently, relying on the log-normal model without combining it with a selection model leads to implausibly large positive abnormal returns.

Changing the assumption about the horizon-dependence of abnormal return and volatility leads to more plausible estimates. In the alternative log-normal model in columns (iv) to (vi), the GPMEs for liquidation returns of -70% and -50% are comparable to those from the SDF model. This result suggests that the alternative model does a better job describing the data compared to the distributional assumptions in columns (i) to (iii).

Note that in Table OA.5 the arithmetic alpha shrinks as we increase the assumed liquidation return. This counter-intuitive feature appears to be driven by the effect that a higher liquidation return lowers the volatility (and hence the Jensen's inequality adjustment term) and dominates the positive effect of the higher liquidation return on γ.

The results in this section demonstrate some of the difficulties that can arise when applying an approach with strong distributional assumptions. How to model the randomness in start-up companies' value processes is an interesting question for further research. At this point, however, without extensive evidence on this issue, it is difficult to set up the log-normal model in a way that is consistent with important features of the data. Our SDF approach circumvents the need to take a stand on these issues.

References

Ang, A., B. Chen, W. N. Goetzmann, and L. Phalippou (2014). Estimating Private Equity Returns from Limited Partner Cash Flows. Working Paper, Columbia University.

Cochrane, J. H. (2005). The Risk and Return of Venture Capital. *Journal of Financial Economics 75*(1), 3 – 52.

Driessen, J., T.-C. Lin, and L. 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, M., C. M. Jones, and M. Rhodes-Kropf (2013). The Price of Diversifiable Risk in Venture Capital and Private Equity. *Review of Financial Studies 26*(8), 1854–1889.

Gompers, P. A. and J. Lerner (1997). Risk and reward in private equity investments: The challenge of performance assessment. *The Journal of Private Equity 1*(2), 5–12.

Hall, R. E. and S. E. Woodward (2007). The Incentives to Start New Companies: Evidence from Venture Capital. Working Paper, NBER.

Hansen, L. P., J. C. Heaton, and A. Yaron (1996). Finite-Sample Properties of Some Alternative GMM Estimators. *Journal of Business and Economic Statistics 14*, 262–280.

Harris, R. S., T. Jenkinson, and S. N. Kaplan (2014). Private Equity Performance: What Do We Know? *Journal of Finance*, forthcoming.

Hochberg, Y. V., A. Ljungqvist, and A. Vissing-Jorgensen (2014). Informational Holdup and Performance Persistence in Venture Capital. *Review of Financial Studies 27*(1), 102–152.

Jenkinson, T., M. Sousa, and R. Stucke (2013). How Fair are the Valuations of Private Equity Funds? Working Paper, Oxford University.

Kaplan, S. N. and A. Schoar (2005). Private Equity Performance: Returns, Persistence, and Capital Flows. *The Journal of Finance 60*(4), 1791–1823.

Korteweg, A. and M. Sorensen (2010). Risk and Return Characteristics of Venture Capital-Backed Entrepreneurial Companies. *Review of Financial Studies 23*(10), 3738–3772.

Ljungqvist, A. and M. Richardson (2003). The Cash Flow, Return and Risk Characteristics of Private Equity. Working Paper, NYU Stern.

Peng, L. (2001). Building a Venture Capital Index. Working Paper, University of Colorado, Boulder.

Phalippou, L. and O. Gottschalg (2009). The Performance of Private Equity Funds. *Review of Financial Studies 22*(4), 1747–1776.

Robinson, D. T. and B. A. Sensoy (2011). Cyclicality, Performance Measurement, and Cash Flow Liquidity in Private Equity. Working Paper, Ohio State and Duke University.

Stucke, R. (2011). Updating History. Working Paper, University of Oxford.

Woodward, S. (2009). Measuring Risk for Venture Capital and Private Equity Portfolios. Working Paper, Sand Hill Econometrics.