RISK AND RETURN IN PRIVATE EQUITY

ARTHUR KORTEWEG*

Marshall School of Business

University of Southern California

Abstract

This chapter is both a primer on estimation methods for assessing risk and return in private equity and a survey of the empirical literature. Most private equity investments are made by private funds that raise capital from institutions and wealthy individuals. Performance data is collected either at the fund-level or at the level of the underlying investments. These data sources have distinct econometric issues, and the chapter describes empirical methods developed for each type of data. The risk and return properties of the small segment of publicly traded vehicles is also considered. The chapter concludes with directions for future research.

Keywords

Private equity, venture capital, buyout, return, performance, risk, risk adjustment, alpha, beta

* I am grateful to Douglas Cumming, Nicola Giommetti, and the handbook editors for helpful comments and suggestions.

1. Introduction

Investments in private equity (PE) have increased dramatically over the last 25 years.[1] The 2020 NACUBO-TIAA Study of Endowments reports that U.S. college and university endowments allocate 22.8% of their assets to leveraged buyout and venture capital strategies, and another 6.9% to private debt and private real estate. Similarly, U.S. pension plans invest nearly 20% in alternative asset classes, which includes private equity, and pension funds in many other developed countries also have double-digit allocations to alternatives (Ivashina and Lerner, 2018).[2]

With the growing role of PE in investors' portfolios, assessing its risk and return has taken on increased importance. This chapter surveys the literature, covering both methodology and empirical results.

Several methodological issues prevent the direct application of standard asset pricing techniques to the private equity setting. Most importantly, returns are generally not observed at a regular frequency, reported fund valuations are often stale or biased in other ways, and portfolio company outcomes are frequently missing for failed investments. Moreover, payoff distributions are highly skewed. The chapter describes methods developed specifically to deal with these problems, distinguishing between approaches for fund-level and investment-level data, since they solve distinct econometric problems.

The empirical literature focuses on venture capital (VC) and leveraged buyout, historically the two dominant strategies within the PE space. Estimates of risk and return vary substantially, depending on the data source, sample period, and methodology used. For example, Figure 1 shows the distribution of capital asset pricing model (CAPM) beta estimates, a measure of systematic risk, for VC and buyout from the literature. For each strategy, the figure shows the estimates based on pre-fee investment-level data and after-fee fund-level data. Most papers find that VC and buyout are riskier than investing in the public stock market index, which has a beta of one by construction. But the point estimates vary widely, with some papers estimating betas less than one and others reporting estimates over 2.5.

[INSERT FIGURE 1 AROUND HERE]

Data quality and coverage have improved significantly in recent years, and researchers have started to adopt similar methodologies, most notably approaches that use stochastic discount factors, such as the public market equivalent metric. This has led to some convergence in estimates. Put differently, the variation in Figure 1 overstates the uncertainty about estimates to some extent, as the papers included in the figure were not cherry-picked based on any judgment regarding the quality of the data or methodology used to arrive at the results, or even whether they estimate the risk of a reasonably similar set of firms or funds. Still, Figure 1 underscores an important result regarding the uncertainty about estimates in PE that permeates the literature.

Despite the uncertainty, the weight of evidence from the PE literature suggests the following main takeaways (in addition to the insights on the level of market risk from Figure 1):

- Relative to a market beta-matched investment in public stocks, the average VC fund earned positive risk-adjusted returns net of fees before the turn of the millennium. Since then, most studies have found average net-of-fee fund returns that are either statistically insignificant or negative.

- The average leveraged buyout fund earned positive net-of-fee returns relative to a market beta-matched public stock portfolio for the entire period since its early growth in the 1980s.
- Beyond market risk, VC looks like an investment in small growth stocks, whereas buyout tends to load on value but does not have a consistently meaningful size loading.
- Liquidity risk is important, but the size of the liquidity premium is still an open question.
- Other risks matter: idiosyncratic risk appears to be priced in VC, both in the time-series and the cross-section. Term structure and credit market risk factors also play a role. However, more work is needed to explore these, and other factors. When additional risk factors are accounted for, the net-of-fee risk-adjusted performance of both VC and buyout is lower, in many cases insignificant and in some cases negative.
- Gross-of-fee risk-adjusted return estimates appear high relative to net-of-fee estimates. The literature has not yet determined whether this difference is explained by fees alone, or whether it stems mainly from differences in data or methodology. Similarly, differences in beta estimates, which appear to be higher in gross-of-fee data, have not been satisfactorily explained.
- Investors' experience in alternative PE vehicles such as co-investments, which have grown dramatically in recent years, depends on the qualities and match between the investor and manager. Private funds of funds in VC have performed on par with direct VC funds, but weaker in buyout.

- Publicly traded PE vehicles do not generate statistically significant risk-adjusted returns (either positive or negative) by most estimates.

Several topics are not covered in this chapter, such as the evidence on return persistence, the large heterogeneity in performance across types of investors (i.e., limited partner or LP) and managers (i.e., general partner or GP), the debate regarding GP skill versus luck, and the LP's portfolio allocation decision. Korteweg and Westerfield (2022) provide an in-depth survey of these questions. The source of any value creation, that is, whether PE creates value at portfolio companies, is also not considered (this is the subject of the chapter by Sørensen and Yasuda in this same handbook volume). Finally, the performance of buyout debt, venture debt, private investments in public equity (PIPEs), angel investments, and entrepreneurs are not covered, although the tools and techniques described in the chapter are applicable to these investments as well.

The chapter is structured as follows. Section 2 describes the institutional setting of the PE industry. Section 3 discusses methods for fund-level data and presents results for main funds. Section 4 turns to the empirical evidence on alternative PE vehicles and funds-of-funds, and section 5 covers publicly traded PE vehicles. Section 6 considers investment-level data and methods, and Section 7 concludes and lists open questions for future research.

2. Institutional Setting

This section provides a brief overview of the PE industry for readers who are unfamiliar with its institutional setting. It may be safely skipped by those versed in the ins and outs of private equity. I present a level of detail that will suffice for the purpose of this chapter without intending to be comprehensive. There are many excellent books and articles with more complete descriptions of the inner workings of private equity, such as Lerner, Leamon, and Hardymon (2012).

2.1. Funds

Traditionally, investors gain exposure to private equity by investing in funds. Funds are pools of capital that are raised and managed by a private equity firm. They are legally structured as limited partnerships with a finite lifetime, typically 10 years, after which they are liquidated.[3] The PE firm professionals serve as General Partners (GPs) and the investors are the Limited Partners (LPs). This structure is common across the different PE subclasses. The key difference between these strategies is the types of assets or securities they invest in. For example, venture capital (VC) funds invest in startup companies, buyout funds buy more seasoned, often struggling, firms (using leverage to help finance the purchase), real estate funds invest in office buildings, warehouses, shopping centers, and other (typically commercial) properties, and private debt funds invest in fixed income products (for example, corporate loans and bonds, or structured credit products. These are also known as private credit funds). There are also funds that invest in other PE funds (secondary funds and fund-of-funds). Sacrificing accuracy for ease of exposition, I will henceforth refer to the fund’s investments as portfolio companies.

Table 1 presents descriptive statistics of PE funds raised since 1980, when private equity funds started to gain rapid popularity due to the passage of the 1979 ERISA “prudent man” rule (which made it feasible for pension funds to invest in PE) and a substantial reduction in the capital gains tax rate that was implemented around the same time. The number of funds, as well as their typical size, varies by strategy. There have been well over 1,000 funds raised in venture capital, buyout, and real estate until the 2019 vintage year, the end of the sample period, in part due to these strategies having been around the longest.[4] Natural resources and especially infrastructure funds are more recent additions to the private equity scene, so naturally there have been fewer funds raised in these spaces over the sample period. The largest funds are raised in buyout and infrastructure, with average (median) fund sizes of $1.3 billion ($500 million) and $2.1 billion ($999 million). Private debt and natural resource funds are also quite large at $868 million ($438 million) and $914 million ($410 million). In contrast, the average fund in venture capital is only $248 million, with a median fund size of $150 million. The heterogeneity in fund sizes across PE classes depends on factors such as the typical size of an individual investment and the targeted number of investments per fund, amongst others.

[INSERT TABLE 1 AROUND HERE]

The fund’s GPs are responsible for putting in all the hard work, from sourcing new investments to negotiating and executing deals, managing the portfolio, and ultimately harvesting the gains. Unlike mutual funds and ETFs, they are active investors. For example, in venture capital and buyout, GPs serve on the board of their portfolio companies and provide a variety of value-added services, including monitoring and corporate governance (Lerner, 1995, Hellmann and Puri, 2002, Baker and Gompers, 2003, Edgerton, 2012, Bernstein, Giroud, and Townsend, 2016, Ewens and Marx, 2018), operational experience and strategic advice (Bharath, Dittmar, and Sivadasan, 2014, Davis et al., 2014, Bernstein and Sheen, 2016, Biesinger, Bircan, and Ljungqvist, 2020, Cohn, Hotchkiss, and Towery, 2020, Fracassi, Previtero, and Sheen, 2022), human resources (Hellmann and Puri, 2002, Amornsiripanitch, Gompers and Xuan, 2019), innovation (Lerner, Sorensen, and Stromberg 2011), and relieving financial constraints (Boucly, Sraer, and Thesmar, 2011, Cohn, Hotchkiss, and Towery, 2020, and Fracassi, Previtero, and Sheen, 2022).

The LPs are investors such as pension funds, endowments, insurance companies, sovereign wealth funds, and wealthy individuals. They are pure capital providers and are not actively involved with running the fund: they do not have any control over individual investment and sale decisions, or how the portfolio assets are managed. Instead, they negotiate a contract with the GPs at the time of fundraising, called the LP Agreement (or simply LPA), which describes each party's rights and responsibilities.

2.2. Fund Cash Flows

An important feature of PE funds is that the money committed by LPs is not immediately transferred to the fund. The LPA allows for an "investment period" (also known as the commitment period), usually 4 to 6 years (3 years for real estate and private debt funds), in which the fund is allowed to make investments in new portfolio companies. LPs hold on to their capital until the GPs identify a desired investment (or when management fees need to be paid; more on fees in the next section). At this point the GPs issue a capital call to the LPs. Capital calls (also referred to as "contributions") are pro rata to fund commitments, for example, an LP who has committed $10 million dollars to a $200 million fund will contribute 5% of the total amount of any capital call. Figure 2 shows the average cumulative capital calls over a fund's life, for an LP with a $10 million commitment, broken down by strategy.[5] Calls are depicted as negative numbers because they are outflows from the LP's perspective. The figure shows that on average, committed capital is gradually called up during the first five years of the fund's life. For an individual fund, the pattern is less smooth than Figure 2 suggests. The time it takes for a fund to be fully called up varies by fund strategy, deal flow, the length of the investment period, and other considerations. There is considerable heterogeneity in the rate of capital calls across vintage years and funds, as shown by the size of the shaded area in the figure, which captures the range between the top and the bottom 5 percent of the distribution.

[INSERT FIGURE 2 AROUND HERE]

After the investment period, the fund enters the "harvest period". GPs' workload drops significantly at this time, as they switch from deal sourcing to managing the fund's investments and realizing exits. If the fund still has "dry powder" (uncalled capital), follow-on investments in existing portfolio companies may still be made, such as participation in new fundraising rounds of previously backed startups in VC.

Exits take the form of acquisitions, usually by strategic acquirers but in some cases by other PE funds, or – in venture capital and buyout – an initial public offering (IPO) of the portfolio company. In some strategies, investments are self-liquidating, for example the maturing of loans in private credit. Not all investments are successful, however, and some may be written off completely. This is especially common in venture capital: Kerr, Nanda, and Rhodes-Kropf (2014) show that of all startup firms in the Venture Economics database that received their first round of early-stage financing between 1985 and 2009, about 55 percent resulted in a loss. Only 6 percent returned more than 5 times their investment, but this small subset of startups accounted for roughly half of the gross return that was realized over this sample period. This observation suggests a high level of skewness in VC returns, a topic I will return to in the next section.

Cash that returns to the fund is distributed to the LPs, after withholding of any fees due to the GPs. Like capital calls, distributions are pro-rata to LPs' commitments. In VC, distributions stem solely from portfolio company exits, since startups do not pay dividends, but in other strategies there may be additional distributions from intermediate cash flows that realize before exit. For example, in buyout, portfolio companies sometimes pay dividends (often through recapitalizations) and they pay a variety of fees to the GPs, some of which are shared with the LPs. In real estate, property rental income generates a steady cash flow stream to the fund.

Figure 2 shows the average after-fee cumulative distributions over a fund's life. Since it takes time for portfolio companies to mature and exit, distributions do not usually accelerate until the third year into the fund, and most tend to come in between years 4 and 9. The shaded area in the figure, indicating the range between the top and bottom 2.5 percent of the distribution, reveals a large degree of heterogeneity across funds and time.

The LPA sometimes allows GPs to reinvest proceeds from early exits. In such cases, after the money is distributed to the LPs, it gets "recalled" (or "recycled"). Global Investment Performance Standards (GIPS), a voluntary set of standards used by investment managers, requires that recalled distributions are both counted as an actual distribution and as additional paid-in capital. As such, a fund with recalled distributions may report cumulative capital calls in excess of its committed fund size. This can be seen in Figure 2, where cumulative capital calls sometimes exceed the LP's $10 million commitment.

Fund cash flow data sometimes report negative capital calls. This usually represents a return of excess capital called, which can happen if the amount invested ended up being lower than the amount originally called, if a deal fell through at the last minute, or if the GP returns excess fees to the LPs. The returned amount can be called again later.

The solid line in Figure 2 traces the typical pattern of cumulative net cash flows (defined as distributions minus capital calls) and reveals the well-known "J-curve" that also shows up in common performance metrics such as cash multiples and IRRs: cumulative net cash flows are negative in the first few years of the fund (the "valley of tears" as some call it), before turning positive later in the fund's life.

The holding time for an individual investment can vary, depending on the type of investment, macro-economic conditions, managerial skill, and luck, amongst others. For example, buyout investments often take 3 to 7 years from initial investment to IPO or acquisition, whereas early-stage VC investments can at times last 10 years or longer (and the average time to liquidity in VC has been growing longer in the past two decades). If the fund still has unrealized investments by the end of its legal lifetime, it can be extended beyond the initial 10-year period. In VC, for example, the LPA often gives the GP the right to trigger two one-year extensions. Any further extensions must be approved by the LPs, which involves some negotiating over management fees. Figure 2 shows that fund extensions are quite common across strategies, as a material amount of distributions come in after year 10 (the figure cuts off at 12 years, but it is not uncommon to see funds that are over 15 years old, especially in early-stage VC). If the fund terminates while there remain unexited investments, these may be sold to another fund, or the securities may be distributed to the LPs.

2.3. Fees

A fund's GPs charge two types of fees to their LPs: a management fee and a profit share called "carried interest" or simply "carry" (also known as "promote" in real estate funds). It is also common for GPs to charge certain fees to portfolio companies, some of which may be shared with LPs.

The management fee is typically 2% per year, with some variation around this number. The percentage fee number is usually applied to committed capital during the investment period, and often switches to net invested capital (the aggregate invested amount of remaining unexited investments) during the harvest period. The percentage fee number may also drop after the investment period. Metrick and Yasuda (2010a) report fund fee statistics for a sample of 238 VC and buyout funds. They compute total management fees over the median fund's lifetime of 17.75% and 12% of committed capital for VC and buyout, respectively. Robinson and Sensoy (2013) find 21.4% and 14.2%, in a larger sample of 837 funds. These fees are paid out of committed capital, so if total management fees add up to 15% of committed capital, a $100 million fund only invests $85 million (if there are no recycled distributions).

Carried interest is the share of profits that GPs earn if the fund returns more than its committed capital to LPs. For VC and buyout, carried interest equals 20% for most funds, whereas in real estate there are usually multiple tiers. In some strategies, such as buyout and real estate, there may be an additional "hurdle rate" (also known as "preferred return" or simply "pref", typically equal to 8% per year on net invested capital) that needs to be returned to LPs before the GPs start to earn carry. Funds with a hurdle often allow for a "catch-up" period where the GPs receive a larger proportion (often 100%) of distributions until they have received their carry share on the fund's profits to date. There is variation across funds in the timing of carried interest payments. European funds typically assess carry on a whole-fund basis, that is, the GP does not earn carried interest until the entire committed capital plus hurdle has been returned to LPs. In contrast, American funds typically compute carry on a deal-by-deal basis. The deal-by-deal structure allows the GPs to earn carry earlier, but if a fund had lackluster performance following early success, the LPs may claw back any excess carry paid, which carries some credit risk for the LPs.[6] For more details on the timing of carried interest, see Litvak (2009) and Hüther, Robinson, Sievers, and Hartmann-Wendels (2020).

Apart from LP fees, some funds also charge fees to their portfolio companies. These fees are especially common in buyout and are described in the management service agreement contract between the GPs and the portfolio company. They include, but are not limited to, transaction fees for investment banking services provided by the GPs to the company at the time of the buyout (and possibly for any add-on acquisitions) and annual advisory fees (also known as monitoring fees). Metrick and Yasuda (2010a), from informal discussions with buyout practitioners, report that one-time transaction fees are typically 1 to 2% of deal value, and annual monitoring fees range from 1% to 5% of the portfolio company's EBITDA. Phalippou, Rauch, and Umber (2018) collect detailed fee data for a sample of 454 large U.S. leveraged buyout deals and find that total fees add up to 1.75% of enterprise value for the average portfolio company. About 70% of fees are rebated to LPs on average, with an interquartile range of 25% to 80%. Since most portfolio companies in buyout are fully owned by the fund, these fees are effectively dividend payments (with a different LP-GP split than carried interest). Their characterization as fees may carry a tax advantage as fee payments are tax deductible for the portfolio company.

Some LPs obtain fee breaks as part of side letters they negotiate with their GPs. Lee (2015) describes the content of side letters anecdotally. In a survey of 249 LPs across 30 countries, Da Rin and Phalippou (2017) find that 63% of large LPs always get side letters, compared to only 17% of small LPs. Begenau and Siriwardane (2021) provide suggestive evidence that carried interest rates differ across LPs invested in the same fund; larger LPs and those that have a longer investment history with the GP earn higher net-of-fee returns.

At present, fund fee data is very difficult to collect. Fee payments cannot be identified in publicly available fund data sets, such as Burgiss, Preqin, and Thomson VentureXpert (formerly called Venture Economics), as they only report cash flows and returns net of fees. Thus, fund data can only inform about risk and return from the LPs' perspective. To say something about GP skill, we need gross-of-fee returns. Looking ahead, it may be possible to get gross-of-fee fund data for a reasonably large and representative set of funds, as pressure to increase transparency in fee payments has been growing in recent years. For example, California Government Code Section 7514.7 requires annual public disclosure of the fees, expenses, and carried interest paid by

California public pension funds – which includes some very large LPs such as CalPERS and CalSTRS – to GPs after January 1, 2017. But as of this writing, the most comprehensive coverage of gross-of-fee returns comes from portfolio company data, which I discuss in section 6.

2.4 Fund sequences

GPs raise a sequence of flagship funds, typically distinguished by Roman numerals (e.g., Sequoia I, Sequoia II, etc.), and may have more than one active fund at any given time. Since the GPs' workload is very high during the investment period, and to avoid agency problems from actively investing in new deals out of multiple funds at the same time, the LPA prohibits GPs from raising a new fund that directly competes with the current fund until a prespecified, substantial percentage of the fund is invested. Note, however, that this does not prevent GPs from raising funds in other geographies or strategies, from raising "opportunity" funds (also known as "select" or "growth" funds) in early-stage VC[7], or alternative vehicles that are offered to select LPs. I discuss alternative vehicles, the number and size of which has grown dramatically in recent years, in more detail in section 4.

Based on the number of PE firms and the number of funds reported in Table 1, the average PE firm has raised between 2 and 4 funds during the sample period, depending on the strategy (e.g., in buyout, 431 firms raised 1,300 funds, for an average of 3.0 funds per firm). Empirically, most GPs raise a new fund every two to four years. Consequently, subsequent funds overlap in time for six or more years. Accounting for this overlap is important for the calculation of correct standard errors of performance metrics (as well as for inference regarding return persistence, but this is not a focal issue of this chapter; see Korteweg and Sørensen, 2017, for a model of overlap and persistence).

3. Fund Data

3.1 Data Structure and Notation

Table 2 shows an example of the type of cash flow data that researchers would typically work with, for a fictitious fund i. Calendar dates are denoted by t, with t_0 and T being the date of the first and last observation for the fund, respectively. For notational simplicity I suppress the dependence of t_0 and T on i. Based on the initial cash flow date of February 1, 2022, researchers would generally consider this fund to be a 2022 vintage year fund, though the GPs likely started raising the fund in 2021. Date T for this fund is December 31, 2024. Note that the fund is not yet liquidated at this time. This may happen because we are at the end of the sample period, or the fund may have stopped reporting after this date.

[INSERT TABLE 2 AROUND HERE]

Cash flows are reported from an LP's perspective and are normalized to a $10 million commitment. Capital calls, $I_{i,k}$, are indexed by k, from 1 to K. Unlike Figure 2, capital calls from here on out are defined as positive numbers (except for recalled distributions, which as considered negative capital calls, as described in the previous section). Thus, from Table 2, $I_{i,1}$, the initial capital call that occurred on February 1, 2022, equals $200,000. Distributions, $D_{i,d}$, are numbered from d = 1 to D.

Our fund only has $D = 2$ distributions: $D_{i,1}$ = \$500,000, which occurred on April 10, 2024, and $D_{i,2}$ = \$1,000,000 on December 20, 2024.

Some performance measures are based on net cash flows, $C_{i,j}$, defined from the LP's perspective as the distribution minus the capital call for each date on which there is a cash flow (either a capital call or a distribution, or both). For example, $C_{i,1}$ in Table 2 is the net cash flow on the date of the first capital call, and since there was no distribution on that date, it is equal to -\$200,000. Net cash flows are numbered from $j = 1$ to J, where $J \leq D + K$, with strict inequality if a capital call and a distribution occur on the same day. This happens for the fund in Table 2, which has $K = 7$ capital calls and $D = 2$ distributions, but only $J = 8$ net cash flows because on April 10, 2024 it had both a capital call and a distribution, which combine into one net cash flow ($C_{i,6} = D_{i,1} - I_{i,6}$ = \$500,000 - \$200,000 = \$300,000). As with fund start and end dates, K, D, and J vary by fund, but I suppress their dependence on the fund index i.

Apart from cash flows, GPs report fund net asset values (NAVs) to their LPs at the end of each quarter, which provide a snapshot of the aggregate value of the fund's ownership position in its portfolio investments, net of any borrowing at the fund level.[8] In other words, the NAV represents the value that has not yet been distributed to LPs. For a liquidated fund, the final NAV is zero. The fund in Table 2 reports a large NAV of \$3.3 million on December 31, 2024, since it is not yet liquidated at the end of the sample period (in fact, it is still in the investment period, having called only \$4.45 million of a \$10 million commitment).

For most performance metrics we need the time difference, τ, between fund inception (t_0) and the date of observed cash flows or NAVs. The second column of Table 2 shows τ for our example fund. The time-since-inception of net cash flow number j, capital call k and distribution d are denoted by $\tau(j)$, $\tau^I(k)$ and $\tau^D(d)$, respectively. Table 2 does not show these numbers explicitly, but they are easily identified: For example, $\tau^I(2)$ in Table 2 – the since-inception time of the second capital call – is 0.37. Like the other timing variables, I suppress the dependence of the τ variables on the fund index, i, to keep notation tractable.

I use the conventional notation R_t for a return from time t-1 to t. That is, R_t = 1.05 represents a 5% return over the period. Lower-case letters denote log-returns, i.e., $r_t \equiv \ln\left(R_t\right)$. For returns over an interval from time t to $t+\tau$, I use the notation $R_{t,\tau}$. For example, column 6 in Table 2 shows the return on the public stock market portfolio from fund inception to the date of each reported cash flow or NAV (these are fictitious market returns for the sake of the example, not the actual returns in the data, which are unknown as of the time of this writing). Stock market and other factor returns data must be collected by the researcher. PE data providers, such as Preqin or Burgiss, usually only provide the capital calls, distributions, and NAVs (as well as some GP and fund characteristics).

3.2 Traditional Private Equity Return Metrics

3.2.1 Definitions

The two traditional metrics of return in private equity are cash multiples and internal rates of return (IRR). Gompers, Kaplan, and Mukhalyamov (2016) survey 79 buyout GPs and find that in over 90% of investments, they use one or both of these measures to evaluate the deal. Moreover, virtually all of them report that either the cash multiple or the IRR is the most important metric to their LPs. Similarly, of 546 VCs surveyed by Gompers et al. (2020), 63% use cash multiples, and 42% use IRRs to evaluate investments. As many as 84% of these VCs think cash multiples are an important benchmark metric for their LPs, and 81% think IRRs are important.

The standard cash multiple used in practice is Total Value to Paid-in Capital (TVPI), computed for fund i as[9]

$$TVPI_i = \frac{\sum_{d=1}^{D} D_{i,d} + NAV_{i,T}}{\sum_{k=1}^{K} I_{i,k}}. \tag{1}$$

For example, as of December 31, 2024, the fund in Table 2 had total capital calls of $4.45 million, total distributions of $1.5 million, and its most recent reported NAV is $3.3 million. Its TVPI is therefore ($1.5m + $3.3m)/$4.45m = 1.08.

The other customary return metric, IRR, is defined as the discount rate (expressed as an annual rate) such that the net present value (NPV) of the fund’s net cash flows discounted back to t_0 (the date of the fund's first cash flow) is zero.

$$\sum_{j=1}^{J} \frac{C_{i,j}}{(1+IRR_i)^{\tau(j)}} + \frac{NAV_{i,T}}{(1+IRR_i)^{T-t_0}} = 0. \tag{2}$$

Based on the 8 net cash flows and the NAV as of December 31, 2024, the fund in Table 2 has an IRR of 4.24%.

Both TVPI and IRR include the most recently reported NAV as a pseudo-distribution, to reflect any undistributed value remaining inside the fund. For fully liquidated funds (for which $NAV_{i,T}$ is zero), both TVPI and IRR are based entirely on cash flows.[10]

Note that the classification of certain cash flows as capital calls or distributions – for example, whether a recalled distribution is accounted for as a negative distribution or as a capital call, or whether a return of excess capital called is a negative capital call or a distribution – matters for TVPI (and more generally, for any performance metric based on ratios of distributions and contributions, such as the Kaplan-Schoar public market equivalent (PME) discussed below). Metrics that are based on net cash flows, such as the internal rate of return (IRR), are invariant to this classification issue.

3.2.2 Data

Table 1 reports descriptive statistics of TVPI and IRR using the data introduced in section 2.1, separated into six PE fund strategies. The average fund TVPI over the 1980 to 2019 sample period ranged from 1.28 for infrastructure funds to 1.89 for VC funds. Infrastructure funds also had the lowest median TVPI at 1.22, whereas buyout had the highest median at 1.61. An important feature of private equity fund TVPIs across strategies is their high level of volatility and (positive) skewness. Venture capital funds especially stand out with a standard deviation of TVPI of 2.43. The high variance is in part due to time variation in average performance, but the lion’s share is due to cross-sectional variation (this is not shown here, but see section 3.7.1 below for more details). Since TVPI is bounded below at zero, it is not surprising that VC exhibits large positive skewness and excess kurtosis (in comparison, the Normal distribution has zero skewness and zero excess kurtosis). To illustrate the degree of skew and kurtosis more clearly, the left column of plots in Figure 3 shows the kernel-smoothed distribution of TVPI, with VC at the top left.[11] The second and higher moments of the other PE strategies are for the most part not quite as striking as in VC, but they are still very large compared to traditional publicly traded asset classes such as stocks and bonds.

[INSERT FIGURE 3 AROUND HERE]

Turning to IRR, the mean (median) ranges from 9.59% (8.78%) for infrastructure to 16.19% (14.00%) for buyout. The shape of the IRR distribution is broadly similar to TVPI, with each strategy exhibiting large variation (again, mostly cross-sectional) and positive skewness and excess kurtosis. Two notable differences are the higher volatility, skewness, and kurtosis of the IRRs of natural resources fund, which are much larger than for TVPI, and the slight negative skew in infrastructure IRRs. Both strategies are relatively small, so these results may simply be due to the small sample of funds. Moreover, infrastructure has only recently grown to become a more common PE strategy, with many funds still in the early stages of their life cycle (as underscored by the fact that only 13% of infrastructure funds have been fully liquidated by year-end 2019, as reported in Table 1). Performance results are likely to change as these funds mature and more exits realize. The recency of most infrastructure funds may also (partly) explain their lower average performance, as performance metrics tend to exhibit a J-curve over a fund's horizon, like the cash flow patterns discussed in section 2.2.

Unfortunately, no currently available private equity fund data set is comprehensive. Kaplan and Lerner (2017) describe the pros and cons of the main PE data sets that are available for purchase to researchers. Harris, Jenkinson, and Kaplan (2014) compare the leading commercially available fund data sets – Cambridge Associates, Preqin, and Thomson Venture Economics (now called VentureXpert) – with Burgiss, a data set that is accessible to academic researchers. They make a strong case for the Burgiss data being the highest quality data set, but they find that performance metrics in Preqin (used for Table 1 and Figure 3) and Cambridge Associates data are qualitatively similar to Burgiss. Preqin does miss a few top performing VC funds that did not accept public money, as their data is primarily sourced from Freedom of Information Act (FOIA) requests. This likely affects the mean performance numbers but should have only a small (if not negligible) effect on medians. In contrast, the Thomson Venture Economics data has a strong downward performance bias compared to the other data sets, due to its failure to update cash flows and NAVs for about 40% of its funds after some point during funds' active lifetimes (Stucke, 2011).

3.2.3 Drawbacks

The traditional performance measures, while informative, suffer from several drawbacks. For example, TVPI ignores the time value of money: One dollar distributed in year 1 of the fund is considered just as valuable as a dollar distributed in year 10. IRR does account for the timing of cash flows, but it has several problems that are well described in any undergraduate-level corporate finance textbook worth its salt: First, an IRR may not exist in some (rare) cases. Second, there can be multiple IRRs. For example, while the IRR for the fictitious fund from Table 2 is unique, if there were an additional distribution of, say, $1.5 million on March 31, 2022, then there would be two IRRs, 42%, and 673%. Each number is a valid IRR, so it is not clear which is the "correct" number to report. Although such a large early distribution is admittedly rather unrealistic, multiple IRRs can happen in more realistic scenarios. A third drawback is that IRR implicitly assumes that distributions can be reinvested at the same rate. Phalippou (2013) constructs an example based on Yale's endowment track record that shows that early high returns can generate IRRs that do not vary much with subsequent, larger investments, even if those later investments have returns that vary substantially. Fourth, the IRR of a portfolio of investments is not the weighted average of the constituents' IRRs (Phalippou, 2008, and Phalippou and Gottschalg, 2009). This means, for example, that the average IRR is not the IRR of the average fund. This is mainly an implementation issue but it is important to be aware that the portfolio IRR should be computed on the aggregated portfolio cash flows. Finally, IRRs can be manipulated to the extent that GPs can time cash flows, for example, by delaying a capital call. This latter concern is grounds for a more general call for research into the development of a manipulation-proof PE performance measure, possibly along the lines of Goetzmann et al. (2007).

The main critique of both the TVPI and IRR performance metrics, as far as this chapter is concerned, is that neither adjusts for risk. For example, consider the two fictitious funds in Figure 4. Both funds have five years of cash flows, and both funds have a TVPI of 2.5 and an IRR of 50.2%. Yet, fund A looks considerably less risky than fund B (in fact, fund A looks very much like a fixed-coupon bond investment). A different benchmark return would seem appropriate for the two funds, yet the TVPI and IRR metrics do not provide any guidance in this respect. Moreover, while Gompers et al. (2020) report that of 546 institutional VCs, 64% say that they adjust their investments' target metrics for risk, it is not clear how those benchmarks are determined.

[INSERT FIGURE 4 AROUND HERE]

3.3 Risk and Return: Factor Models

At this point it is worthwhile to briefly review the standard asset pricing approach to assessing risk and return as it has been developed in the academic literature over the last 60 years. In this approach, one estimates risk and risk-adjusted returns by regressing a security *i*'s excess return (its return minus the risk-free rate, R_t^F, over the period) on a set of *G* risk factors, whose returns are grouped in a vector $F_t = [F_{1t}, F_{2t}, \ldots, F_{Gt}]'$,

$$R_{i,t} - R_t^F = \alpha_i + \beta_i' F_t + \varepsilon_{i,t}. \tag{3}$$

For example, the CAPM has only one risk factor, the excess return on the market portfolio, $F_t = R_t^M - R_t^F$.[12] The elements of the vector of factor loadings, β_i, represents the exposure of the security to each risk factor. These loadings represent systematic, undiversifiable risk exposures (as opposed to the residual, $\varepsilon_{i,t}$, component of returns, which represents diversifiable, or idiosyncratic risk). The security's benchmark portfolio is the portfolio that invests β_i in the factors, and therefore has the same (priced) risk exposures as security *i*. If the security is priced correctly, and if all factors are traded excess returns that capture all relevant sources of risk, then the intercept, α_i (sometimes referred to as risk-adjusted return), is equal to zero. A positive alpha means that an investor can achieve an excess return over the benchmark portfolio, on average.

It is tempting to take a similar regression approach in PE. However, there are few publicly traded funds for which returns can be measured on regular intervals, and the ones that are traded are a selected subset of the fund universe (see section 5). An alternative for private funds could be to use fund IRRs (minus the risk-free rate over the fund's life) on the left-hand side of equation (3), and annualized returns to the factor portfolios over the same horizon as each fund on the right-hand side. While this is a cross-sectional regression across funds, the variation in vintage years allows for variation in the factor returns needed to estimate betas. For example, Kaplan and Schoar (2005) estimate CAPM betas by regressing fund IRRs on the average annual return to the S&P 500 stock index over the first five years of the fund. Assuming all funds have the same risk exposure, they estimate a beta of 1.23 for VC and 0.41 for buyout, meaning that VC is riskier than the public stock market (which has a CAPM beta of 1 by construction), and buyout is less risky. As Ilmanen, Chandra, and McQuinn (2020) point out, it is problematic to compare IRRs, which are money-weighted returns in the sense that periods with higher cash flows get more weight, with factor returns, which are time-weighted such that each period has equal weight. Put differently, this procedure uses different definitions of returns on the left and right-hand sides of equation (3), and the estimated betas are therefore not true factor loadings. In simulations, Axelson, Sørensen, and Strömberg (2014) show that these types of regressions produce downward biased beta estimates. A better approach may be to regress fund IRRs on the IRRs of portfolios that mimic the capital calls and distributions of the PE fund but invest in the risk factors instead (see, for example, Ljungqvist and Richardson, 2003, for the construction of such matched investment portfolios). At present, the econometric properties of such regressions are unknown. However, in unreported work I find that this alternative approach still produces biased beta estimates, and that the bias is larger if expected returns vary over time or if cash flows are highly correlated with factor returns.

Another alternative that appears feasible, at least at first pass, is to use the quarterly fund NAVs together with the fund's cash flows to construct a quarterly return series. The next section explores this idea.

3.4 NAV-based Returns

An intuitively appealing way to define a return on fund *i* for quarter *t* is to use the reported NAVs and the fund's capital calls and distributions over the quarter,

$$R_{i,t} \equiv \frac{NAV_{i,t} + C_{i,t}}{NAV_{i,t-1}}. \tag{4}$$

Here, $C_{i,t}$, is the total net cash flow (distributions minus contributions) over quarter *t*. This definition is subtly different from $C_{i,j}$ in section 3.1 above, where *j* indicated the fund's cash flow number in the sequence of cash flows from 1 to *J*. Formally, the relation between the two cash flows can be expressed as

$$C_{i,t} = \sum_{j:t_0+\tau(j)\in(t-1,t]} C_{i,j} . \tag{5}$$

The final three columns in Table 2 shows the quarterly NAVs and net cash flows (based on a $10 million commitment), and the resulting return for the fictitious fund from section 3.1. For example, in the quarter ending June 30, 2024, there was a \$500,000 distribution and two capital calls of \$200,000 and \$50,000, for a net cash flow of $C_{i,t}$ = \$250,000 (note that researchers have to construct these quarterly net cash flows themselves, data providers do not report $C_{i,t}$). With a NAV of \$4.6 million at the start of the quarter, and an ending NAV of \$4.5 million, the fund return for the quarter was $R_{i,t} = (\$4.5\text{m} + \$0.25\text{m}) / \$4.6\text{m} = 1.033$, or 3.3%.

With the returns from Table 2 we can estimate a factor model, such as the CAPM: Assuming for simplicity that the risk-free rate is equal to zero (such that returns and excess returns are equal), the fund's beta is 1.22, and the intercept is -23% per quarter. However, as there are only 11 return observations, the standard errors of these coefficients are very large, at 0.66 and 67%, respectively, with beta barely significantly different from zero at the 10% level. To bring down the standard errors, one could estimate a common alpha and beta for a group of funds or even an entire strategy, such as VC. These pooled estimates are common across research methodologies (although a few papers estimate fund-level parameters, usually with the help of some additional identifying assumptions). When estimating a common beta across funds, it is important to correct the standard errors for cross-correlations, as residuals in the same calendar quarter are highly correlated.

To build intuition, it is worth highlighting a few characteristics of the quarterly fund returns as defined by equation (4). First, note that $R_{i,t}$ is not defined for the very first quarter of the fund's life, since the starting NAV is zero. Second, in any quarter in which there are no capital calls or distributions, the return is simply the percentage change in NAV. The first quarter of 2023 in

Table 2 is an example: The fund's NAV grows by $550,000, from $2.2 million to $2.75 million, which is a 25% return. Third, distributions are analogous to dividends in public stocks: They represent a transfer of value out of the fund. For example, if a fund with a $100 NAV makes a $10 distribution, its NAV drops to $90. If there are no other cash flows, then its return is ($90 + $10) / $100 = 0%. Fourth, capital calls for investments work like equity issues by public firms, in the sense that they raise the fund's NAV, just like an equity issue raises the company's asset value. For instance, in the first quarter of 2024, the fund in Table 2 made a $500,000 capital call, which raised the fund's NAV by exactly the same amount. This could happen because new investments were still valued at cost at the end of the quarter, other investment marks remained unchanged, and the capital call was not used for management fees. If any part of the capital call is applied to management fee payments, then this lowers the return to the LP, as the money that is called up by the GP is not invested in any portfolio asset: a single $10 capital call in the quarter that is used for fees on a $100 beginning-of-quarter NAV yields a -10% return to the LP (assuming the NAV did not change over the quarter, i.e., portfolio company valuations remained the same). However, capital calls for fees could result in anomalous returns, especially early in the fund's life when NAV is low. For example, suppose a fund started only recently and has made few investments, such that current NAV is $1 million. Let's also suppose that the GP makes a $2 million capital call for fees in the current quarter, and no other activity takes place. Leaving the portfolio company valuations unchanged, the fund return for the quarter is ($1m - $2m)/$1m = -1, that is, a -200% return. This is a rather stark example, and it is not clear how often this actually happens in the data, but it is not inconceivable.

Two other complications with quarterly fund return series should be noted here. While most funds report NAVs at the end of the calendar quarter, not all of them do: Some GPs report at the end of January, April, July, and October, while other report in February and every three months thereafter. This complicates the comparison of returns across funds. A second complication is that, depending on the data source, some NAVs may be missing, such that returns cannot be computed. To the extent that these are random occurrences, it may not bias factor model estimates, but there is likely some selectivity as to which NAVs do not show up in the data.

Rather than computing returns at the fund level, one can define a return at the aggregate level, R_t, using the sum of NAVs and aggregate calls and distributions. This is equivalent to a NAV-weighted portfolio of all N_t existing funds,

$$R_t = \sum_{i=1}^{N_t} w_{i,t-1} R_{i,t} = \sum_{i=1}^{N_t} \frac{NAV_{i,t-1}}{\sum_{j=1}^{N_t} NAV_{j,t-1}} R_{i,t} = \frac{NAV_t + C_t}{NAV_{t-1}}, \tag{6}$$

defining the aggregate $NAV_t \equiv \sum_{i=1}^{N_t} NAV_{i,t}$, and the aggregate net cash flow $C_t \equiv \sum_i^{N_t} C_{i,t}$. An advantage of aggregate returns, compared to fund-level returns, is that the first quarter of each fund can be used in the construction of the aggregate index, by setting the beginning-of-period NAV of newly raised funds equal to zero (such that $NAV_{t-1} \equiv \sum_{j=1}^{N_{t-1}} NAV_{j,t-1} = \sum_{j=1}^{N_t} NAV_{j,t-1}$).

Another advantage is that anomalous fund returns due to capital calls for fees do not appear in the aggregate return series.

Examples of aggregate return series constructed in the spirit of equation (6) are the Cambridge Associates Private Investments Benchmarks, the Burgiss Global Private Capital Returns, and the State Street Private Equity Index.[13] To illustrate the properties of the time series of aggregate returns, Figure 5 graphs the quarterly returns for the Cambridge Associates data for U.S. venture capital, U.S. buyout, and global real estate funds, which are freely available online for the period ending in the third quarter of 2020.[14] Table 3 reports descriptive statistics of the annualized returns. The average annualized returns are 17.06%, 14.17%, and 9.40% for venture capital, buyout, and real estate, respectively. These numbers are not so different from the IRR numbers in Table 1, even though this is not an apples-to-apples comparison as Tables 1 and 3 use different return definitions, as well as different time series and fund coverages. Pushing the IRR comparison further, despite these critiques, we see that the index returns have much lower volatility, skewness, and kurtosis than fund IRRs. The lower volatility is a result of diversification, and the "normalization" of index returns relative to individual fund returns is a natural consequence of the Central Limit Theorem. Still, there is considerable skewness and kurtosis left in VC returns even at the aggregate level. Buyout and real estate returns have, if anything, a slight negative skew in the aggregate, while kurtosis remains noticeable, especially for real estate.

[INSERT TABLE 3 AROUND HERE]

[INSERT FIGURE 5 AROUND HERE]

For assets class returns with little or no skewness we can consider their Sharpe ratio – the ratio of the mean excess return to its standard deviation – ratios as a metric of the risk and return relationship. For a mean-variance investor, who does not care about higher moments such as skewness or kurtosis, a higher Sharpe ratio means a higher compensation for the level of risk taken. If there is no correlation among asset classes, then the Sharpe is closely related to optimal portfolio weights for a mean-variance investor (I do not explore the portfolio allocation question here, but see Korteweg and Westerfield, 2022, for an in-depth treatment). Table 3 shows that VC, buyout, and real estate had Sharpe ratios of 0.65, 1.15, and 0.83 respectively, over the 1995Q1-2020Q3 sample period. In comparison, the public stock market had a Sharpe ratio of 0.52 over the same period. As such, the private asset classes appear to be quite attractive investments. However, this is a premature conclusion, as there are several issues with the index returns that complicate matters.

While the PE index returns may serve as a benchmark to compare performance against (although representativeness, which may vary by data provider, is a concern), they are very difficult to realize in practice. There is currently no *investable* PE index, so investors must build their own portfolios of funds. Apart from the largest LPs, most institutional investors commit to only a couple of PE funds per year, given that GPs generally require a capital commitment of at least $5 to $10 million from a single LP (e.g., DeLuce, 2020). Since the cross-sectional variation in fund performance is extremely high, it is not feasible to achieve the level of diversification of the index with only a few funds. A second investability issue is that the index returns, when cumulated over time, presume that any cash distributions are reinvested. However, PE funds are closed-end, so one can only invest in new funds that are raising at the time, and these GPs may not be representative of the overall universe of active funds. Moreover, the commitments to new funds will be called up slowly over the next few years, and there is limited liquidity in the sense of being able to sell PE investments (or get out of future commitments some other way) if needed.

Apart from issues of investability, the PE returns constructed using NAV data inherit a number of drawbacks related to the use of NAVs. The next two sections delve into the concerns surrounding the use of NAVs and NAV-based returns.

3.5 Staleness in NAVs

Thus far I have worked under the assumption that a fund’s NAV is equal to the market value of the PE portfolio. Current accounting standards (ASC Topic 820, formerly known as FAS 157, which has been effective since 2007) require that the NAV represents the fair valuation of portfolio companies. In funds that invest in publicly traded assets, such as most mutual and hedge funds, it is straightforward to compute the portfolio’s market value because the constituents’ current valuations can be easily looked up. But for PE portfolio investments (so-called “Level 3” assets) there is currently no market to mark to. GPs, their consultants, and auditors, therefore have substantial leeway in determining the NAVs they report. GPs usually rely on recent deals and comparable assets to determine valuations, but this still leaves a fairly large degree of subjectivity and necessarily relies on historical data.[15] Before 2007 it was quite common practice to leave portfolio company valuations at cost. GPs tended to be conservative in marking up portfolio company valuations while being quicker to mark them down if they thought values had dropped (Anson, 2002, 2007). Since ASC 820/FAS 157, valuations have become more accurate (Jenkinson et al, 2020, Easton et al., 2021). Still, overall, Brown et al. (2019) find that NAVs remain on average conservative. Moreover, price staleness operates beyond the PE fund level: Agarwal et al. (2022) report that mutual fund family valuations of privately-held firms (mostly late-stage startups) are not updated in 46% of quarters, and it takes an average of 2.3 quarters for prices to update (the median is 2 quarters). Thus, prices tend to be stale, reflecting a mix of current and historical valuations.[16]

Price staleness causes the observed time series of NAVs to behave like a moving average that smooths out the "true" underlying fund value. This causes autocorrelation in NAV-based returns, as well as downward biased estimates of the true fundamental return volatility, risk factor loadings, and correlations with other asset classes, and upward biased estimates of Sharpe ratios and alphas (see the literatures on return smoothing in real estate, e.g., Geltner, 1991, 1993, Fisher et al., 1994, and in hedge funds, e.g., Asness et al., 2001, Getmansky et al., 2004, Bollen and Pool, 2008, Cassar and Gerakos, 2011, Cao et al., 2017).

The empirical autocorrelation in NAV-based PE returns is indeed high. Table 3 shows that, for VC and real estate, the one-quarter autocorrelations in the Cambridge Associates returns are 0.60 and 0.62, respectively. It drops gradually over longer lags and reaches zero (effectively) by lags of four to five quarters. Buyout is less persistent, but the one-quarter autocorrelation is still as high as 0.32. It also takes about four quarter lags to drop close to zero. By comparison, the one-quarter autocorrelation in the public stock market over the same sample period is -0.06. Turning to factor loadings, Panel A of Table 4 shows that the market model beta estimated from equation (3) for VC, buyout, and real estate are 0.579, 0.465, and 0.166, respectively. These beta estimates appear surprisingly low, and much lower than the risk of the public market portfolio (which is one by construction). After all, venture capital invests in startups that in many cases do not yet have a marketable product or service, surely a risky proposition. As shown by Berk et al. (2004), even if this risk is often of a technical nature and idiosyncratic, the investment and financing decisions that must be made with respect to the startup result in a strong systematic risk exposure even at the earliest stages. Buyout is a highly levered equity investment, often in a publicly traded firm that is taken private by means of the buyout. Unless the bought-out companies have extremely low underlying asset risk, it is difficult to believe the PE fund's market risk loading would be as low as 0.5. Indeed, it turns out that estimated market betas are substantially higher when NAV staleness is corrected for. The next five subsections discuss various proposed ways to make this correction.

[INSERT TABLE 4 AROUND HERE]

3.5.1 Longer-horizon returns

The previous section revealed that autocorrelations in NAV-based returns disappear at longer lags. Thus, a first proposed solution to the staleness problem is to use longer-horizon returns. Emery (2003) presents suggestive evidence that longer-horizon returns can indeed mitigate the staleness problem, by showing that the correlation between NAV-based PE index returns and public market returns is markedly higher when using annual returns instead of quarterly returns.

In implementing this solution, a choice must be made whether to use nonoverlapping annual returns (e.g., treating each full calendar year return as a separate observation) or a rolling window of overlapping returns (e.g., considering the return from the start of 1991 Q1 to the end of Q4 as one observation, the return from the start of 1991 Q2 to the end of 1992 Q1 as the next observation, etc.). The nonoverlapping return series has the advantage that it is easy to work with, because the residuals should be uncorrelated as long as the return is calculated over a long enough horizon. In contrast, in the rolling window approach, standard errors must be corrected for overlap-induced autocorrelation, which generally requires additional assumptions about the nature of the residual autocorrelation. The Newey-West and Hansen-Hodrick estimators are classic examples.

The key disadvantage of using nonoverlapping returns is that few observations remain. Even VC and buyout, the strategies that have been around the longest, did not really become meaningfully large until the mid-1980s. Thus, by 2021, there are only around 35 nonoverlapping annual return observations. The low number of observations is especially troublesome for (co)variance estimates (including betas), which, unlike means, are more precisely estimated at higher frequencies (Merton, 1980).

A drawback to using longer-horizon returns, whether overlapping or not, is that the timing of cash flows within the period becomes more important as the chosen horizon grows longer. Equations (4) and (6) ignore such timing, which may be acceptable for quarterly returns, but more sophisticated return computations should be employed for longer horizons (for example, the above-mentioned methods used by Cambridge Associates, State Street, and Burgiss; see footnote 13). Unfortunately, in many cases these techniques revert to IRR-type calculations, the very issue we tried to avoid in the first place.

Ultimately, the question of the optimal horizon remains unanswered. The trade-off is between mitigating the impact of staleness and having enough observations for meaningful inference (and keeping the intra-period cash flow timing issue in check). The answer likely varies by strategy and the level of disaggregation (for example, early-stage VC may have a different degree of staleness from VC overall). It may even vary at the individual fund-level, and may vary over time. For the aggregate VC, buyout, and real estate series in Table 3, annual returns appear to be a reasonable choice for the 1995 to 2020 sample period.

3.5.2 Unsmoothing

A rich literature in real estate, which shares many of the same concerns regarding staleness of valuations, and hedge funds, where return smoothing is common, revolves around methodologies to unsmooth a return series. Examples of unsmoothing algorithms are in Geltner (1991, 1993), Fisher et al. (1994), Getmansky et al. (2004), and Couts et al. (2020).

Unsmoothing algorithms rely on a specified relation between the observed, smoothed, returns and the unobserved returns based on nonstale fundamental valuations. For example, Geltner (1991) assumes that the observed return is a weighted average of the current and its *p* past fundamental returns,

$$R_t = \theta_0 R_t^* + \theta_1 R_{t-1}^* + \ldots + \theta_p R_{t-p}^*, \tag{7}$$

where R_t are the observed returns and R_t^* the fundamental returns. The θ's are all fractions, and they sum to one. Equation (7) implies that R_t follows an infinite-order autoregressive (AR) process, but if $\theta_0 > \theta_p$ for all $p > 0$, then the higher-order terms quickly tend to zero, and the process can be approximated with a finite-order AR that includes the first *k* empirically meaningful lags of R_t,

$$R_t = \phi_1 R_{t-1} + \phi_2 R_{t-2} + \ldots + \phi_k R_{t-k} + e_t . \tag{8}$$

The residual $e_t = \theta_0 R_t^*$, so we can rewrite (8) to recover the fundamental return

$$R_t^* = \frac{R_t - \phi_1 R_{t-1} - \phi_2 R_{t-2} - \ldots - \phi_k R_{t-k}}{\theta_0} . \tag{9}$$

One additional assumption is needed to identify θ_0. One choice is to set $\theta_0 = 1 - \sum_{j=1}^{k} \phi_j$, which equates the true means of R_t and R_t^* (since the θ's sum to one, this equality of means must hold). An alternative strategy, proposed in Geltner et al. (2003), is to calibrate the volatility of R_t^* to a specific number, but this approach is rather ad-hoc, as there is little guidance what that number should be.

Korteweg and Westerfield (2022) apply the Geltner (1991) unsmoothing algorithm to the Cambridge Associates quarterly aggregate VC and buyout returns from 1998 to 2020, using *k* equal to 2. They show that unsmoothing dramatically changes many characteristics of the return series. Return volatility is magnified by a factor of approximately 2.5. Sharpe ratios drop by two thirds, from 0.48 to 0.16 in VC and from 0.93 to 0.31 in buyout. Higher moments (skewness, kurtosis) are closer to zero after unsmoothing, but the change is not very large. For VC, the correlation with public stock returns goes up by about 25% (from 0.42 to 0.51), whereas for buyout, the correlation increases only marginally. Finally, by design, the unsmoothed return series are not autocorrelated.[17] For other applications of unsmoothing methods in private equity, see Goetzmann et al. (2019), Aliaga-Díaz et al. (2020), and Brown et al. (2021b).

A drawback of many unsmoothing methods is that the specification of the observed returns process is typically based on a statistical model, rather than being microfounded with an economic model of price staleness. The resulting unsmoothed returns series can be sensitive to model specification. Moreover, parameter estimates tend to have large standard errors, and can be unstable over time. Imposing additional structure from an economic model has the potential to alleviate some of these issues.

A well-founded model of price staleness should also explain how fund-level staleness aggregates up. For example, if fund-level observed returns are a weighted average of past fundamental returns, as in equation (7), then aggregate-level returns follow the same process only if portfolio weights are fixed over time. Couts et al. (2020) present evidence that traditional unsmoothing techniques applied at a granular level can still leave substantial autocorrelation in aggregate returns.

A final, fundamental concern with many unsmoothing algorithms is that they aim to produce an uncorrelated return series. This is a reasonable assumption in efficient markets, where money can be moved around easily and costlessly, such that fundamental returns are likely (close to) independent over time. In PE, however, nonstale returns could still exhibit material autocorrelation, given the limited arbitrage opportunities due to the illiquid nature of PE investments. The model in Geltner (1993) is an example of an algorithm that does allow for autocorrelation in unsmoothed returns.

3.5.3 Dimson correction

As early as the 1960s, researchers in public equities recognized that nonsynchronous trading of securities could have a material effect on risk loading estimates. Nonsynchronous trading is like staleness in that a security's price does not reflect its current market value if there is no trade in the current period and its historical price is reported instead. Two key papers in this literature are Scholes and Williams (1977) and Dimson (1979). The latter is the focus of this section, by virtue of it having been applied to private equity in the literature.

Dimson develops an easily applicable method of correcting a security's beta estimates if observed expected returns are a weighted average of current and past realized fundamental returns. This specification is almost identical to equation (7), the key difference being that the left-hand side is an expectation. In its general form, if both the security and the risk factors are subject to staleness, the algorithm involves adding leads and lags of the factor returns to the model in equation (3). The corrected beta estimates are the sum of the loadings of all leads and lags, including the contemporaneous loading, for each factor. Since commonly used risk factors are not stale, it is not necessary to include leads, and the Dimson regression becomes

$$R_{i,t} - R_t^F = \alpha_i + \sum_{s=-S}^{0} \beta_{i,s}^{'} F_{t+s} + \varepsilon_{i,t}. \tag{10}$$

Note that if the returns to security *i* do not suffer from staleness, and are serially uncorrelated, then the lagged factor loadings are all zero, and the model simplifies to the standard factor model of equation (3).

Table 4 Panel B shows an application of the Dimson method to the quarterly Cambridge Associates returns for the VC, buyout, and real estate strategies. The regressions use the CAPM and include 6 lags of the market factor, although only the first few lags (3 for VC and buyout and 4 for real estate) are statistically significant. The bottom row of Panel B shows the resulting Dimson beta estimate for each strategy, which is the sum of the contemporaneous and lagged betas. The VC beta estimate is 1.658, almost three times the size of the estimate of 0.579 that results from using only the contemporaneous factor return (see Panel A). For buyout and real estate, the Dimson beta estimates are 0.851 and 0.713, compared to 0.465 and 0.166 based on contemporaneous returns only. Given the higher beta estimates, it is not a surprise that alpha estimates are lower: Whereas alphas are positive and statistically significant for all strategies in the standard contemporaneous factor return model in Panel A, for both VC and real estate the null hypothesis of zero alpha can no longer be rejected in the Dimson method in Panel B. For buyout, the alpha drops from 1.9% to 1.1% per quarter but it remains statistically significant.

Table 5 summarizes the PE literature that has applied the Dimson method. Most papers use quarterly NAV-based aggregate returns. Exceptions are Woodward and Hall (2004), who use a monthly VC index constructed from portfolio company data (Metrick and Yasuda, 2010b, also report results for this same index, albeit over a different sample period, as well for the NAV-based Cambridge Associates index); Anson (2007), who uses fund-level returns; Bilo et al. (2005) and Phalippou (2010), both of whom uses publicly traded PE vehicles (see section 5 for a description of these papers), Boyer et al. (2022a), who use a proprietary data set of secondary market transactions, and Agarwal et al. (2022), who use returns of individual startup company securities held by mutual fund families. The number of lags used in the Dimson regressions ranges from 1 to 8 quarters.

[INSERT TABLE 5 AROUND HERE]

Estimated CAPM betas for VC range from 0.9 to 2.2, depending on the data set, sample period and factor used. For broad-based stock market indices such as the S&P 500, Wilshire 5000, or the CRSP value weighted index, the betas are centered around 1.2 to 1.4. Loadings tend to be closer to one in regressions of VC returns on the Nasdaq and Russell 2000. This result is to be expected, as these factors are closer proxies to startup firms, that is, small to medium-sized growth firms (based on the zero to positive loading on the small-minus-big (SMB) factor and the negative loading on the high-minus-low book-to-market (HML) factor in the Fama and French, 1993, model estimates in Metrick and Yasuda, 2010b, Ewens et al., 2013, and Agarwal et al., 2022, amongst others; see Table 5). For buyout, CAPM betas for NAV-based PE returns are in the range of 0.7 to 1.0 when using a broad public stock index for the market portfolio. Loadings on the Nasdaq and Russell 2000 are lower, but these are less relevant benchmarks for buyout funds, which tend to invest in mid- to large-sized value firms (as indicated by the insignificant loading on SMB and the positive loading on HML in Ewens et al., 2013 (see also Stafford, 2022, for further evidence that, historically, buyout loads on value). Using the Russell 1000 index of large stocks as the factor, Anson (2007) finds a loading of 0.7. Finally, Phalippou (2010) reports a higher beta of 1.5 for publicly traded buyout funds, and Boyer et al. (2022a) estimate a buyout beta of 1.8 for an index constructed using secondary market prices (sections 5 and 3.5.5 below dive deeper into publicly traded vehicles and secondary PE market data, respectively).

Estimated alphas for VC are generally close to zero, with estimates centered around roughly 0.5% per quarter, whereas buyout alpha estimates are higher, with most estimates around 1.2% to 1.4% per quarter. This pattern of low VC alphas and higher buyout alphas is one that is echoed in many other studies, across different methodologies.

Dimson uses a probabilistic model of trading to motivate his method, but this causes some tension with the specification of the return process that is the foundation for the beta correction. The trading model implies that expected observed prices are a weighted average of current and past (unobserved) fundamental prices. This weighted-averaging property carries over to expected observed capital gains, defined as the change in expected (log) observed prices. However, it does not apply to total returns that include cash flows, since capital calls and distributions are observed when they realize. Hence, more research is needed to work out the microfoundation of the empirical model.

Finally, the Dimson correction is subject to a similar concern as many of the unsmoothing methods in that it assumes that fundamental returns are serially uncorrelated. This is not necessarily a natural assumption in PE, as argued in the previous section.

3.5.4 Imputation and filtering

One of the earliest proposed solutions for the staleness issue in the PE literature was to simply impute any missing prices. Gompers and Lerner (1997) pioneered this approach. Using data on portfolio company holdings for a single large PE firm, E.M. Warburg, Pincus & Co., they mark each unobserved company valuation to market in each quarter by using the return to an equal-weighted index of publicly traded firms in the same three-digit Standard Industrial Classification (SIC) industry from the time of the last observed value. They then aggregate all valuations and cash flows, and compute the quarterly returns on Warburg, Pincus's portfolio for the period 1972 to 1997 (this calculation is analogous to equation (6), using the portfolio's cash flows and valuations). They estimate a CAPM beta of 1.4 and an alpha of 2.0% per quarter. For the Fama and French (1993) model they find a loading of 1.3 on the market factor, 0.8 on SMB, and 0.1 on HML, and an alpha of 1.7% per quarter. The *t*-statistics on the alphas are 1.8 and statistically significant at the 10% level (for both models). Note that these factor loadings cannot be attributed to one single strategy, since Warburg, Pincus holds a diversified mix of VC, buyout, and other investments. This, and the fact that the loadings are estimated for a single firm, make it difficult to draw strong comparisons with other papers.

A drawback of the above imputation technique is that the mark-to-market value is only as good as the proxy that is used to update valuations. For example, in VC and buyout, there is an inherent tension in using public company returns as a proxy for private company returns, since being publicly traded is an endogenous decision, and these firms may be inherently different. A second drawback is that portfolio holdings information is needed to impute a fund's NAV (or the aggregate PE market). Portfolio company data is not part of standard fund data sets, and thus more difficult to obtain, especially if they need to be linked to specific funds.

Two papers develop methods to filter PE returns without the need for portfolio company data. Brown et al. (2021a) develop a state-space model to filter "true" valuations of individual funds.

They assume that the evolution of (latent) true fund valuations follow a factor model, and that observed NAVs are noisy observations of the current and past true valuations (and thus, observed NAVs exhibit staleness). While their method imposes a high degree of structure on the problem, an interesting aspect is that they estimate fund-specific alpha and beta estimates. For a data set of funds raised between 1983 and 2008, they report an average (median) VC fund CAPM beta of 1.2 (1.1), with an annual alpha of 0.1% (–2.6%) per year. The average (median) buyout fund has a beta of 1.1 (1.0) and an annual alpha of 1.5% (2.1%). They find a high degree of variation in these parameters across funds: for VC, the first and third quartile betas are 0.9 and 1.4, and alphas range from –11.4% to +6.7%. For buyout the beta range is from 0.9 to 1.3, with alphas ranging from –5.5% to +9.6%.

The second paper, by Ang et al. (2018), use Bayesian Markov chain Monte Carlo methods to filter a time series of realized aggregate PE returns from cash flows of funds raised between 1994 and 2008 (with cash flows until 2015). They then regress these returns on various sets of factors to estimate factor loadings and risk-adjusted returns. For VC (buyout) they report a CAPM beta of 1.8 (1.3), and Fama-French three-factor loadings of 1.7 (1.2), 0.8 (0.5) and –0.6 (0.6) for the market, SMB, and HML factors, respectively (the HML loading is statistically insignificant for VC, and the size loading is insignificant in buyout). For the CAPM and three-factor models the VC alpha is close to zero and insignificant, whereas it is positive and significant for buyout. For other models (discussed in section 3.7.3) the VC alpha varies between –3% to –5% per year (-4% and +2% for buyout). The Ang et al. method does not use observed NAVs as an input to their model. As such, the paper is closer to the stochastic discount factor methods that are discussed in section 3.7.

Finally, there is a small literature on imputing investment-level valuations in venture capital. Most of this literature revolves around constructing an aggregate VC index by adopting some form of repeat-sales regression (RSR), extended to account for the fact that failure is unobserved for many startups. Related work filters portfolio company valuations using a state-space model. I discuss these methodologies and the relevant literature in sections 6.2 and 6.3, which deal specifically with investment-level data.[18]

3.5.5 Secondary market prices

A final option for dealing with NAV issues is to simply replace them with actual market prices of sales of fund stakes by LPs. This seems like a sensible solution, but there are some complications. The secondary market for LP stakes only started to become material after the turn of the millennium, so the observed time series is quite short. The market underwent massive growth in the wake of the global financial crisis of 2008, when many LPs needed to reduce their exposure to private equity (Hege and Nuti, 2011). Since then, the market has continued to mature, and it is in the process of transforming from pure liquidity provision to a widely accepted means for portfolio rebalancing. Despite the substantial growth, the market is still very illiquid, and transaction costs are high. Most sales are conducted in an auction format and can take up to six months to finalize.

Nadauld et al. (2019) collect data on fund transactions from a leading intermediary. They report that the average purchase price as a percentage of NAV in 2008 and 2009 was 78.1% and 54.4%, respectively, indicative of the high need for liquidity by sellers during this crisis period (but possibly also a sign of reluctance by GPs to mark down NAVs). Average purchase prices for the years 2010 to 2014, the end of their sample period, fluctuated between 82.2% and 93.2% of NAV. There was substantial variation in prices paid within a given year, with a standard deviation of around 32%.

Boyer et al. (2022b) construct transactions-based PE indices for VC and buyout between 2006 and 2017, using a Heckman selection model to account for the fact that sellers choose which funds to sell. The resulting CAPM betas and volatilities are higher than NAV-based indices, and alphas are lower. They estimate a VC beta of 1.0, and a buyout beta of 2.3 to 2.4 (depending on whether they use price or size-weighting across funds). Alphas are negative but statistically insignificant for both strategies. On the one hand, the transactions-based betas are considerably higher than NAV-based betas of 0.3 and 0.5, which they estimate on contemporaneous returns.[19] On the other hand, compared to Dimson-corrected betas in Table 5, the VC beta is on the lower end of the estimated range, whereas the buyout beta is considerably higher. However, the large buyout beta appears to be largely driven by the crisis: Eliminating the years 2008 and 2009 results in a transactions-based VC beta of 0.7 and a buyout beta of 0.9. The latter is more in line with Dimson buyout beta estimates. Whether or not it is legitimate to exclude the financial crisis is up for debate. On the one hand it is a dangerous practice to selectively exclude legitimate data. On the other hand, perhaps the large discounts that are the root cause of the high estimated betas (as transacted prices tumbled relative to reported NAVs), are due to the selected subset of LPs that were in a bind and willing to sell at almost any price (a subtly different selection problem than the fund selection problem mentioned above). But these prices, especially when realized in a thin market (Hege and Nuti, 2009, point out that the secondary market even temporarily froze in early 2009), may not be representative of the fundamental value of the transacted funds' portfolio companies for the vast majority of LPs that did not suffer a (large) liquidity shock and were not looking to get out of their commitments. Albuquerque et al. (2018) find evidence that liquidity provision in the secondary market is an important driver of the variation in discounts to NAV, so perhaps these types of illiquidity features are better captured by a separate aggregate liquidity factor; section 3.7.3 discusses liquidity premia estimates from liquidity factors. A final complication with secondary market prices that should be mentioned, is that transactions often involve portfolios of funds, and it is at times difficult to assign prices to the individual funds that make up the portfolio.

With the continued growth in the secondary market, the transactions-based approach is a promising avenue for future research, not in the least because it can also circumvent some other concerns with NAVs besides staleness, as the next section explains.

3.6 Other Concerns with NAVs

As is evident from the prior section, the issues surrounding staleness of NAVs occupies a fair amount of space in the PE literature, although there is some debate whether it's a bug or a feature.[20] However, besides staleness, there are other three other issues with using NAVs as proxies for fund market values: manipulation, fees, and, for VC specifically, the use of post-money valuations to compute NAVs.

Regarding manipulation, on the one hand there is evidence that some GPs may strategically manipulate reported NAVs, especially those who are underperforming and want to raise another fund (e.g., Phalippou and Gottschalg, 2009, Cumming and Walz, 2010, Jenkinson et al., 2013, Barber and Yasuda, 2017, Chakraborty and Ewens, 2018, Brown et al., 2019, Jenkinson et al., 2020, Smith et al., 2022). On the other hand, Hüther (2021) does not find evidence of inflated interim valuations at the deal-level, but instead, that valuation peaks before fundraising are the result of poorer investment opportunities that are also observed for non-fundraising firms with similarly-aged funds.

A second concern is that NAVs are usually reported as the value of the fund's investments, gross of fees. But fund cash flows are reported net of fees in commercial data sets. Therefore, performance measures that mix cash flows and NAVs in their calculation, such as the NAV-based returns above, are difficult to interpret and compare. One piece of suggestive evidence of the difference between gross and net of fee NAVs can be found in a sample of publicly traded funds of funds in Jegadeesh et al. (2015), who report an average discount of market prices to NAV of 11.8%, measured one year after IPO. However, Jegadeesh et al. show that there are other reasons why funds may trade at a discount. Moreover, public funds-of-funds are not representative of private primary funds that make up the bulk of the PE population (see section 5 for details), so this number is best treated as an extremely speculative suggestion that the difference may be sizeable. Another dimension that is important for risk and return estimates, for which there is no existing evidence that I know of, is how strongly (if at all) the difference between gross and net of fee NAVs varies with factor returns.

For VC exclusively, a third issue is that portfolio company valuations are often marked to the post-money valuations of new financing rounds, when they occur. Post-money valuations are computed as the invested amount in a round divided by the percentage of equity that the investor owns upon conversion to common stock (the security sold is usually some form of convertible preferred equity). For example, a $2 million investment for 10% of the equity upon conversion would translate to a post-money valuation of $20 million. As pointed out in Metrick and Yasuda (2010b, p.318), Gornall and Strebulaev (2020), and Ewens, Gorbenko, Korteweg (2022), post-money valuations are not equal to fundamental (market) valuations, as they erroneously assume that all investors own the same security (founders, for example, usually own common stock without any of the preferred terms that VCs typically receive, whereas Chernenko et al. (2021) show that mutual fund investors tend to negotiate for stronger downside protections). Post-money valuations usually overstate the fundamental valuation, and the difference can be large (Gornall and Strebulaev estimate an average of 48% for a sample of unicorns, and Ewens, Gorbenko, and Korteweg estimate a 20% difference for a representative first-round contract). How this affects risk and return estimates is more subtle, as it depends on the correlation between contract terms and risk factor realizations. As far as I am aware, there is no existing empirical evidence regarding the strength of these correlations.

Table 6 presents a crude overview of the potential of the various methods from section 3.5 in dealing with the above problems, possibly with modification. To be clear, these are my unverified speculations. To my knowledge, no extant research validates the claims from the table (in fact, it is not even established how first-order some of these problems are).

[INSERT TABLE 6 AROUND HERE]

As far as NAV manipulation is concerned, imputation/filtering methods may be able to correct for manipulation by using information from public markets, or, if manipulation is transitory, discipline imposed from the time-series of valuations. Methods that base filtered NAVs on discounted future fund cash flows (e.g., Ang et al., 2018, Brown et al., 2021a) may also work. Secondary market prices can help if buyers see through any manipulation (Brown et al., 2019, show results that suggest they might). I do not see a way how the other methods (long-horizon returns, unsmoothing, and Dimson correction) can get around the issue. On this note, it is worthwhile reiterating the earlier call for more research into the development of a manipulation-proof PE performance metric.

For the problems with fees and post-money valuations, a solution using filtering methods might be possible but is not obvious. One potential solution is to base filtered NAVs on discounted future fund cash flows, especially if they filter PE returns directly as in Ang et al. (2018). To the extent observed NAVs are used as observations in the filtering process, the observation equation should recognize that the deviation of NAV from true values might have a permanent component. More research is needed here to show whether such solutions can work. A more intuitive solution to both problems, at least in principle, is the use of secondary market prices. The market price paid for a fund in a secondary transaction should account for the expected carried interest and other fees that may need to be paid on profits from portfolio company exits (in fact, this may be one of the reasons why there is a discount to NAV in the first place). Similarly, the market price should reflect the fundamental value of a VC fund’s portfolio companies, not the (biased) post-money valuation.

3.7 Stochastic Discount Factors

Given the myriad issues with NAVs, many performance metrics either avoid using them altogether, or they use them as little as possible. For example, IRR and TVPI, the two most common metrics used in practice, rely solely on cash flows when applied to liquidated funds. For a not-yet-liquidated fund, both measures only use the last observed NAV, ignoring all previous NAVs in the fund's history. However, as mentioned in section 3.2.3, a key drawback to IRR and TVPI is that they ignore risk. A natural way to account for risk is to consider the discounted cash flows of a fund, where the discount reflects the riskiness of the cash flows. If a fund is fair investment, then the net present value (that is, the sum of the discounted future capital calls and distributions) should be zero in expectation.

A popular way to perform discounting is to use a stochastic discount factor (SDF). I do not cover SDFs in detail (Cochrane, 2005b, is a great read on this topic), but I briefly illustrate the underlying intuition and mechanics. Consider the following, stylized two-period model. An investor wants to maximize the expected utility of today's (time *t)* and next period's (time *t*+1) consumption. Given a time-separable, increasing and concave utility function $U(X)$ in consumption, X, and a time preference $\rho < 1$ that measures the degree to which an investor prefers receiving a certain payoff of \$1 today versus \$1 tomorrow, the choice problem is

$$\max_{X_t, X_{t+1}} U(X_t) + \rho E_t\left[U(X_{t+1})\right]. \tag{11}$$

The investor has no sources of income but has wealth, W_t, out of which she can consume. Anything not consumed today can be invested in a risky security that costs P_t per unit today and returns R_{t+1} next period. The return is a random variable as of time *t*. From the first-order condition of this optimization problem it follows that[21]

$$E_t\left[M_{t,1}R_{t+1}\right]=1, \tag{12}$$

with the SDF

$$M_{t,1}=\rho\frac{U'\left(X_{t+1}\right)}{U'\left(X_t\right)}. \tag{13}$$

Equation (12) is the key pricing equation underlying much of modern asset pricing. The expression can also be written in prices by multiplying both sides by P_t:

$$P_t=E_t\left[M_{t,1}\left(P_{t+1}+C_{t+1}\right)\right], \tag{14}$$

where C_{t+1} is next period's net cash flow. Note that in equation (14) I swapped the two sides of the equation from equation (12), to emphasize that today's price is the (conditional) expected payoff multiplied by the SDF. As such, equation (14) clarifies how the SDF takes care of discounting future payoffs.

Intuitively, the SDF can be thought of as a close cousin of state prices. State prices are prices of (fictitious) assets that pay exactly $1 in one state of the world, at one specific future time (here: next period). Bad states of the world have high state prices (and a high realization of the SDF): A payoff (here: $P_{t+1}+C_{t+1}$) will be highly valued when consumption (here: X_{t+1}) is relatively low, such that marginal utility is high. Conversely, good states of the world have low state prices (and low realizations of the SDF).[22] Another way to illustrate this intuition is to rewrite the expectation on the right-hand side of equation (14) as

$$
\begin{aligned}
P_t &= E_t\left[M_{t,1}\right]E_t\left[P_{t+1}+C_{t+1}\right]+Cov_t\left[M_{t,1},P_{t+1}+C_{t+1}\right] \\
&= \frac{E_t\left[P_{t+1}+C_{t+1}\right]}{R_{t+1}^F}+Cov_t\left[M_{t,1},P_{t+1}+C_{t+1}\right].
\end{aligned}
\tag{15}
$$

The second line uses the fact that the risk-free rate is known at time *t*, such that

$E_t\left[M_t R_{t+1}^F\right]=E_t\left[M_t\right]R_{t+1}^F=1$. The first term in equation (15) is thus the value of the payoff discounted at the risk-free rate, and the second term is a risk adjustment that is positive if the payoff is positively correlated with the SDF. In other words, assets whose payoffs are positively correlated with SDF realizations (that is, assets that pay out in bad times), are more desirable and therefore have higher prices.

SDF pricing also applies to longer-horizon payoffs. For example, the present value of a single cash flow that realizes two periods from today is

$$
\begin{aligned}
P_t &= E_t\left[M_{t,1}P_{t+1}\right]=E_t\left[M_{t,1}E_{t+1}\left[M_{t+1,1}C_{t+2}\right]\right]=E_t\left[E_{t+1}\left[M_{t,1}M_{t+1,1}C_{t+2}\right]\right] \\
&= E_t\left[M_{t,1}M_{t+1,1}C_{t+2}\right],
\end{aligned}
\tag{16}
$$

Where the last step invokes the law of iterated expectations. Thus, the two-period SDF is $M_{t,2}=M_{t,1}M_{t+1,1}$, and more generally, for *h* periods, $M_{t,h}=\prod_{s=0}^{h-1}M_{t+s,1}$. Also note that $M_{t,0}=1$, since cash flows that occur immediately require no discounting.

3.7.1 Public Market Equivalent

A PE metric that discounts cash flows, popular in academia and increasingly so in practice, is the public market equivalent (PME). Building on earlier work by Long and Nickels (1996), it is defined by Kaplan and Schoar (2005) as[23]

$$PME_i = \frac{\sum_{d=1}^{D} D_{i,d} \Big/ R^M_{t_0,\tau^D(d)} + NAV_{i,T} \Big/ R^M_{t_0,T-t_0}}{\sum_{k=1}^{K} I_{i,k} \Big/ R^M_{t_0,\tau^I(k)}} . \tag{17}$$

A comparison with equation (1) readily reveals PME to be a discounted TVPI, using the realized public stock market return as the discount rate.[24] To demonstrate the mechanics, consider the example fund in Table 2. The April 10, 2024, distribution of $500,000 is discounted at the 9.5% market return since fund inception to a value of $456,621. The December 20, 2024, distribution's discounted value is $1,000,000/1.244 = $803,859. Together with a discounted final NAV of $3,300,000/1.250 = $2,640,000, the numerator of equation (17) is $3,900,480. The denominator is the discounted value of the seven capital calls, which add up to $4,101,269. The PME of this fund is therefore $3,900,480 / $4,101,269 = 0.95.

One (common) interpretation is that a PME greater (smaller) than one means that the fund performed better (worse) than a hypothetical alternative strategy that invested the fund's capital calls in the public stock market. Put differently, the fund is benchmarked against the public stock market. In the example above, since the PME is less than one, an investor would have been better off investing the fund's capital calls in the public stock market.

Table 1 shows descriptive statistics of Kaplan-Schoar PME for various PE strategies, using the S&P 500 as the proxy for the public stock market. Over the sample period, VC and real estate had a PME centered just below one (at a mean of 0.96 for both strategies, a median of 0.98 for VC and 0.97 for real estate). Private debt looked very similar with a mean PME of 1.01 and a median of 0.98. Natural resources and infrastructure had PMEs centered well below one. In contrast, the mean and median PME for buyout were 1.11 and 1.10. Like IRR and TVPI, the variation in PME is very high (especially in VC) and there is a large positive skewness (except for infrastructure) and excess kurtosis. The final column in Figure 3 shows the kernel-smoothed distribution of PME for the various strategies.

Many papers report PMEs, primarily for VC and buyout, using different data sources and time periods. Table 7 presents a representative (but not comprehensive) overview of these estimates. The table focuses on main PE vehicles (sections 4 and 5 discuss alternative vehicles, co-investments, and publicly traded vehicles).

[INSERT TABLE 7 AROUND HERE]

Table 7 shows that the mean VC PME from Burgiss data tends to be higher (on the order of 1.2 to 1.4) than the Preqin-based estimates from Table 1 (and other papers). The medians are comparable between the two data sets. This finding is consistent with Preqin data missing a few top-quartile VC firms, which affects the mean but not the median. Buyout results are similar between Burgiss and Preqin. The Venture Economics numbers for both buyout and VC funds tend to be lower, due to the performance bias discussed in section 3.2.2.

There is some degree of time-variation in average PMEs. Many VC fund vintages from the 1990s performed very well as they benefited from the internet boom: Harris et al. (2014) report an average (median) PME of 1.99 (1.26) for the decade. The first decade of the 2000s has been poor, however, as average PMEs dopped to 0.91 (0.84). Performance appears to be improving with the 2010 to 2015 vintages (Harris et al., 2022), although many of these funds are still far from finished at the time of this writing. Buyout fund performance has been more stable over time (see also Figures 1 and 2 in Harris et al. (2014) for a more detailed time-series of VC and buyout PMEs by vintage). Notwithstanding this time-series variation, and the fact that there is a cyclical component to cash flows (Robinson and Sensoy, 2016), most of the high PME variance is explained by the cross-section (that is, variation in PMEs across funds of the same vintage year). This dominant role of cross-sectional variation in PE fund returns stands in contrast to other managed funds that primarily invest in public securities, such as mutual funds.

To make a more convenient comparison with alpha estimates from factor models, Gredil et al. (2022) develop an annualized version of the PME, called direct alpha. Although Gredil et al. derive it somewhat differently, direct alpha can be computed as the (fixed, annualized) rate of return that needs to be added to the benchmark return to make the PME equal to one. While useful, one caveat to this measure is that it presumes that one can reinvest at the same alpha, similar to the reinvestment problem with IRR.[25]

If the PME is interpreted as a benchmarking metric, then the choice of benchmark is important, as it is supposed to capture similar risk as the investment it is compared against. Under this interpretation, using the public stock market is equivalent to assuming that the CAPM holds, and that the asset being evaluated has a beta of one. While this may be a sensible choice for buyout funds based on CAPM beta estimates that are indeed often close to one (see Figure 1), it may not be an appropriate choice for other PE strategies. For example, VC beta estimates are typically far above one, and comparing private debt to a public equity investment is a stretch even for risky loans. In recognition of this, some authors choose different benchmark portfolios to compute the

PME. Ljungqvist and Richardson (2003) show PME results with the Nasdaq as the discount rate for distributions, which could be a natural choice for VC (see section 3.5.3) and using the risk-free rate to discount capital calls (they call their metric the Profitability Index, but the calculation is the same as the PME). Harris et al. (2014), Phalippou (2014), Robinson and Sensoy (2013, 2016), and Hüther et al. (2020) use a levered version of the PME, essentially (exogenously) choosing the market beta at which to compute the discount rate. Phalippou (2014) also uses a variety of size and value-sorted portfolios. Gredil (2022) uses publicly traded industry sub-index returns. Munday et al. (2018) study private credit fund PMEs using four different benchmark indices: a high-yield index, a levered loan index, a publicly traded BDC index, and the Cliffwater direct lending index. They find PMEs just below one, except for the levered-loan PME, which is 1.14. Finally, Driessen et al. (2012) take the choice of discounting rate one step further and estimate the factor model parameters needed to make the PME equal to one, and apply their approach to a sample of VC and buyout funds.[26]

A drawback of the benchmarking interpretation as described above is that in most cases the choice of benchmark is rather ad hoc, without any quantitative guidance on whether the appropriate risk-matching was achieved. An alternative interpretation of PME, specifically, as an application of SDF pricing with log-utility investors, avoids the benchmark choice issue altogether. [27] Under this interpretation, PME is a valid measure of risk-adjusted performance without any assumption or restriction on beta(s). To see this, first note that an investor with log utility over wealth has an SDF that equals the reciprocal of the return on her wealth portfolio. Using the market portfolio as a proxy for the wealth portfolio,[28]

$$M_{t,h} = \frac{1}{R_{t,h}^{M}}. \tag{18}$$

Now consider the expected value (as of fund inception) of the PME numerator in equation (17). If we include all future distributions until liquidation, then the final NAV is zero, and the expected numerators is $E_{t_0}\left[\sum_{d=1}^{D} D_{i,d} \Big/ R^{M}_{t_0,\tau^D(d)}\right]$. Switching the expectation and summation, and using equation (18) yields $\sum_{d=1}^{D} E_{t_0}\left[M_{t_0,\tau^D(d)} D_{i,d}\right]$, which is the sum of the present values of all future distributions. By the same logic, the expected value of the denominator is the sum of the present values of all capital calls. Thus, the ratio of the two is how much money we expect to receive, in present value terms, over the life of the fund, for a present value investment of one dollar (importantly, this is not quite the *expected* PME, due to Jensen's inequality). For not-yet-liquidated funds, the final reported NAV is included and discounted at the SDF. This is correct, and the interpretation does not change, if the NAV indeed represents the true net asset value of the fund as of its reported date, but it is of course subject to the same issues with NAV as discussed above.

Under the SDF-pricing interpretation, no assumption on the beta of PE is needed, as the SDF properly values each cash flow depending on the state of the world in which it realizes. This implies that beta may even vary from fund to fund without invalidating the PME metric. For the same reason, the fact that the timing of cash flows may be endogenously chosen by the GP is also not a problem per se. Moreover, the SDF pricing model leaves the distribution of payoffs unrestricted, which is important given the skewed and fat-tailed distributions common in PE. Of course, this does mean that we must assume that the log utility function is correctly specified, which implies that we are also making an assumption about which risk factors matter to investors, as well as their level of risk aversion (I will return to this below).

To summarize, we have two valid but different interpretations of PME: either as a benchmarking exercise that assumes the investment has a CAPM beta of one (without assuming a specific utility function other than what's needed for the CAPM to hold), or as a ratio of present values under the assumption of log-utility investors (but without any restriction on beta).

3.7.2 Generalized PME

Korteweg and Nagel (2016) introduce the generalized PME (GPME), which, as the name implies, is a generalization of PME to an arbitrary SDF:

$$GPME_i = \sum_{j=1}^{J} M_{t_0,\tau(j)} C_{i,j} + M_{t_0,T-t_0} NAV_{i,T} \tag{19}$$

Besides allowing for a generic SDF, the difference with PME is that the above definition is based on net cash flows (including any final NAV as a pseudo-cash flow), rather than a ratio of distributions to contributions. To allow for comparisons across funds, cash flows (as well as the final NAV) can be scaled by the fund's commitment size.[29]

Defining the performance metric on net cash flows has several advantages over the ratio definition. First, the expected GPME is simply the expected NPV of a \$1 commitment, earned over the lifetime of the fund. Thus, unlike the ratio definition, the average GPME in the data has a well-defined interpretation. Note that the key threshold number to compare the GPME against is zero (not one, which is the threshold number in the ratio definition). A second advantage is that the GPME of a portfolio of funds is simply the weighted average of the constituent GPMEs, with weights proportional to capital committed (this is not true for the ratio definition). Third, the net cash flow definition allows for the comparison of two GPMEs through their difference: an “excess” GPME is a GPME itself, namely the GPME of the difference in cash flow streams and final NAVs (see, for example, Lerner et al., 2020, for an application of excess PMEs). Finally, the classification of negative capital calls and distributions is not an issue for the net cash flow definition, unlike the ratio definition (for the same reasons as those discussed in section 3.2.1).

In their empirical application, Korteweg and Nagel (2016) specify an SDF that is exponential-affine in the market return:

$$M_{t,1} = \exp(a - b r_{t+1}^{M}), \tag{20}$$

with r_{t+1}^{M} the one-period log return on the market from time t to $t+1$. This is the SDF for a power utility investor with risk aversion equal to b. It rules out arbitrage opportunities as the SDF is positive in every state of the world, and it compounds nicely for longer horizons,

$$M_{t,h} = \prod_{s=0}^{h-1} M_{t+s,1} = \prod_{s=1}^{h} \exp(a - b r_{t+s}^{M}) = \exp(ah - b \sum_{s=1}^{h} r_{t+s}^{M}) = \exp(ah - b r_{t,h}^{M}), \tag{21}$$

where $r_{t,h}$ is the h-period log-return from time t to $t+h$.

Korteweg and Nagel estimate the SDF parameters to price the public market and risk-free rate.[30] Applying this SDF to a sample of 545 VC funds from Preqin, they estimate a GPME of –0.103, which is statistically significantly different from zero at the 5% level, taking into account the uncertainty about the estimated SDF parameters as well as the cross-sectional correlations between funds due to their overlap in time. This estimate indicates that per $1 commitment, the average VC fund lost a present value of about 10 cents over its lifetime. This loss is driven by the post-1998 vintages in their sample, as the earlier vintages have a GPME that is statistically indistinguishable from zero.

Several recent papers estimate GPME for various investments. Hüther et al. (2021) report positive GPME estimates for buyout funds. Andonov et al. (2021) and Jeffers et al. (2021) estimate a negative GPME for infrastructure and impact funds, respectively. Giacoletti and Westrupp (2018) apply the GPME model to real estate investments. Cordell et al. (2021) estimate multi-factor GPMEs for collateralized loan obligations with a few different SDFs, including the Fama and French (1993) three-factor model and an intermediary asset pricing model.

Note that the SDF implied by the PME is a special case of equation (20) with $a = 0$ and $b = 1$.[31] The SDF parameter point estimates in Korteweg and Nagel are 0.09 and 2.65, and the log-utility model is statistically rejected. Still, it is worthwhile to dig deeper into the differences between PME and GPME. To fix ideas, suppose a fund makes a $1 capital call today and invests it in a portfolio company that is liquidated next year at a return of $R^F + \alpha + \beta\left(R^M - R^F\right)$. In other words, the investment pays off according to the CAPM plus a risk-adjusted return α (I suppress time subscripts to avoid clutter). For simplicity, there is no idiosyncratic return (this could be easily added but does not change the intuition), these are the fund's only cash flows, there are no fees, and the fund is fully liquidated when the investment pays off. The expected PME of this fund is

$$\begin{aligned} E\left[PME\right] &= E\left[\frac{R^F + \alpha + \beta\left(R^M - R^F\right)}{R^M} / 1\right] \\ &= \left(R^F(1-\beta) + \alpha\right) \cdot E\left[1 / R^M\right] + \beta \end{aligned} \tag{22}$$

The first line uses the fact that there is only one capital call (of $1) in the denominator, which is nonrandom as it occurs today, so that there is no Jensen inequality problem with taking expectations of ratios. If the β of the investment were equal to 1, then the expected PME simplifies to $\alpha \cdot E\left[1/R^M\right]+1$. If, in addition, α were zero, then PME equals one. This is not a surprise, because in this case the portfolio firm simply earns the market return.

Next, suppose we have an SDF that prices R^M and R^F, i.e., $E\left[MR^M\right]=E\left[MR^F\right]=1$. Then,

$$\begin{aligned} E\left[GPME\right] &= -1+E\left[M\cdot\left(R^F+\alpha+\beta\left(R^M-R^F\right)\right)\right] \\ &= -1+R^F(1-\beta)E\left[M\right]+\alpha E\left[M\right]+\beta E\left[MR^M\right] \\ &= \frac{\alpha}{R^F} \end{aligned} \tag{23}$$

where the last line uses the fact that the SDF prices the risk-free rate, such that $E\left[M\right]=1/R^F$. Gupta and Van Nieuwerburgh (2021) point out that it is natural to discount alpha, which in this example realizes next year, at the risk-free rate. The important takeaway from equation (23) is that the value of beta does not matter for the expected GPME. Therefore, a levered investment in stocks (which increases beta) will be evaluated correctly by GPME.[32] A benchmarking interpretation to GPME is then that a GPME above zero means that there is a component of PE payoffs that cannot be replicated by a (possibly levered) investment in stocks and bonds (as the SDF prices these two assets in this example; more generally, the replication allows for investments in all of the factors priced by the SDF).

As one might intuitively expect, if investors have log utility, then PME and GPME are equivalent (such that PME equals one when GPME equals zero). This happens because under log preferences, $E\left[1/R^M\right]=1/R^F$, so that $E\left[PME\right]=1+\alpha/R^F$ regardless of the value of beta. If, however, log utility does not hold, beta will enter the expected PME, and leverage will no longer be properly accounted for. With the recent increase in subscription lines of credit usage by PE firms (Schillinger et al., 2019, and Albertus and Denes, 2020), this leverage issue becomes more problematic, as LPs often do not know how levered the fund is, so it becomes near impossible to determine what beta to use to construct a discount rate in a "levered PME" type exercise.

To illustrate the empirical importance of the log-utility restriction of PME, Table 2 in Korteweg (2019) shows its impact on effective discount rates used by PME, for each decade between the 1960s and the 2010s, and for various levels of beta. While PME works perfectly well when beta equals one, for an investment with a beta of 2 (which is in the range of estimates for VC) the PME-implied benchmark return can be off by as much as 10% per year or more. While the PME-implied benchmark rate is usually too low for investments with a beta above one, this is not necessarily true (for example, in the 1970s and the first decade of the 2000s the implied benchmark was too high). The reason for this result is that log-utility restricts both the risk-free rate (to $1/E\left[1/R^M\right]$) and the implied market risk premium (because it restricts the coefficient of risk aversion to one; see Korteweg and Nagel, 2016, for explicit expressions under lognormal returns). When beta equals one, the mismatch between the PME-implied and the actual risk-free rate and market premium happen to cancel out, but this is not generally true. In contrast, the GPME with the SDF of equation (20) and an additional assumption that returns are lognormal, implies a CAPM in logs:

$$\log E\left[R_{i,t+1}\right]=r^F+\beta\left(\log E\left[R_{t+1}^M\right]-r^F\right). \tag{24}$$

Thus, the GPME can properly benchmark any asset that conforms to this model (but note that the lognormal returns assumption is not necessary for GPME to hold, giving it an advantage over factor models, especially in PE). This duality generalizes to multiple factors, that is, $M_{t,1} = \exp(a - b' f_{t+1})$ corresponds to a multi-factor model in log returns.

Since SDF pricing works on expectations, per equation (12), the GPME method works well when taking averages over large samples (especially in the time series, to achieve more variation in the factor realizations). A downside is that the GPME for any individual fund can be very noisy. To address this, Korteweg and Nagel (2022) develop an extension to the GPME method that produces more accurate individual fund-level estimates of risk-adjusted performance.

A final important point is that Gredil (2022) shows evidence that VC and buyout GPs have skill in timing the public equity market, but any such type of factor timing skill is not recognized as outperformance by (G)PME, to the extent that the factors are priced by the SDF. While factor timing skill is valuable, it is better implemented through other, more liquid, securities with lower transaction costs, such as stocks and bonds, when possible. Rather, the (G)PME metric is geared toward identifying whether GPs possess skill that is specific to PE and cannot be realized elsewhere.

3.7.3 Other SDFs

With a machinery to use generic SDFs in place, we can turn our attention to exploring what risk factors besides the market may be in the SDF, or, in the dual representation, factor models. This is a recent but growing area of research in PE. Loadings on the Fama and French (1993) SMB and HML factors were described in sections 3.5.3 and 3.5.4 above, so I will not revisit these results here. Ang et al. (2018) apply the Fama and French (2015) 5-factor model and find that VC funds do not have statistically significant loadings on the profitability or investment factors. Buyout funds have a significant positive loading on the profitability factor but do not load on investment.

Gredil et al. (2020a) estimate two consumption-based asset pricing models: the external habit and the long-run risks model. Results are mixed. They find small VC outperformance after the turn of the millennium, although this is not reliably statistically significant across specifications. For buyout they cannot reject the null hypothesis of zero outperformance for most specifications.

Illiquidity is a major feature of PE, but the size of the illiquidity premium is a priori unclear. On the one hand, most investors in PE have long horizons and should be well positioned to handle liquidity shocks. On the other hand, not only is there illiquidity in the realization of distributions, LPs are also exposed to funding liquidity shocks due to GPs having the discretion to call capital at any time (Lerner and Schoar, 2004), and penalties for defaulting on capital call are stiff (e.g., Litvak, 2004, and Banal-Estañol et al., 2017). Franzoni et al. (2012), Ang et al. (2018), and Hüther et al. (2021) find that buyout funds load positively and significantly on the Pástor and Stambaugh (2003) illiquidity factor, with an illiquidity premium of about 3% per year. Buchner (2016) and Ang et al. (2018) find insignificant loadings for VC. Metrick and Yasuda (2010b) also estimate an insignificant loading when using the Cambridge Associates VC index returns, but a small positive loading for the Sand Hill Econometrics index (a VC index). Agarwal et al. (2022) also do not find any reliably positive loadings on the liquidity factor for mutual fund investments in private firms. However, the Pástor and Stambaugh factor may not be capturing the type of illiquidity that PE is exposed to, as it was developed to capture order-flow illiquidity in publicly traded stocks. A small literature considers the PE illiquidity premium from a portfolio allocation model perspective, but this falls outside the scope of this review. The interested reader can find a survey of these papers in Korteweg and Westerfield (2022).

Idiosyncratic risk may be important in PE. The potential benefits to diversification are large given the large cross-sectional variation in returns, but agents are under-diversified and risks are imperfectly shared (Opp, 2019): GPs hold a sizeable amount of wealth in their own funds, for incentive-provision reasons (Ewens et al., 2013), and a typical LP commits to only three PE funds per year (Gredil et al., 2020b). Ewens et al. find supporting evidence that GPs in VC price their idiosyncratic risk exposure into the deals they negotiate with portfolio companies. In the cross-section of GPs, those with higher idiosyncratic risk earn higher expected returns. Peters (2018) takes a time-series perspective and shows that VC loads positively on aggregate idiosyncratic volatility shocks. This implies lower expected returns to VC, since investors prefer securities that pay off when idiosyncratic risk spikes unexpectedly (put differently, this risk carries a negative risk premium). Notably, the role of idiosyncratic risk for buyout returns has thus far been unexplored in the literature.

Gupta and Van Nieuwerburgh (2021) introduce term structure factors, factors related to inflation, GDP growth, and price-dividend ratios and dividend growth rates of various public equity portfolios into the SDF. The introduction of term structure effects seems especially important for private equity funds, given their long investment horizon, and even more so for buyout given its reliance on leverage (see also Giommetti and Jørgensen, 2021). Overall, Gupta and Van Nieuwerburgh estimate negative average risk-adjusted profits for buyout, VC, real estate, and infrastructure funds. I discuss their empirical methodology, which is based on SDF pricing but implemented somewhat differently than GPME, in the next section.

Hüther et al. (2021) use a five-factor SDF with credit market factors, calibrated to price the loans and bonds of leveraged buyout deals. Using this SDF to price buyout PE fund cash flows using the GPME methodology (they call this the "Credit Market Equivalent"), the authors find positive but statistically insignificant risk-adjusted performance.[33]

Korteweg et al. (2021) move away from a representative agent model in which there is one unique SDF, and instead use individual U.S. public pension funds' performance to evaluate the benefits of investing in PE. They find no reliable evidence that pension plans either over or under-allocated to PE, with the possible exception of buyout funds, which have historically looked attractive.

There are also a few potential risk factors that have not yet received much attention in the PE literature. First, while skewness in payoffs and returns is well-known, I am not aware of any work that quantifies (co-)skewness and its associated risk premium in PE (e.g., Harvey and Siddique, 2000). Is skewness mostly idiosyncratic or is there a large systematic component? Second, research in public equities suggests that many common risk factors, such as the value, profit, and investment factors, may be subsumed by a single cash flow duration factor (e.g., DeChow et al., 2004, Chen and Li, 2020, Gonçalves, 2021, Gormsen and Lazarus, 2021). Third, there is some suggestive evidence that there may be a PE-specific risk factor (Korteweg and Sørensen, 2010, Ang et al., 2018), but the nature and size of this potential factor has not yet been investigated.

Finally, generalizations such as time-varying SDF loadings (corresponding to time-varying risk premia) and using conditioning information in computing the expected GPME are as-of-yet left unexplored.

3.7.4 Critiques

I end this section on fund data with two critiques that future research can hopefully resolve. First, the use of SDFs and their close counterpart, factor models, is ubiquitous in asset pricing research, but little attention has been paid to how appropriate their application is in PE. Most importantly, a key underlying assumption to the derivation of both SDF pricing and factor models is that investors can make very small adjustments to their portfolios at very low transaction costs. While reasonable for public equities and other securities, this is clearly violated in PE. How does this affect implementations of the SDF/factor model approach in PE?

The second critique revolves around an empirical regularity documented by Martin (2012), which may impact the performance of the (G)PME method: for long horizons (i.e., large h), while the expectation of $M_{t,h}R_{t,h}$ equals one, most sample paths have realizations that are close to zero (mainly because M often ends up close to zero) and a few paths have very high outcomes (i.e., really bad times that coincide with high returns). A potential peso problem is that small samples may not include those high realizations, resulting in sample averages that can be significantly below one (or zero, in the case of GPME). Thus, while the (G)PME approach of using SDF realizations is consistent and unbiased, the distribution of realized (G)PME is skewed and may result in estimated performance that is below the mean in many cases. To be fair, the Martin result is for very long horizons, on the order of a century or longer. The effect for PE funds, which operate on the order of 10 to 20 years may be small, as suggested by simulations in Korteweg and Nagel (2022).

To the extent that the above peso problem does pose a problem, the approach by Gupta and Van Nieuwerburgh (2021) is a potential solution, although it comes at the cost of assuming additional structure. Their two-step method first constructs replicating portfolios of each PE fund, using a series of zero-coupon bonds and dividend and capital gain "strips" that pay off at a single date. The second step is to specify and calibrate an asset pricing model that prices these strips. This model requires not only an SDF, but also a high-dimensional vector autoregression model for the state variables, with specific distributional assumptions. The benefit of this method is that it avoids SDF realizations and goes straight to expectations. The value of any non-replicable PE fund cash flows, minus any difference in the cost of constructing the replicating portfolio and the PE capital committed, is then the risk-adjusted profit of the fund. Note that this risk-adjusted profit measure gives credit to the GP for market-timing capital calls, unlike (G)PME.

4. Alternative Vehicles and Funds of funds

The empirical results described in the previous section focus on flagship PE funds. Since the turn of the millennium, an increasing proportion of investments in PE are being made in vehicles outside of the main funds: In the 1980s, less than 10% of capital invested in PE went to alternative vehicles, compared to nearly 40% by 2017 (Lerner et al., 2020).

Alternative vehicles can be broadly classified as GP-directed and discretionary investments, depending on the degree of LP control. GP-directed vehicles (also called parallel or sidecar funds) leave the key decision powers with the GP. They generally make the same investments and dispositions as the main fund, their main purpose being to accommodate tax, regulatory, or other requirements of certain LPs. Sidecar funds with a lower fee structure than the main fund, and continuation funds (an increasingly popular way to keep exposure to portfolio companies when the original investing fund is at the end of its lifecycle) can also be included in this category.

In discretionary investments, the LP has a say in which investments are made. In some instances, select LPs are offered the opportunity to co-invest alongside the GP in a particular portfolio company, in other cases the GP sets up a separate fund. These funds may be in the form of a pledge fund where LPs decide on transactions on a deal-by-deal basis, or a co-investment or overage fund that is raised at the same time as the main fund. Finally, LPs sometimes engage in solo investments, in which they source and invest in a deal alone. These solo transactions are the playground of larger LPs that have the critical mass to support a dedicated PE team to generate deal flow and perform due diligence and post-transaction monitoring.

Lerner et al. (2020) report that the number of GP-directed and discretionary vehicles (including solo investments in previously PE-financed companies) are roughly equal, but the former represent about 25% more capital ($50 billion versus $38 billion).

Standard commercial data sets, as of this writing, either do not contain alternative funds and solo investments, or they are not well identified or classified. As such, the empirical literature on the performance of alternative vehicles is very small. The earlier key papers focused on co-investments and solo investments. Fang et al. (2015), in proprietary data from seven large LPs, find little evidence of net-of-fee outperformance. Solo transactions, while having PMEs above one, underperform fund benchmarks, especially in VC. For co-investments, PMEs are lower than the associated funds. This result is surprising, especially for solo deals (given that LPs pay no fees), and the authors say this is indicative of adverse selection. However, Braun et al. (2020) find no evidence of adverse selection in gross-of-fee PMEs from deal-level co-investments data in VC and buyout, sourced from CapitalIQ. While these contrasting results appear puzzling, the explanation could simply be that the two papers have very different data sets, representing distinct slices of the population. Lerner et al. (2020) use custodial data from State Street, which is more comprehensive than the data in prior studies. It also covers private debt in addition to VC and buyout. They find that performance in alternatives depends on the LPs' and GPs' qualities, and the match between LP and GP. While the average alternative vehicle has a lower PME than its corresponding main fund, the alternatives of top GPs tend to outperform. Top GPs offer preferential access to LPs that have high bargaining power (for example, due to their reputation as measured by past performance, or the size of their checkbook), and these LPs tend to earn higher PMEs in their alternative vehicle investments. The biggest difference in access is in discretionary vehicles, which on average tend to perform the best. Finally, alternatives have performed better in the years since the global financial crisis of 2008, a period that was not well covered by prior work.

Funds of funds are a potentially attractive investment for small investors. GPs typically require a minimum capital commitment, so that it is difficult to build a diversified portfolio of funds with a small PE budget. By pooling their money with other small investors in a fund of funds, LPs can achieve a higher level of diversification, and share the due diligence and monitoring expenses of running a PE portfolio. The ability to make larger commitments may also give the fund of funds opportunities to invest with high-quality GPs, which smaller investors may not be able to access. The main drawback is that they pay an additional layer of fees to the fund-of-funds manager. Since most fund performance studies are focused on direct fund investments, fund-of-funds are usually excluded from the data, so less is known about their performance. Harris et al. (2018) study the performance of fund-of-funds specifically. They main result is that VC funds of funds earn an average PME of 1.16, on par with the performance of direct VC fund investing, whereas the average PME of buyout funds of funds is 1.14, which is below the average PME of investing in direct funds.

Like funds of funds, secondary funds (which buy LP stakes in existing funds) are usually excluded from data sets that study fund PE fund performance. As described in section 3.5.5, the secondary market in LP stakes has grown dramatically since the global financial crisis of 2008, and it is undergoing a transformation from pure liquidity provision to LPs to a more common portfolio rebalancing mechanism that encompasses a wider set of participants. Secondary funds performed extremely well in the aftermath of the financial crisis, when there were ample opportunities to buy fund stakes at deep discounts to NAV, but returns have come down since (Nadauld et al., 2019). It will be interesting to see how performance in this market develops as it matures.

5. Publicly Traded Vehicles

A small slice of the PE space is traded publicly, either over the counter or on stock exchanges. With regularly observed market prices, one can assess risk and return using traditional methods, thereby avoiding many of the issues that plague private funds.

There are two main types of publicly listed vehicles: management companies and funds. Listed management companies, such as Blackstone and KKR, provide investors the opportunity to share in the fees and carried interest income earned by GPs, and their risk-return profile should therefore be interpreted as approximating GPs' exposures in PE funds. Listed funds, on the other hand, can be thought of as traded net-of-fee LP stakes. These are dominated by broad, evergreen funds-of-funds portfolios, such as 3i Group, that are invested across the PE spectrum but also often include sizeable stakes in publicly traded securities. Additionally, in the U.S., many traded funds are business development corporations (BDCs), which invest primarily in debt securities such as high-yield loans to mid-cap sized firms, and Small Business Investment Companies (SBICs), a type of leveraged VC fund that is more heavily regulated than limited partnerships but receives favorable borrowing rates due to government guarantees (in fact, many BDCs have SBIC subsidiaries). Most publicly traded funds are listed in Europe and concentrate in the U.K.

due to favorable tax rules for Investment Trusts and Venture Capital Trusts. Several indices of listed PE are widely available, such as the S&P Listed Private Equity Index, the Société Générale Privex index, the ALPS-RedRocks Global Listed Private Equity index, and the LPX Listed Private Equity Index Series. It is not uncommon for these indices to be used as PE investment benchmark returns, even though they usually include both the listed management companies and funds, despite their likely very different risk and return profiles.

Historically, the universe of publicly traded vehicles is extremely small. Martin and Petty (1983) identify 37 VC vehicles that traded publicly at some time during the 1970s, but only 17 had price data for their chosen sample period of 1974 to 1979. Their sample is further reduced to 11 vehicles, eight SBICs and three VC firms, due to the fact that many are only sparsely traded. Similarly, Brophy and Gunther (1988) identify 12 publicly traded vehicles for 1981 to 1985, all SBICs and BDCs, and all of which are traded over the counter (not on a centralized exchange).

In conjunction with the growth of PE as an asset class, the number of listed vehicles has expanded since the mid-1980s, and many institutions now have an allocation to listed PE. For example, Cumming et al. (2011) survey 171 European institutional investors, and report that 34% (primarily smaller and private institutions) have listed PE in their mandates. Bilo et al. (2005) identify 287 listed vehicles between 1986 and 2003, of which they analyze 114 that they find to be sufficiently liquid to plausibly estimate factor models. They estimate a CAPM beta of 1.2 for a value-weighted buy-and-hold strategy of listed PE, with an annualized alpha of –1.2%. An equal-weighted buy-and-hold strategy had a beta of 0.7 and an alpha of –0.1%, suggesting that the smaller vehicles performed better over the sample period. A fully rebalanced equal-weighted strategy performed remarkably well (with an alpha of 10.2% and a beta of 0.6). Applying the Dimson (1979) correction to account for staleness arising from thin trading raises the betas to around one but does not materially change the alphas (see Table 5 for the Dimson estimates, reported in the buyout column because this is likely the dominant strategy, although Bilo et al. do not provide information on the breakdown of strategies). But the apparent gains from the rebalancing strategy could well be wiped out by transaction costs, as accounting for bid-ask spreads alone reduces the annualized mean return by 8.3% (from 16.0% to 7.7%) over the sample period.

Phalippou (2010) identifies 19 publicly traded buyout funds and funds-of-funds in the CapitalIQ database as of 2008 that have at least 200 consecutive weeks of return data and for whom a zero weekly return is observed in Datastream for no more than once per 20 weeks. Using Dimson regressions with 10 weeks' lags, he reports an average beta across funds of 1.5 and an average annualized alpha of 6%.

Jegadeesh et al. (2015) provide a more detailed breakdown of risk and return across fund types and strategies. They analyze a sample of 129 traded funds and 24 traded funds-of-funds from 1994 to 2008. They separate the sample into VC and buyout vehicles, depending on their declared focus (for funds) or the strategy that has the greater asset allocation (for funds-of-funds). They find that VC and buyout funds have CAPM betas of 1.2 and 0.9 with respect to the MSCI World Index, and fund-of-fund betas are 1.0 and 0.7. In the Fama and French (1993) model augmented with a momentum factor (Carhart, 1997), all vehicles load positive on the small size factor and on the value factor (with insignificant loadings in VC), and all except buyout fund-of-funds load negatively on momentum. The fact that the loadings do not differ much across the types of listed vehicles highlights the mixed nature of their investments. Alpha estimates across factor models and vehicles range between –0.4% and +0.2% per month, with partnerships toward the upper end of this range and funds-of-funds at the lower end, but none are significantly different from zero.

McCourt (2018) studies 134 publicly listed PE funds, separated by strategy (VC, buyout, mezzanine, and fund-of-funds). His sample period spans 1995 to 2015, making this the only paper in the listed-vehicles literature that include post-financial crisis years. He estimates a market beta of around one for most strategies, except VC, which has a beta of 1.4. All strategies have a positive loading on SMB (close to 0.6 for most except VC, which has a higher loading of 1.1), and a negative loading on momentum of around –0.1. The loadings on HML are mixed: It is positive for buyout and mezzanine, negative for VC, and insignificant for fund-of-funds. This is consistent with the value and growth investing nature of buyout and VC, respectively, with funds-of-funds being a mixed bag. While alphas are statistically insignificant, the point estimates are positive for all strategies (around 0.3% per month for all but VC, which is 0.1%). Using tools from the mutual fund literature, such as cross-sectional bootstraps and false discovery rates, he finds evidence of that some managers in buyout and mezzanine funds are skilled, but not in VC, while the evidence for funds-of-funds is mixed.

While publicly traded vehicles are a useful alternative lens for assessing PE risk and return, the estimates should not be directly compared to those of private funds. First, the mixed investment strategies and highly diversified nature of listed vehicles makes it difficult to attribute risk loadings and risk-adjusted returns to any single strategy. The relative homogeneity in estimated risk loadings across types of listed PE vehicles and (inferred) strategies is consistent with the lines indeed being blurred here. Moreover, BDCs and SBICs, while also investing in startups, own very different securities with very different risk-return profiles than private VC funds, as described above. Second, public listing is an endogenous choice, and there is likely selection bias in the characteristics of PE vehicles that are traded. Selection on size is one obvious example, but there are bound to be many other variables that drive listing choice. Although both Jegadeesh et al. (2015) and McCourt (2018) argue that their samples are representative of the private fund space, future research should explore a formal model of the listing decision to analyze this question in more depth. Third, the pressures and incentives of having permanent public capital may change GPs' incentives and result in different risk and return profiles of the underlying investment portfolios compared to unlisted, private PE vehicles.

As a final point, many of the listed funds, though publicly traded, are still quite illiquid, raising the question how accurately market prices and returns reflect underlying valuations. Moreover, the permanence of public capital means that there is no forced final liquidation of the portfolio, so that misvaluation can persist for even longer periods of time than in the private limited partnerships. The fact that payout policies of the listed vehicles tend to be highly smoothed also does not help to alleviate this concern. Thus, market returns of listed funds may not capture all the underlying (co)variation of PE returns. Several papers in the traded vehicles literature use Dimson corrections to deal with this, but more work is needed to establish whether this is sufficient to obtain unbiased factor model estimates.

6. Investment Level Data

Analyzing returns at the investment level has two distinct advantages over fund-level data. First, investment-level data are reported before fees, allowing for direct assessment of GP skill. Second, one could estimate risk and return for specific industries, geographies, cohorts, and other characteristics, whereas funds are a mix of multiple investments across a broad set of attributes (this is true even for specialized funds such as early-stage U.S. VC funds). One could even assess performance for individual PE firm partners based on the portfolio companies they are associated with (see, for example, the VC partner fixed effect estimation in Ewens and Rhodes-Kropf, 2015).

Estimates of risk factor loadings and risk-adjusted returns differ between investment-level and fund-level data. Figure 1 shows that for VC, the average (median) CAPM beta estimate across papers that use investment data is 1.8 (1.9), compared to 1.7 (1.4) from (net-of-fee) fund data. For buyout, the investment-level beta is 2.0 (2.2), versus 1.3 (1.1) for funds. Risk-adjusted returns are also materially higher in investment-level data. Fees are a natural explanation for the lower net-of-fee risk-adjusted returns, although the quantitative effect has not been analyzed directly.[34] For betas, the effect of fees is more subtle. Consider carried interest: the division of profits from investments implies that there is a degree of risk-sharing between GPs and LPs. As a result, the LPs' exposure is less sensitive to market (up)swings than the underlying portfolio of investments.[35] On the other hand, management fees are paid according to a predetermined schedule, and so they might function somewhat akin to the way leverage raises a company's equity beta.[36] Ultimately, the effect of fees on betas is still an open question.

While the differences in alpha and beta estimates could be consistent with the effect of fees, they can also stem from differences in data sources (which cover different parts of the population), sample periods, and methodologies. Hüther et al. (2020) is a unique paper in this respect, because the authors have data from one single (large) LP who collected both pre-fee and post-fee cash flows for 85 U.S. VC funds raised between 1992 and 2005 that invested in 3,552 portfolio companies. This allows them to compute fund performance both before and after fees. The difference between the average (median) gross-of-fee and net-of-fee PME relative to the Nasdaq ranges from 0.23 to 0.41 (0.17 to 0.33), depending on vintage years. The difference is larger when PME is high, consistent with carried interest payments being higher when funds perform well. Unfortunately, the sample is rather small, and the paper does not consider the difference in gross and net betas (this is not its focal interest).

Besides sampling variation, beta estimates could differ because many fund beta estimates are based mostly or entirely on cash flows while investment-level estimates tend to rely on total returns that include both cash flows and valuations (for example, round-to-round returns in VC). In essence, the question is whether the difference is due to differences in cash flow and discount rate betas? Boyer et al. (2022a) report lower risk-adjusted fund returns after accounting for discount rate risk using secondary market transactions data, compared to using cash flows only. On the one hand, this suggests that discount rate risk indeed matters, and that "total" beta (including both cash flow and discount rate risk) of PE funds is higher than cash flow beta alone. On the other hand, their results further widen the gap between (lower) fund-level and (higher) investment-level risk-adjusted returns.

Finally, in VC, the heavy use of post-money valuations in investment-level data could distort results in a way that fund cash flows do not (apart from NAVs, if used).[37] With all these potential explanations floating around, future research should identify the relative importance of the various channels and reconcile the investment-level and fund-level estimates.

The remainder of this section considers an array of performance measures for investment-level data. A key issue in many of these data sets, especially in VC, is that data are nonrandomly missing, and a large part of this section is devoted to ways of dealing with this problem.

6.1 Cash flow-based Performance Measures

Investment-level performance can be computed based on each deal's cash flows, either for the portfolio company as a whole or as experienced from a single GP's perspective. Based on these cash flows one can compute the standard performance metrics, such as TVPI, IRR, and (G)PME, for each investment. In VC, the cash flows for a specific portfolio company are the investments made in each round of financing, and the ultimate payoff. The payoff can be measured as the value at exit (IPO or acquisition) or when the GP sells its stake in the startup (this may be post-IPO, as VCs often hold on to some of their stake in the company post-IPO, see Basnet et al., 2020, and Jenkinson et al., 2021). In buyout, there may be cash flows before exit, for example from portfolio company dividend payments and monitoring fees charged to the portfolio company by the GPs (as described in section 2.3).

Most papers that report deal-level TVPIs, IRRs, or GPMEs consider buyout investments only. Overall, they find that gross-of-fee buyout performance has been consistently high over time.

For a sample of 25 public-to-private buyout deals from the early 1980s, Kaplan (1989) reports an average (median) total nominal return of the buyout capital, both the debt and equity claims, of 127.9% (111.3%), realized over an average period of 2.7 years. As this return is measured from the initial buyout date until exit, with no observed intermediate cash flows, this corresponds to a TVPI multiple of 2.28 (2.11). Adjusted for the contemporaneous market return, the average (median) return is 41.9% (28.0%). The returns to the equity investors (that is, the PE fund) are significantly higher.

Groh and Gottschalg (2011) collect data from private placement memoranda (PPMs) of buyout funds. A PPM contains the full history of deals for a GP, and include intermediate cash flows such as dividends, but there may be selection regarding the institutions that are willing to share their information. For a total of 133 U.S. buyout deals that took place between 1984 and 2004, the average (median) annual IRR is 50.1% (35.7%), earned by GPs over an average holding period of 3.8 years.

Franzoni et al. (2012) have data on 4,403 completed buyouts that were closed between 1975 and 2006, sourced from CEPRES. The data, which are voluntarily disclosed by GPs, primarily cover buyouts in the U.S. and Europe. Like Groh and Gottschalg, the data also contain the exact cash flows, including any intermediate dividends, of each deal. They report a value-weighted mean TVPI of 2.52 and a modified IRR of 19%.[38]

Four papers report buyout PMEs. Guo et al (2011) compute a market- and risk-adjusted return that is essentially a PME calculation but expressed as a percentage return. Their discount rate is a CAPM-based expected return estimate. For 90 public-to-private buyouts of U.S. firms between 1990 and 2006, they estimate that buyout investors earned an average (median) CAPM-benchmarked PME of 1.63 (1.41). However, these results are based on deals for which post-buyout data is available, raising success bias concerns. Acharya et al. (2013) compute PMEs using the sector return as the discount rate and report an average (median) of 1.88 (1.37) across 395 large buyout deals, based on proprietary deal-level data. Braun et al. (2017) report a median TVPI of 1.5 and PME of 1.3 for a large sample of 12,541 buyout investments between 1974 and 2013. Their data represent the near-complete deal history of GPs that raised capital from three large fund-of-fund managers. For the 7,568 fully realized deals in their sample, the median TVPI and PME numbers are 1.9 and 1.4, respectively. Finally, Brown et al. (2020) have a large data set of portfolio companies collected by Burgiss. Across 15,095 buyout deals they report an average (median) TVPI of 2.14 (1.56) and an average (median) PME with respect to the S&P 500 of 1.45 (1.06).

Brown et al. (2020) have a large data set of 31,206 VC deals, showing an average (median) TVPI of 2.13 (1.00), and an average (median) PME of 1.35 (0.71). Not surprisingly, the outcomes of VC deals are significantly more volatile compared to buyout deals. In comparison, Hüther et al. (2020) report gross-of-fee VC fund PMEs (with respect to the Nasdaq) between 0.85 and 1.19 (0.85 and 1.06), across buckets of vintages. In line with the literature, pre-1998 vintages are at the high end of this range and post-1999 vintages are at the low end.

6.2 Index Methods

An alternative to the pure cash flow-based approach is to incorporate additional valuation data. In particular, in VC, the post-money valuation of the startup (defined in section 3.6) is observed when it raises a new financing round, providing intermediate indications of value over the investment horizon. I start by describing a two-step approach that first constructs an aggregate price index, and then estimates a factor model on the index returns.

6.2.1 Repeat-sales regression

Repeat-sales regressions are a standard approach to impute missing valuations, borrowed from the real estate literature (e.g., Bailey et al., 1963, Case and Shiller, 1987, 1989), where they are traditionally used to compute home price indices such as the S&P/Case-Shiller index. Its key intuition can be conveyed succinctly with a stylized example (see OECD et al., 2013, for a thorough yet accessible treatise). Suppose we have valuations for three startups (named A, B, and C) as shown in Table 8. Startup A raised money in year 1 at a valuation of $1 million and had another financing round in year 2 at $1.5 million. Company B raised its first round at a value of $4 million in year 2 and at $5 million in year 3, and C had funding rounds at $2 million and $5 million valuations in years 1 and 4. As is typical in VC, none of the companies paid dividends over the sample period.

[INSERT TABLE 8 AROUND HERE]

RSR assumes that all startups have the same expected return over the same period. Since startup A had a 50% return from year 1 to year 2, we can impute company C's expected value to be $3 million in year 2 (shown in square brackets in Table 8 to indicate this is an imputed value). For year 3, we impute the value of A and C to have grown by 25% from year 2, which is the observed return on B over this period. The one-year return from year 3 to year 4 can now be found from startup C, which grew from an imputed value of $3.75 million to an observed value of $5 million, or 33%.

Doing these imputations by hand is easy in this simple example, but quickly becomes intractable in large samples with many startups and time periods. Fortunately, it is possible to formulate the problem in regression form. Assume the following process for the natural logarithm of log-valuations of startup *i*,

$$v_{i,t}^{PRE} = v_{i,t-1}^{POST} + \phi_t + \varepsilon_{i,t}, \tag{25}$$

where $v_{i,t-1}^{POST}$ is the log of the post-money valuation (as explained in section 3.6) at time *t-1*, and $v_{i,t}^{PRE}$ is the log of the pre-money valuation at time *t*. The pre-money valuation for a financing round is simply the post-money valuation (in the same round) minus the invested amount (if no capital infusion occurs, then pre- and post-money valuations in the same period are equal). The ϕ_t parameter is the one-period log-return on the repeat-sales index, which is common to all startups. The residual $\varepsilon_{i,t}$ is assumed to be mean zero.

The purpose of computing the log-value change from post-money to pre-money is to measure the actual return experienced by an investor. To see this, note that a capital infusion mechanically raises the (post-money) value of the company, but does not change the investor's wealth, just like a share issuance for cash in itself does not change a company's share price. For example, consider a startup with a \$100 post-money valuation at t-1. Suppose that no value was created from *t-1* to *t*, and the company raises \$50 in equity at time t. The post-money valuation at *t*, which is measured right after the cash infusion, is then \$150, and the pre-money valuation is \$100. Thus, the return measured as the post-money at *t*-1 to pre-money at *t* correctly reflects the investor's actual return over the period, which is zero.[39] Note that the resulting index represents an investor's buy-and-hold experience, not the aggregate value of all startups (to get the latter index, only post-money valuations should be used).

The log-return over a horizon τ, is then

$$v_{i,t+\tau}^{PRE} - v_{i,t}^{POST} = \sum_{s=1}^{\tau} v_{i,t+s}^{PRE} - v_{i,t+s-1}^{POST} = \sum_{s=1}^{\tau} \phi_{t+s} + \sum_{s=1}^{\tau} \varepsilon_{i,t+s} \ . \tag{26}$$

Equation (26) suggests that we can estimate the ϕ parameters from a regression of observed log-returns (e.g., round-to-round returns or round-to-exit returns) on a set of indicator variables, one for each period except the very first one. The indicators for a given startup equal one if the period is in between sales, and zero otherwise (the period at which the investment is first made is also zero, but the period in which the next value is observed equals one). For example, the regression for the Table 8 example is (assuming that the first and second values for each firm are post-money and pre-money valuations, respectively, and there are no intermediate capital infusions)

$$\begin{bmatrix} \ln(1.5) - \ln(1) \\ \ln(5) - \ln(4) \\ \ln(5) - \ln(2) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} \phi_2 \\ \phi_3 \\ \phi_4 \end{bmatrix} + \begin{bmatrix} \eta_2 \\ \eta_3 \\ \eta_4 \end{bmatrix}. \tag{27}$$

The numbering of ϕ's and η's starting at 2 is a result of dropping the very first period of valuations. The η's are sums of ε's, per equation (26), and assumed to be mean zero and orthogonal to the indicator matrix, analogous to the standard ordinary least squares (OLS) identification assumption (we will revisit this assumption soon). The OLS estimates of the ϕ's are 0.4055, 0.2231, and 0.2877. Exponentiating the estimates produces the 50%, 25%, and 33% returns that we found above.[40] Any additional financing rounds of the same startups can be added as separate return observations.

OLS estimation is efficient if the residuals in equation (27) have equal variance across observations. But the η's are sums of one-period residual log-returns $\varepsilon_{i,t}$ from equation (25). If the ε's have constant variance, σ^2, and if they are uncorrelated over time, then the η corresponding to a return over τ periods has variance $\tau\sigma^2$. In other words, the residuals in equation (27) are heteroskedastic. The weighted least squares (WLS) estimator is therefore more efficient than OLS, weighing each observation by the reciprocal of the time between sales. In the example, the weight matrix is a diagonal matrix in which the first two observations receive a weight of 1, but the third observation receives a weight of 1/3 (equivalently, one can multiply each observation by the square root of these weights and run OLS on the transformed variables). However, the real estate literature has found that residual variances do not scale linearly with τ, and do not approach zero as τ shrinks (e.g., Case and Shiller, 1987, Giacoletti, 2021, Sagi, 2021). This result is attributed to house-specific shocks that do not scale with time between sales, for example due to search frictions. Axelson et al. (2014) find a similar result for buyout investments, and I suspect a similar phenomenon applies to startups. Another critique of OLS and WLS estimation, and indeed much of the literature on investment-level returns, is that cross-correlations in the error terms are assumed to be zero. But, contemporaneous ε's are likely to be positively correlated, and to the extent that return horizons overlap, so are the η's. While the OLS and WLS estimators are still unbiased and consistent, reported standard errors that ignore these correlations are too low. Unfortunately, there is no easy clustering fix to allow for cross-correlations that is consistent with the underlying model. See Korteweg and Nagel (2016) for a standard error estimator that accounts for the return overlap.

The assumption that all startups have the same expected return (the same ϕ's) can be relaxed by allowing for covariates (e.g., Korteweg and Sørensen, 2016). For example, each industry can have its own price index, allowing for different expected returns in, say, biotech versus software. This approach is related to Gompers and Lerner (1997), who use publicly-traded industry returns to impute the unobserved valuations of portfolio companies (see the description in section 3.5.4).

6.2.2 Selection bias

An important challenge that needs to be dealt with in commercial investment-level data sets is success bias. This is especially important in VC, as the final outcome for many startups remains unobserved. This can be seen in Figure 6, which shows a pie chart of outcomes for 18,237 startups that were first funded between 1987 and 2005. By the end of the sample period, 10.4% has gone public, 23.4% were acquired, 15.9% had failed, and 50.4% had an unknown outcome. Some of these companies were still private at the end of the sample period, because not enough time had passed since first funding (a right-censoring problem). However, many firms had their last funding round years before the sample ended: in a different sample from the 2013 National Venture Capital Association Yearbook, as many as 35% of 11,686 startups that were first funded between 1991 and 2000, still had an unknown outcome by 2010. While these firms are likely out of business, this event was never recorded. As such, they are sometimes referred to as zombie firms, as they are dead but still appear to be alive.[41] The reason for this success bias is that IPOs and large acquisitions are fairly easy to observe from filings and other disclosures, whereas failures and small acquisitions (many of which occur at a loss to the original investors) are not widely publicized. Thus, round-to-round or round-to-exit returns, which require at least two valuation observations for a given firm, are more likely to be observed for startups that are ultimately successful.

[INSERT FIGURE 6 AROUND HERE]

The consequence of the selection problem is that the RSR index returns are upward biased.[42] Econometrically, selection shows up as a correlation between the error terms and the covariates in equation (27): firms that have a positive return shock (η) are more likely to be observed, that is, they are more likely to have ones in the indicator matrix. This correlation between residuals and explanatory variables violates the regression identification assumption and results in biased estimates of the RSR returns (ϕ).[43]

Peng (2001) addresses the sample selection problem by estimating separate RSR indices for successes and failures with known outcomes. For the zombie firms (with unknown outcomes), he uses a nonparametric model to assign a probability of success to each firm in each period. He then distributes the value of the zombie companies across the success and failure indices, according to the estimated probabilities. Finally, he combines the two indices based on their relative NAVs. Peng then estimates factor loadings by regressing the monthly index returns on the excess return of the S&P 500. On a sample of 5,634 startups that were first funded by VCs between 1987 and 1999, he estimates a market beta of 1.3 and a statistically insignificant alpha of –0.2% per month (and this is based on standard errors that too low since the second step ignores the estimation error in the index; more on this below). Using annual returns instead, he finds a beta of 2.4 and an alpha of –0.9% per year. This result is rather puzzling, since there is no obvious reason why the beta should change with measurement frequency, and why alpha does not scale with horizon.

Woodward and Hall (2004) use an RSR model and impute missing valuations after the last observed round for a company as a function of company characteristics and the amount of time that has passed since the last observed funding event. Unfortunately, the paper is not very detailed in how exactly this imputation is performed.[44] A more fleshed out implementation of the same idea appears in Hwang et al. (2005), who follow Heckman (1979) and consider sample selection as an omitted-variables problem. They employ a two-step method. The first step is an ordered probit model that estimates the probability of observing a funding round of each firm in each period, as a function of observable characteristics. The second step uses the inverse Mills ratio of the first step to correct the RSR for selection bias. In an updated version of Peng's (2001) VC data, with companies funded between 1987 and 2003, they estimate an S&P 500 market beta of 0.6 and a statistically insignificant quarterly alpha of 0.9%. Relative to the Nasdaq, the alpha is 1.0%. Like most other papers in the literature, they find higher alphas in the pre-2000 period, of 3.5% per quarter.[45]

Some commercial indices also recognize the sample selection issue. For example, the Refinitiv Venture Capital Research Index (formerly the Thompson Reuters Venture Capital Research Index) estimates the value of each company over time, estimating missing values and interpolating values between financing rounds.[46] It uses a Heckman correction for missing rounds and assigns failures as having occurred one year after final observed round if the startup is currently listed as defunct.[47]

6.3 Round-to-Round Returns

The two-step approach of the previous section produces both an index (the first step) and factor model estimates (the second step). If we are only interested in the factor model estimates, we can skip the index and directly use the returns from one valuation observation to the next (e.g., from deal inception to exit, or, for VC, between financing rounds). This alternative yields proper standard errors, unlike applications of the index approach, which typically ignore estimation error in the index when estimating the factor model.

To estimate factor loadings directly from returns, replace the value process in equation (25) with

$$v_{i,t}^{PRE} = v_{i,t-1}^{POST} + r_t^F + \delta + \beta' f_t + \varepsilon_{i,t} \,. \tag{28}$$

Equation (28) simply replaces the index return, ϕ_t, with a log-factor model, where f_t are (log) factors and β is a vector of loadings. The log risk free rate, r_t^F, accounts for the fact that factor models are specified on excess returns, so that δ is a risk-adjusted log-return. Note the similarity of this model and the log-factor model implied by the exponential-affine SDF model with lognormal returns from Korteweg and Nagel (2016) that was described in section 3.7.2. The counterpart to the longer-horizon log return in equation (26) is

$$v_{i,t+\tau}^{PRE} - v_{i,t}^{POST} = \sum_{s=1}^{\tau} r_{t+s}^{F} + \tau\delta + \beta' \sum_{s=1}^{\tau} f_{t+s} + \sum_{s=1}^{\tau} \varepsilon_{i,t+s} \ . \qquad (29)$$

The δ and β parameters can thus be estimated by regressing observed portfolio company returns (from round-to-round or round-to-exit) in excess of the log risk-free rate (over the same period) on the return horizon, τ (which varies by observation), and the log factor returns (also over the same period). Analogous to RSR, observations can be weighted to improve efficiency of the estimator in light of heteroskedastic residuals, and cross-correlations in the error term due to overlapping observations should be accounted for in computing standard errors. The δ and β estimates can be transformed from the log model to arithmetic alpha and factor loadings using the formulas in footnote 5 in Cochrane (2005a) for the market model and footnote 1 in Korteweg and Sørensen (2010) for the multifactor model (in the remainder of this section I will use alpha to refer to the arithmetic alpha).[48]

An application of the above methodology is in Axelson et al. (2014). They have data on all 2,075 investments made by a single buyout fund of funds. Running a WLS regression on equation (29), they estimate a log-CAPM annual delta of –4.6% and beta of 2.4. The annual arithmetic return alpha implied by the model is 8.6%. Allowing for jumps in values (that may arise from the illiquid nature of buyouts), does not significantly impact the beta estimate, but increases the arithmetic alpha to 16.3% per year. Axelson et al. rationalize their beta, which is at the higher end of the estimated range (see Figure 1), from a Modigliani-Miller calculation based on an unlevered asset beta of 0.66 and a leverage ratio that is typical for buyouts.

Axelson et al. (2014) argue that selection bias is less of a concern in buyout compared to VC, and they therefore do not make any adjustments for selection bias. Papers in VC do grapple with the issue. Ewens et al. (2016) adjust for selection by reweighing ultimate outcomes (IPO, acquisition, or known failure) in their observed-returns sample to match the exit rates in their full sample that includes companies with missing valuation data. Using returns from initial capital infusion to exit for 13,035 VC investments between 1990 and 2008, they estimate a log-CAPM beta of 2.3. Their estimated δ is –10.5% per year, which maps to an arithmetic annual risk-adjusted return of 45.6%. For the Fama and French 3-factor model, they estimate a market factor loading of 2.2, an SMB loading of -0.4, and an insignificant HML loading. Korteweg and Nagel (2016) also follow the reweighting approach to estimate the average GPMEs for individual VC deals, and find estimates in the order of 0.52 to 0.61, depending on what they assume about the returns to known liquidations without liquidation valuations. Given an average time between round of about one year, this implies an approximately 52% to 61% annual return.

Two papers, Cochrane (2005a) and Korteweg and Sørensen (2010), take a more structural approach to the selection problem. Both papers use equation (29) and add a selection equation that models the probability of observing a financing round. Cochrane (2005a) considers the log-CAPM and models the selection as a function of the startup's underlying value only. Korteweg and Sørensen (2010) formulate the problem as a state-space model, allowing for multiple risk factors and for additional covariates in the selection equation. They also point out that the empirical problem is more complicated than the standard Heckman (1979) setup, which assumes that periods with unobserved valuations are uninformative about the valuation process. In reality, because valuations are dynamically linked over time through equation (29), these periods are still useful in estimating the parameters of the price process. Cochrane (2005a) estimates a selection-corrected log-CAPM beta of 1.9 with respect to the S&P 500, and an annual arithmetic alpha of 32% (39% if using the Nasdaq instead, with a beta of 1.4). In contrast, without selection correction the arithmetic annual alpha is a mind-boggling 462%. Korteweg and Sørensen (2010) estimate a log-CAPM beta of 2.8, and an alpha that's more or less equal to Cochrane's, at 3.3% per month. Splitting the time-series, they find that the positive alpha is primarily coming from the internet boom of 1994-2000, when the monthly alpha was 5.8%. From 2000 to 2005, the end of their sample period, the alpha was negative at –2.6%. For the Fama and French 3-factor model they estimate a market factor loading of 2.3, an SMB loading of 1.1, and an HML loading of –1.6, implying (consistent with the literature) that startup companies look like small-growth investments. In an extension of their model, they find that results are robust to allowing for startup-specific alphas and betas (see their appendix A.5).

Buchner (2020) deals with selection by explicitly modeling the cash flow process of each investment. He estimates a VC beta of 2.6 and an annual alpha of 8.9% using a sample of 6,380 startups from 1980 to 2005. For a sample of 4,418 buyout deals he reports a beta of 2.2 and an annual alpha of 7.0%. He finds a similar time-series pattern of VC alphas as Korteweg and Sørensen (2010). Consistent with the literature, risk-adjusted performance for buyouts has been consistently strong throughout the sample period.

Finally, estimates of idiosyncratic volatility for startups (that is, the standard deviation of $\varepsilon_{i,t}$ in equation (28)) are of independent interest for a variety of applications, such as contingent claims valuation models that price the securities purchased by VCs (e.g., Metrick and Yasuda, 2010b, Gornall and Strebulaev, 2020). Reported volatility estimates range from just below 90% to over 140% per year: Cochrane (2005a) estimates 86% per year, Korteweg and Sørensen (2010), 142% (41% per month), and Fleckenstein and Longstaff (2020), using a very different sample and methodology, report 95%.

7. Conclusion

Over the last quarter century, significant progress has been made in our understanding of risk and return in PE. At the same time, much remains to be discovered and understood. A key challenge for the literature is to uncover the set of relevant risk factors and their loadings (betas), including gaining a better understanding of the impact of the various dimensions of illiquidity in PE and their impact on risk premia. A second unanswered question is whether PE payoffs can be spanned by public securities, or if there are components to payoffs that are unique to the PE space (and if so, whether these are priced factors or pure alpha)? Third, work is needed to further explore the time-variation in factor loadings and risk-adjusted returns, both in calendar time and over the life of a fund or portfolio company. Fourth, reconciling the pre- and post-fee risk and return estimates is important, especially given the apparent large differences in risk-adjusted returns. Finally, further research is needed into the extent of performance manipulation by GPs and the development of a manipulation-proof risk-adjusted return measure, if one such measure exists.

This chapter has largely been silent on the literature that documents the large heterogeneity in returns experienced by different types of LPs. Performance also appears to be persistently different across GPs. While there has been some recent work towards individual fund benchmarking, much is left to be done to better understand these differences, both methodologically and empirically. This in turn can help answer questions related to the source(s) of persistence in GP and LP performance, whether managers have styles, how contracts between GPs and LPs and between GPs and portfolio companies affect performance, and how PE fits into a portfolio of (public) stocks, bonds, and other securities.

To end on a positive note, the recent convergence in the literature toward the use of stochastic discount factor methods, and the continuing improvements in data quality and coverage, hold promise that significant progress can be made towards answering these questions.

References

Acharya, V.V., Gottschalg, O.F., Hahn, M., Kehoe, C., 2013. Corporate governance and value creation: Evidence from private equity. Rev. Fin. Stud. 26, 368–402.

Agarwal, V., Barber, B., Cheng, S., Hameed, A., Yasuda, A., 2022. Private company valuations by mutual funds. Rev. Finance, forthcoming.

Albertus, J.F., Denes, M., 2020. Private equity fund debt: Capital flows, performance, and agency costs. Work. Pap., Carnegie Mellon Univ.

Albuquerque, R., Cassel, J., Phalippou, L., Schroth, E., 2018. Liquidity provision in the secondary market for private equity fund stakes. Work. Pap., Boston College, Univ. Oxford, and EDHEC.

Aliaga-Díaz, R., Renzi-Ricci, G., Ahluwalia, H., Grim, D.M., Tidmore, C., 2020. The role of private equity in strategic portfolios. Vanguard Research.

Amornsiripanitch, N., Gompers, P.A., Xuan, Y., 2019. More than money: Venture capitalists on boards. J. Law, Economics., and Organization 35, 513–543.

Andonov, A., Kräussl, R., Rauh, J.D., 2021. The subsidy to infrastructure as an asset class. Rev. Financ. Stud. 34, 3880–3994.

Ang, A., Chen, B., Goetzmann, W.N., Phalippou, L., 2018. Estimating private equity returns from limited partner cash flows. J. Finance 73, 1751–1783.

Anson, M.J.P., 2002. Managed pricing and the rule of conservatism in private equity portfolios. J. Priv. Equity 5, 18–30.

Anson, M.J.P., 2007. Performance measurement in private equity: another look. J. Priv. Equity 10, 7–21.

Asness, C.S., Krail, R.J., Liew, J.M., 2001, Do hedge funds hedge? J. Portfolio Manag. 28, 6–19.

Axelson, U., Jenkinson, T., Strömberg, P., Weisbach, M., 2013. Borrow cheap, buy high: The determinants of leverage and pricing in buyouts. J. Finance 68, 2223–2267.

Axelson, U., Sørensen, M., Strömberg, P., 2014. Alpha and beta of buyout deals: A jump CAPM for long-term illiquid investments. Work. Pap., London School of Economics, Dartmouth, and Stockholm School of Economics.

Bailey, M.J., Muth, R.F., Nourse, H.O., 1963. A regression method for real estate price index construction. J. Am. Stat. Assoc. 58, 933–942.

Baker, M., Gompers, P.A., 2003. The determinants of board structure at the initial public offering. J. Law. Econ. 46, 569–598.

Banal-Estañol, A., Ippolito, F., Vicente, S., 2017. Default penalties in private equity partnerships. Work. Pap., Univ. Pompeu Fabra and Univ. Carlos III Madrid.

Barber, B., Yasuda, A., 2017. Interim fund performance and fundraising in private equity. J. Financ. Econ. 124, 172–194.

Basnet, A., Pukthuanthong, K., Turtle, H., Walker, T., 2020. VC ownership post-IPO: When, why, and how do VCs exit? Work. Pap., Concordia Univ., Univ. Missouri, and Colo. State Univ.

Begenau, J., Siriwardane, E., 2021, How do private equity fees vary across public pensions? Work. Pap., Stanford Graduate School of Business.

Berk, J.B., Green, R.C., Naik, V., 2004. Valuation and return dynamics of new ventures. Rev. Financ. Stud. 17, 1–35.

Bernstein, S., Giroud, X., Townsend, R., 2016. The impact of venture capital monitoring. J. Finance 71, 1591–1622.

Bernstein, S., Sheen, A., 2016. Operational consequences of private equity buyouts: Evidence from the restaurant industry. Rev. Financ. Stud.29, 2387–2418.

Bharath, S., Dittmar, A., Sivadasan, J., 2014. Do going-private transactions affect plant efficiency and investment? Rev. Financ. Stud.27, 1929–1976.

Bilo, S., Christophers, H., Degosciu, M., Zimmermann, H., 2005. Risk, returns, and biases of listed private equity portfolios. Work. Pap., Univ. Basel.

Bisinger, M., Bircan, C., Ljungqvist, A., 2020. Value creation in private equity. Swedish House of Finance Research Paper No. 20-10.

Bollen, N.P.B., Pool. V.K., 2008. Conditional return smoothing in the hedge fund industry. J. Financ. Quant. Anal. 43, 267–298.

Boucly, Q., Sraer, D., Thesmar, D., 2011. Growth LBOs. J. Financ. Econ. 102, 432–453.

Boyer, B.H., Nadauld, T.D., Vorkink, K., Weisbach, M.S., 2022. Discount Rate Risk in Private Equity: Evidence from Secondary Market Transactions. J. Finance, forthcoming.

Boyer, B.H., Nadauld, T.D., Vorkink, K., Weisbach, M.S., 2022. Private equity indices based on secondary market transactions. Work. Pap. Brigham Young Univ. and The Ohio State Univ.

Braun, R., Jenkinson, T., Schemmerl, C., 2020. Adverse selection and the performance of private equity co-investments. J. Financ. Econ. 136, 44–62.

Braun, R., Jenkinson, T., Stoff, I., 2017. How persistent is private equity performance? Evidence from deal-level data. J. Financ. Econ. 123, 273–291.

Brophy, D.J., Guthner, M.W., 1988. Publicly traded venture capital funds: Implications for institutional "fund of funds" investors. J. Bus. Venturing 3, 187–206.

Brown, G.W., Ghysels, E., Gredil, O., 2021a. Nowcasting net asset values: The case of private equity. Work. Pap. Univ. North Carolina and Tulane Univ.

Brown, G.W., Hu, W., Kuhn, B.-K., 2021b. Private investments in diversified portfolio. Work. Pap. Univ. North Carolina and Burgiss.

Brown, G.W., Gredil O., Kaplan, S.N., 2019. Do private equity funds manipulate returns? J. Financ. Econ. 132, 267–297.

Brown, G.W., Harris, R.S., Hu, W., Jenkinson, T., Kaplan, S.N., Robinson, D., 2020. Private equity portfolios: A first look at Burgiss holdings data. Work. Pap., Univ. North Carolina, Univ. Virginia, Burgiss, Univ. Oxford, Univ. Chicago, Duke Univ.

Buchner, A., 2016. Risk-adjusting the returns of private equity using the CAPM and multi-factor extensions. Finance Res. Lett. 16, 154–161.

Buchner, A., 2020. The alpha and beta of private equity investments. Work. Pap., Univ. Passau.

Buchner, A., Stucke, R., 2014. The systematic risk of private equity. Work. Pap., Univ. Passau and Univ. Oxford.

Chakraborty, I., Ewens, M., 2018. Managing performance signals through delay: Evidence from venture capital. Manag. Sci. 64, 2875–2900.

Cao, C., Farnsworth, G., Liang, B., Lo, A., 2017. Return smoothing, liquidity costs, and investor flows: Evidence from a separate account platform. Manag. Sci. 63, 2049–2395.

Carhart, M.M., 1997. On persistence in mutual fund performance. J. Finance 52, 57–82.

Case, K.E., Shiller, R.J., 1987. Prices of single-family homes since 1970: new indexes for four cities. N. Engl. Econ. Rev., 46–56.

Case, K.E., Shiller, R.J., 1989. The efficiency of the market for single-family homes. Am. Econ. Rev. 79, 125–137.

Case, B., Quigley, J., 1991. The dynamics of real estate prices. Rev. Econ. and Stat. 73, 50–58.

Cassar, G., Gerakos, J., 2011. Hedge funds: Pricing controls and the smoothing of self-reported returns. Rev. Financ. Stud. 24, 1698–1734.

Cederburg, S., Stoughton, N., 2020. Discretionary NAVs. Work. Pap., Univ. Arizona and Vienna Univ.

Chen, S., Li, T., 2020. A unified duration-based explanation of the value, profitability, and investment anomalies. Work. Pap., Southw. Univ. Financ. Econ. and City Univ. Hong Kong.

Chernenko, S., Lerner, J., Zeng, Y., 2021. Mutual funds as venture capitalists? Evidence from unicorns. Rev. Financ. Stud. 34, 2362–2410.

Cochrane, J.H., 2005. The risk and return of venture capital. J. Financ. Econ. 75, 3–52.

Cochrane, J.H., 2005. Asset Pricing, 2nd ed. Princeton Univ. Press, New Jersey.

Cochrane, J.H., 2021. Portfolios for long-term investors. Work. Pap., Stanford Univ.

Cohn, J.B., Hotchkiss, E., Towery, E., 2020. The motives for private equity buyouts of private firms: Evidence from U.S. corporate tax returns. Work. Pap., Univ. Texas Austin, Boston College and Univ. Georgia.

Cordell, L., Roberts, M.R., Schwert, M., 2021. CLO performance. Work. Pap., Fed. Reserve Bank of Philadelphia and Univ. Pennsylvania.

Couts, S., Gonçalves, A., Rossi, A., 2020. Unsmoothing returns of illiquid funds. Work. Pap., Univ. Southern Calif., Univ. North Carolina, and Univ. Arizona.

Cumming, D., Fleming, G., Johan, S.A., 2011. Institutional investment in listed private equity. European Fin. Manag. 17, 594–618.

Cumming, D., Walz, U., 2010. Private equity returns and disclosure around the world. J. Intern. Bus. Stud. 41, 727–754.

Da Rin, M., Phalippou, L., 2017. The importance of size in private equity: Evidence from a survey of limited partners. J. Finan. Intermediation 31, 64-76.

Davis, S.J., Haltiwanger, J., Handley, K., Jarmin, R., Lerner, J., Miranda, J., 2014. Private equity, jobs, and productivity. Amer. Econ. Rev. 104, 3956–3990.

DeChow, P.M., Sloan. R.G., Soliman, M., 2004. Implied equity duration: A new measure of equity risk. Rev. Acc. Stud. 9, 197–228.

DeLuce, A., 2020, Private equity fees and terms study. Callan Institute Research Study.

Dimson, E., 1979. Risk measurement when shares are subject to infrequent trading. J. Financ. Econ. 7, 197–226.

Driessen, J., Lin, T.-C., Phalippou, L., 2012. A new method to estimate risk and return of nontraded assets from cash flows: The case of private equity funds. J. Financ. Quant. Anal. 47, 511–535.

Easton, P.D., Larocque, S., Sustersic Stevens, J., 2021. Private equity valuation before and after ASC 820. J. Inv. Manag. 19, 105–135.

Edgerton, J., 2012. Agency problems in public firms: Evidence from corporate jets in leveraged buyouts. J. Finance 67, 2187–2213.

Emery K., 2003. Private equity risk and reward: assessing the stale pricing problem. J. Priv. Equity 6, 43–50.

Espen Eckbo, B., Lithell, M., 2021. Merger-driven listing dynamics. Work. Pap., Dartmouth College and Norwegian Sch. Econ.

Ewens, M., 2009. A new model of venture capital risk and return. Work. Pap., UC San Diego.

Ewens, M., Gorbenko, A., Korteweg, A., 2022. Venture capital contracts. J. Financ. Econ. 143, 131–158.

Ewens, M., Jones, C.M., Rhodes-Kropf, M., 2013. The price of diversifiable risk in venture capital and private equity. Rev. Financ. Stud. 26, 1854–1889.

Ewens, M., Marx, M., 2018. Founder replacement and startup performance. Rev. Financ. Stud. 31, 1532–1565.

Ewens, M., Rhodes-Kropf, M., 2015. Is a VC partnership greater than the sum of its partners? J. Finance 70, 1081–1113.

Ewens, M., Rhodes-Kropf, M., Strebulaev, I., 2016. Insider financing and venture capital returns. Work. Pap., Caltech, MIT, and Stanford Univ.

Fama, E.F., French, K.R., 1993. Common risk factors in the returns on stocks and bonds. J. Financ. Econ. 33, 3–56.

Fama, E.F., French, K.R., 2015. A five-factor asset pricing model. J. Financ. Econ. 75, 3–52.

Fama, E.F., MacBeth, J.D., 1974. Long-term growth in a short-term market. J. Finance 29, 857–885.

Fang, L., Ivashina, V., Lerner, J., 2015. The disintermediation of financial markets: Direct investing in private equity. J. Financ. Econ. 116, 160–178.

Fisher, J.D., Geltner, D.M., Webb, R.B., 1994. Value indices of commercial real estate: A comparison of index construction methods. J. Real Estate Finance and Econ. 9, 137–164.

Fleckenstein, M., and Longstaff, F.A., 2020. Private equity returns: Empirical evidence from the business credit card securitization market. Work. Pap., Univ. Delaware and UCLA.

Fracassi, C., Previtero, A., Sheen, A., 2022. Barbarians at the store? Private equity, products, and consumers. J. Finance 77, 1439–1488.

Franzoni, F., Nowak, E., Phalippou, L., 2012. Private equity performance and liquidity risk. J. Finance 67, 2341–2373.

Gatzlaff, D.H., Haurin, D.R., 1997. Sample selection bias and repeat-sales index estimates. J. Real Estate Finance Econ. 14, 33–50.

Gatzlaff, D.H., Haurin, D.R., 1998. Sample selection and biases in local house value indices. J. Urban Econ. 43, 199–222.

Geltner, D.M., 1991. Smoothing in appraisal-based returns. J. Real Estate Finance and Econ. 4, 327–345.

Geltner, D.M., 1993. Estimating market values from appraised values without assuming an efficient market. J. Real Estate Res. 8, 325–346.

Geltner, D.M., MacGregor, B.D., Schwann, G.M., 2003. Appraisal smoothing and price discovery in real estate markets. Urban Stud. 40, 1047–1064.

Getmansky, M., Lo, A.W., Makarov, I., 2004. An econometric model of serial correlation and illiquidity in hedge fund returns. J. Financ. Econ. 2004, 529–609.

Giacoletti, M., 2021. Idiosyncratic risk in housing markets. Rev. Financ. Stud. 34, 3695–3741.

Giacoletti, M., Westrupp, V., 2018. The risk-adjusted performance of asset flippers. Work. Pap., Univ. Southern Calif. and AQR.

Giommetti, N., Jørgensen, R., 2021. Risk adjustment of private equity cash flows. Work. Pap., Copenhagen Bus. Sch.

Goetzmann, W. N., 1992. The accuracy of real estate indices: repeat sale estimators. J. Real Estate Finance Econ. 5, 5–53.

Goetzmann, W. N., Gourier, E., Phalippou, L., 2019. How Alternative Are Alternative Investments? The Case of Private Equity Funds. Work. Pap., Yale School of Management.

Goetzmann, W.N., Ingersoll, J., Spiegel, M., Welch, I., 2007, Portfolio performance manipulation and manipulation-proof performance measures. Rev. Financ. Stud. 20, 1503–1546.

Goetzmann, W.N., Peng, L., 2006. Estimating house price indexes in the presence of seller reservation prices. Rev. Econ. Stat. 88, 100–112.

Gompers, P.A., Gornall. W., Kaplan, S.N., Strebulaev, I.A., 2020. How do venture capitalists make decisions? J. Financ. Econ. 135, 169–190.

Gompers, P.A., Kaplan, S.N., Mukharlyamov, V., 2016. What do private equity firms say they do? J. Financ. Econ. 121, 449–476.

Gompers, P.A., Lerner, J., 1997. Risk and reward in private equity investments: the challenge of performance assessment. J. Priv. Equity 1, 5–12.

Gonçalves, A.S., 2021. The short duration premium. J. Financ. Econ. 141, 919–945.

Gormsen, N.J., Lazarus, E., 2021. Duration-driven returns. Work. Pap., Univ. Chicago and MIT.

Gornall, W., Strebulaev, I.A., 2020. Squaring venture capital valuations with reality. J. Financ. Econ. 135, 120–143.

Gredil, O., 2022. Do private equity managers have superior information on public markets? J. Financ. Quant. Anal. 57, 321–358.

Gredil, O., Griffiths, B.E., Stucke, R., 2022. Benchmarking private equity: The direct alpha method. Work. Pap., Tulane Univ., Ares Manag. Corp., and Warburg Pincus LLC.

Gredil, O., Sørensen, M., Waller, W., 2020a. Evaluating private equity performance using stochastic discount factors. Work. Pap., Tulane Univ. and Dartmouth College.

Gredil, O., Liu, Y., Sensoy, B., 2020b. Diversifying private equity. Work. Pap., Tulane Univ., Purdue Univ., and Vanderbilt Univ.

Groh, A.P., Gottschalg, O., 2011. The effect of leverage on the cost of capital of US buyouts. J. Bank. Finance 35, 2099-2110.

Guo, S., Hotchkiss, E. S., Song, W., 2011. Do buyouts (still) create value? J. Finance 66, 479–517.

Gupta, A., van Nieuwerburgh, S., 2021. Valuing private equity investments strip by strip. J. Finance 76, 3255–3307.

Hakansson, N.H., 1971. Toward a general theory of portfolio choice. J. Finance 26, 857–884.

Hall, R.E., Woodward, S.E., 2010. The burden of nondiversifiable risk of entrepreneurship. Am. Econ. Rev. 100, 1163–1194.

Harris, R.S., Jenkinson, T., Kaplan, S.N., 2014. Private equity performance: What do we know? J. Finance 69, 1851–1882.

Harris, R.S., Jenkinson T., Kaplan, S.N., Stucke, R., 2018. Financial intermediation in private equity: How well do funds of funds perform? J. Financ. Econ. 129, 287–305.

Harris, R.S., Jenkinson T., Kaplan, S.N., Stucke, R., 2022. Has persistence persisted in private equity? Evidence from buyout and venture capital funds. Work. Pap., Univ. Virginia, Univ. Oxford, Univ. Chicago, and Warburg Pincus.

Harvey, C.R., Siddique, A., 2000. Conditional skewness in asset pricing tests. J. Finance 55, 1263–1295.

Haurin, D.R., Hendershott, P.H., 1991. House price indexes: Issues and results. Real Estate Econ. 19, 259–269.

Heckman, J.J., 1979. Sample selection bias as a specification error. Econometrica 47, 153–161.

Hege, U., Nuti, A., 2011. The private equity secondaries market during the financial crisis and the "valuation gap." J. Priv. Equity 14, 42–54.

Hellmann, T., Puri, M., 2002. Venture capital and the professionalization of start-up firms: Empirical evidence. J. Finance 57, 169–197.

Higson, C., Stucke, R., 2012. The performance of private equity. Work. Pap., London Bus. Sch. and Univ. Oxford.

Hüther, N., 2021. Do private equity managers raise funds on (sur)real returns? Evidence from deal-level data. Work. Pap., Indiana Univ.

Hüther, N., Robinson, D.T., Sievers, S., Hartmann-Wendels, T., 2020. Paying for performance in private equity: Evidence from venture capital partnerships. Manag. Sci. 66, 1756–1782.

Hüther, N., Schmid, L., Steri, R., 2021. Credit market equivalents and the valuation of private firms. Work. Pap., Indiana Univ., Univ. Southern Calif. and Univ. Luxermbourg.

Hwang, M., Quigley, J.M., 2004. Selectivity, quality adjustment and mean reversion in the measurement of house values. J. Real Estate Finance Econ. 28, 161–178.

Hwang, M., Quigley, J.M., Woodward, S.E., 2005. An index for venture capital, 19872003. Contrib. Econ. Anal. Policy 4, 1–43.

Ilmanen, A., Chandra, S., McQuinn, N., 2020. Demistifying illiquid assets: Expected returns for private equity. J. Altern. Investments 22, 1–15.

Imbierowicz, B., Rauch, C., 2021. The pricing of private assets: Mutual investments in 'unicorn' companies. Work. Pap., Deutsche Bundesbank and Amer. Univ. Sharjah.

Ivashina, V., Lerner, J., 2018. Looking for alternatives: Pension investments around the world, 2008 to 2017. Work. Pap., Harvard Univ.

Jeffers, J., Lyu, T., Posenau, K., 2021. The risk and return of impact investing funds. Work. Pap. Univ. Chicago and Yale Univ.

Jegadeesh N., Kräussl R., Pollet J.M., 2015. Risk and expected returns of private equity investments: Evidence based on market prices. Rev. Financ. Stud. 28, 3269–3302.

Jenkinson, T., Jones, H., Rauch, C., 2021. Long goodbyes: Why do private equity funds hold onto public equity? Work. Pap., Univ. Oxford and Amer. Univ. Sharjah.

Jenkinson, T., Sousa, M., Stucke, R., 2013. How fair are the valuations of private equity funds? Work. Pap., Univ., Oxford.

Jenkinson, T., Landsman, W.R., Rountree, B.R., Soonawalla, K., 2020. Private equity net asset values and future cash flows. The Accounting Rev. 95, 191–210.

Jud, G.D., Seaks, T.G., 1994. Sample selection bias in estimating housing sales prices. J. Real Estate Res. 9, 289–298.

Kaplan, S.N., 1989. The effects of management buyouts on operating performance and value. J. Financ. Econ. 24, 217–254.

Kaplan, S.N., Lerner, J., 2017. Venture capital data: Opportunities and Challenges, in: Haltiwanger, J., Hurst, E., Miranda, J., Schoar, A. (Eds), Measuring Entrepreneurial Businesses: Current Knowledge and Challenges. Univ. Chicago Press, Chicago, pp. 413–431.

Kaplan, S.N., Schoar, A., 2005. Private equity performance: Returns, persistence, and capital flows. J. Finance 60, 1791–823.

Kerr, W.R., Nanda, R., Rhodes-Kropf, M., 2014. Entrepreneurship as experimentation. J. Econ. Perspectives 28, 25-48.

Korteweg, A., 2019. Risk adjustment in private equity returns. Ann. Rev. Financ. Econ. 11, 131–152.

Korteweg, A., Nagel, S., 2016. Risk-adjusting the returns to venture capital. J. Finance 71, 1437–1470.

Korteweg, A., Nagel, S., 2022. Risk-adjusted returns of private equity funds: A new approach. Work. Pap., Univ. Southern Calif. and Univ. Chicago.

Korteweg, A., Panageas, S., Systla, A., 2021. Evaluating private equity from an investor's Perspective. Work. Pap., Univ. Southern Calif., and UCLA.

Korteweg, A., Sensoy, B., 2023. How unique is VC's American history? J. Econ. Lit. 61, 274–294.

Korteweg, A., Sørensen, M., 2010. Risk and return characteristics of venture capital–backed entrepreneurial companies. Rev. Financ. Stud. 23, 3738–3772.

Korteweg, A., Sørensen, M., 2016. Estimating loan-to-value distributions. Real Estate Econ. 44, 41–86.

Korteweg, A., Sørensen, M., 2017. Skill and luck in private equity performance. J. Financ. Econ. 124, 535–562.

Korteweg, A., Westerfield, M., 2022. Asset allocation with private equity. Foundations and Trends in Finance 13, 95–204.

Kraus, A., Litzenberger, R., 1975. Market equilibrium in a multiperiod state preference model with logarithmic utility. J. Finance 30, 1213–1227.

Lee, A., 2015. How LPs are driving private equity fundraising. International Financ. Law Rev., published July 8, 2015.

Lerner, J., 1995. Venture capitalists and the oversight of private firms. J. Finance 50, 301–318.

Lerner, J., Leamon, A., Hardymon, F., 2012. Venture Capital, Private Equity, and the Financing of Entrepreneurship, first ed. John Wiley & Sons, New Jersey.

Lerner, J., Mao, J., Schoar, A., Zhang, N.R., 2020, Investing outside the box: Evidence from alternative vehicles in private equity. Work. Pap., Harvard Business School Finance Working Paper No. 19-012.

Lerner. J., Schoar, A., 2004. The illiquidity puzzle: Theory and evidence from private equity. J. Financ. Econ. 72, 3–40.

Lerner, J., Sørensen, M., Strömberg, P., 2011. Private equity and long-run investment: The case of innovation. J. Finance 66, 445–477.

Litvak, K., 2004. Governance through exit: Default penalties and walkaway options in venture capital partnership agreements. Willamette Law Rev. 40, 771–778.

Litvak, K., 2009. Venture capital limited partnership agreements: Understanding compensation arrangements. Univ. of Chicago Law Rev. 76, 161–218.

Ljungqvist, A., Richardson, M., 2003. The cash flow, return and risk characteristics of private equity. Work. Pap., New York Univ.

Long, J.B., 1990. The numeraire portfolio. J. Financ. Econ. 26, 29–69.

Long, A.M., Nickels, C.J., 1996. A private investment benchmark. Work. Pap., Univ. Tex. Invest. Manag. Co.

Martin, I., 2012. On the valuation of long-dated assets. J. Pol. Econ. 120, 346–358.

Martin, J.D., Petty, J.W., 1983. An analysis of the performance of publicly traded venture capital companies. J. Financ. Quant. Anal. 18, 401–498.

McCourt, M., 2018. Estimating skill in private equity performance using market data. Work. Pap., Univ. of Melbourne.

McKenzie, M., Janeway, W., 2008. Venture capital fund performance and the IPO market. Work. Pap., Univ. Cambridge.

Merton, R.C., 1980. On estimating the expected return on the market: An exploratory investigation. J. Financ. Econ. 8, 323–361.

Metrick, A., Yasuda, A., 2010a. The economics of private equity funds. Rev. Financ. Stud. 23, 2303–2341.

Metrick, A., Yasuda, A., 2010b. Venture Capital and the Finance of Innovation, 2nd ed. John Wiley & Sons, New Jersey.

Munday, S., Hu, W., True, T., Zhang, J., 2018. Performance of private credit funds: A first look. J. Altern. Inv. 21, 31–51.

Munneke, H.J., Slade, B.A., 2000. An empirical study of sample-selection bias in indices of commercial real estate. J. Real Estate Finance Econ. 21, 45–64.

Munneke, H.J., Slade, B.A., 2001. A metropolitan transaction-based commercial price index: A time-varying parameter approach. Real Estate Econ. 29, 55–84.

Nadauld, T.D., Sensoy, B.A., Vorkink, K., Weisbach, M.S., 2019. The liquidity cost of private equity investments: Evidence from secondary market transactions. J. Financ. Econ. 132, 158–181.

OECD, Eurostat, International Labour Organization, International Monetary Fund, The World Bank, United Nations Economic Commission for Europe, 2013. Repeat sales methods, in: Handbook on Residential Property Price Indices, Eurostat, Luxembourg, pp. 66–71.
Peng, L., 2001, Building a venture capital index. Work. Pap., Univ. Colorado Boulder.

Opp, C.C., 2019. Venture capital and the macroeconomy. Rev. Fin. Stud. 32, 4387–4446.

Ouyang, S., Yu, J., Jagannathan, R. 2021 Return to venture capital in the aggregate. NBER Work. Pap. 27690.

Pástor, L., Stambaugh, R.F., 2003. Liquidity risk and expected stock returns. J. Pol. Econ. 111, 642–685.

Peters, R.H., 2018. Volatility and venture capital. Work. Pap., Tulane Univ.

Phalippou, L., 2008. The hazards of using IRR to measure performance: The case of private equity. J. Performance Measurement 12, 55–66.

Phalippou, L., Gottschalg, O., 2009. The performance of private equity funds. Rev. Financ. Stud. 22, 1747–1776.

Phalippou, L., 2010. Risk and return of private equity: An overview of data, methods, and results, in: Cumming, D. (Ed.), Private Equity: Fund Types, Risks and Returns, and Regulation. John Wiley & Sons, New Jersey, pp. 257–282.

Phalippou, L., 2013. Yale's endowment returns: Case study in GIPS interpretation difficulties. J. Altern. Investments 15, 97–103.

Phalippou L., 2014. Performance of buyout funds revisited? Rev. Finance 18, 189–218.

Phalippou, L., Gottschalg, O., 2009. The performance of private equity funds. Rev. Financ. Stud. 22, 1747–1776.

Phalippou, L., Rauch, C., Umber, M., 2018. Private equity portfolio company fees. J. Financ. Econ. 129, 559–585.

Robinson, D.T., Sensoy, B.A., 2013. Do private equity fund managers earn their fees? Compensation, ownership, and cash flow performance. Rev. Financ. Stud. 26, 2760–2797.

Robinson, D.T., Sensoy, B.A., 2016. Cyclicality, performance measurement, and cash flow liquidity in private equity. J. Financ. Econ. 122, 521–543.

Roll, R., 1973. Evidence on the "growth-optimum" model. J. Finance 28, 551–566.

Roll, R., 1977. A critique of the asset pricing theory's tests part I: On past and potential testability of the theory. J. Financ. Econ. 4, 129–176.

Rubinstein, M., 1976. The strong case for the generalized logarithmic utility model as the premier model of financial markets. J. Finance, 551–571.

Sagi, J., 2021. Asset-level risk and return in real estate investments. Rev. Financ. Stud. 34, 3647–3694.

Schillinger, P., Braun, R., Cornel, J., 2019. Distortion or cash flow management? Understanding credit facilities in private equity funds. Work. Pap., Techn. Univ. München and Blackrock.

Schlingemann, F.P., Stulz, R.M., 2022. Have exchange-listed firms become less important for the economy? J. Financ. Econ. 143, 927–958.

Scholes, M., Williams, J., 1977. Estimating betas from nonsynchronous data. J. Financ. Econ. 5, 309–327.

Shanken, J., 1992. On the estimation of beta-pricing models. Rev. Financ. Stud. 5, 1–33.

Smith, E.E., Smith, J.K., Smith, R.L., 2022. Bias in the reporting of venture capital performance: The disciplinary role of FOIA. Rev. Corp. Fin. 2, 493–525.

Sørensen, M., Jagannathan, R., 2015. The public market equivalent and private equity performance. Financ. Analysts J. 71, 43–50.

Stafford, E., 2022. Replicating private equity with value investing, homemade leverage, and hold-to-maturity accounting. Rev. Financ. Stud. 35, 299–342.

Stucke, R., 2011. Updating history. Work. Pap., Univ. Oxford, Oxford, UK.

Wallace, N.E., Meese, R.A., 1997. The construction of residential housing price indices: A comparison of repeat-sales, hedonic-regression, and hybrid approaches. J. Real Estate Finance Econ. 14, 51–73.

Woodward, S.E., 2009. Measuring risk for venture capital and private equity portfolios. Work. Pap., Sand Hill Econ., Palo Alto, CA.

Woodward, S.E., Hall, R.E., 2004. Benchmarking the returns to venture. National Bureau of Economic Research Work. Pap. #10202.

Table 1: Descriptive Statistics for Private Equity Funds

This table shows descriptive statistics for a sample of closed private equity funds with a North American geographic focus for vintage years ranging from 1980 to 2019. Funds of funds, secondary funds, co-investments, and venture debt funds are excluded. Each column represents a different private equity strategy. *Fund size* is in millions of U.S. dollars. *Number of firms* is the count of the number of unique PE firms in the sample. The same firm may raise multiple funds. *TVPI* is the total value to paid-in capital multiple, *IRR* is the internal rate of return, and *PME (KS)* is the Kaplan-Schoar (2005) public market equivalent using the S&P 500 as the benchmark. Reported fund sizes are averages, with medians in parentheses. Fund performance statistics are computed across all funds and vintages for each strategy, based on cash flow and Net Asset Value data reported as of year-end 2019. *Exc. Kurtosis* is the kurtosis of the distribution in excess of the Normal distribution (which has kurtosis of 3). Source: Preqin.

	VC	Buyout	Real Estate	Private Debt	Natural Resources	Infra-structure
Number of funds	1,470	1,300	1,435	723	291	144
Fund size ($m)	248	1,333	574	868	914	2,067
	(150)	(500)	(269)	(438)	(410)	(999)
Liquidated (%)	46	38	38	29	33	13
Number of firms	569	431	380	244	92	65
Fund performance:						
TVPI						
Mean	1.89	1.80	1.42	1.41	1.55	1.28
Median	1.38	1.61	1.38	1.30	1.30	1.22
St. Dev.	2.43	1.23	0.51	0.50	1.05	0.51
Skewness	7.81	8.46	0.81	2.78	4.18	1.29
Exc. Kurtosis	92.51	134.78	3.12	14.88	25.38	4.64
IRR (%)						
Mean	14.79	16.19	11.79	11.19	14.94	9.59
Median	9.22	14.00	11.58	10.02	9.54	8.78
St. Dev.	37.64	21.04	14.90	13.15	42.73	16.87
Skewness	5.91	4.85	0.89	1.55	9.77	-0.68
Exc. Kurtosis	58.68	63.72	33.11	39.71	126.56	16.15
PME (KS)						
Mean	0.96	1.11	0.96	1.01	0.94	0.86
Median	0.98	1.10	0.97	0.98	0.86	0.90
St. Dev.	0.42	0.32	0.26	0.19	0.30	0.24
Skewness	1.25	0.30	0.83	1.09	1.72	-1.27
Exc. Kurtosis	6.38	0.33	8.62	2.30	5.05	2.01

Table 2: Fund Data for a Fictitious Private Equity Fund

Fund data for a fictitious private equity fund, indexed by i. The first column shows the cash flow dates, t, where dates of the first and last observations, t_0, and T, are February 1, 2022, and December 31, 2024. Column two shows years elapsed since t_0. The third and fourth columns are capital calls, I (indexed by k from 1 to K), and distributions D (indexed by d from 1 to D), respectively. Column five shows the net cash flow from the Limited Partner's perspective, $C = D - I$, indexed by j from 1 to J. Column six is the public stock market return since t_0 ($R^M_{t_0,\tau}$, where a 5% return is shown as 1.050). The final three columns show quarterly fund data: the end-of-quarter reported Net Asset Value (NAV) and its quarterly net cash flow and fund return (R), based on the net cash flow and the change in NAV over the quarter. All cash flows (capital calls, distributions, and net cash flows) and NAVs are in thousands of U.S. dollars, normalized to a $10 million commitment. Cash flows are net of fees paid to the fund's General Partners.

Date	Years since t_0	Cash flow data			Stock market return	Quarterly data		
		Capital calls	Distri-butions	Net cash flow		Net Asset Value	Net cash flow	Fund return
t	τ	$I_{i,k}$	$D_{i,d}$	$C_{i,j}$	$R^M_{t_0,\tau}$	$NAV_{i,t}$	$C_{i,t}$	$R_{i,t}$
Feb 1, 2022	0	200	0	-200	1			
Mar 31, 2022	0.16	0	0	0	1.020	190	-200	N/A
Jun 16, 2022	0.37	100	0	-100	1.022			
Jun 30, 2022	0.41	0	0	0	1.017	250	-100	0.789
Jul 16, 2022	0.45	1,900	0	-1,900	1.026			
Sep 30, 2022	0.66	0	0	0	1.054	2,200	-1,900	1.200
Dec 31, 2022	0.91	0	0	0	1.098	2,200	0	1.000
Mar 31, 2023	1.16	0	0	0	1.159	2,750	0	1.250
Jun 20, 2023	1.38	1,500	0	-1,500	1.178			
Jun 30, 2023	1.41	0	0	0	1.023	3,800	-1,500	0.836
Sep 30, 2023	1.66	0	0	0	1.044	4,100	0	1.079
Dec 31, 2023	1.91	0	0	0	1.106	4,100	0	1.000
Feb 1, 2024	2.00	500	0	-500	1.110			
Mar 31, 2024	2.16	0	0	0	1.062	4,600	-500	1.000
Apr 10, 2024	2.19	200	500	300	1.095			
Apr 20, 2024	2.21	50	0	-50	1.108			
Jun 30, 2024	2.41	0	0	0	1.167	4,500	250	1.033
Sep 30, 2024	2.66	0	0	0	1.222	4,300	0	0.956
Dec 20, 2024	2.88	0	1,000	1,000	1.244			
Dec 31, 2024	2.91	0	0	0	1.250	3,300	1,000	1.000
Total		**4,450**	**1,500**	**-2,950**			**-2,950**	

Table 3: Descriptive Statistics for PE Index Returns

This table reports descriptive statistics for Cambridge Associate quarterly private equity index returns data and the public stock market index from the first quarter of 1995 until the third quarter of 2020. The private equity indices in the first three columns are the U.S. Venture Capital, the U.S. Private Equity, and the Global Real Estate index, respectively. The public stock market index is the value-weighted return of all U.S. incorporated firms that are listed on the NYSE, AMEX, or NASDAQ (sourced from Kenneth French's data library). The mean, standard deviation, skewness, and excess kurtosis (relative to the Normal distribution) are reported as annualized numbers, in % (e.g., 5 represents a 5% annualized return). Sharpe ratios are annualized. Autocorrelations are reported for lags ranging from 1 to 6 quarters and are computed on quarterly returns. Sources: https://www.cambridgeassociates.com/private-investment-benchmarks/ and https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

	Venture Capital	Buyout	Real Estate	Public Stocks
Returns (annualized, in %):				
Mean	17.06	14.17	9.40	11.46
St. Dev.	23.19	10.42	8.80	17.59
Skewness	1.75	-0.31	-1.07	-0.28
Exc. Kurtosis	5.67	0.53	3.09	0.15
Sharpe ratio (ann.)	0.65	1.15	0.83	0.52
Autocorrelation:				
1 quarter	0.60	0.32	0.62	-0.06
2 quarters	0.48	0.24	0.56	0.04
3 quarters	0.32	0.12	0.32	0.02
4 quarters	0.04	0.06	0.31	-0.04
5 quarters	-0.04	0.00	0.08	0.12
6 quarters	-0.08	0.02	-0.03	0.03

Table 4: CAPM Estimates for PE Index Returns

This table shows Capital Asset Pricing Model (CAPM) estimates for the quarterly Cambridge Associates venture capital, buyout, and real estate index returns from the first quarter of 1995 until the third quarter of 2020. The regression model is

$$R_{i,t} - R_t^F = \alpha_i + \sum_{s=-S}^{0} \beta_{i,s} \left[R_{t+s}^M - R_{t+s}^F \right] + \varepsilon_{it},$$

where *i* refers to the venture capital, buyout, or real estate index described in Table 3. The excess market return on the right-hand side is from Kenneth French's data library (see Table 3 for details). The intercept is in % per quarter (e.g., 5 means 5% per quarter). The columns labeled "Coeff." report regression coefficients, and the "s.e" columns show Newey-West standard errors that are robust to heteroskedasiticy and autocorrelation in the residuals. Panel A shows results for S= 0, i.e., regressions of index excess returns on contemporaneous market excess returns. Panel B uses S = 6 quarters, i.e., the right-hand side contains the contemporaneous market excess return and its first 6 lags. The bottom row shows the Dimson beta estimates, which are the sum of the contemporaneous and lagged betas. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

	VC		Buyout		Real Estate	
	Coeff.	s.e.	Coeff.	s.e.	Coeff.	s.e.
Panel A: Contemporaneous factor return only (S = 0)						
$\beta_{i,0}$	0.579	0.153***	0.465	0.044***	0.166	0.089*
α_i	2.368	1.097**	1.909	0.320***	1.405	0.751*
Panel B: Dimson correction (S = 6 quarters)						
$\beta_{i,s}$:						
s = 0 (contemp.)	0.577	0.163***	0.476	0.034***	0.173	0.079**
s = -1	0.212	0.105**	0.143	0.032***	0.124	0.056**
s = -2	0.235	0.112**	0.090	0.029***	0.108	0.051**
s = -3	0.263	0.097***	0.084	0.024***	0.111	0.050**
s = -4	0.229	0.192	0.051	0.033	0.116	0.050**
s = -5	0.078	0.170	0.012	0.030	0.059	0.038
s = -6	0.064	0.091	-0.006	0.041	0.023	0.038
α_i	-0.052	1.186	1.140	0.328***	0.379	0.877
$\sum_{s=-6}^{0} \beta_{i,s}$	1.658	0.327***	0.851	0.087***	0.713	0.236***

Table 5: Factor Model Estimates in Private Equity using the Dimson (1979) Method.

This table summarizes the literature that estimates factor models in private equity using the Dimson (1979) method. The “Level” column shows whether data is observed at the aggregate/index level (“Aggr.”), fund level (“Fund”) or individual security level (“Sec.”). “Lags” shows the number of lags in the Dimson regression. “Model” reports the factor model used, where CAPM is the Capital Asset Pricing Model, FF3 stands for Fama and French (1993), and FF3+PS is the Fama-French (1993) model augmented with the Pástor-Stambaugh (2003) liquidity factor. The return series used as the market factor is in “Market factor,” where CRSP-VW is the CRSP value-weighted market index, and KF is the market factor from Kenneth French’s data library. Alphas are in % and at the same frequency as the data (in “Freq.”), except Bilo et al. (2005), Metrick and Yasuda (2010b), Phalippou (2010) and Stafford (2022), who report annualized alphas. For FF3 the “ β ” column shows loadings on the market / Small-minus-Big (SMB) / High-minus-Low (HML) factors, augmented with the liquidity factor loading as the final number for the FF3+PS model. If the same paper estimates multiple models, only the differences in specifications are shown in the additional rows (blank cells mean no change). Data source abbreviations: CA = Cambridge Associates; CRSP = Center for Research in Securities Prices; D = Datastream; LPS = LP Source; P = Preqin; R = Thomson Reuters; SEC = Securities and Exchange Commission; SHE = Sand Hill Econometrics; SMI = Secondary market intermediary (unnamed); VE = Venture Economics. N.R. stands for “Not Reported.”

Paper	Data source	Level	Sample period	Freq.	Lags	Model	Market factor	VC		Buyout	
								α (%)	β	α (%)	β
Woodward and Hall (2004)	SHE	Aggr.	1987-2002	Month	13	CAPM	Nasdaq	N.R.	0.9	N.R.	N.R.
Bilo et al. (2005)	D	Aggr	1986-2003	Week	3	CAPM	MSCI World	N.R.	N.R.	$-0.1^{*,a}$	1.1[a]
								N.R.	N.R.	$10.4^{*,b}$	1.0[b]
Anson (2007)	VE	Fund	1985-2005	Quarter	3	CAPM	S&P 500	0.2	1.4	0.8	0.7
							Nasdaq	0.5	1.1	1.2	0.4
							Russell 1000	0.9	1.4	1.2	0.7
							Russell 2000	1.5	0.9	1.3	0.6
Woodward (2009)	CA	Aggr	1996-2008	Quarter	5	CAPM	Wilshire 5000	0.5	2.2	1.4	1.0
Metrick and Yasuda (2010b)	CA	Aggr	1981-2008	Quarter	8[c]	FF3 +PS	KF	0.1^{*}	2.0/1.0/-1.5/ 0.2	N.R.	N.R.
	SHE	Aggr	1989-2008	Month	24[c]			-2.1^{*}	1.6/-0.1/-0.7/ 0.3	N.R.	N.R.

Phalippou (2010)	D	Fund	1993-2008	Week	10	CAPM	Local index	N.R.	N.R.	6*	1.5
Ewens et al. (2013)	VE+P+LPS	Aggr	1980-2007	Quarter	4	CAPM	CRSP-VW	–0.2	1.2	1.2	0.7
						FF3		0.5	1.1/–0.1/–0.9	0.9	0.8/0.1/0.2
Peters (2018)	CA	Aggr	N.R.	Quarter	5	CAPM	CRSP-VW	0.6	1.4	N.R.	N.R.
Agarwal et al. (2022)	CRSP +SEC+R	Sec.	2010-2018	Quarter	2	CAPM	N.R.	0.5	1.3	N.R.	N.R.
						FF3		0.7	1.0/1.5/–1.0	N.R.	N.R.
Boyer et al. (2022a)	SMI	Aggr	2006-2018	Quarter	1	CAPM	KF	N.R.	N.R.	-2*	1.8
	Burgiss				3					4*	0.8
	Preqin				3					4*	0.7
Stafford (2022)	CA	Aggr	1996-2014	Quarter	3	CAPM	KF	N.R.	N.R.	5.0*	0.8
	Burgiss									4.5*	0.9
	Preqin									5.2*	0.7

*Alpha is annualized; [a]Based on equal-weighted buy-and-hold returns; [b]Based on equal-weighted weekly rebalanced returns; [c]Reported lags are for the market factor (the other factor loadings are estimated using half the number of lags used for the market factor).

Table 6: Issues with Fund Net Asset Values and Proposed Solutions

This table shows whether techniques proposed in the PE literature may be able to resolve the main concerns with using reported net asset values for fund performance evaluation (with "Y" for yes, and "N" for no). The columns show the concerns, which are described in detail in the text (sections 3.5 and 3.6). The rows list the various techniques (see section 3.5.1 through 3.5.5 for details). Most techniques were developed specifically to deal with the staleness issue, but some show promise to resolve other problems as well.

	Staleness	Manipulation	Fees	VC post-money valuations
Longer-horizon returns	Y	N	N	N
Unsmoothing	Y	N	N	N
Dimson correction	Y	N	N	N
Imputation/filtering	Y	Y	Y	Y
Secondary market prices	Y	Y	Y	Y

Table 7: Public Market Equivalent Estimates of Private Equity Funds

This table shows a representative (but non-comprehensive) overview of estimates of the Public Market Equivalent (PME) of Kaplan and Schoar (2005) from the private equity literature. Results for funds of funds, secondary funds, venture debt funds, co-investments, alternative vehicles, and publicly traded vehicles are excluded (to the extent they are excluded or separately reported in the original papers). The results shown are for active funds (not only liquidated funds, to the extent that these results are separately reported). All PME calculations use the return on the S&P 500 as the discount rate, unless noted in the "Notes" column (where CRSP-VW stands for the CRSP value-weighted market index). The "Vintage" column shows the fund vintage years covered by the sample in the paper. "Table" is the source table number in the cited paper that contains the PME numbers. "EW" and "VW" stand for equal-weighted and value-weighted mean, respectively. The "Other PME" column shows results for PE strategies other than venture capital (VC) or buyout (see the "Notes" column for the name of the strategy). Data source abbreviations: B = Burgiss; CA = Cambridge Associates; CP = California Public Employees' Retirement System; IFD = Impact Finance Database; LP*n* =sourced from *n* limited partners (LPs); P = Preqin; SS = State Street; St = StepStone SPI; VE = Venture Economics. N.R. stands for "Not Reported."

Paper	Data source	Vintage	Table	VC PME			Buyout PME			Other PME			Notes
				EW	VW	Median	EW	VW	Median	EW	VW	Median	
Kaplan and Schoar (2005)	VE	1980 -1994	II	0.96	1.21	0.66	0.97	0.93	0.80				
McKenzie and Janeway (2008)	LP2	1980 -2002	VI	1.98	N.R.	1.21							
Phalippou and Gottschalg (2009)	VE	1980 -1993	2B	N.R.	0.88	N.R.	N.R.	0.96	N.R.				
Higson and Stucke (2012)	CA +CP	1980 -2008	VI				1.22	1.23	1.13				
Axelson et al. (2013)	P	1987 -2010	VI				1.36	N.R.	1.35				CRSP-VW
Harris et al. (2014)	B	1984 -2008	III	1.36	1.45	1.02	1.22	1.27	1.16				
Phalippou (2014)	P	1993 -2010	I				1.03	1.19	0.99				CRSP-VW
Robinson and Sensoy (2016)	LP1	1984 -2005	2	1.06	N.R.	0.84	1.19	N.R.	1.09				
Brown et al. (2019)	B+St	1971 -2016	1	1.19	N.R.	1.11	1.21	N.R.	1.18				CRSP-VW

Gredil et al. (2020b)	P	1983-2012	I	1.08	N.R.	0.82	1.14	N.R.	1.09				
Harris et al. (2022)	B	1984-2015	1	1.29	N.R.	N.R.	1.18	N.R.	N.R.				
Lerner et al. (2020)	SS	1985-2010	App. B	1.44	1.19	1.09	1.21	1.20	1.15				
Andonov et al. (2021)	P	2002-2018	3	0.99	1.02	0.95	1.05	1.06	1.01	0.93	0.93	0.95	Infrastructure funds
	B	2002-2018								0.89	N.R.	0.91	Infrastructure funds
Brown et al. (2021a)	B	1983-2008	1	1.2	N.R.	0.82	1.1	N.R.	1.02				CRSP-VW
Jeffers et al. (2021)	IFD+P	1999-2016	1	0.88	N.R.	0.82				0.78	N.R.	0.77	Impact funds; CRSP-VW

Table 8: Repeat-Sales Regression Illustration

This highly stylized example illustrates the intuition behind repeat-sales regressions (RSR). The table shows the valuations for three fictitious startups (A, B, and C) at the time of their fundraising rounds. The numbers in square brackets are the imputed valuations that the RSR approach would assign to these startups in the periods where no valuations are observed. The final row shows the RSR index return. Note that there is no imputed valuation for startup B in year 1, as the company only raises its first round of funding in year 2.

Year	1	2	3	4
A	1	1.5	[1.875]	[2.5]
B		4	5	[6.67]
C	2	[3]	[3.75]	5
RSR return	N/A	50%	25%	33%

Figure 1
Distribution of Capital Asset Pricing Model (CAPM) beta point estimates from the literature on venture capital (VC) and leveraged buyouts (BO). The figure shows separate box plots for gross-of-fee and net-of-fee estimates (typically from individual investment-level and fund-level data, respectively) for each strategy. The bottom and top of each box are the first and third quartiles of the distribution. The center line is the median. The mean is marked with an "x". The distributions of VC gross and net of fees betas is based on 10 and 11 estimates, respectively. The buyout gross and net of fees distributions are based on 5 and 12 estimates. The box plots include estimates from the following papers (each paper contributes at most one beta to each distribution, but some papers contribute estimates to multiple plots): Gompers and Lerner (1997); Peng (2001); Woodward and Hall (2004); Bilo et al. (2005); Cochrane (2005a); Hwang et al. (2005); Anson (2007); Ewens (2009); Woodward (2009); Korteweg and Sørensen (2010); Phalippou (2010); Groh and Gottschalg (2011); Driessen et al. (2012); Franzoni et al. (2012); Ewens et al. (2013); Axelson et al. (2014); Buchner and Stucke (2014); Jegadeeh et al. (2015); Ewens et al. (2016); Ang et al. (2018); Peters (2018); Buchner (2020); Brown et al. (2021a); Agarwal et al. (2022); Boyer et al. (2022b); Stafford (2022).

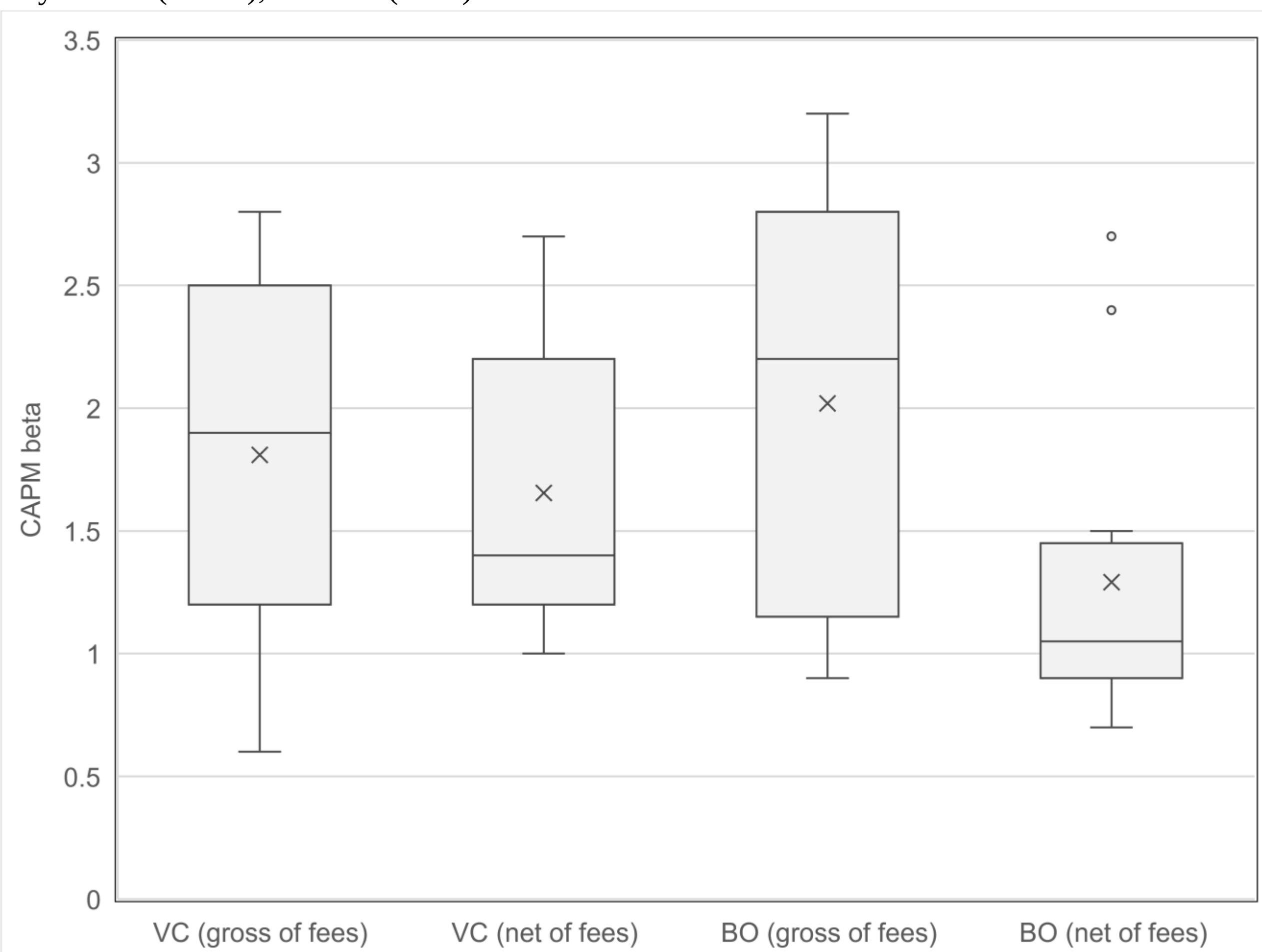


Figure 2

Cash flow patterns for PE funds by strategy, for 1980 to 2005 vintages with a North American geographic focus for which cash flows are reported in Preqin. Funds of funds, secondary funds, co-investments, and venture debt funds are excluded. Cash flows are available until the end of 2019, they are normalized to a $10 million capital commitment to each fund, and are net fees to the General Partners. The dotted red line represents average cumulative capital calls since fund inception (in millions of U.S. dollars, shown as negative numbers) across funds that are active at a given number of years since fund inception (on the horizontal axis). The dashed blue line shows average cumulative distributions to LPs, and the solid black line is the average cumulative net cash flow (cumulative distributions minus capital calls). The shaded areas show the range between the 5th and 95th percentiles of the cross-fund distribution of cumulative capital calls and distributions. The distribution for net cash flows is omitted for readability. Source: Preqin.

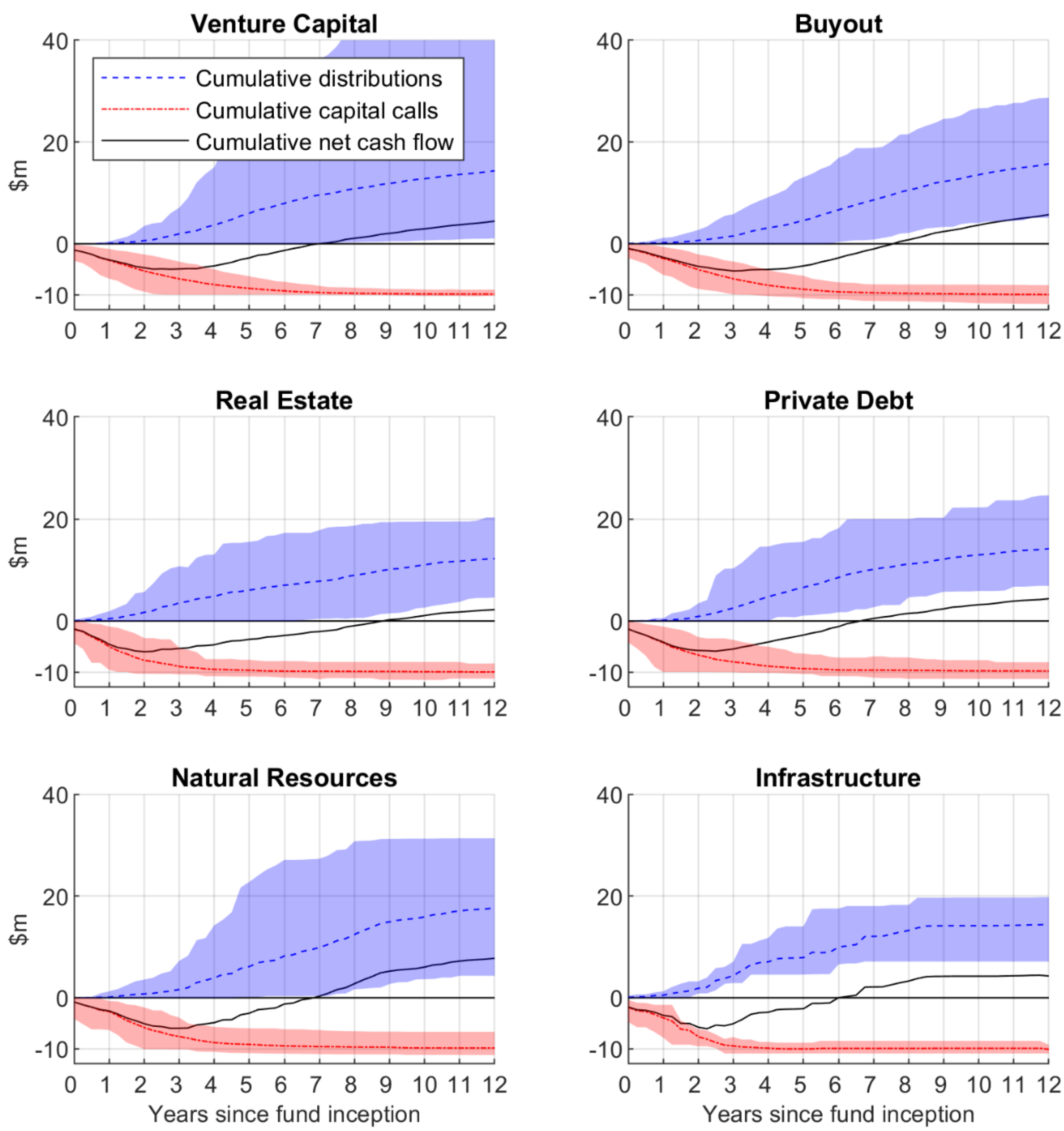


Figure 3

This figure shows the Epanechnikov kernel smoothed distribution of three performance measures for private equity funds: total value to paid-in capital (*TVPI*), internal rate of return (*IRR*), and the Kaplan-Schoar (2005) public market equivalent (*PME (KS)*). The sample of funds is described in Table 1, except that vintages after 2014 are excluded to ensure that each fund has at least 5 years of data (performance is computed on cash flows and net asset values reported by the end of 2019). Each row represents a different private equity strategy. The distributions of the performance metrics for a given strategy include all funds and vintages for that strategy.

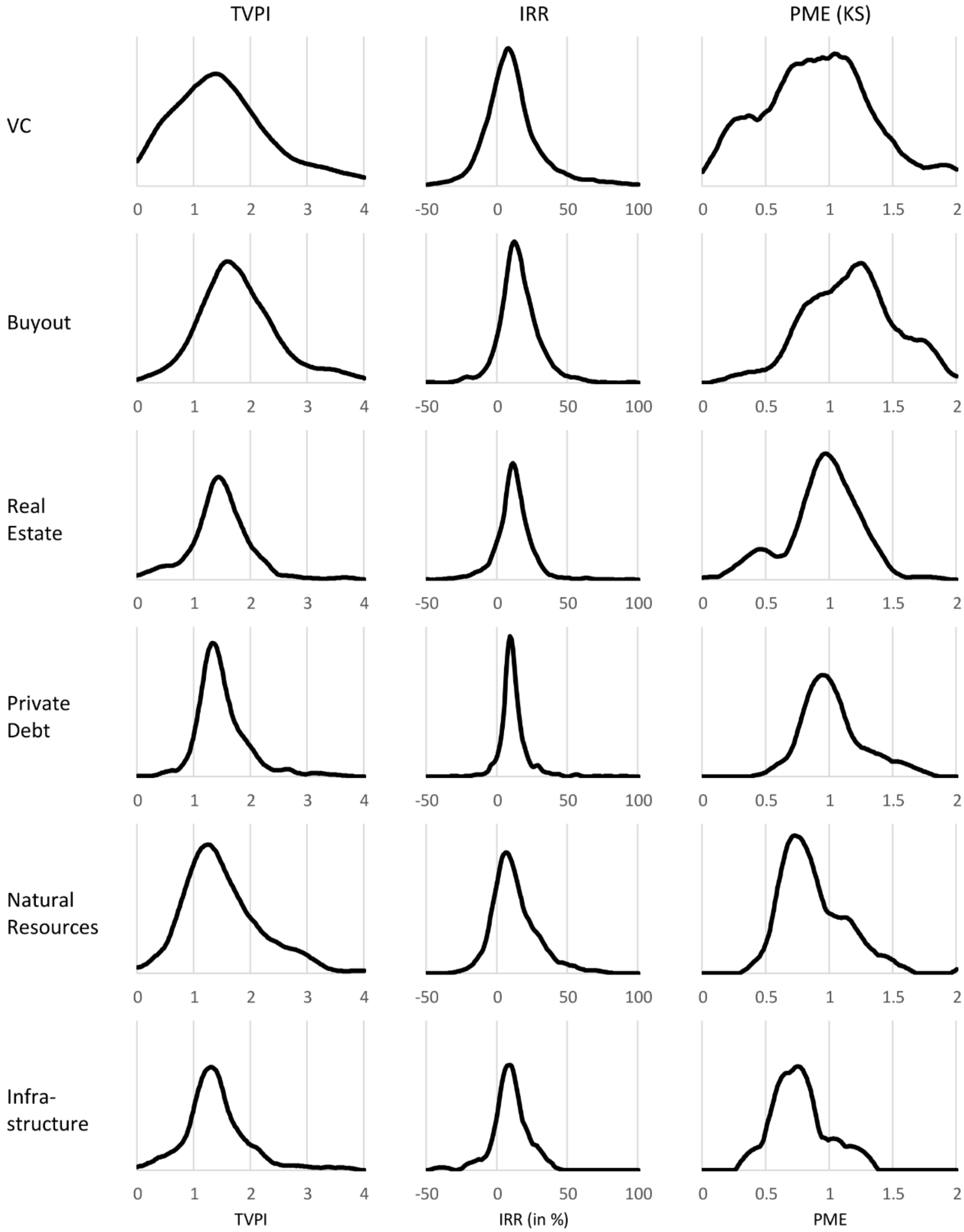
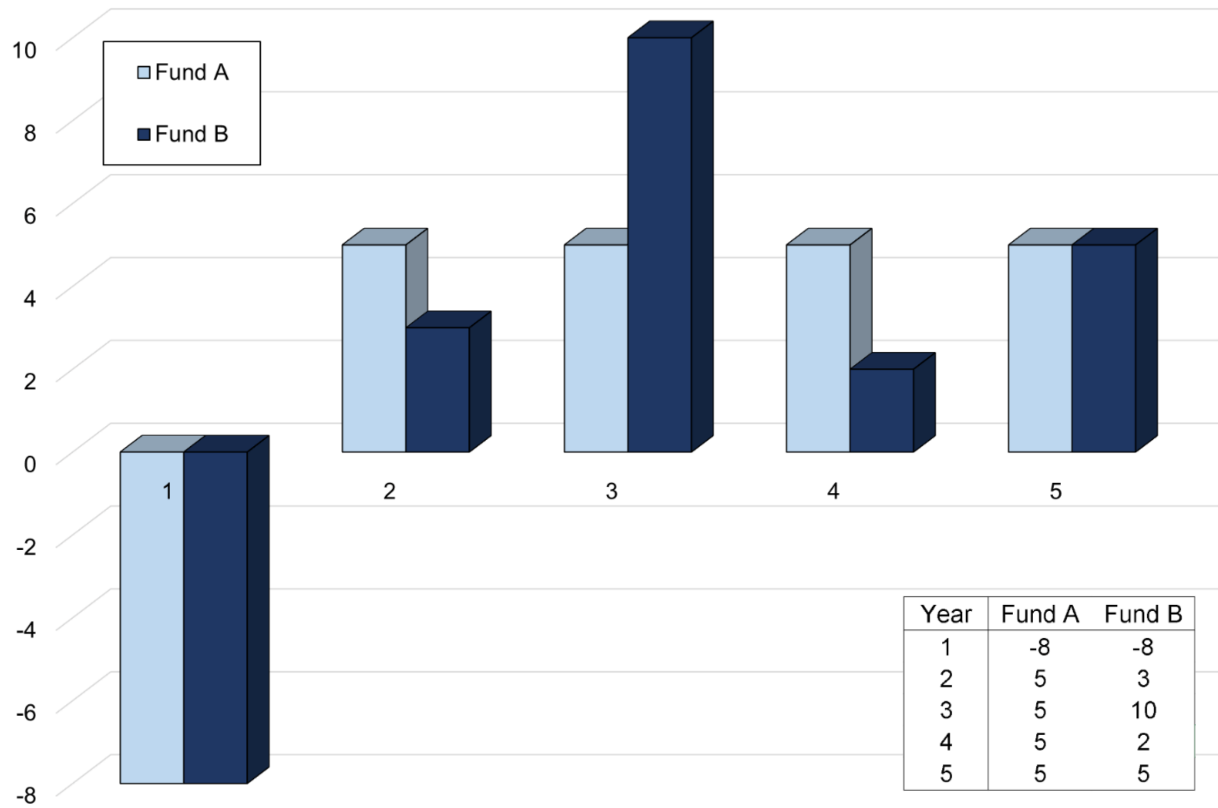


Figure 4

Example of cash flows for two fictitious private equity funds, called A and B, to illustrate differences in riskiness. Years are on the horizonal axis, and annual cash flows are on the vertical axis. The table in the bottom right corner shows the cash flow numbers. Both funds have a total value to paid-in capital multiple of 2.5, and an internal rate of return of 50.2%.



Year	Fund A	Fund B
1	-8	-8
2	5	3
3	5	10
4	5	2
5	5	5

Figure 5

This figure shows the times series of quarterly returns (in %) of the Cambridge Associates private equity indices from the first quarter of 1995 until the third quarter of 2020. See Table 3 for a detailed description of the data. The light grey line represents the VC fund index, the dark blue line is the buyout fund index, and the dashed black line is the private real estate fund index.

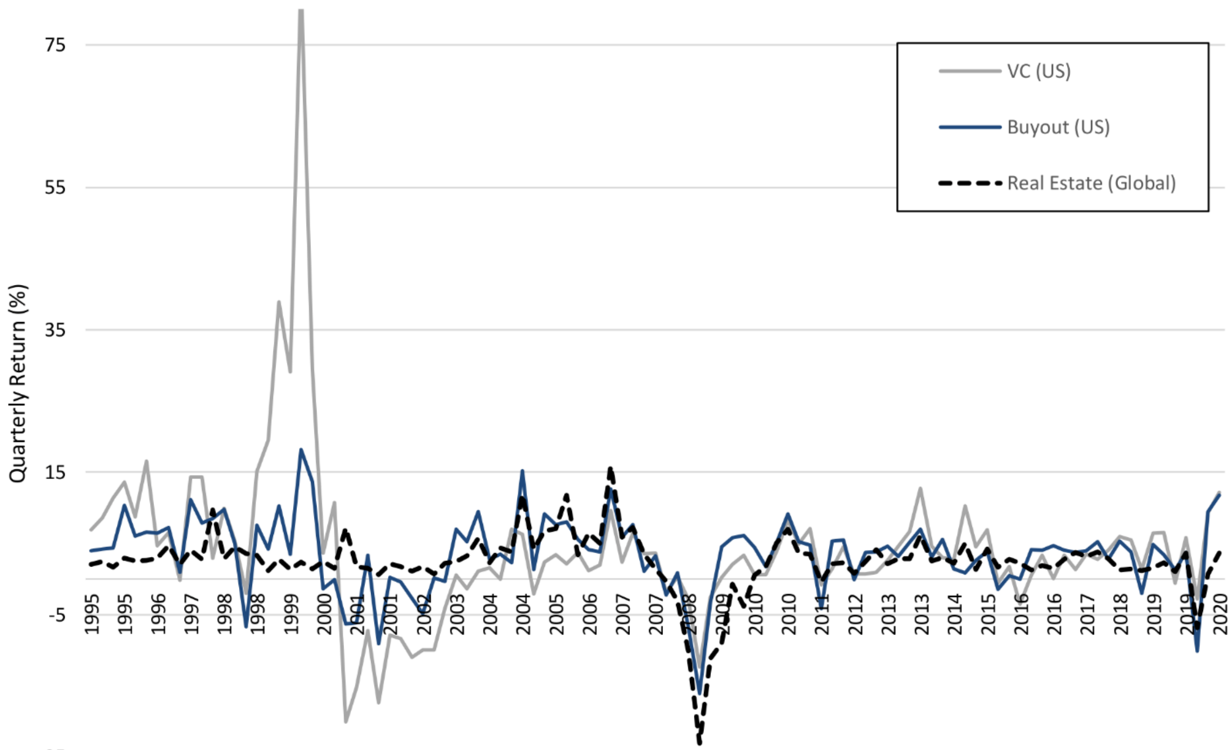
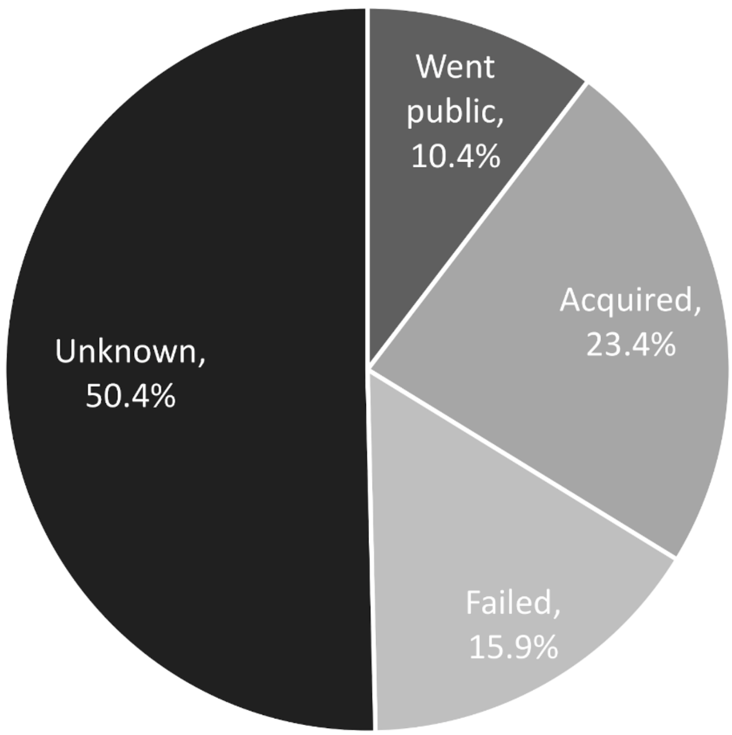


Figure 6

Pie chart of the outcomes of a sample of 18,237 startup companies that were first funded between 1987 and 2005. Source: Korteweg and Sørensen (2010).



Footnotes

[1] I define private equity to include all forms of PE investments, including leveraged buyout, venture capital, private real estate, private debt, infrastructure, and others.
[2] Whereas private markets have grown in importance, publicly traded firms now represent a smaller proportion of the economy than they did in the 1970s (Schlingemann and Stulz, 2022). However, most of the decrease in the prominence of public markets occurred prior to the rapid acceleration in private equity investments that started in the 1990s (see also Espen Eckbo and Lithell, 2021).
[3] The (limited) partnership fund structure for highly risky investments has a long history. See, for example, Korteweg and Sensoy (2023), and references therein.
[4] Vintage year can be defined in various ways. A common definition used by researchers is the year of the fund's first capital call to LPs (capital calls are explained below). Even though fundraising typically takes 6 months to a year from first raise to final close, funds often have multiple closings so that GPs can start calling capital and invest before the final close.
[5] Figure 2 uses the sample of funds in Preqin for which cash flows are available. This is a subsample of the funds included in Table 1 (see Section V of the Internet Appendix of Korteweg and Nagel, 2016, for a detailed comparison of the two samples). A second difference is that the sample for Figure 2 does not use any post-2005 vintages, to allow for enough years of cash flow data to construct the figure, whereas Table 1 uses vintages all the way through 2019.
[6] To limit the risk of GP clawbacks not getting repaid, the LPA sometimes arranges for early carry to be paid into an escrow account.
[7] Opportunity funds are funds raised by early-stage VCs to make later-stage investments in successful portfolio companies. These types of funds, as well as continuation funds (which purchase portfolio company investments from GPs' prior funds), have become more popular in recent years as more startup firms stay private longer, increasing their capital needs and time to liquidity.
[8] Many funds have lines of credit. The amounts borrowed are traditionally small relative to the size of the fund and are used for transaction cost reasons by allowing the GP to combine multiple small capital calls into one, or to cover for any capital calls that are slow to arrive so that a deal does not fall through. Recently, some GPs have started to use credit lines more aggressively, possibly distorting performance metrics (Schillinger et al., 2019, and Albertus and Denes, 2020). See section 3.7.2 for more discussion on the impact of leverage on performance measurement.
[9] At the deal level, the cash multiple is usually referred to as Multiple on Invested Capital (MOIC). For all intents and purposes, MOIC is the same as TVPI.
[10] For not-yet-liquidated funds, TVPI is sometimes split into Distributed-to-Paid-In Capital (DPI) and Residual-Value-to-Paid-In Capital (RVPI), where the former only uses distributions in the numerator, and the latter only uses the most recently reported NAV. TVPI is the sum of DPI and RVPI. For a fully liquidated fund, RVPI equals zero so that TVPI = DPI.
[11] In the initial few years of a fund, performance measures can be extremely volatile and are often primarily driven by NAVs. To provide a more informative picture, the sample used to construct Figure 3 excludes post-2014 vintages, such that each fund has at least 5 years of performance data (which is collected through the end of 2019). Other than the earlier vintage cutoff, the sample is identical to the sample used in Table 1.
[12] Formally, the CAPM does not have an intercept (α_i), and the model in equation (3) with the excess market return as the sole risk factor is known as the "market model". In empirical work in PE, researchers almost always include an intercept and use the CAPM and market model labels interchangeably. I will follow this practice, at the risk of upsetting theorists.
[13] The Cambridge Associates, Burgiss, and State Street index returns are slightly more sophisticated than equation (6) suggests. CA and State Street compute an IRR for the quarter to account for the timing of cash flows within the quarter. Burgiss uses the Modified Dietz approach, which has the same goal. Differences in the timing of constituent fund NAV reports within the quarter could similarly be accounted for, although published index descriptions are not clear in whether this is in fact done.
[14] https://www.cambridgeassociates.com/private-investment-benchmarks/, accessed March 18, 2021.

[15] See, for example, the "Valuation of Portfolio Company Investments of Venture Capital and Private Equity Funds and Other Investment Companies - Accounting and Valuation Guide" published by the American Institute of Certified Public Accountants (AICPA). To underscore that valuation is difficult even for relatively mature PE investments, Cederburg and Stoughton (2020), Agarwal et al. (2022), and Imbierowicz and Rauch (2021) show that there is large cross-sectional dispersion in valuations of PE investments by mutual funds, even within the same portfolio company and the same security.

[16] In index returns that use aggregated fund NAVs (e.g., equation (6)), there is some additional staleness because not all funds report at end of a calendar quarter, as mentioned in section 3.4. But, this particular source of staleness is relatively short-lived – operating on the span of one quarter – unlike the staleness due to conservative portfolio company valuations considered here.

[17] Korteweg and Westerfield use the equal-means assumption for θ_0. They find that if they use the volatility-calibration method instead, estimates of means and variances change, but higher moments (skewness, kurtosis) remain unchanged, and Sharpe ratios are only marginally affected.

[18] Another imputation technique is proposed by Woodward (2009), who shows how to invert the Dimson (1979) regression to impute valuations. Since this is a by-product of the Dimson approach, I do not separately discuss this method here.

[19] Boyer et al. (2022a) update the results for buyout funds in Boyer et al. (2022b), extending the sample to 2018. They estimate a transactions-based buyout beta of 1.8. The updated results also include Dimson beta estimates of NAV-based returns from Burgiss and Preqin data, as shown in Table 5.

[20] Some long-horizon investors may in fact be attracted to the PE space exactly because of the lack of mark-to-market pricing. As Cochrane (2021, p.8) puts it: "Why do so many institutions, like our endowments, prize assets like private equity, venture capital and real estate with no clear market values? Well, perhaps they like those assets precisely because the assets are hard to mark to market, easy to just pay out 5% of a made-up value, not to sell in a panic, and not fire the asset manager based on an irrelevant price […]." Nevertheless, one still needs good estimates of alphas and betas to estimate the distribution of *long-term* payoffs, as the mark-to-market issue is a transitory one. Moreover, interim values undoubtedly matter for investors who suffer liquidity shocks and need to get out of their position.

[21] Proof of equations (12) and (13): Let ϕ be the number of units of the risky security bought. Since $X_t = W_t - \phi P_t$ and $X_{t+1} = \phi P_t R_{t+1}$, the choice problem (11) can be rewritten as $\max_{\phi} U(W_t - \phi P_t) + \rho E_t\left[U(\phi P_t R_{t+1})\right]$. The first-order condition is $-U'(X_t)P_t + \rho E_t\left[U'(X_{t+1})P_t R_{t+1}\right] = 0$. Divide by $U'(X_t)P_t$ and reorganize to find the solution.

[22] See Sørensen and Jagannathan (2015) for a more in-depth description of this state-price analogy.

[23] Other PME-type definitions have been proposed, such as PME+ and modified PME (see, for example, Gredil et al., 2022, for a description and critique of these alternatives). The Kaplan and Schoar (2005) version is the most commonly used definition, especially in the literature, and the one that is most closely linked to theory (as explained below).

[24] The PME is sometimes computed as a ratio of future (rather than discounted) values. These calculations are equivalent. This can be shown by substituting $R^M_{t_0,\tau} = R^M_{t_0,T-t_0} / R^M_{t_0+\tau,T-\tau}$ into the PME formula of equation (16), yielding the ratio of future values $PME_i = \left[\sum_{d=1}^{D} D_{i,d} \cdot R^M_{t_0+\tau^D(d),T-\tau^D(d)} + NAV_{i,T}\right] / \left[\sum_{k=1}^{K} I_{i,k} \cdot R^M_{t_0+\tau^I(k),T-\tau^I(k)}\right]$.

[25] Other ways of calculating an annualized excess return are the index comparison method by Long and Nickels (1996), PME+, and modified PME. See Gredil et al. (2022) for more details and a critical comparison with direct alpha.

[26] This description is a bit of a misrepresentation: Driessen et al. (2012) do not use the Kaplan and Schoar (2005) PME. Instead, they use an NPV formula similar to equation (19) below and estimate the market model parameters that make this NPV equal to zero. The intuition is the same, however. Another way to think of their approach is that it modifies the fund IRR calculation in equation (2) by substituting IRR_i with a factor model.

[27] Log-utility has a long history in (financial) economics. See, amongst others, Hakansson (1971), Roll (1973), Fama and MacBeth (1974), Kraus and Litzenberger (1975), Rubinstein (1976), and Long (1990).

[28] Using the public stock market as a proxy for the wealth portfolio is common in empirical applications but invites the well-known Roll (1977) critique typically applied to tests of the CAPM, which says that the model may be rejected not because log-utility is the wrong assumption but because the proxy is poor. It is especially noteworthy that the proxy portfolio does not include any private equity investments. In the benchmarking interpretation, on the other hand, using the public stock market return is perfectly fine if that is indeed the (risk-matched) benchmark portfolio that we want to compare PE investments against.

[29] Note the deceptive similarity between equation (19) and the IRR calculation in equation (2). The GPME would indeed be zero if M is equal to 1/(1+IRR), but it is important to realize that this is not a proper SDF because it varies from fund to fund.

[30] To be precise, Korteweg and Nagel estimate the SDF on portfolios of public equities and Treasuries that match investments made in the private equity funds in their sample. While this approach should be asymptotically equivalent to simply weighing each time period equally, in samples with short time series it helps to concentrate weight in periods where the exposure to PE is highest and it is therefore most important to price factor realizations correctly. Also, per the results in Shanken (1992), it is better to estimate the SDF over the same period as the PE investments, rather than over a longer historical sample period.

[31] If the PME is implemented using net cash flows, as in equation (19), remember that in this Korteweg-Nagel type PME, zero is the counterpart to the PME = 1 threshold in the Kaplan-Schoar PME.

[32] A more accurate modeling of a levered strategy shows that the alpha is in fact magnified by leverage: suppose the investor borrows *x* dollars (where *x* could be negative) on top of their own $1 and invests the full $(1+*x*) in an asset that pays the market rate of return plus an additional alpha. Assuming that the investor can borrow at the risk-free rate, then $E[GPME] = -1 + E\left[M \cdot \left\{(1+x)\left(R^M + \alpha\right) - xR^F\right\}\right] = (1+x)\frac{\alpha}{R^F}$. Korteweg and Nagel (2016) show this "alpha magnification" in an exercise with artificially levered VC funds. They also show that PME does not properly account for leverage (see their Table III).

[33] A tangentially related paper by Fleckenstein and Longstaff (2020) uses secondary market data of securitized business credit card targeted to entrepreneurs. Calibrating a Merton-style model to this data they find an expected return on small entrepreneurial companies of 14% and a market beta of around 1.2.

[34] Metrick and Yasuda (2010a) and Stafford (2022) simulate the effect of fees on fund performance, but to my knowledge, no paper has analyzed the effect of a specific fund's (actual) fees on its performance and compared this with the gross and net-of-fee estimates in the literature. Robinson and Sensoy (2013) and Hüther et al. (2020 come closest. The former have fund-specific fee data and net-of-fee fund performance, but they do not have gross-of-fee data (without portfolio company level data, reverse-engineering the gross-of-fee fund returns is not possible in their data). The latter have both gross and net-of-fee data, but only for a small subset of funds, and their analysis focuses on the investment level, which is discussed in section 6.

[35] A more technical explanation is that carried interest can be thought of as a call option on the portfolio of investments that LPs have written to GPs. The factor loading of carried interest is the option's elasticity times the corresponding factor loading of the underlying portfolio of investments. Since the elasticity of a call option is above one, the beta of carried interest is higher than that of the underlying portfolio. As a result, the net-of-fee beta should be lower than the gross-of-fee beta. The strength of the impact of carried interest on betas may depend on the timing of carried interest payments (see Hüther et al., 2020). See also Metrick and Yasuda (2010a) and Ang et al. (2018) for a discussion of carried interest as an option.

[36] Additionally, since there is some flexibility on the part of the GP regarding the timing of capital calls to pay management fees, this could lead to some degree of systematic risk in these fees. In practice, this risk is likely to be very low.

[37] I refer the reader back to section 3.6 for a more detailed discussion of post-money versus fundamental valuations.

[38] The modified IRR allows for the specification of a reinvestment rate, making it less sensitive to the reinvestment assumption implied by the traditional IRR.

[39] Footnote 18 of Korteweg and Nagel (2016) describes a more elaborate example of the importance of computing post- to pre-money returns. Note also that equation (25) is consistent with a log-version of the fund NAV return calculation in equation (4) if NAVs are interpreted as post-money valuations (portfolio company values are usually equated to post-money valuations when financing rounds occur). With no dividends (i.e., no distributions), the

return in equation (4) equals the pre-money valuation (post-money valuation minus invested amount, i.e., capital calls) at time *t* divided by the post-money valuation at *t-1*.

[40] There is more than one way to set up the RSR. The more traditional approach is to set the indicators to -1 in the period of the first observed valuation, and to +1 at the second observed valuation date, with zeros in between (and excluding the very first period, as above). The left-hand side is the same as in equation (27), and the regressor matrix is [1 0 0; -1 1 0; 0 0 1]. The estimated parameters are now the log price index (not the log-return on the price index), but the end result is identical. This formulation is slightly more efficient when there are many startups, due to the multiplication and inversion of sparser matrices, but it is one step removed from returns. For clarity of the exposition, which focuses on returns, I have chosen the alternative setup shown in the main text.

[41] Other than zombie firms, the set of firms with unknown outcomes includes firms that have only recently started and haven't had enough time to reach an exit (a right-censoring issue), and "lifestyle" companies that are still in business but haven't had to raise funding in years. The latter are not likely to produce much value to the VC investors, even if they provide a means of income to the founders.

[42] Success bias is also a potential concern for the cash flow-based measures discussed in section 6.1, if failures are nonrandomly missing. Unlike commercial data sets, proprietary investment-level data often include the entire universe of deals that an LP or GP invested in, with final outcomes, and are therefore not subject to the same success bias. However, since these data are often sourced from one or a few LPs or GPs, selection bias can still be an issue if these investors are not representative of the population. Fund-level performance measures are also not usually subject to success bias at the fund level, since all investments (both successes and failures) will be accounted for by the time funds are liquidated. But, there may be selection bias in which funds are included in the data set.

[43] The existence of sample selection bias in RSR was first suggested in real estate by Haurin and Hendershott (1991). Wallace and Meese (1997) show empirical evidence. A number of papers in real estate apply standard selection models to the problem (e.g., Jud and Seaks, 1994, Gatzlaff and Haurin, 1997, 1998, Munneke and Slade, 2000, 2001, Hwang and Quigley, 2004, and Goetzmann and Peng , 2006).

[44] Hall and Woodward (2010) and Ouyang et al. (2021) appear to use a similar imputation method as in Woodward and Hall (2004), and provide more details on the algorithm: based on a subset of data without missing financing round data, they fit a regression on observed round data (primarily the round number, amount raised) and other covariates (e.g., Hall and Woodward include the increase in the Wilshire 2000 index in the preceding two years). They then use the fitted model to impute the missing round data in the full sample. Neither paper applies RSR: Hall and Woodward (2010) are interested in the returns to entrepreneurs, and Ouyang et al. compute GPME performance statistics.

[45] Hwang et al. (2005) also explore a hybrid model that combines RSR with a hedonic model (hybrid models in real estate have been proposed, amongst others, by Case and Quigley, 1991). A hedonic price index can help to mitigate sample selection bias in real estate indices by allowing for the inclusion of single-sale homes, which may be systematically different from homes that have sold multiple times. In hedonic models, values are imputed using estimated shadow prices on a range of observed characteristics. A pure hedonic model is not particularly useful in PE, as very few portfolio company characteristics are usually observed, and most of their value is in difficult-to-observe and unique intellectual property and growth options.

[46] See https://www.refinitiv.com/content/dam/marketing/en_us/documents/methodology/venture-capital-research-index-methodology.pdf for a description of the index construction methodology.

[47] While not uncommon, setting values for deemed-liquidated companies to zero at some exogenously-chosen time after the last observed round is problematic, as the exact timing of exit can matter quite a bit for estimates of risk and return.

[48] Since the RSR value process is also specified in logs, the index calculation also needs an adjustment to compute arithmetic returns (see Goetzmann, 1992).