Working Paper 20-073

How Do Private Equity Fees Vary Across Public Pensions?

Juliane Begenau
Emil Siriwardane

Harvard Business School

How Do Private Equity Fees Vary Across Public Pensions?

Juliane Begenau
Stanford University

Emil Siriwardane
Harvard Business School

Working Paper 20-073

Copyright © 2020, 2021, and 2022 by Juliane Begenau and Emil Siriwardane.

Working papers are in draft form. This working paper is distributed for purposes of comment and discussion only. It may not be reproduced without permission of the copyright holder. Copies of working papers are available from the author.

Funding for this research was provided in part by Harvard Business School.

How do private equity fees vary across public pensions?*

Juliane Begenau† Emil Siriwardane‡

March 2022

Abstract

We study how investment fees vary within private-capital funds. Net-of-fee return clustering suggests that most funds have two tiers of fees, and we decompose differences across tiers into both management and performance-based fees. Managers of venture capital funds and those in high demand are less likely to use multiple fee schedules. Some investors consistently pay lower fees relative to others within their funds. Investor size, experience, and past performance explain some but not all of this effect, suggesting that unobserved traits like negotiation skill or bargaining power materially impact the fees that investors pay to access private markets.

*Special thanks go to our discussants Arthur Korteweg, Yael Hochberg, and Stijn Van Niewerburgh. We are grateful to Jonathan Berk, Ted Berk, John Campbell, Josh Coval, Peter DeMarzo, Mark Egan, Steve Grenadier, Debbie Gregory, Paul Gompers, Daniel Green, Robin Greenwood, Ben Hebert, Victoria Ivashina, Josh Lerner, Hanno Lustig, Chris Malloy, Ludovic Phalippou, Monika Piazzesi, Josh Rauh, Claudia Robles-Garcia, Amit Seru, Andrei Shleifer, Erik Stafford, Ilya Strebulaev, Per Stromberg, Adi Sunderam, Boris Vallee, and Luis Viceira, as well as seminar participants at Harvard, the NBER CF SI 2020, the NBER LTAM 2021, the Red Rock Conference, the Virtual Corporate Finance Conference, the University of Washington, the University of Southern California, Stanford, and the Swedish House of Finance for their comments and discussions. Sid Beaumaster, Ruying Gao, Francesca Guiso, Rick Kaser, Elizaveta Savinova, and Yuan Tian provided excellent research assistance. We are also indebted to Maeve McHugh and Michael Smyth at Preqin for their help compiling the dataset for our analysis. First draft is from November 2019.

†Begenau: Stanford GSB and NBER and CEPR. E-mail: begenau@stanford.edu

‡Siriwardane: Harvard Business School and NBER. E-mail: esiriwardane@hbs.edu

1 Introduction

Investment in private markets through vehicles like private equity and venture capital has boomed over the past two decades and now exceeds $7.3 trillion (McKinsey, 2021). Much of this growth has been fueled by an influx of capital from defined-benefit public pension plans across the globe (Ivashina and Lerner, 2018). As more public money has flooded into these asset classes, so too have calls for increased transparency around the structure and operations of private market funds (SEC, 2014). Proponents of this view argue that the limited partnership agreements (LPAs) governing fees and investment terms in private markets are complex and opaque. Further complicating matters, LPAs are often modified by so-called side letter agreements that are privately negotiated between individual investors (LPs) and fund managers (GPs). This environment makes it difficult to answer basic questions about costs in private markets. For instance, are fees set uniformly within most funds, and if not, by how much do they vary? What factors determine variation in fee policies across GPs and funds? Which LPs pay lower fees? These are the types of questions that we seek to answer in this paper.

The main challenge we face is that LPAs and associated side-letters are almost never observed by fund outsiders.[1] We circumvent this issue by instead studying the *net-of-fee* returns of different investors in the same private market fund. Intuitively, if investors pay different fees in a fund, they will earn different net-of-fee returns. This observation provides us with an opportunity to characterize fee policies without observing the full contract between LPs and GPs. Of course, in practice, there are other factors that could induce within-fund variation in net-of-fee returns (e.g., measurement error), and we are careful to account for these alternatives throughout our analysis. An added advantage of our approach is that it translates any fee differences into terms that ultimately matter for investors, namely returns. This is especially useful because LPs cannot easily forecast the fees they will pay in private market funds due to the complexity of LPAs and side letters (State Comptroller SEC letter, 2015).

[1]The LPA is visible and agreed upon by all investors. However, side letters supersede the LPA and can grant some investors more favorable terms than others (Morgan Lewis, 2015; Toll and Centopani, 2017). The visibility of side letters depends on a number of factors (e.g., most-favored nation status). We discuss these details in Section 2.1.2.

Figure 1 illustrates our empirical strategy by plotting the cumulative net-of-fee return earned by two investors in the same fund. Returns at each point in time equal cumulative distributions received per dollar of investment (DVPI). In our data, DVPI is measured on an after-fee basis because contributions into the fund include all fees (e.g., management fees) and distributions by the fund are net of all fees (e.g, performance fees or carry). Capital for this particular fund was obtained in a single fundraising round, which means that both investors entered the fund at the same time. The two investors earn identical returns for the first ten years of the fund's life, but a wedge emerges thereafter. By the end of the fund's life, the orange investor has earned \$1.62 per dollar invested compared to \$1.54 for the blue investor. We exploit this type of within-fund variation in net-of-fee returns to assemble four key sets of results about private market fund fees. Our analysis uses a novel panel dataset of cashflows at the LP-fund level that covers \$438 billion of investments made by 218 U.S. public pensions in 2,400 private market funds.

First, we show that investors' net-of-fee returns in the typical fund are clustered together, as opposed to being continuously distributed across investors (see Figure 3 for an example). Using unsupervised machine learning techniques (Jain et al., 1999; Rousseeuw, 1987), we document that 36% of funds in our sample have one DVPI cluster or tier, 61% have two tiers, and the remainder have three or more tiers. These decompositions are also robust to unsupervised learning techniques that are specifically designed to separate true clusters from classical measurement error (Tibshirani et al., 2001), thus alleviating concerns that we spuriously detect clusters due to noise in our return data. We interpret the overall patterns of return clustering as evidence that investors in most funds are grouped into one of two tiers of fees.

Next, we estimate how much fees vary across investor tiers in the average fund. Fees in private markets can be decomposed into several categories, such as management fees, performance-based fees (carry), organizational fees, and portfolio company fees (Phalippou et al., 2018). The panel nature of our data allows us to estimate variation in the first two categories. Intuitively, within a fund, the standard deviation (or dispersion) of management fees should grow linearly with fund age and dispersion in carry should grow linearly with performance. Building on this logic, we estimate that the average within-fund dispersion in management and carry are 91 basis points and 5.8%, respectively.[2] For comparison, the levels of management fees and carry in private equity are generally thought to be around 200 basis points and 20%, respectively (Gompers and Lerner, 1999). These estimates also differ widely across asset classes. The dispersion in management fees within the average private equity (PE) fund is 83 basis points, but is only 42 basis points in the average venture capital (VC) fund. In terms of carry, we estimate that dispersion in the average PE fund equals 3.3% and is 0.5% for VC, the latter of which is not statistically different from zero. Our estimates further indicate that infrastructure, private debt, and real estate funds generally have more within-fund dispersion in both types of fees compared to PE funds. We also develop a placebo test by exploiting the fact that performance fees are contingent on a minimum level of fund performance. Accordingly, in all asset classes, we confirm that dispersion in carry is not detectable in unprofitable funds.

Our third set of results characterizes the types of funds and fund managers that use multiple fee structures. We start by showing that GPs tend to use a consistent fee policy across all of their funds. Moreover, the identity of the law firm employed by each fund is a key determinant of why some use multiple fee structures, a finding that likely reflects the importance of law firms in assembling LPAs and side letters. The use of multiple fee structures also varies strongly across asset class. Notably, VC funds are far less likely to tier investors than other asset classes – 48 percentage points (pp) less likely than infrastructure and 33 pp less likely than private equity. This evidence accords with the estimates of within-fund dispersion in management fees and carry that we discuss above. While these results point to a somewhat static component of fee policy that depends on GP identity or asset class, there is also a dynamic component that evolves with the fundraising environment. In particular, we document a robust negative relationship between investor demand and the use of multiple fee schedules. We proxy for demand using past GP performance (Berk and Green, 2004), GP experience, and fund subscription status. For example, we find that funds managed by GPs without an established track record – defined as having raised less than three funds – are 13

[2]These estimates are in line with the menu-model that Bain Capital has offered to its investors in recent years (Zuckerman and Or, 2011; Markham, 2017).

pp more likely to tier investors. Similarly, undersubscribed funds are 12 pp more likely to use multiple fee structures.[3] These correlations support the notion that GPs who face low demand are more likely to offer some LPs fee breaks, perhaps as a way to attract more capital commitments (e.g., via signaling effects). Lastly, we test whether the use of placement agents explains variation in fee policy. This test is motivated by the fact that some pensions have restrictions on the use of placement agents when investing in private markets. If these pensions seek exceptions from paying placement agent fees, then the funds in which they invest should be more likely to exhibit fee dispersion. Consistent with this hypothesis, funds that use placement agents are 13 pp more likely to use multiple fee schedules.

Finally, we document that some public pensions tend to pay relatively lower fees across all of their funds. Within a given fund, we categorize an investor as being top-tier in terms of fees if it earns above-median net-of-fee returns for the majority of the fund's life. An F-test from a fixed-effects regression comfortably rejects the null hypothesis that top-tier fee status is randomly assigned to investors in each fund. The rejection of random tier assignment is driven by the wide observed distribution of "pension effects": the 95th percentile pension outperforms in 67% of its funds while the 5th percentile pension outperforms in only 13% of its funds. This finding supports the idea that some investors consistently select or are offered the best fee structure in their respective funds, at least in terms of ex-post performance. We provide further evidence that part of these pension effects are driven by selective matching between LPs and GPs (e.g., relationships).

There are several possible reasons why some pensions could consistently pay lower fees than others when investing in private markets.[4] For instance, fund managers could offer fee reductions to pensions who lower the cost of raising a fund, either by drawing in other investors or by providing larger amounts of capital. Consistent with this intuition, pensions that are large in overall size are 14 to 17 pp more likely to be in the lowest-fee tier for the average fund. Similarly, those that

[3] Undersubscribed funds are those whose fundraising efforts fall short of the initially planned fund raising goal.

[4] Some funds allow LPs to deploy additional capital to specific portfolio companies in a fund at lower or no cost (so-called co-investment rights, Fang et al. (2015)). These types of special-purpose vehicles are a small part of public pensions' portfolios during our sample and are listed as separate entities, which allows us to exclude them entirely from our analysis. See Section 5.4 for a complete discussion.

contribute more capital to a fund are also more likely to be in the lowest fee tier. While this suggests that size is an important determinant of top-tier status (Da Rin and Phalippou, 2017; Clayton, 2020), we show that other pension characteristics that proxy for investor sophistication correlate with tier assignment. Specifically, we find that pensions in low-fee tiers tend to be better governed, more experienced, and have high past performance. Nonetheless, even after controlling for all of these observable characteristics, there are still a subset of pensions who consistently outperform others within their respective funds. We interpret this as evidence that unobservable traits related to negotiation and contracting skill materially impact the fees that investors pay in private market funds.

We make no claims that fees are the sole reason why net-of-fee returns may vary across investors in the same fund. In Section 5, we use several complimentary approaches to investigate how alternatives sources of return variation would impact our main results, focusing on broad four channels: (i) measurement error; (ii) differences in entry timing, either due to multiple fundraising rounds or secondary transactions; (iii) accounting practices that vary across LPs; and (iv) differences in gross-of-fee returns across investors in the same fund (e.g., co-investments). In one approach, we directly file Freedom of Information Act requests to probe whether accounting practices are likely to differ across pensions. In another, we analyze subsamples of the data in which other sources of within-fund return dispersion are unlikely or not possible. For example, we study funds that were raised in a single round, which rules out the possibility that variation in contribution schedules across LPs causes mechanical variation in returns within these funds. In all cases, we find robust support for our overall characterization of fees in private market funds.[5]

We end the paper with two exercises that highlight the aggregate implications of our analysis. In the first, we compute how much better off pensions would be had each been in the top-tier of fees within their respective funds. Our estimates suggest that the aggregate amount of forgone

[5]Much of our analysis is based on regressions in which the outcome variable is an indicator for whether a fund uses multiple fee schedules or whether an investor is in the best fee tier in a fund. These indicators could in principle inherent any measurement error in returns. However, even in this case, the regression coefficients – and hence our characterization of how fee policies vary across funds or investors – will still not be statistically biased. A similar logic holds when considering our other main results and alternative sources of return dispersion (See Section 5).

capital due to fees in our sample equals \$4.3 per \$100 invested or \$19 billion in total. In the second, we explore the extent to which measures of aggregate performance depend on whether fund-level returns are measured using investors in the best or worst tier of fees. In line with our other results, we find that aggregate private equity performance is much more sensitive to this choice than venture capital. We view both exercises as more suggestive in nature because they are sensitive to all potential sources of within-fund return variation, as well as the composition of investors in our sample (e.g., we do not observe university endowments).

The primary contribution of this paper is to document and characterize fee dispersion within private market funds. Fee dispersion is a natural equilibrium outcome given the contracting environment in our setting. In particular, the existence of side letters means that fee determination in a fund can be viewed through the lens of search and bargaining models (Burdett and Judd, 1983; Bester, 1988; Duffie et al., 2005; Allen et al., 2019). These models generally predict that consumers or investors will pay different prices for the same good, either due to simple cost-based pricing, search and negotiation frictions, contract complexity, or heterogeneity in LP sophistication (Salop and Stiglitz, 1977; Gabaix and Laibson, 2006).[6] This prediction has been confirmed in many market settings, including those for health care (Sorensen, 2000; Grennan, 2013), automobiles (Goldberg, 1996), financial securities (Eisfeldt et al., ming), residential mortgages (Allen et al., 2019), and mutual funds (Hortaçsu and Syverson, 2004). Our results extend these empirical studies to the private-capital market. In addition, our estimates of within-fund dispersion in management fees are comparable to Hortaçsu and Syverson (2004), who show that management fees for S&P 500 index funds range from 10 to 268 basis points. Private market funds are far more complex and opaque than S&P 500 index products, so if anything, one would expect higher levels of dispersion in our setting (Gabaix and Laibson, 2006).

This study also contributes to prior research on the private equity industry. Much of the research on fees in private investment vehicles has focused on the level of fees (Gompers and Lerner, 2010;

[6]It may seem puzzling that there are only a small number of fee schedules in practice, since LPs in the same fund could theoretically have unique schedules via side letters. However, from the perspective of GPs, negotiating and implementing several fee structures is costly (Morgan Lewis, 2015), thereby limiting the number of contracts in equilibrium. These costs are one of the motivations for standardizing LPAs (ILPA, 2019).

Metrick and Yasuda, 2010), the type of fees (Phalippou et al., 2018), or on across-fund variation in fees (Robinson and Sensoy, 2013). While others have alluded to fee differences within funds (Da Rin and Phalippou, 2017; Toll and Centopani, 2017; Clayton, 2020), we believe our paper is the first to systematically study this phenomenon in a large sample of funds. Our results reveal important features of the fund-formation process and suggest that GPs vary considerably in how they set investment terms with their LPs. The notion of fee tiers within funds is also consistent with recent studies showing that GPs differentiate among investors through co-investments and other special purpose vehicles (Lerner, Mao, Schoar, and Zhang, 2018; Fang, Ivashina, and Lerner, 2015; Braun, Jenkinson, and Schemmerl, 2019).[7] Additionally, our finding that some pensions consistently receive better terms aligns with the findings of Lerner et al. (2018), who show that GPs often offer these bespoke investments to only a preferred set of investors.

The remainder of the paper is organized as follows. In Section 2, we discuss the structure and contracting environment for funds, as well as our main sources of data. Section 3 documents the existence of return clusters in most funds, estimates the size of fee dispersion across investor tiers, and characterizes the types of funds that use multiple fee schedules. Section 4 analyzes whether and why some LPs are consistently in the top tier of fees in their respective funds. In Section 5, we probe the robustness of our findings to other sources of return variation within funds. Section 6 explores some aggregate implications of fee dispersion and concludes. Additional details and results are available in an online appendix.

2 Institutional Background and Data

We begin with background on the structure of private market vehicles, the contracting environment between LPs and GPs, and the legal mechanisms through which LPs in the same fund may pay different fees. We also provide details on our primary analysis sample.

[7] GPs may prefer to differentiate among LPs using fee structures instead of co-investments if some LPs cannot easily co-invest. This is likely the case for many U.S. public pensions (see Section 5.4).

2.1 Private Market Investment Vehicles

2.1.1 Basic Fund Structure and Lifecycle

The focus of this paper is public pension investments into private market vehicles, namely private equity (PE). A typical PE fund has two types of investors, the general partner (GP) and the limited partners (LPs). The GP manages the fund and usually contributes about 1-5% of its own capital to the fund. The bulk of the fund's capital therefore comes from LPs, who are entities like pensions, endowments, and family offices. At the beginning of a fund's lifecycle, GPs secure capital commitments from LPs, after which capital is formally "called" from the LPs. Some of this called capital is invested by the GP, while the rest is used to pay management fees and other fund expenses that we discuss below. Investments are held for several years before they are liquidated. Conditional on fund returns exceeding a predetermined hurdle rate, the GP then withholds a portion of the investment proceeds as a performance fee (or "carry") before issuing distributions back to the LPs. In most cases, each LP's pre-carry claim on distributions is proportional to its capital commitment, meaning gross-of-fee distributions (per dollar of commitment) should be equal across LPs. From start to finish, funds typically have a total lifespan of ten to fifteen years.

2.1.2 Contracting Environment and Types of Fees

Investment into private market vehicles is governed by a private contract, the limited partnership agreement (LPA), between the GP and LPs. Generally speaking, the GP and the LPs privately negotiate the terms of the LPA, including the expenses borne by LPs, tax treatment of fund income, the ability of the GP to unilaterally amend the LPA, and the degree to which the GP is indemnified through the partnership.

The LPA dictates and governs four broad types of expenses that ultimately determine the returns of LPs: (i) management fees, which are typically a percentage of committed capital; (ii) performance-contingent fees or carry; (iii) fund and organizational expenses; and (iv) portfolio company fees. Portfolio company fees are paid to the GP by the firm in which the partnership invests and in many cases the LPA requires this income to be shared with LPs though fee offsets or rebates (Phalippou et al., 2018). We provide more detail on the nature of all of these expenses in Internet Appendix B.5.

LPAs can also be used to create multiple investor classes. For instance, tax-exempt investors like pensions can opt to be separated from taxable investors in order to minimize tax burdens for both groups.[8] Funds may also allow LPs to choose from a menu of fees, with each choice representing a different investor class. Bain Capital is one notable example of a GP who recently shifted to a menu-model, offering investors a choice to pay 1% management fee and 30% carry or 2% management fee and 20% carry (Zuckerman and Or, 2011; Markham, 2017).

Though the LPA is visible and agreed upon by all LPs in the fund, its terms are often superseded by additional agreements (so-called "side letters") that are negotiated bilaterally between the GP and individual LPs.[9] Side letters can alter many aspects of the original LPA, such as reporting requirements by the GP, explicit modifications of fees, or exemptions from paying certain fund expenses (e.g., placement agent fees). They can also establish provisions for "most favored nation" (MFN) status, which under certain conditions allow an LP to view and select the terms of side letters that have been offered to other LPs.[10]

The use of side letters to modify LPAs is widespread. Under confidentiality agreements, we obtained a subset of LPAs for 91 funds in our sample. Within this subset, 75% indicated that the GP has the sole discretion to enter into a side letter with any LP. Moreover, the LPAs explicitly stated that any such side letter supersedes the LPA and may confer rights and benefits to an LP that are not granted to others. LPAs also varied widely in terms of both side letter transparency and access. 54% had no language requiring GPs to treat LPs equally in terms of access to side letter provisions, and in the majority of these cases, GPs were not even required to notify LPs when they

[8] LPAs for tax-exempt investors can allow capital to flow through blocker corporations that improve tax efficiency. According to several large LPs and our read of LPAs, public investors may have to opt in or negotiate for these types of tax optimization services because they are not always treated as tax-exempt by default.

[9] In a survey of global LPs, Da Rin and Phalippou (2017) report that 59% of LPs "always negotiate contract terms".

[10] The nature of MFNs and carve outs that apply to them vary across LPAs. Some MFNs will automatically confer the benefits of all other side letters. Others give LPs the ability to opt into side letter provisions granted to LPs of a similar size and within a fixed window (e.g., one month after close). See Toll and Centopani 2017, Chart 2.32.

entered to a side letter. 32% indicated that side letter provisions would be offered to LPs who were deemed similar by the GP, though the definition of similar was generally vague. Only 13% of LPAs required all LPs to be notified of side letters and given the right to opt into their terms, provided that elections were made within 20-60 days of the fund's close.

In principle, side letters or similar LPA amendments are therefore one mechanism through which economic terms could differ across investors in the same fund. Ideally, we could directly explore this channel if we observed the side letters associated with each fund in our data, though this is difficult in practice due to the private nature of these contracts. Nonetheless, industry surveys of GPs indicate that nearly 50% use side letters to offer some investors more favorable terms (Toll and Centopani, 2017, Chart 2.31). This is likely a lower bound because such surveys are self-reported and GPs have little incentive to reveal their pricing strategy. We provide further background on how the contracting environment (e.g., LPAs, side letters, MFNs) could give rise to fee dispersion in Internet Appendix B.

2.2 Data sources and sample definition

2.2.1 Data Sources

We obtain investment performance data from Preqin, a data provider that specializes in alternative assets markets. Preqin's data on private market investments is sourced primarily from Freedom of Information Acts (FOIA) requests of public pensions and legally-required annual reports.[11] The Preqin data covers funds from vintage year 1990 onward and contains cash-flow data on LP-level investment into individual funds. We specifically observe the amount of committed capital by the investor in the fund, the amount of capital that has been "called" from the investor (i.e., actual contribution amounts), and the amount of capital that has been distributed back to the investor by the fund. These variables are all reported in cumulative terms. Importantly, distributions are

[11]To encourage the same reporting standards across investors and funds, Preqin provides detailed guidelines on submitting performance data in their FOIA requests. After data is submitted to Preqin, the information is reviewed internally and, when possible, is cross-referenced against as many different sources as possible. Further details on Preqin's collection process can be found in Preqin's Private Capital Performance Data Guide. https://docs.preqin.com/reports/Preqin-Private-Capital-Performance-Data-Guide.pdf

reported net of performance-based fees that are withheld by the GP (e.g., carry), and contributions are inclusive of fixed fees such as management fees that are calculated as a percentage of an LP's committed capital. This means that investment multiples are net of fees. We also observe the net asset value (NAV) of each investor's current investments in the fund. For a given investor in a fund, the NAV reflects the market value of investments that have not yet been liquidated, net of any performance fees earned by the GP.

2.2.2 Sample Definition

The sample and variables that we use are taken directly from Begenau et al. (2020). In that companion paper, we discuss a variety of quality control filters that we apply to the raw Preqin data in order to ensure that the resulting cash flow variables are comparable across investors in the same fund.[12] To keep the current paper self-contained, we summarize the main features of our approach below.

The raw data file by Preqin has roughly 750,000 observations and is unique at the level of data source, LP, fund, and date. To be included into our sample, we require a complete set of non-missing identifiers in terms of investor, fund, fund manager, date and fund vintage, as well as non-missing information regarding an LP's contribution, distribution, commitment size, and fund net-asset-value. In addition, we require cash flows to be denominated in USD and focus on LPs who are U.S. public pension funds.[13] This choice eliminates an potential issues that currency conversion may have on our analysis of within-fund returns. There are 376,394 observations that remain after applying these filters and deleting duplicates.

In addition to these basic sample filters, we drop any source-investor-fund cell containing a negative contribution or distribution that is too large to plausibly reflect a fee offset.[14] These cases

[12] The latest version of the companion paper can be found here. To keep our analysis as transparent as possible, we have also posted the code we use to clean the Preqin data here.

[13] The vast majority of investors in our data are U.S. public pension funds (83%) and UK public pension funds (7%). Other investor types in our dataset include public university endowments, government agencies, insurance companies, foundations, and private sector pensions.

[14] Overall, negative contributions and distributions occur in 4% and 1.5% observations, respectively. As is standard in the literature, we retain most of these observations, dropping only those that appear implausibly large to be a fee offset (see Begenau et al., 2020). Fee offsets may reflect, among other things, monitoring fees that are passed from

are incredibly rare and only affect 0.25% of observations. To be conservative, we also drop any LP from the sample if more than 2.7% of their observations have any potential quality issue (e.g., implausibly large negative distribution), which only affects 15 LPs and 0.86% of total commitments in the sample. We discuss the choice of these cutoffs in detail in Begenau et al. (2020).

After applying these additional filters, we retain fund-quarter cells in which there are at least two LPs reporting cash flows, since our focus is on within-fund variation in returns. We also drop all funds that are related to multi-strategy investment (only 2 funds), co-investment, or secondary sales.[15] This leaves us with $233,526$ observations that are unique at the investor-fund-quarter level (p,f,t). We refer to this data set as the master sample.

For some of our subsequent analysis, we condense the data to the investor-fund (p,f) level. Specifically, for each fund f, we denote the set of available dates as $t = 1,...,T$ and $N(t)$ as the number of investors observed at time t. For each fund, we then select the date s such that $s = \underset{t \geq T-20}{\operatorname{argmax}} \; N(t)$. In words, we focus on dates that are within five years of the last observed date for each fund. Within this set, we then select the date that contains the most number of investors in the fund. This approach allows enough time for any fee differences in the fund to appear in cash flows (e.g., carry), while still retaining as many investors as possible. We refer to this condensed data as the core sample.

2.2.3 Summary Statistics

Table 1 describes the core sample. Panel A shows that we have 9,830 investor-fund (p,f) level observations, covering 218 unique pension funds (LPs), 856 unique fund managers (GPs), and 2,400 funds. By asset class, 4,465 of observations are investments in Private Equity funds, 1,956 are in Venture Capital, 1,888 are in Real Estate, 1,215 are in Private Debt, and 306 are in Infrastructure. In total, the core sample covers \$543 billion of commitments and \$438 billion of contributions.

Panel B of Table 1 presents summary statistics of the core sample. The average fund is observed

portfolio companies back to LPs. See Internet Appendix B.5 for more details on the types of fund income that can lead to fee offsets.

[15] We identify co-investment funds based on their category type in Preqin and if their listed name includes “Co-”. We identify secondary transactions in a similar manner.

is 9 years since its final close and contains 4 investors. The average commitment size is $55 million or 5% of total fund size. Pensions in our sample vary widely in overall size, with the average pension managing $24 billion in total assets and the largest managing $354 billion.

2.2.4 Definition of Returns

We measure returns using standard industry return multiples. Specifically, we define the realized cash multiple for investor p in fund f at time t as:

$$r_{pft} \equiv \frac{\text{Cumulative Distribution}_{pft}}{\text{Cumulative Contribution}_{pft}}. \tag{1}$$

We refer to this as a realized multiple because it only reflects distributions that have been paid by the fund to LPs. In practice, it is commonly referred to as the distributed value to paid-in capital ratio or DVPI. Similarly, we define the total multiple on invested capital as:

$$r^T_{pft} \equiv \frac{\text{NAV}_{pft} + \text{Cumulative Distribution}_{pft}}{\text{Cumulative Contribution}_{pft}}. \tag{2}$$

Compared to DVPI, this measure reflects both remaining net asset value (unrealized value) and realized distributions. It is commonly referred to as the total value to paid-in capital ratio or TVPI.

LPs also report internal rates of return (IRRs) to Preqin. However, we primarily analyze return multiples and instead use IRRs only for robustness tests. The main reason is that reported IRRs are missing for 20% of the observations in our data. We have also computed IRRs ourselves based on observed cash flows, though in most cases these cannot be used to measure within-fund return variation. This is because most funds do not have a fully balanced panel across investors and IRRs are very sensitive to the timing of cash flows. The unbalanced nature of our panel does not preclude us from comparing DVPI or TVPI across investors in the same fund for the same quarter, since contributions and distributions are reported on a cumulative basis and, unlike IRRs, these return measures do not depend on the timing of individual cash flows as long as the accumulated amount is correct.

3 Characterizing Fee Dispersion

We begin this section by documenting that net-of-fee returns vary across investors in the typical fund. This variation is driven by the presence of return clusters or tiers, which we interpret as evidence that GPs offer LPs a limited number of fee structures. We then propose a method to estimate the size of fee differences across investor tiers, one that exploits both the cross-sectional and time-series variation in LP contributions and distributions in the same fund. Our approach allows us to differentiate between fees and expenses that are contingent on performance (e.g., carry) and those that scale with commitments and time (e.g., management fees). We also characterize the funds and GPs that are more likely to employ multiple fee structures.

It is important to note upfront that fees are one of several possible sources of within-fund return dispersion. We investigate these alternative channels in Section 5 and argue that they should not materially impact the analysis in this section.

3.1 Within-Fund Return Dispersion and Clustering

3.1.1 Baseline Evidence

Figure 2 shows that within-fund variation in returns is a ubiquitous feature of the data. To construct the graph, we first compute the standard deviation (or dispersion) σ_f of returns within each fund f in the core sample. Positive values of σ_f indicate that returns within fund f differ across investors. We then plot the distribution of σ_f across funds, broken out by fund vintage. Each panel of the plot shows the across-fund distribution of σ_f for different return measures. Regardless of fund age or how returns are measured, there is clear return variation within many funds.

In the subset of funds that exhibit return dispersion, returns tend to cluster together, as opposed to being continuously distributed (e.g., as if drawn from a uniform distribution). Figure 3 illustrates this type of clustering for an anonymous fund in our sample. The plot simply shows the distribution of DVPI across investors in the fund on a single date. We focus on DVPI since it is based solely on realized cash flows. This particular fund has 16 investors for whom we have a full panel of returns.

Moreover, it was raised in a single round, which guarantees that all investors entered the fund at the same time. The distribution of DVPI in the fund exhibits two distinct clusters, with a majority of investors earning \$1.14 per dollar invested and a smaller subset earning \$1.195.

To assess whether this example generalizes, we measure the degree of return clustering (or tiering) in each fund using unsupervised machine learning techniques, namely k-means clustering (Steinhaus, 1956; MacQueen, 1967; Jain et al., 1999). The basic idea of k-means clustering is to assign observations in a dataset to one of k cluster centers. The cluster centers are not known ex-ante and are chosen to minimize the squared distance of each observation to its nearest cluster center. There are a variety of approaches used by the machine learning literature to select the optimal number of clusters k. One popular method is to select the k-means model with the highest Silhouette score $\bar{s}_k$ (Rousseeuw, 1987). Formally, suppose one fits a k-means model to the data. For each observation i, define $a(i)$ as the average Euclidean distance to all other points in the same cluster and $b(i)$ as the average distance to all points in the nearest (or neighboring) cluster. The Silhouette value for a individual observation i is then defined as:

$$s_k(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}}$$

$s_k(i)$ is by construction bounded between -1 and 1, with values near 1 indicating that observation i is well-matched to its cluster (Rousseeuw, 1987). The overall Silhouette score $\bar{s}_k$ for a given k-means model is the average $s_k(i)$ over all observations.

For each fund f, we determine the number of DVPI clusters based on Silhouette scores for each date t, denoted by $Tiers_{ft}$. We round DVPIs to two decimals before estimating $Tiers_{ft}$. The number of clusters $Tiers_f$ in fund f is then defined as the time-series average of $Tiers_{ft}$, rounded to the nearest integer. As we discuss in Section 5.2.1, funds that raise capital in multiple rounds may mechanically have dispersion in contribution rates – and thus DVPI – early in their life. To mitigate these mechanical effects, we also exclude the first year of each fund's life when computing tiers. Panel A of Figure 4 plots the distribution of $Tiers_f$ across all funds. According to this procedure, 36% of funds have one DVPI tier, 61% have two tiers, and the remainder have three or more tiers. Hence, the clustering example shown in Panel A of Figure 3 appears to be a more general phenomenon in our sample of funds.[16] In Internet Appendix C.2, we also show that clustering in DVPI occurs due to clustering in both contributions and distributions. We build on this observation in Section 3.2 when decomposing fees into fixed and performance fees.

3.1.2 Robustness to Other Cluster Selection Methods

For robustness, we also determine the number of DVPI clusters in each fund using the Gap statistic approach (Tibshirani et al., 2001). The Gap statistic is based on the null hypothesis that the data is drawn from a single-cluster distribution, such as the uniform distribution over the range of observed values. Using this null distribution, we simulate returns and fit a sequence of k-means clustering models on the simulated data. The optimal k is then chosen based on whether the observed fit of a k-means model is sufficiently far from the fit of the same model applied to the simulated data. Thus, if the clusters are "loose" in the sense that they would occur reasonably often even if the data were drawn from a single-cluster distribution, the Gap statistic will not reject the null and conclude that the optimal $k = 1$. Put differently, if all investors truly did earn the same return but reported their returns with noise, the Gap statistic will push us towards finding that most funds have a single return cluster. In contrast, cluster selection based on Silhouette scores may be more sensitive to noise in smaller samples. We describe the Gap statistic approach in more detail in Internet Appendix C.1.

Panel A of Figure 5 compares the fraction of funds that have two DVPI clusters under the Gap statistic and Silhouette score methods. The first thing to note from the plot is that the majority of funds (55%) have two DVPI tiers according to the Gap statistic. This is not a mechanical result, as Panel B of the figure shows that 66% of funds with at least five investors have two return clusters. Figure 5 does show that the Silhouette method identifies more two DVPI-tier funds compared to the Gap statistic. In our full sample, 61% of funds have two DVPI tiers using Silhouette scores

[16] We also document similar clustering patterns using TVPI and IRR in Internet Appendix Figure IA7.

and this number rises to 73% in funds with at least five investors. The difference between the two selection methods is driven entirely by the fact that the Gap statistic is more likely to classify funds as having a single DVPI tier, which is unsurprising given its design. However, the fact that both approaches deliver comparable results indicates that return clustering is a robust feature of the data.

Our subsequent analyses rely heavily on an indicator θ_f for whether a fund has multiple investor tiers. We define this indicator based on whether a fund has multiple clusters in either call rates (contributions-to-commitments), distribution rates (distributions-to-commitments), or DVPI. Using call rates to define θ_f is helpful because younger funds whose DVPI is near zero may appear to have a single DVPI cluster, even if there is clustering in call rates. By this definition, 72% of funds have multiple investor tiers (i.e., $\theta_f = 1$) according to the Silhouette score approach and 69% do according to the Gap statistic. Moreover, the two approaches agree on which funds have multiple investor tiers in 96% of cases.[17] For completeness, we also confirm that all of our main conclusions are robust when using the Gap statistic to define θ_f in Internet Appendix C.1.

Arguably the easiest way to measure clustering would be based on unique values. This simple approach still suggests meaningful within-fund return clustering. In funds with at least five investors, 68% of funds have three or less DVPI tiers.[18] This cuts strongly against the notion that returns are continuously distributed in the typical fund, and again, our focus on funds with at least five investors means there is nothing mechanical driving this result. Using unique values to determine DVPI clusters is less accurate if fees are not the only source of within-fund return variation. For example, any measurement error – a concern for all empirical work – would lead counts of unique values to overstate the number DVPI clusters. The unsupervised machine learning techniques that we employ are precisely designed to separate true clusters in returns from any such noise. With that said, for the purpose of defining whether a fund has multiple investor tiers (θ_f), the unique value approach agrees with Silhouette scores for 99% of funds and with the Gap

[17]There are 197 funds that have a single DVPI tier according to Silhouette scores, yet are classified as having multiple tiers when we use call rate clusters to define the indicator variable θ_f. Indeed, these funds have a larger fraction of investor-quarter observations where DVPI is less than 0.01 compared to other funds (37% vs 21%).

[18]For each fund-quarter we count unique DVPI values d_{ft}. The number of clusters based on unique values is then the time-series average of d_{ft}, rounded to the nearest integer. This mirrors our use of the Silhouette and Gap methods.

statistic approach for 95% of funds.

The observed patterns of return dispersion and clustering suggest to us that the majority of funds group investors into one of two fee tiers. We favor this interpretation for a few reasons. First, fee dispersion is a natural equilibrium outcome given the contracting environment in private equity funds. In particular, the existence of side letters means that fee determination in a fund can be viewed through the lens of search and bargaining models (Burdett and Judd, 1983; Bester, 1988; Hortaçsu and Syverson, 2004; Duffie et al., 2005; Allen et al., 2019). These models generally predict that consumers or investors will pay different prices for the same good, either due to simple cost-based pricing, search and negotiation frictions, contract complexity, or heterogeneity in LP sophistication (Salop and Stiglitz, 1977; Gabaix and Laibson, 2006). Second, negotiating and implementing many different fee structures is costly from the perspective of a GP (Morgan Lewis, 2015). Any such costs should serve to limit the number of effective fee structures that arise in equilibrium, thereby leading to a small number of return clusters empirically. We discuss these theoretical considerations further in Internet Appendix B.6.

For the remainder of the paper, we therefore assume that our estimated return clusters can be used to proxy for fee tiering in a given fund. We explore other potential mechanisms for return dispersion and clustering (e.g., accounting differences across LPs) in Section 5 and, when applicable, we will flag how any such mechanisms would impact our subsequent analysis.

3.2 Estimates of the Size of Fee Dispersion

While it appears that private equity funds often use multiple fee structures, it remains unclear how large fees differ across investor tiers. Fee structures could differ along several dimensions, including management fees and associated fee offsets, fund and organizational expenses, taxes, or performance fees. Though we do not observe exact contractual differences, the panel nature of our data allows us to decompose differences in fees into “fixed fees” that scale with commitment and time, and variable fees that scale with performance. To see why, recall that contributions into a fund include capital that is invested by the GP, management fees, fund expenses, and portfolio-company fee offsets (Phalippou et al., 2018). Defining (capital) call rates as the cumulative about of contributions per dollar of commitment, we can decompose the call rate for investor p in fund f at time t into three terms:

$$\text{call-rate}_{p,f,t} = i_{p,f,t} + m_{p,f} \times t + \varepsilon^{m}_{p,f,t}. \tag{3}$$

$i_{p,f,t}$ is defined as the cumulative amount of capital that has been invested by the GP per dollar of commitments. $m_{p,f}$ denotes any fees that are charged on an annual basis as a percentage of investor p's commitment size (e.g., management fees). $\varepsilon^{m}_{p,f,t}$ is a residual term that captures any fund expenses, portfolio company fees, or measurement error. $i_{p,f,t} = i_{f,t}$ will be constant across LPs if capital is invested by the GP on a pro rata share of commitments. Under this assumption, we can write the within-fund standard deviation of call rates in fund f at time t as:

$$p^{\sigma}_{ft} = m^{\sigma}_{f} \times t + \varepsilon^{\sigma}_{f,t}, \tag{4}$$

where m^{σ}_{f} is the within-fund dispersion in fixed fees and $\varepsilon^{\sigma}_{f,t}$ is the dispersion in the residual. This equation says that dispersion in call rates should be linearly related to age (measured in years) with a slope of m^{σ}_{f}. Panel A of Figure 6 confirms a strong linear relationship using a binned scatter plot. The plot pools data on all funds within their first five years of life, as this is the period when management fees are typically charged as a percent of committed capital. We further restrict our attention to funds that we classify as having multiple investor tiers as in Section 3.1.2. The slope of 91 basis points (standard error of 8) reflects the average within-fund dispersion in fixed fees across all funds. As a point of comparison, the level of management fees in private equity is generally thought to be around 200 basis points (Gompers and Lerner, 1999). This dispersion estimate is also unlikely to be biased by other sources of return dispersion like LP-specific accounting or measurement error, as it unlikely that either could generate a linear relationship between call-rate dispersion and age (see Section 5.1 for further discussion).

Table 2 repeats this analysis within each asset class. The first and second columns of the table show the point estimates of m_f^σ and their associated standard errors, which are clustered by fund. Estimates of m_f^σ are based on funds that have multiple tiers as defined in Section 3.1. Within-fund dispersion in fixed fees varies strongly across asset classes. Fixed fees within infrastructure funds vary on average by 103 basis points, whereas they vary by only 42 basis points in venture capital funds. The low dispersion in venture capital funds is consistent with the analysis we will present in Section 3.3.

Much like contributions, we can decompose net-of-fee distributions (per dollar of commitment) d as follows:

$$d_{pft} = g_{pft}\left[1 - c_{pf} \times h_{pft}\right] + \nu_{pft}$$

where g_{pft} is the unobserved gross-of-fee distribution process, c_{pf} is the rate of net-of-tax performance fees, h_{pft} is an indicator variable for whether fund f has cleared p's hurdle rate, and ν_{pft} is a residual term (e.g., measurement error). Under the assumption that gross returns $(g_{pft} = g_{ft})$ and hurdle rates $(h_{pft} = h_{ft})$ are equal across investors, then we can write the within-fund standard deviation in distribution rates as:

$$d_{ft}^\sigma = c_f^\sigma g_{ft} h_{ft} + \nu_{ft}^\sigma \tag{5}$$

where c_f^σ is the within-fund dispersion in performance fees and ν_{ft}^σ is the dispersion in the residual. Intuitively, dispersion in performance fees should widen linearly with performance, so long as the fund is above its hurdle rate. This intuition implies that the relationship between dispersion in distribution rates and gross performance should have a specific functional form, one that resembles the payoff of a call option. In the region where funds are below their hurdle rate, d_{ft}^σ should be insensitive to performance. In the region where funds are above their hurdle rate, d_{ft}^σ should be linearly increasing in performance with a slope equal to c_f^σ.

Panel B of Figure 6 confirms this call-option pattern empirically. To construct the binned scatter plot, we pool over all funds and proxy for each fund's gross fund performance using its maximum TVPI. Based on industry conventions, we define funds as profitable enough to charge performance fees if they have a TVPI of at least 1.09 and an IRR of at least 9%. The binned scatter plot also partials out vintage fixed effects to control for any potential age-specific effects on performance fees. In the region where funds are defined as unprofitable, the slope estimate is -0.4 and is not statistically different from zero. In the region where funds are profitable, the slope of indicates that performance fees vary by $c = 5.8\%$ (standard error equals 0.9%) in the average fund. Once again, it seems hard to imagine that anything other than dispersion in carry or its tax treatment could generate this precise pattern.

The middle columns of Table 2 present estimates of c^{σ} by asset class and their associated standard errors, which are again clustered by fund. We include fund vintage fixed effects to account for aggregate fund performance. These estimates are based only on funds with multiple tiers (see Section 3.1) that are profitable enough to charge performance fees. Private debt and real estate have the largest standard deviation in performance-based fees at 6.8% and 6.4%, respectively. On the other hand, we estimate that carry dispersion in venture capital funds equals 0.5%. Private equity funds lie in the middle of these two extremes with an average within-fund dispersion of 3.3%. In all asset classes except venture capital, we reject the null that the average within-fund dispersion in variable fees equals zero.

Dispersion in performance rates should only manifest once a fund has cleared its hurdle rate. We exploit this observation to implement a series of placebo tests in the last three columns of Table 2. The table reports the point estimate of a regression of dispersion in distribution rates on fund performance in the subsample of unprofitable firms. As expected, the estimated slopes are generally small in all asset classes and none are statistically different from zero.

To summarize, venture capital funds appear to have the lowest average dispersion in both fixed and performance fees. Within private equity funds, the average dispersion in fixed and performance fees equal 83 bps and 3.3%, respectively, which is consistent with public reports on the menu-model that Bain Capital has used in recent years (Markham, 2017). Infrastructure, private debt, and real estate funds tend to have larger within-fund dispersion in both types of fees. As further validation for our estimation approach, dispersion in performance fees is not detectable in unprofitable funds.

3.3 Which funds use multiple fee structures?

We now investigate the types of funds and GPs that are more likely to use multiple fee structures. Our focus is on an indicator variable for whether fund f has multiple investor tiers in either DVPI, call rates, or distribution rates. In what remains, all discussion of tiering or tiering propensity is based on this indicator. We define tiers in each variable using the Silhouette method described in Section 3.1. We restrict our analysis to funds that are at least one year old to give time for potential fee differences to materialize in the data.

Table 3 presents a set of OLS regressions of a fund's propensity to tier investors on various correlates. In column (1), we regress the tier-indicator on a full set of GP fixed effects. An F-test of the null of no GP effects is strongly rejected. This result is driven by the fact that 44% of GPs use a consistent fee policy across all of their funds. In column (2), we test for the presence of law firm fixed effects using information provided by Preqin on the identify of law firms used by each fund. Lawyers naturally play a large role in the fund formation process because GPs rely heavily upon them to form LPAs and side letters. Indeed, in our analysis of a limited set of LPAs (Section 2.1.2), there was considerable heterogeneity in terms of the length and language of the contracts. To the extent that law firms consistently differ in their approach to LPAs and side letters, heterogeneity in tiering policy should be tied to heterogeneity in law firm usage across funds. We find support for this hypothesis in column (2), which shows that we can comfortably reject the null of no law-firm fixed effects. The existence of GP and law-firm fixed effects is most naturally attributable to fee policy, as it is hard to imagine how other sources of return dispersion would generate these fixed effects.

The remaining columns of Table 3 map tiering propensities to observable characteristics. In all specifications, we include two sets of fixed effects. The first set is based on the number of investors in the fund and is designed to eliminate any bias that could be induced from measurement error in our classification of tiers. For example, measurement error could lead to an upward bias in the number of tiers for funds with many investors. The second set of fixed effects is on based on the deciles of each fund's size, which accounts for the fact that fund size and contracting terms are determined endogenously. We use robust standard errors to account for the heteroskedasticity inherent in linear probability models.

In columns (3)-(5), we investigate the relationship between tiering behavior and proxies for LP demand. In column (3), we build off of Berk and Green (2004) by measuring demand with GP past performance. By their logic, GPs with a track record of strong returns face higher demand because investors use observed performance to update about unobserved GP skill. We compute past performance using the average quartile ranking of all funds raised by a GP raised prior to each fund f's final close date. When aggregating, we exclude funds that are less than four years old because quartile rankings are less informative in the early stage's of a fund's life (Brown et al., 2019). Quartile rankings are measured as of the closing date of each fund to ensure that the information was available to investors at that time.[19] Funds in quartile 1 are the best performing fund and funds in quartile 4 are the worst performing. Column (3) shows that GPs in the worst quartile of past performance are roughly 17 pp more likely to use multiple fee structures than those in the best quartile. This effect is measured with precision and is meaningful given that roughly 40% of funds use a single fee structure. Similarly, column (4) shows that undersubscribed funds – defined as those whose final close size was below their initial target – are 12 pp ($t = 4.83$) more likely to tier investors.[20] GPs without an established track record may also face less demand for capital commitments. Accordingly, column (5) shows that GPs who are raising their first, second, or third funds in a given strategy are 13 pp ($t = 6.63$) more likely to tier investors.[21] Overall, there appears to be a fairly robust and negative correlation between measures of investor demand and the likelihood that a GP uses multiple fee structures.

This correlation informs the types of theories that can explain fees in private-market funds. It is natural to think that a high-skilled GP may have monopoly power over pricing, since LPs view

[19] See this report by Preqin for the methodology.

[20] Relatedly, we find that funds that take longer to raise capital also have a higher propensity to tier. Each extra month of fundraising is associated with roughly 1 pp increase in the likelihood of tiering. This is a fairly large effect given that the standard deviation of the time between fund launch and final close is about 9 months. We omit this specification from Table 3 because fundraising time is missing for 85% of funds.

[21] There is a sharp cutoff in the propensity to tier after three funds, which suggests it takes three funds for a GP to establish a track record. This result makes sense when considering that fund performance is not known with confidence for several years. See Internet Appendix C.4.1 for more details.

them a scare good with no close substitutes. In principle, these GPs could then perfectly price discriminate by charging each possible LP their exact reservation value, yet this would counterfactually predict that high-skilled GPs use more fee structures. One possible opposing force is that eliciting individual reservation values lengthens the fundraising process. This is more costly for talented GPs because it delays them from deploying their investment skill.[22] Relatedly, LPs may be more likely to agree on the value of an established GP, which further limits the net benefits of perfect price discrimination. Conversely, GPs without an established track record of high performance may face more disagreement about their value and find that their initial fee offering does not generate sufficient demand. In response, it may be optimal to offer discounts to some LPs, especially if their commitment will then attract more hesitant investors (Toll and Centopani, 2017, p. 29). GPs who can easily source capital commitments may also charge a single high price if public pensions are prohibited from paying fees above a certain threshold (e.g., 3-and-30), perhaps due to public scrutiny and headline risk (Dyck et al., 2018).

In column (6), we test whether funds that use placement agents are more likely to tier investors. We obtain this information from Preqin and assume that missing entries mean that no placement agent was used.[23] There are two reasons why we expect placement agent usage to be positively associated with tiering. First, placement agents are used by GPs to secure capital commitments from LPs. GPs who can easily raise capital may not employ placement agents and, for reasons discussed above, these GPs are also less likely to tier. Second, it is common practice for placement agent fees to be charged to the fund, meaning LPs bear a large portion of their cost. This practice has come under public scrutiny in recent years, in part due to several high-profile scandals in which payments were made by placement agents to public pension employees in return for capital commitments. In response, many pensions now ban investment in funds that use placement agents or refuse to pay placement agent fees, the latter of which can be accomplished using exceptions in side letters.[24] Consistent with this view, column (6) shows that the use of placement agents is

[22] Anecdotally, investors in famous funds are often given a take-it-or-leave-it offer at a single high price. This allows the GP to close the fund quickly and focus on investments.

[23] 67% of funds in our sample do not have any placement agent listed.

[24] For example, the Pennsylvania Public School Employee's Retirement System (PSERS) has a public policy that

associated with a 13 pp ($t = 6.83$) increase in the likelihood of tiering. We view this point estimate as a lower bound because our treatment of missing entries should bias it towards zero.

In column (7), we explore whether a fund's propensity to tier investors varies across asset classes. All point estimates are relative to private equity because it is the omitted category in the regression. Venture capital funds are far less likely to use multiple investor tiers than other asset classes – 48 pp less likely than infrastructure and 33 percentage points less likely than private equity.[25] Both point estimates are measured with statistical precision. This result accords with our finding that within-fund dispersion in management and carry is substantially lower in venture capital funds (Section 3.2). It is also consistent with the observation that the venture capital industry was an early adopter of standardized LPA and side letter provisions (Robbins, 2019).

In sum, we find that some GPs consistently use multiple fee classes for their LPs. For instance, managers of venture capital funds are far less likely to tier their LPs. At the same time, GP tiering policy has a dynamic component that evolves with market conditions. Funds are more likely to use multiple fee structures if they are unsubscribed, managed by GPs with a short or poor track record, or use placement agents.

4 Do some investors consistently pay lower fees?

In this section, we show that LPs who outperform in one fund are more likely to outperform in their other funds. This pattern is consistent with the idea that some LPs are able to consistently select or obtain the best terms in their respective funds, at least ex-post. Furthermore, matching between GPs and LPs appears to be an important component of this persistent outperformance. We then analyze the characteristics of the LPs that are most likely to outperform and estimate how much these traits explain their persistent outperformance.

---

states they will not make any payments to placement agents through their funds (PSERS, 2020).

[25]These results are supported by our analysis of LPAs from Section 2.1.2. In this limited subset, virtually all funds whose LPAs contain side letter language have multiple return tiers. Moreover, venture capital (VC) funds are 5% less likely to include side letter language in their LPAs. Within the funds whose LPAs have side letter language, VC funds are 12% more likely to allow all LPs to view any side letters and 22% more likely to allow all LPs to opt in to any side letter provisions.

4.1 Pension-effects

4.1.1 Baseline estimates

In light of our evidence on within-fund tiering of investors, we now analyze whether some investors are more likely to be in the top tier of performance in all of their funds. Specifically, within each fund, we define an indicator variable y_{pf} based on whether investor p has above-median TVPI for the majority of fund f's life. We construct y_{pf} using medians based on our finding that most funds have two tiers of investors. In this sense, y_{pf} can be interpreted as a measure of whether p is a top-tier investor in fund f. We then assess the persistence of within-fund performance across funds using the following regression:

$$y_{pf} = \lambda_a + \alpha_p + \varepsilon_{pf}, \tag{6}$$

where λ_a are age fixed effects as measured by vigintiles of fund f's age and α_p denotes an investor fixed effect. Controlling for fund age allows us to better isolate pensions who truly have dominated fee contracts from those who trade off high fixed fees for low variable fees or vice versa. For instance, pensions who trade off the two fee components may outperform early in a fund's life but not later. When estimating (6), we exclude funds that we classified as having a single investor tier as in Section 3.1.2.

Under the null hypothesis of no pension effects, the estimated α's should not be statistically distinguishable from each other. In other words, if pensions are randomly assigned to fee tiers in each fund, then we should not be able to reject an F-test that the α's are jointly equal to each other. In Table 4, we report the number of pension effects K, the F-tests and their associated p-values based on the core sample. When moving from rows (1) to (3), we conduct the F-test for whether the α's are jointly equal based on funds that are at least one, four, and eight years old, respectively. In all cases, the estimated F-statistic is large enough that we reject a null of no pension effects with a p-value of less than 0.01.

The standard approach to conducting F-tests like those in Table 4 rely on parametric assumptions to test the null of no pension effects. As a robustness check, we calculate non-parametric

p-values based on a permutation test where we: (i) randomly assign return paths to investors within each fund f; (ii) calculate the simulated value of y_{pf}; (iii) run regression (6); and (iv) recalculate the F-statistic from the test of equality across α's. We repeat this procedure 500 times to generate an simulated distribution of F-statistics, after which we compute a non-parametric p-value based on where the actual F-statistic falls in this distribution. We denote the p-values based on these permutation tests as p^*. Reassuringly, we again reject the null of no pension effects.

Though the preceding F-tests provide a statistical sense of the size of pension-effects in our data, they do not easily convey the economic magnitude of such effects. To get a better sense, we compute the distribution of the estimated pension effects and compare it to the distribution implied by random assignment of returns (i.e., fees) within each fund. Because the true distribution of pension effects $\alpha's$ differs from the estimated distribution due to sampling error, we adjust the estimated pension effects using an Empirical Bayes method (Morris, 1983). Let $\hat{\alpha}$ denote the vector of estimated α's based on regression (6). Using Casella (1992, Eqs. 7.11 and 7.13), we can calculate the empirical Bayes estimate $\widetilde{\alpha}$ as

$$\widetilde{\alpha} = \bar{\alpha} + \max\ (1 - B, 0) \times (\hat{\alpha} - \bar{\alpha}),$$

where $\overline{\alpha}$ is the average of the estimated fixed effect vector $\hat{\alpha}$ and

$$B = \frac{1}{F}\left(\frac{K-1-2}{K-1}\right)$$

is a shrinkage coefficient. The F in the formula for the shrinkage coefficient B corresponds to the F-statistic from the joint test that the $\hat{\alpha}$ are equal (as reported in Table 4). Intuitively, the F-statistic is larger (and B smaller) when the pension-effects are estimated with more precision, and in turn, the Bayes estimate does not shrink $\hat{\theta}$ as much towards its mean.

Panel A of Figure 7 visualizes the rejection of the F-test. The orange line shows the distribution of pension effects under the random assignment of return paths (i.e. contracts) within each fund. The distribution of observed pension effects $\widetilde{\alpha}$ (in blue) has much fatter tails: the 95th percentile pension outperforms in 67% of its funds while the 5th percentile pension outperforms in only 13% of its funds. These fat tails are the reason why we reject the null of random assignment.[26]

4.1.2 LP-GP Effects

Search and bargaining models often predict that relationships between LPs and GPs factor into equilibrium pricing. We investigate whether selective matching between GPs and LPs relates to an LPs net-of-fee return outperformance by augmenting regression (6) as follows:

$$y_{pfg} = \lambda_a + \eta_{pg} + \varepsilon_{pfg} \tag{7}$$

where y_{pfg} is the indicator of investor's p relative outperformance in fund f managed by fund manager g. We define the indicator as above. As before, λ_a are the vigintiles of fund f age. The new term in the regression is η_{pg}, which are LP-GP fixed effects. The LP-GP effects η_{pg} measure the outperformance of investor p in funds managed by GP g. If, for instance, some investors receive better terms than others in funds managed by a specific set of GPs, then we should reject an F-test of the joint significance of the η's.

Table 4 reports F-statistics and their associated p-values from testing whether the η's are jointly equal. We reject the null of no LP-GP effects (η's) when using parametric p-values and non-parametric p-values based on the permutation tests described in Section (4.1.1). This evidence suggests that matching between LPs and GPs is important for understanding why some pensions consistently outperform others when investing in the same fund.

4.2 Observable Pension Characteristics

In this section, we map the pension effects to observable pension characteristics to better understand why some pensions consistently outperform others in their respective funds. We proceed in

[26] In Internet Appendix C.6.1, we further show that a pension's tier-assignment in one fund can be predicted by its assignment in other funds. These out-of-sample tests suggest that pensions differ in negotiation skill or bargaining power in a way that consistently impacts investment costs in all of their funds.

two steps. First, we replace the pension effects in regression (6) with characteristics related to size and investor sophistication. Second, we show that characteristics explain some, but not all of the observed pension effects. This suggests that unobservable traits like negotiation skill or bargaining power materially impact the fees pensions pay in funds.

4.2.1 Observable Characteristics X_{pf}

We characterize the pensions that tend to pay lower fees by replacing pension fixed effects with observable characteristics X_{pf} in regression (6):

$$y_{pf} = \mu_f + \lambda_a + \beta \mathbf{X}_{pf} + \varepsilon_{pf}, \tag{8}$$

where y_{pf} is the indicator of investor's p relative outperformance in fund f as described above, and μ_f is a fund fixed effect and λ_a are the vigintiles for fund age as before. We include fund fixed effects in the regression to ensure that β is identified using within-fund variation.

We consider the following set of observable characteristics X_{pf}. For each investor p in fund f, we compute p's share of the total fund as their commitment amount divided by the total fund size. We include each investor's share of the fund to account for potential returns to scale when raising capital. For example, one might expect that GPs might reduce fees for investors that account for a larger fraction of the fund, as this would then free up the GP to focus on optimizing the investment portfolio instead of raising capital.

Due to information asymmetries about manager skill, signaling effects are likely to be important for GPs when they raise a fund. For example, if a GP secures a capital commitment from a large and well-known pension, then other pensions may be more willing to commit capital to the fund. We code investor p as "Large" if its total assets under management are over \$100 billion at the time of fund f's launch, a designation that is reserved for easily recognizable pensions in our data. In addition to the potential signaling effect that they may have on fund raising, large investors are also more likely to possess the ability to deploy large amounts of capital quickly, so size is likely related to the economies to scale in fund raising discussed above.

We include three variables that capture the experience and potential negotiation skill of each investor in private markets. A priori, it is plausible to think that skill in fee negotiation improves as investors become more experienced in the nature of private market investment vehicles. Motivated by the existence of LP-GP effects in Section 4.1.2, we measure relationship strength using the number of matches between an LP and GP. Specifically, for each LP-GP pair, we count the number of funds that are managed by general partner g in which p has invested. We use the full dataset to compute this measure because we want to capture settings where a GP reduces fees for investor p in fund f in expectation that the investor will invest in future funds raised by the GP. The second variable captures how well an investor's prior funds have performed. Arguably, LPs skilled at manager selection are also skilled in contract negotiations. We measure each investor p's past fund performance as the average quartile ranking of its active funds at the time of fund f's close, where quartile rankings mimic those used in Section (3.3). The third variable we use is an indicator for whether investor p was an early private equity investor based on having invested in PE prior to 2008.

The last set of variables that we include are related to pension governance.[27] We include board size to account for any potential coordination problems that may cause larger boards to sub-optimally negotiate fee contracts. In addition, Andonov et al. (2018) find that pension boards with more state officials are more likely to make poor investment decisions in private equity, likely due to distortions from political considerations. Motivated by that finding, we include the percent of each pension's board that is elected by plan beneficiaries. The idea here is that board members who are elected by plan beneficiaries will be more cognizant of fees when approving PE allocations. For each pension and fund pair (p, f), the total number of board members and the percent of elected members are both measured as of fund f's vintage year.

Each column of Table 5 shows estimates of regression (8) for funds that are of a minimum age (1, 4, and 8 years). The first row reveals a robust relationship between commitment size and top-tier fee status. In the largest sample of funds (column 1), the point estimate of 0.96 pp ($t = 4.80$)

[27] We are grateful to Josh Rauh for sharing the data on pension board composition from Andonov et al. (2018). The data ends in 2013, so we extend it to 2018 to better match our sample.

means that an LP who commits 10 pp of additional capital (relative to the fund's total size) is 10 pp more likely to be in the best fee tier.[28] LPs who are large by overall AUM are also 14-17 pp more likely to be in the best tier of fees, even after controlling for their commitment size. The point estimates are statistically significant at conventional levels in all specifications. This finding is consistent with a mechanism by which large and well-known LPs are given fee discounts for drawing in additional capital commitments.

Measures of investor sophistication and experience also appear to play an important role in determining an investor's net-of-fee return performance. For instance, each additional fund between the LP and the GP increases the likelihood of being a top-tier investor by around 0.70 pp, though the point estimate is not always precisely measured. This finding relates to our evidence of LP-GP matching in Section 4.1.2. Investors with past success in manager selection are 10-13 pp more likely to be top-tier investors in their funds. Recall that we measure past success for each investor based on the average quartile ranking of its previous funds. Similarly, pensions who were early investors in private equity are roughly 7-9 pp more likely to outperform others in their funds. Finally, there is also some evidence that better pension governance improves the probability of being a top-tier investor. Pensions whose boards have 10 percent more elected members are about 2 percent more likely to be top-tier investors. For context, the standard deviation of percent of elected board members is 26% in our sample. On the other hand, there does not appear to be a stable relationship between board size and top-tier investor status.

4.2.2 How much of pension-effects are due to observables?

In this section, we assess how much observable characteristics account for the observed pension effects documented in Section 4.1. There are three steps to adjust pension effects for observable characteristics. First, we take the full panel of raw returns, as measured by TVPI, and regress it on fund-by-quarter fixed effects and the full set of covariates from Section 4.1.1 interacted with

[28] We also find that LPs who have committed above a fixed cutoff are all given a discount in 63% of funds that have multiple management fee tiers. While this suggests size is an important determinant of management-fee tier assignment (Da Rin and Phalippou, 2017; Clayton, 2020), it is also natural to expect that other factors like signaling, sophistication, and experience play a role. See our subsequent results and Internet Appendix C.5.1 for more discussion.

vigintiles for fund age. We use the raw level of returns because it allows us to flexibly control for covariates. Second, we use the residuals from the regression to compute whether investor p is a top-tier investor in fund f, denoted by $\tilde{y}_{pf}$, as in Section 4.1. Third, we estimate pension effects based on this characteristic-adjusted tier assignment $\tilde{y}_{pf}$ and apply the empirical Bayes procedure described in Section 4.1.1.

Panel B of Figure 7 plots the distribution of the resulting characteristic-adjusted pension effects as a dashed-green line, as well as the observed pension effects (characteristic unadjusted) as a blue line. The blue line is identical to Panel A. The plot shows that characteristics do account for some of the observed pension effects. The characteristic-adjusted distribution is shifted to the left because some of the differences in contract terms (taken as given in Panel A) are absorbed by observable characteristics like size. If all pensions earned the same within-fund return after characteristic adjustments, then no pension would outperform any other and the entire mass of the green-dashed line would be at zero. Nonetheless, the figure also shows that there are still many LPs who consistently over- or underperform after controlling for characteristics. For example, the far right tail of the characteristic-adjusted pension effects is at 80% and the far left tail is at 10%. Based on the set of covariates that we consider, this evidence suggests that a subset of pensions consistently outperform others in their funds for reasons that are orthogonal to pension size, share of a fund's commitments, or past experience in funds. We interpret this as evidence that unobserved traits like negotiation skill or bargaining power meaningfully impact the fees that pensions pay in funds.

5 Other Sources of Within-Fund Return Dispersion

We now investigate how other sources of within-fund return dispersion would impact the preceding results, focusing on four broad channels: (i) measurement error; (ii) differences in entry timing, either due to multiple fund raising rounds or secondary transactions; (iii) accounting practices that vary across LPs; and (iv) differences in gross-of-fee returns across investors in the same fund. All of these channels likely contribute to the within-fund dispersion shown in Figure 2, and we make no claims that fees are the sole determinant. Still, as we argue below and in Internet Appendix A, these other sources should not confound our overall conclusions about the nature of fee dispersion in private equity.

5.1 Measurement Error

Measurement errors are one simple reason why returns in our data could differ across LPs in the same fund. Given that the data is sourced primarily via FOIA requests, these errors could occur when Preqin transcribes the FOIA data that they receive from LPs. To gauge the size of this channel, we created our own dataset by filing FOIAs directly with a sixty-five of the pensions in our sample. We chose these pensions based on the funds with the most observed dispersion and the LPs whose performance was the most extreme relative to others in their respective funds. In the vast majority of cases (~97%), the data from our direct FOIA was identical to the Preqin data. For the small number of cases where the data did not perfectly match, the size of the deviations was economically small. Internet Appendix A.1.1 contains the full results of this audit, including the exact language of our FOIA requests. Measurement error could also occur if LPs report erroneous data in their FOIA replies to Preqin. While this is certainly possible, most public pensions are audited annually, which in principle should reduce the occurrence of random reporting errors over time.

More importantly, any measurement error in the data should not materially bias our results. For one, the machine learning techniques that we use in Section 3.1 are designed to be robust to noise, yet all consistently find evidence of return clustering. If a fund truly has a single fee structure and returns are just measured with noise, we should not be able to reliably detect multiple clusters. Measurement error should also not bias our estimates of within-fund dispersion in management and carry (Section 3.2). To see why, recall that our estimate of carry dispersion relies on a call-option relationship between dispersion in distributions and fund performance. It is hard to imagine how measurement error would generate this precise pattern in the data and hence bias our estimates.[29] Even if measurement error leads to some misclassification of multi-tier funds, our characterization of the types of funds and GPs that tier investors should also be minimally affected (Section 3.3). The simple reason why is that our outcome variable – an indicator for whether a fund has multiple return tiers – would then be measured with error, which should not bias the regression coefficients in Table 3. This argument breaks down if some types of funds are more likely to have reporting errors than others (e.g., private equity versus venture capital). However, we see no reason to expect this to be the case. A similar logic implies that measurement error should not bias our finding that some LPs are more likely to outperform others when investing in the same fund (Section 4).

5.2 Differences in Entry Timing

LPs in the same fund could earn different returns if they enter or exit at different times. We now explore two channels that could generate variation in entry and exit: multiple commitment rounds and secondary market transactions.

5.2.1 Multiple Fund Raising Rounds

GPs often secure capital in multiple fundraising rounds. In this case, investors who commit at different rounds (e.g., first vs second) may initially have different contribution schedules. Thus, even in the absence of fee dispersion, returns may differ across LPs due to variation in the timing of their cash flows. Returns will also cluster based on the fundraising round in which LPs commit capital.

There are several reasons why this mechanical timing-induced dispersion does not drive our results. We present the complete argument in Internet Appendix A.2 and summarize the main points here. First, and most importantly, 63% of funds in our sample were raised in a single round. Timing-induced dispersion is not possible in these funds, yet 57% still have multiple return tiers.

[29] If in each period, LPs input contributions with random error, then dispersion in *cumulative* contributions will grow linearly with age and bias our management fee estimates upward. This explanation is unlikely given the presence of clusters in contributions (see Internet Appendix C.2).

Moreover, we show in Internet Appendix A.6 that all of our main conclusions regarding the size and nature of fee dispersion hold if we only analyze single-round funds.

In addition, funds generally have equalization methods (so-called "true ups") that adjust the contributions of LPs to account for differences in entry timing. For example, LPs who join at later rounds may initially make larger contributions in order to catch them up with those that committed in earlier rounds. This process ensures that future distributions are dispersed as if all LPs had committed at the same time, as we assume in Section 3.2. An implication of this process is that tiering rates in contributions should be driven by data early in a fund's life. However, we find that 93% of funds have the exact same number of contribution tiers after excluding the first five years of their data. In practice, small differences in contributions can persistent after equalization is complete because late-closing LPs will often make a one-time interest payment to those who committed earlier (e.g., LIBOR plus a spread). This suggests that dispersion in cumulative contributions should decrease over time, though the results in Section 3.2 indicate the opposite. Instead, the positive linear relationship between dispersion in contributions and fund age is consistent with dispersion in management fees.

5.2.2 Secondary Market Transactions

Secondary market transactions involve the sale of an existing stake in a fund. The market for such transactions has grown in recent years and is often used by LPs for portfolio rebalancing or liquidity management (Nadauld et al., 2019). Much like multiple commitment rounds, a secondary transaction may induce variation across LPs in entry or exit times, thereby leading to return dispersion. While secondary sales and purchases are certainly possible, we argue in Internet Appendix A.3 that these transactions are infrequent in our sample. We also argue in Internet Appendix A.7 that, to the extent that they are present, secondary market transactions should not meaningfully impact our estimates of within-fund dispersion in fees (Section 3.2), our characterization of funds that tier (Section 3.3), or our characterization of the LPs that are consistent in the top-tier of fees (Section 4.2).

5.3 Accounting Practices

Public pensions that invest in private capital vehicles have no legally mandated accounting standards, which could lead some LPs in our sample to report cash flows or NAVs differently than others. Any such differences could cause returns to differ across LPs in the same fund. Nonetheless, there is a simple institutional reason why accounting differences are unlikely to be a large source of within-fund return variation. GPs typically send their LPs a quarterly report that summarizes the current state of their investments in the fund. The detail of these reports varies substantially across GPs, but all generally provide a running total of distributions and an estimate of the liquidation value of currently-held investments. According to several large LPs and Preqin, the content of these reports is generally used to satisfy any FOIA requests. In our data, this means that accounting practices should vary across funds, not within funds. With that said, it is still possible for specific LPs to adjust the cash flows and NAVs contained in the investment reports for FOIA requests. We now discuss two specific variables for which any such adjustments are most likely to occur.

Net asset values (NAVs) Fund NAVs measure the estimated value of each LPs share in the fund in the event of an orderly liquidation. LPs could systematically differ in how they report NAVs if some deduct expected performance fees (carry) that would be charged by the GP. If this were the case, these LPs would consistently report lower TVPIs in their respective funds and generate the dispersion observed in Panel B of Figure 2. However, the presence of sizable within-fund variation in DVPI (Panel A, Figure 2) suggests that NAV-accounting is not the primary source of net-of-fee return variation within funds.

Recallable (or Recyclable) Capital Within-fund dispersion in DVPI (or TVPI) could also arise due to differences in how LPs account for recallable capital. Recallable (or recyclable) capital refers to proceeds from liquidated investments that can be reinvested by the GP. The terms of this reinvestment are specified by so-called recycling provisions, which prescribe the amount and horizon over which recallable capital can be deployed. According to the CFA Institute's Global

Investment Performance Standards (GIPS), LPs should account for recallable capital by recording a new distribution and a new contribution equal to amount that is being recalled. To the extent that LPs do not follow GIPS standards, they could instead net out recallable capital, recording no new distribution and no new contribution. In a given fund, these two approaches could lead to the appearance of multiple clusters in DVPI (and TVPI). To gauge the potential strength of this channel, in Begenau et al. (2020) we filed FOIA requests to a subset of pensions about their accounting practices. Reassuringly, 100% of respondents reported that they conform to GIPS standards. In addition, IRRs are invariant to the accounting of recallable capital, yet we observe similar patterns of variation and clustering when measuring returns with IRRs (Panel C, Figure 2 and Internet Appendix C.3).

More generally, it is useful to consider how heterogeneity in accounting practices could bias our main conclusions. This source of return dispersion should not materially bias our estimates of fee dispersion or our characterization of the funds that use multiple fee schedules (Section 3). Moreover, the fact that funds with multiple return clusters tend to be of a certain type (e.g., use placement agents) cuts against the view that LP-specific accounting is responsible for return variation. The pension effects documented in Section 4.1.1 could in principal be driven by LP-specific accounting conventions. However, NAV reporting or recallable capital accounting are unlikely to drive this result because we still observe pension effects when measuring returns with DVPI and IRR (Panel A of Internet Appendix Figures IA9 and IA10).

5.4 Differences in the gross-return exposure

Net-of-fee returns could differ across investors in the same fund if they have different gross (or pre-cost) exposure to the fund. There are main two mechanisms through which this could occur in practice: (i) co-investment vehicles and (ii) LP-specific restrictions on investment.

5.4.1 Co-investment

Co-investment vehicles allow LPs to augment their exposure to the "main" fund by allocating additional capital towards a particular deal or set of deals (see Fang et al. (2015) for institutional details). These structures are related to so-called side-car or parallel fund vehicles (Lerner et al., 2018). To see why co-investments structures could generate net-of-fee return dispersion, consider a fund in which only investor A has the ability to co-invest. LP in the fund are otherwise equal in terms of their commitment size and all investment terms, namely fees. Further suppose that investor A combines the returns on its co-investment portfolio and the main fund when responding to FOIA requests and reporting to Preqin. If the co-investment vehicle tilts more towards certain portfolio companies relative to the main fund, then investor A's reported net-of-fee return will differ from other LPs. The resulting dispersion would be even larger if the co-investment vehicle had reduced cost structure compared to the main fund, as it often does in practice.

There are several reasons why co-investment vehicles are not the primary source of within-fund return variation that we observe empirically. First, and most importantly, we exclude any funds that Preqin classifies as a co-investment vehicle from our analysis. We expect this classification to be relatively accurate because LPs generally list co-investment vehicles as a separate fund when reporting performance to Preqin. For example, "Fortress Investment Fund IV" and "Fortress Investment Fund IV - co-investment" appear as two separate funds and we drop the latter. Moreover, for several of the largest LPs in our data, we have manually compared the co-investments that are reported on their websites and annual reports against the data in Preqin. In all cases, we found that cash flows from co-investments were indeed listed separately in the Preqin data.

Second, while co-investment vehicles have been increasing in popularity in recent years, they have not been a large part of public pension PE investments for most of our sample (1990-2018). Based on data from CEM Benchmarking, a provider of benchmarking services for thousands of global pensions, Beath et al. (2014) find that less than 5% of U.S. public pensions had any co-investments in PE as of 2014.[30] Smaller pensions may be less able or inclined to co-invest because

[30] Preqin (2014) finds that "relatively few LPs are being offered co-investment rights by GPs in the Limited Partner-

it requires the internal infrastructure to evaluate individual portfolio companies and then deploy capital on relatively short notice. Even for larger pensions, co-investments are not yet a large portion of their portfolios. For instance, in 2019, CalSTRS – the second largest pension fund in the U.S. – reported that less than 5% of its PE portfolio was through co-investments (CalSTRS, 2019).[31]

Third, as part of our data-quality audit (Internet Appendix A.1.1), we asked pensions via FOIA if they utilized any special investment arrangements such as a side-car deals or co-investments. The vast majority responded that they had no such arrangement. For the few cases that affirmed co-investment arrangements, we confirmed that these co-investment relationships were reported separately and therefore not included in our analysis.

Taking a step back, one may also wonder why GPs would use multiple fee structures in the main fund when it can differentiate between investors using co-investments (Phalippou, 2017). A GP may prefer the former if offering co-investment rights to some LPs is costly. This seems plausible in the context of public pensions. For example, many U.S. pensions cannot deploy capital for co-investment (or any investment) without explicit approval from an investment committee or the pension fund's board. These committees and boards often meet infrequently (e.g., once a quarter) and cannot always form a consensus, leading to potentially costly holdup problems from the perspective of the GP. The fact that public pensions do not yet have large co-investment portfolios is consistent with the existence of such holdup costs.

5.4.2 Investor-Specific Mandates

Another reason why the gross-return may deviate for some investors in a fund is what we call investor-specific mandates. One prominent example that has boomed in popularity in recent years are so-called environmental, social, and governance (ESG) restrictions. These restrictions mean that one investor might restrict investment into portfolio companies based on ESG criteria (e.g.,

ship Agreement," despite strong interest from LPs for such rights. The survey further states that "there seems to be some contradiction between the attitudes towards and the actual co-investment activity occurring."

[31] Co-investment by CalPERS – the largest U.S. pension fund – was relatively infrequent prior to 2011, when it launched a dedicated co-investment program (CalSTRS, 2019). The program was suspended in 2016.

firms with large carbon footprints). Any such restrictions will naturally cause returns to differ across investors in the same fund.

To the best of our knowledge, data on investor-specific restrictions are not available for the funds in our sample. However, the National Association of State Retirement Administrators (NASRA) reports that relatively few U.S. pension plans incorporate ESG in their investment process, though some of the larger U.S. pensions have started to do so more in recent years (NASRA, 2018). Motivated by this evidence, we exclude large LPs (those with AUMs over $100 billion) and compute within-fund return standard deviation for funds launched prior to 2010. Internet Appendix Figure IA3 shows that average level of dispersion is marginally lower for this sample of funds and LPs. Assuming this sample is less biased by investor-specific mandates or co-investment, the figure therefore suggests that gross-return differences are not the primary source of within-fund return variation that we observe empirically. We further explore the issue of ESG-mandates below.

5.5 Robustness Sample

In Internet Appendix A.6, we study funds that were raised in a single round and do not purchase oil and gas firms. The first restriction ensures there is no mechanical return dispersion due to differences in commitment timing (Section 5.2.1). The second reduces the likelihood of LP-specific ESG mandates and is based on industry designations from Preqin. Even in this subsample, the majority of funds (66%) continue to have multiple return tiers. Moreover, we show that the average size of management and carry dispersion (Section 3.2), the types of funds and GPs that use multiple fee structures (Section 3.3), and the types of LPs who tend to be top-tier in fees (Section 4.2) are all comparable to what we find in the full sample. These results provide some comfort that our overall characterization of fee dispersion is not confounded by other potential sources of return dispersion. We further discuss how these alternative sources impact the interpretation of our results in Internet Appendix A.7.

6 Implications and Conclusion

6.1 Fees and Foregone Capital

Section 3.2 estimates how much management and performance fees vary in the average fund. This exercise, however, provides an incomplete description of how investment costs vary across LPs. The reason is that realized costs depend on performance. For example, consider a single fund in which one group of LPs pays 20% carry and another pays 30%. The difference in total investment costs between the two groups will be larger if the fund performs well. Thus, as an alternative way to gauge the size of fee dispersion, we now measure how much capital LPs have forgone due to fees on an ex-post basis.

Formally, let r_{pfa} equal the TVPI of investor p in fund f at age a and define $\widetilde{r}_{fa}$ as the average return of LPs who are in the top-tier of fees. As in Section 4.1, top-tier LPs are defined as those with above-median returns. We then compute the amount of forgone capital Γ for each investor as $\Gamma_{pfa} = K_{pfa} \max(\widetilde{r}_{fa} - r_{pfa}, 0)$, where K_{pfa} equals cumulative contributions. Γ can be thought of as the amount of incremental capital each investor would have if it were in the best tier of fees. This capital instead flows to the GP or subsidizes LPs who pay more favorable fees.

To aggregate, we first compute Γ for all investor-fund-age observations. We then group observations based on fund age a, sum the total amount of forgone capital in each group, and scale the total by the sum of contributions in the group. For instance, for $a \in [2,3)$, we retain observations where funds are between their second and third year of life.[32] Within this subset, we then compute Γ and scale it by the total amount of contributions in the subset. This procedure delivers a measure of forgone capital per dollar of contributions for each point in the fund lifecycle. It is similar in spirit to J-curves, which show IRRs at each point in the fund lifecycle. We therefore refer to it as a forgone capital or forgone return curve.

Before proceeding further, there are a few considerations to keep in mind when interpreting our measure of forgone capital Γ. First, it will reflect all potential sources of return dispersion, not just

[32] Funds may therefore be from different vintages but are at the same point in their lifecycle.

fees. Our choice of the benchmark return $\widetilde{r}_{fa}$ should minimize the impact of these other sources, but it is still best to view Γ as an upper bound on true forgone capital. Second, large levels of forgone capital Γ are not necessarily an indication of suboptimal fee choices by LPs. For example, an LP who wants access to a high-skilled GP may find it optimal to pay higher fees, even relative to other investors in the fund.[33] Third, our notion of forgone capital assumes that fund profits are large enough such that all LPs could have feasibly been in the best observed tier of fees (i.e. earn the benchmark return). This seems plausible on its face, but we have no way of confirming it empirically. Finally, for each fund, the benchmark return $\widetilde{r}_{fa}$ is based on the performance of only U.S. public pensions. If other unobserved LPs in the fund (e.g., endowments) pay lower fees ex-post, then we will understate the true amount of forgone capital due to fees.

Figure 8 shows forgone capital curves for several different subsamples. In all plots, we group observations based on deciles of fund age and compute forgone capital per dollar of contributions as described above. The *x*-axis in the plots is the average age for funds in each decile. Panel A shows forgone returns for the full sample of funds. Forgone capital exceeds \$6 per \$100 invested toward the end of the fund lifecycle. It seems to accelerate after funds turn five years old, a pattern that is likely driven by differences in carry taking effect.

Panel B of Figure 8 compares forgone return curves across fund types. Based on our analysis in Section 3.3, we label funds as likely to tier if they are outside of VC, use a placement agent, and are undersubscribed. We then compute the forgone return curve for this set of funds and compare it to all other funds. Unsurprisingly, funds that use multiple fee structures have larger forgone returns throughout the fund lifecycle. Towards the end of their life, forgone capital in these funds approaches \$11 per \$100 of investment, nearly double the amount of forgone capital in other funds. The gap between the two curves accelerates late in the fund lifecycle, again suggesting that performance fees are important for understanding cost heterogeneity in private equity.

Panel C of Figure 8 plots forgone return curves by investor type. We assign LPs into one of

[33] Moreover, LPs who appear to be in a favorable fee tier ex-post may not necessarily have been so ex-ante. For example, it is not clear ex-ante whether a contract with a 1% management and 30% performance fee dominates one with 2% and 20%.

three tiers based on the fraction of fund f_p in which they are a top-tier (low-fee) investor. This measure derives directly from the indicator used in Tables 4 and 5. Bottom-tier LPs are those whose f_p is in the bottom 5th percentile of all investors. Top-tier LPs are those who are in the top 5th percentile. The difference in forgone returns between the two groups is wide, particularly at later stages in the fund cycle. With better fee tier assignment, bottom-tier investors would have earned over \$15 per \$100 invested in late-stage funds, whereas top-tier investors could have earned a modest \$3 per \$100 invested.

Using the core sample, we also construct a single aggregate value of forgone capital. As a reminder, the core sample contains data on a single quarter of data in the later stages of each fund's life (Section 2.2.2). When measuring returns using TVPI, forgone capital equals \$5.6 per \$100 invested or \$25 billion in total. If we instead use DVPI to measure returns, forgone capital equals \$4.3 per \$100 invested or \$19 billion in total. Because DVPI reflects only realized distributions, this is perhaps our most accurate estimate of forgone capital. By comparison, according to data from the Center for Retirement Research at Boston College, the public pensions in our sample paid \$131 billion in total investment expenses from 2001 to 2018. This figure should be viewed as a lower bound because most pensions do not include performance fees when reporting investment expenses. In IRR terms, forgone capital equals 74 bps per year, which is fairly large compared to annual management fees that are on the order of 200 bps.

6.2 Implications for performance benchmarks

The existence of fee dispersion within private equity funds also has potential implications for measures of aggregate performance. The reason why is that vendors like Preqin typically use data from a single LP when reporting fund-level measures of performance. This means that any measure of aggregate performance will necessarily reflect the selection process used by each vendor. Indeed, based on data provided to us directly by Preqin, we find that very large LPs (by AUM) are more likely to be used when reporting fund returns, especially when they commit more capital to a fund.[34]

We gauge the size of these potential selection effects as follows. First, we match each PE and VC fund in our sample to the Preqin performance dataset that is available via Wharton Research Data Services (WRDS), a common data source used for academic research. We focus on PE and VC funds because they are the only ones available on WRDS. Second, for each fund f and quarter q, we compute the minimum and maximum observed DVPI in our dataset, as well as the associated DVPI reported in WRDS. Third, for each fund f, we retain the last available observation that occurs in the ninth year of the fund's life. This facilitates comparison of aggregate performance across vintages and ensures we have an ample amount of funds in each year. Finally, we aggregate returns by taking a weighted average of DVPI across funds in each vintage, where weights are determined by fund size.

Figure 9 shows aggregate DVPI for both PE and VC. The solid lines in the graph correspond to aggregate WRDS-based returns and the dotted lines correspond to the minimum and maximum-based returns. Assuming that all within-fund return dispersion is driven by fees, the graph therefore provides an indication of how fees impact aggregate performance. Panel A shows a meaningful impact in the context of PE. For example, for vintage 2007 funds, aggregate DVPIs based on high-fee LPs, WRDS, and low-fee LPs are 0.95, 1.04, and 1.13, respectively. In stark contrast to PE, Panel B of Figure 9 shows that aggregate VC performance is largely unaffected by within-fund fee dispersion. For 2007 VC funds, aggregate DVPIs based on high-fee LPs, WRDS, and low-fee LPs are 1.18, 1.22, and 1.23, respectively. The contrast between PE and VC is a reflection of the fact that VC funds are far less likely to tier investors (Section 3.3).

It is important to note that the returns in Figure 9 are not directly comparable to returns that have been studied in previous work on aggregate performance. This is because Figure 9 is based on funds that are in both WRDS and our sample, the latter of which contains only funds that have at least two investors. Consequently, it is difficult to compare our aggregate performance statistics

[34] Preqin is 16% ($t = 9.37$) more likely to report returns for very large LPs, define as those with over \$100 billion of total AUM. In terms of commitment size, a 1% increase in fund-share is associated with a 2% ($t = 8.71$) increase in the likelihood of being reported. The unconditional likelihood of selection is 31%.

to those based on other data providers, as the composition of funds across sources is likely to be quite different.[35] We therefore view Figure 9 as providing suggestive evidence that within-fund fee dispersion can, in some cases, distort measures of aggregate performance.

6.3 Conclusion

This paper shows that fees vary within the typical private market investment vehicle, often by meaningful amounts. We further document that fee policies differ across GPs, and depend on the supply and demand of capital at the time of fund raising. Given the size of our sample – we study $438 billion of investments made by 218 public pensions into 2,400 funds managed by 856 GPs – our broad characterization of fees is likely to apply more broadly in private markets, though we may actually understate the size of fee dispersion if other investor types (e.g., endowments) are more favored than public pensions.

Our analysis can be extended in several dimensions. For instance, we find that some pensions have consistently paid higher fees relative to others in their funds. But this does not necessarily imply that these pensions are behaving suboptimally, since some pensions may have traded low management fees for higher carry and simply gotten unlucky ex-post. One way to address this issue is to apply the logic of our estimator for average within-fund dispersion in management and carry at the fund level. With fund-level estimates of fee parameters, it should be possible to determine whether some pensions chose truly dominated contracts ex-ante. A related question is whether some pensions optimally pay higher relative fees to gain access to more skilled GPs. With additional data, it would also be interesting to study whether favorable fee schedules are compliments or substitutes to co-investment rights. Overall, these sorts of analyses would allow for a careful decomposition of within-fund fee dispersion into supply-side (e.g., cost-based pricing) and demand-side (e.g., LP search frictions) factors, which is critical for understanding the welfare and policy implications of fee dispersion in private markets.

[35] For instance, there are 838 funds in the WRDS-Preqin data that have at least nine years of data but are not contained in our sample.

References

Allen, J., R. Clark, and J.-F. Houde (2019). Search frictions and market power in negotiated-price markets. *Journal of Political Economy 127*(4), 1550–1598.

Andonov, A., Y. V. Hochberg, and J. D. Rauh (2018). Political Representation and Governance: Evidence from the Investment Decisions of Public Pension Funds. *The Journal of Finance*.

Beath, A., C. Flynn, and J. MacIntosh (2014). How implementation style and costs affect private equity performance. *Rotman International Journal of Pension Management 7*(1), 1–5.

Begenau, J., C. Robles-Garcia, E. N. Siriwardane, and L. Wang (2020). An empirical guide to investor-level private equity data from preqin. *Working paper*.

Berk, J. B. and R. C. Green (2004). Mutual fund flows and performance in rational markets. *Journal of political economy 112*(6), 1269–1295.

Bester, H. (1988). Bargaining, search costs and equilibrium price distributions. *The Review of Economic Studies 55*(2), 201–214.

Bowden, A. J. (2014, May). Spreading sunshine in private equity. https://www.sec.gov/news/speech/2014–spch05062014ab.html.

Braun, R., T. Jenkinson, and C. Schemmerl (2019). Adverse selection and the performance of private equity co-investments. *Journal of Financial Economics*.

Brown, G. W., O. R. Gredil, and S. N. Kaplan (2019). Do private equity funds manipulate reported returns? *Journal of Financial Economics 132*(2), 267–297.

Burdett, K. and K. L. Judd (1983). Equilibrium price dispersion. *Econometrica: Journal of the Econometric Society*, 955–969.

CalSTRS (2019). *CalSTRS Investment Committee Report*. CalSTRS.

Casella, G. (1992). Illustrating empirical bayes methods. *Chemometrics and intelligent laboratory systems 16*(2), 107–125.

Clayton, W. W. (2020). The private equity negotiation myth. *Yale Journal on Regulation 37*, 67.

Da Rin, M. and L. Phalippou (2017). The importance of size in private equity: Evidence from a survey of limited partners. *Journal of Financial Intermediation 31*, 64–76.

Duffie, D., N. Gârleanu, and L. H. Pedersen (2005). Over-the-counter markets. *Econometrica 73*(6), 1815–1847.

Dyck, A., P. Manoel, and A. Morse (2018). Outrage and compensation in public pension performance. Working Paper.

Eisfeldt, A., B. Herskovic, S. Rajan, and E. N. Siriwardane (Forthcoming). Otc intermediaries. *Review of Financial Studies*.

Fang, L., V. Ivashina, and J. Lerner (2015). The disintermediation of financial markets: Direct investing in private equity. *Journal of Financial Economics 116*(1), 160 – 178.

Gabaix, X. and D. Laibson (2006). Shrouded attributes, consumer myopia, and information suppression in competitive markets. *The Quarterly Journal of Economics 121*(2), 505–540.

Goldberg, P. K. (1996). Dealer price discrimination in new car purchases: Evidence from the consumer expenditure survey. *Journal of Political Economy 104*(3), 622–654.

Gompers, P. and J. Lerner (1999). An analysis of compensation in the u.s. venture capital partnership. *Journal of Financial Economics 51*(1), 3 – 44.

Gompers, P. A. and J. Lerner (2010). What drives venture capital fundraising? *Brookings Papers on Economic Activity*, 149.

Grennan, M. (2013). Price discrimination and bargaining: Empirical evidence from medical devices. *American Economic Review 103*(1), 145–77.

Hortaçsu, A. and C. Syverson (2004). Product differentiation, search costs, and competition in the mutual fund industry: A case study of s&p 500 index funds. *The Quarterly Journal of Economics 119*(2), 403–456.

ILPA (2019). Ilpa principles 3.0: Fostering transparency, governance and alignment of interests for general and limited partners.

Ivashina, V. and J. Lerner (2018). Looking for alternatives: Pension investments around the world, 2008 to 2017. Working Paper.

Jain, A. K., M. N. Murty, and P. J. Flynn (1999). Data clustering: a review. *ACM computing surveys (CSUR) 31*(3), 264–323.

Lerner, J., J. Mao, A. Schoar, and N. R. Zhang (2018). Investing outside the box: Evidence from alternative vehicles in private capital. MIT Sloan working paper.

MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. In *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, Volume 1, pp. 281–297. Oakland, CA, USA.

Markham, I. (2017). *Bain keeps two investor class structure for Fund XII*. Private Funds CFO.

McKinsey (2021). Mckinsey global private markets review 2021. Technical report, McKinsey & Company.

Metrick, A. and A. Yasuda (2010). The economics of private equity funds. *The Review of Financial Studies 23*(6), 2303–2341.

Morgan Lewis (2015). *Venture Capital & Private Equity Funds, Deskbook Series*. Morgan Lewis.

Morris, C. N. (1983). Parametric empirical bayes inference: theory and applications. *Journal of the American statistical Association 78*(381), 47–55.

Nadauld, T. D., B. A. Sensoy, K. Vorkink, and M. S. Weisbach (2019). The liquidity cost of private equity investments: Evidence from secondary market transactions. *Journal of Financial Economics 132*(3), 158–181.

NASRA (2018). *ESG - Environmental, Social And Governance*. NASRA.

Phalippou, L. (2017). *Private Equity Laid Bare*. CreateSpace Independent Publishing Platform.

Phalippou, L., C. Rauch, and M. Umber (2018). Private equity portfolio company fees. *Journal of Financial Economics 129*(3), 559–585.

Preqin (2014). *The State of Co-Investments*. Preqin.

PSERS (2020, March). Placement agent policy.

Robbins, W. (2019, November). Client alert: Ilpa releases model limited partnership agreement. Online.

Robinson, D. T. and B. A. Sensoy (2013, 11). Do Private Equity Fund Managers Earn Their Fees? Compensation, Ownership, and Cash Flow Performance. *The Review of Financial Studies 26*(11), 2760–2797.

Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics 20*, 53–65.

Salop, S. and J. Stiglitz (1977). Bargains and ripoffs: A model of monopolistically competitive price dispersion. *The Review of Economic Studies 44*(3), 493–510.

SEC (2015). *Standardized Private Equity Fee Disclosure*. SEC.

Sorensen, A. T. (2000). Equilibrium price dispersion in retail markets for prescription drugs. *Journal of Political Economy 108*(4), 833–850.

Steinhaus, H. (1956). Sur la division des corps matériels en parties. *Bull. Acad. Polon. Sci 1*(804), 801.

Tibshirani, R., G. Walther, and T. Hastie (2001). Estimating the number of clusters in a data set via the gap statistic. *Journal of the Royal Statistical Society: Series B (Statistical Methodology) 63*(2), 411–423.

Toll, D. M. and P. Centopani (2017). Pe/vc partnership agreements study, 2018-2019.

Zuckerman, G. and A. Or (2011). *Bain Capital Lowers Its Fees*. WSJ.

Table 1: Summary Statistics for Core Sample

Panel A: Funds, GPs, and LPs by Asset Class

	Full Sample	Infrastructure	Private Debt	Private Equity	Real Estate	Venture Capital
Funds	2,400	70	272	964	496	598
GPs	856	36	108	349	194	248
LPs	218	82	127	194	147	117
N	9,830	306	1,215	4,465	1,888	1,956

Panel B: Core Sample Characteristics

	Mean	Stdev	Min	p25	p50	p75	Max
Fund Age (years)	9	6	-1	4	8	12	27
Commitment ($ mm)	55	82	0	12	30	74	1,600
Percent of Fund	5.0	6.8	0.0	1.0	2.7	6.5	99.9
Investors per Fund	4	3	2	2	3	5	29
AUM ($ bn)	23.84	39.11	0.05	2.49	8.81	28.35	354.00

Notes: Panel A reports the number of unique funds, GPs, LPs, and total observations in the core sample, as well as breakdowns by asset class. The core sample contains a single cross-section of LPs that are observed on the same date in each fund and is therefore unique at the LP-fund level (see Section 2.2.2). Panel B reports summary statistics for the core sample. Fund age is the number of years since each fund's final close date. Percent of fund is defined as commitment size over total fund size. AUM is the total assets under management for each LP and is measured in the vintage year of each fund. DVPI is defined as cumulative distributions divided by contributions. TVPI equals DVPI plus the reported liquidation value of any remaining investments in the fund, scaled by cumluative contributions. Internal rates of return (IRRs) are reported by Preqin and are missing for 20% of the 9,830 observations in the core sample.

Table 2: Within-Fund Dispersion in Fixed and Performance Fees

	Mgmt (bps)		Carry (%)		Carry Placebo		
	m	$se(m)$	c	$se(c)$	u	$se(u)$	p-value
Infrastructure	103	22	5.5	2.7	4.2	2.5	0.10
Private Debt	99	18	6.8	1.6	1.7	1.6	0.30
Private Equity	83	7	3.3	0.8	0.4	0.6	0.50
Real Estate	73	10	6.4	1.5	1.2	1.1	0.27
Venture Capital	42	7	0.5	0.3	-0.4	0.6	0.48

Notes: This table presents estimates of the standard deviation in fixed fees (e.g., management fee, m) and performance-based fees (e.g., carry rates c) for the average fund in each asset class indicated by the row label, along with a placebo test (u) of our carry estimator in a subsample of unprofitable funds. Dispersion in fixed fees (m) is estimated via the following regression: $p^{\sigma}_{ft} = a + m \times age_{ft} + \varepsilon_{ft}$, where p^{σ}_{ft} is the within-fund standard deviation in fund f's capital call rate and age_{ft} is its age in years at time t. The call rate equals the fraction of committed capital that has been called for investment. We estimate the regression for funds that are not older than five years. Within the set of profitable funds, dispersion in performance-based fees (c) is estimated via the following regression: $d^{\sigma}_{ft} = \alpha + c \times \widetilde{r}_{ft} + \varepsilon_{ft}$, where d^{σ}_{ft} is the within-fund standard deviation of distributions-to-commitments and $\widetilde{r}_{ft}$ is the within-fund maximum of TVPI for fund f at time t. α is a set of fixed effects for fund vintage. The placebo test for carry estimates the same regression $d^{\sigma}_{ft} = \alpha + u \times \widetilde{r}_{ft} + \varepsilon_{ft}$ in the subset of unprofitable funds where carry dispersion should not be detectable ($u = 0$). We define profitable funds as those with: (i) a TVPI above 1.09 and (ii) an IRR above 9%. All regressions are weighted by the average number of investors in each fund. The columns under the "Carry Placebo" header report u, the standard error of u, and the p-value from the test of the null that $u = 0$. In all cases, standard errors are clustered by fund. All estimates are based on funds in the master sample (see Section 2.2.2) that were determined to have multiple investor tiers by the Silhouette score approach (Rousseeuw, 1987) described in Section 3.1.

Table 3: Characteristics of Funds that Use Multiple Fee Structures

	Dependent Variable: $100 \times 1(\text{Tiers}_f > 1)$						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Quartile of GP's Prior Funds			4.21**				
			(2.24)				
Undersubscribed Fund				12.04**			
				(4.83)			
Fund Number 1-3					12.95**		
					(6.63)		
Uses Placement Agent						12.76**	
						(6.83)	
Infrastructure							15.05**
							(4.45)
Private Debt							5.00*
							(1.86)
Real Estate							5.98**
							(2.75)
Venture Capital							-33.04**
							(-13.36)
GP FE	x						
Law-Firm FE		x					
p(FEs = 0)	0.00	0.00					
Adjusted R^2	0.31	0.10	0.05	0.05	0.04	0.04	0.15
N	1,916	950	1,163	1,498	2,260	2,256	2,260

Notes: This table reports OLS estimates from a linear probability model of the likelihood that a fund has multiple investor tiers. We define an indicator variable for whether a fund has multiple tiers based on the k-means clustering analysis and Silhouette score approach (Rousseeuw, 1987) described in Section 3.1. The dependent variable in the regression is this indicator variable multiplied by 100. For each fund f run by GP g, the quartile of g's prior funds is defined as the average performance quartile of all g's funds that were raised at least four years before f's final close. Quartile rankings are from Preqin and measured as of the time of close and higher values correspond to worse performance, i.e., the bottom quartile is coded as 4 while the top quartile is coded as 1. Funds are undersubscribed if their final close size is below their target fund size. Number of funds raised by the GP is measured as of the time of f's close and includes the current fund. Use of placement agent is based on information from Preqin. We assume missing entries mean no placement agent is being used. Asset class designations are also from Preqin. Columns (1) and (2) respectively include only a GP fixed effect and a law-firm fixed effect. The regressions in column (3)-(7) include: (i) a fixed effect based on the average number of investors in the fund over its lifetime, rounded to the nearest integer; and (ii) a fixed effect based on the decile of the fund's size. In column (7), buyout funds are the omitted asset class in the regression. Heteroskedasticity-robust t-statistics are reported below point estimates.

Table 4: Do Some LPs Consistently Pay Lower Fees?

	LP Effects					LP-GP Effects				
Age Min.	F	p	p^*	K	N	F	p	p^*	K	N
1	3.67	<0.01	<0.01	182	7,210	1.68	<0.01	<0.01	1,580	4,460
4	3.53	<0.01	<0.01	159	5,487	1.71	<0.01	<0.01	1,142	3,084
8	2.60	<0.01	<0.01	135	3,518	1.75	<0.01	<0.01	678	1,789

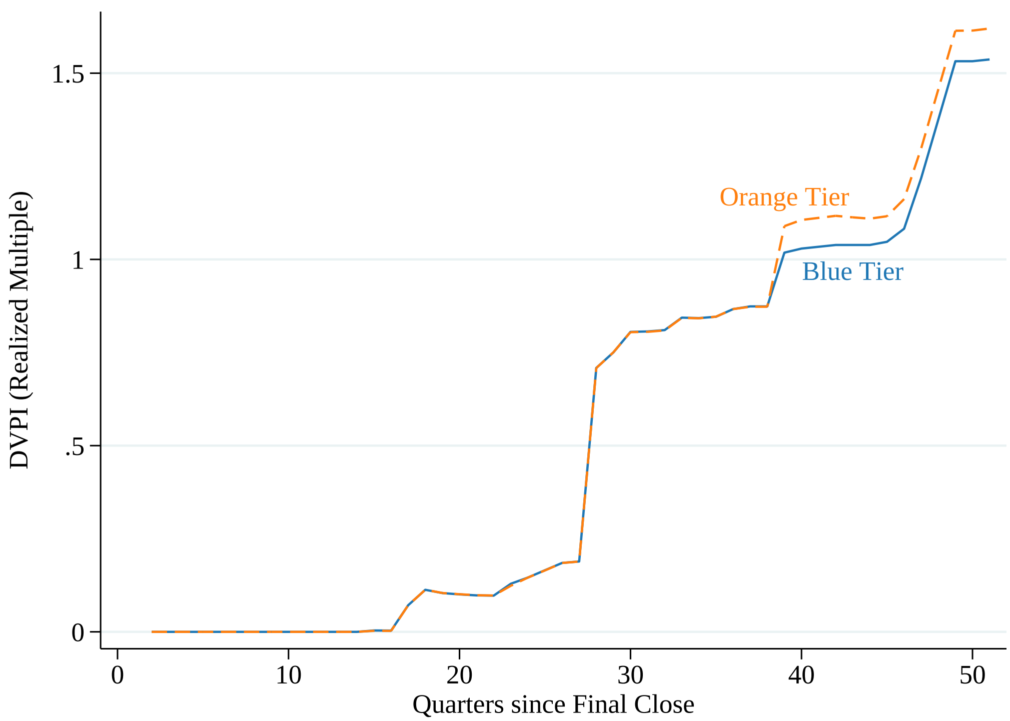
Notes: This table is based on the following regression: $y_{pf} = \lambda_a + \alpha_k + \varepsilon_{pf}$, where y_{pf} is an indicator variable that equals 1 if p has above median returns in fund f. We determine whether p has above median returns in fund f based on whether it is above median on average over the life of the fund. Returns are measured using TVPI. λ_a are fixed effects based on vigintiles of fund f's age. $\alpha_k = \alpha_p$ are fixed effects for LPs (p) for the results in the columns on the left and $\alpha_k = \alpha_{gp}$ are LP-GP fixed effects for the results in the columns on the right. The table shows the F-statistic, the p-value, and a nonparametric p-value of the null hypothesis that the α_p jointly equal zero. To generate the nonparametric p-value (p^*), we randomly assign return paths within each fund, compute y, run the regression, and retain the F-statistic. We do so 500 times then generate p^* by comparing the actual F-statistic to the simulated distribution of F-statistics. We repeat the analysis using DVPI and IRR to measure returns in Internet Appendix C.6.2. All estimates are based on funds in the core sample (see Section 2.2.2) that were determined to have multiple investor tiers by the Silhouette score approach (Rousseeuw, 1987) described in Section 3.1.

Table 5: The Determinants of Top-Tier LPs

Dependent Variable:	$100 \times 1(p \text{ is Top-Tier in } f)$		
	(1)	(2)	(3)
Percent of Fund	0.96**	0.73**	0.85**
	(4.80)	(3.16)	(3.12)
Large Pension (AUM)	13.97**	16.67**	15.64**
	(4.97)	(5.10)	(3.78)
LP-GP Fund Count	0.71**	0.52	0.71
	(2.25)	(1.30)	(1.13)
Quartile of LP's Prior Funds	-9.51**	-10.89**	-12.52**
	(-3.31)	(-3.34)	(-2.93)
Early PE Investor	8.44**	9.43**	6.75*
	(3.44)	(3.24)	(1.83)
Elected Board Members (%)	0.12**	0.16**	0.25**
	(2.79)	(3.04)	(3.44)
Board Size	0.61**	0.34	-0.10
	(2.80)	(1.29)	(-0.28)
Fund Age Min. (yrs)	1	4	8
R^2	0.11	0.11	0.13
N	5,050	3,438	1,688

Notes: This table reports OLS estimates from a linear probability model of the likelihood that an investor p is a top-tier investor in fund f. An investor is defined as top-tier if it has above-median returns (TVPI) in fund f for the majority of its life. This indicator variable is multiplied by 100 in the regression. Percent of fund is defined as p's commitment relative to total fund size. Large Pension (AUM) is an indicator if p has AUM over \$100 billion. LP-GP Fund Count is the number of funds between p and the manager of fund f, measured over our full sample. The variable Quartile of LP's Prior Funds measures the average performance quartile ranking of p's funds that were active at the time of fund f's close, conditional on at least four years of performance history. Quartile rankings come from Preqin and are measured as of each fund's close date. The bottom quartile is coded as 4, meaning higher values correspond to worse performance. We define an indicator variable for whether p was an early investor in private markets if its first entry into the dataset is before 2008. Elected Board Members (%) is the percent of p's board that is elected by members or the general public, measured at the time of f's close. Board size equals the number of board members at the same point in time. All regressions include a fund fixed effect and fixed effects based on vigintiles of fund age. t-statistics are reported below point estimates and are based on standard errors that are clustered within each investor-vintage cell. All estimates are based on funds in the core sample (see Section 2.2.2) that were determined to have multiple investor tiers by the Silhouette score approach (Rousseeuw, 1987) described in Section 3.1.

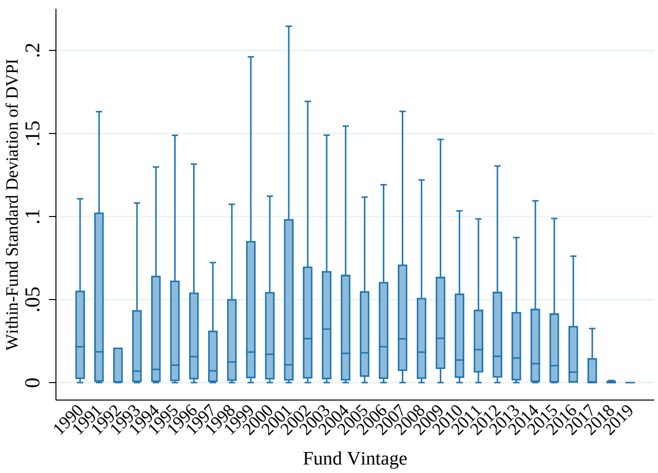
Figure 1: Example Cashflow Profiles of Investors in the Same Fund



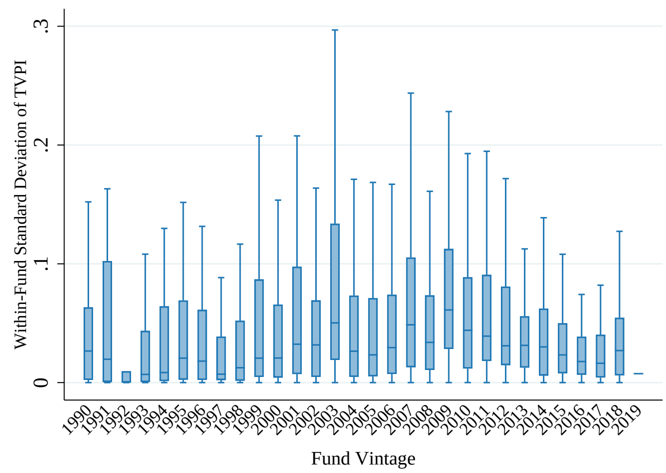
Notes: This plot shows the time-series evolution of DVPI for two anonymous investors in the same fund. DVPI is defined as the cumulative amount of distributions, scaled by the cumulative amount of contributions. We are not able to identify individual funds or investors per our data-sharing agreement with Preqin.

Figure 2: The Prevalence of Within-Fund Dispersion in Net-of-Fee Returns

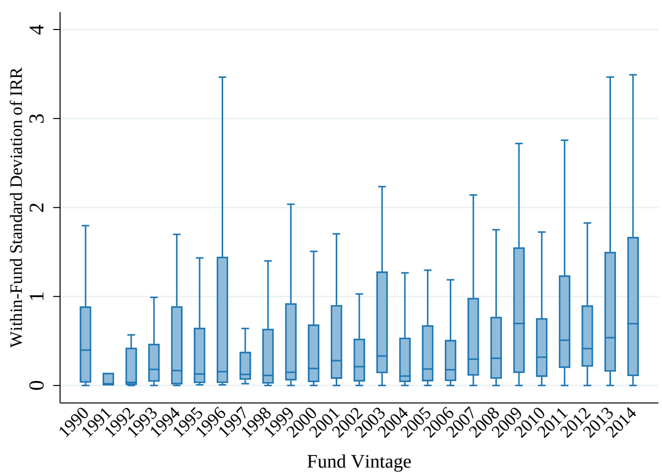
Panel A: DVPI (Realized Multiple)



Panel B: TVPI (Total Multiple)

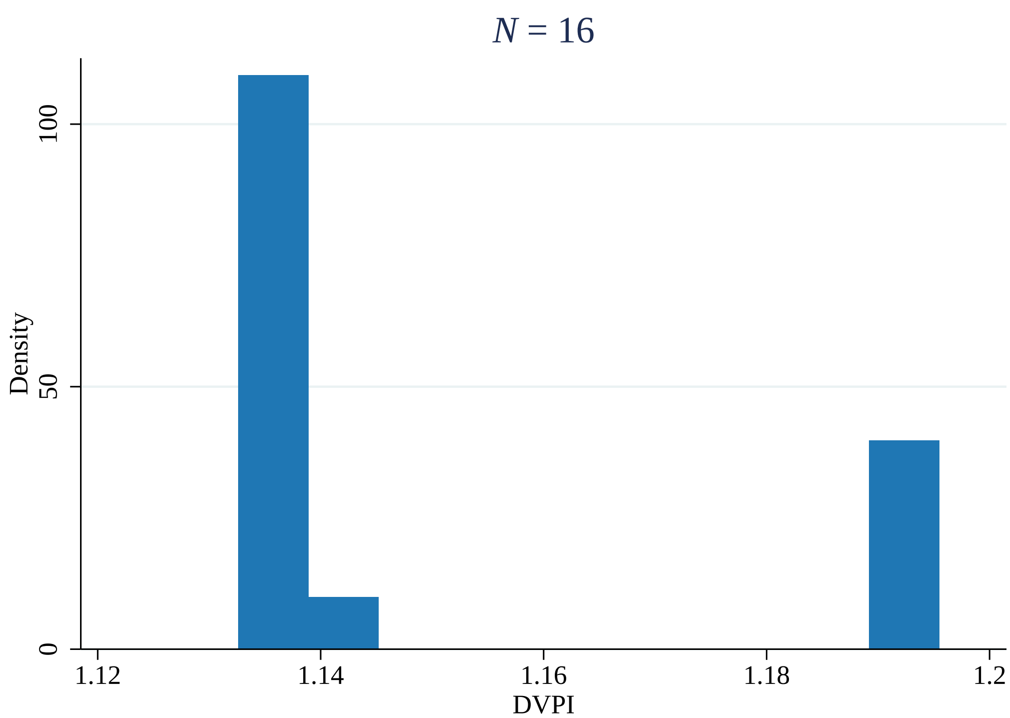


Panel C: Reported IRR



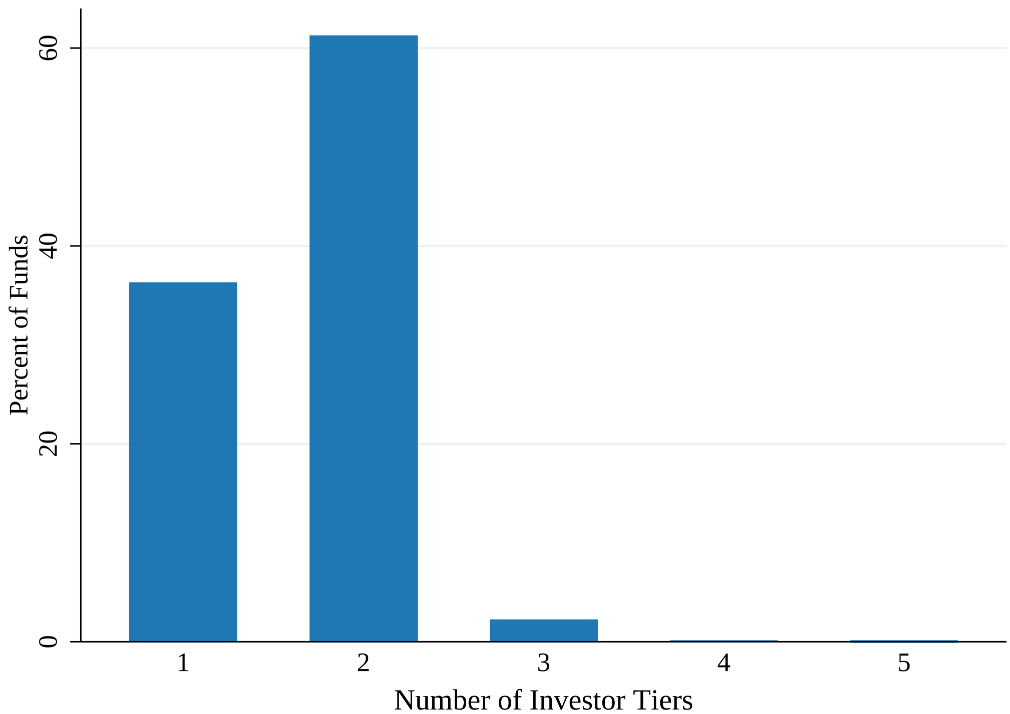
Notes: This plot shows the *across-fund* distributions of the *within-fund standard deviation* of net-of-fee returns across funds by fund vintage. To construct the plot, we first compute the standard deviation (or dispersion) σ_f of returns within each fund f in the core sample. See Section 2.2.2 for the definition of the core sample. Positive values of σ_f indicate that returns within fund f differ across investors. We then plot the distribution of σ_f across funds, broken out by fund vintage. Each panel of the plot shows the across-fund distribution of σ_f for different return measures. Panel A uses DVPI, which is defined as the cumulative amount of distributions, scaled by the cumulative amount of contributions. Panel B uses TVPI, which is defined as the cumulative amount of distributions plus any remaining net-asset-value, scaled by the cumulative amount of contributions. And Panel C uses IRRs, which are those reported by Preqin and are expressed in percentage points. For IRRs, we exclude vintage years after 2015 since IRRs are typically unstable early in a fund's life-cycle.

Figure 3: Example of Clustering in DVPI



Notes: This plot shows the distribution of net-of-fee returns as measured by DVPI for 16 investors in the same fund at a fixed point in time. The fund was closed in a single fundraising round, which means all investors entered it at the same time. We are not able to identify individual funds or investors per our data-sharing agreement with Preqin.

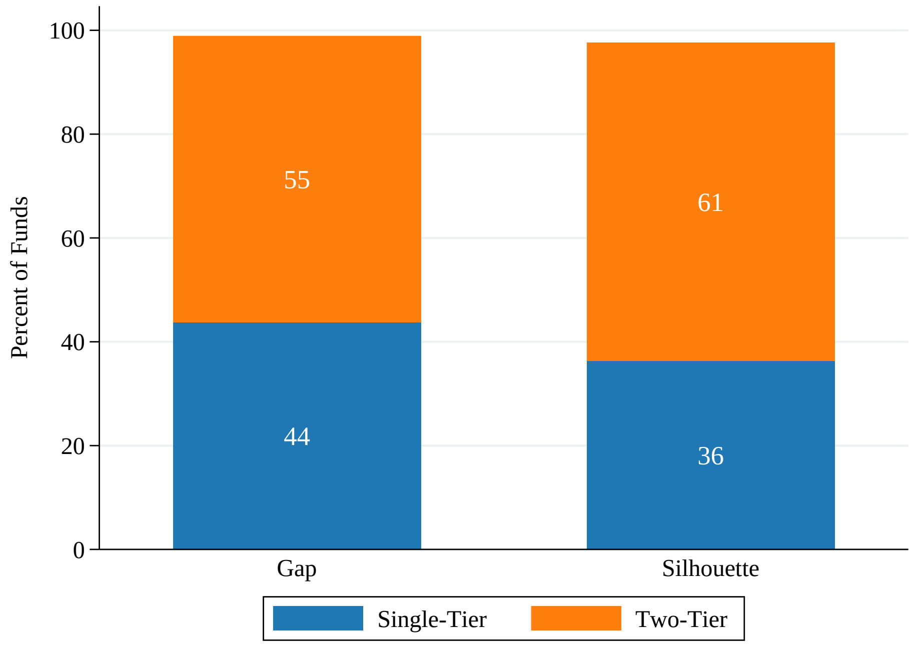
Figure 4: Distribution of Return Cluster Counts Across Funds



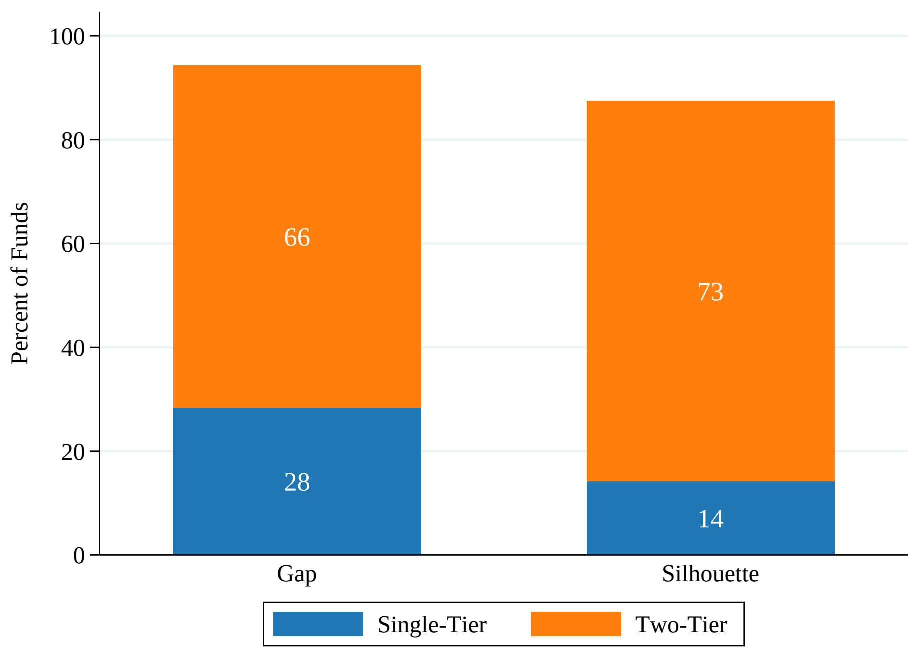
Notes: This plot shows the distribution of within-fund clusters (investor tiers) in net-of-fee returns (DVPI) across funds. For each fund f and date t, we compute the number of clusters in DVPI using a k-means clustering analysis, where the number of clusters is chosen based on Silhouette scores (Rousseeuw, 1987). The number of clusters at the fund level is defined as the average number of clusters over each fund's life, rounded to the nearest integer. See Section 3.1 for more details.

Figure 5: Robustness of Clustering Analysis

Panel A: All Funds

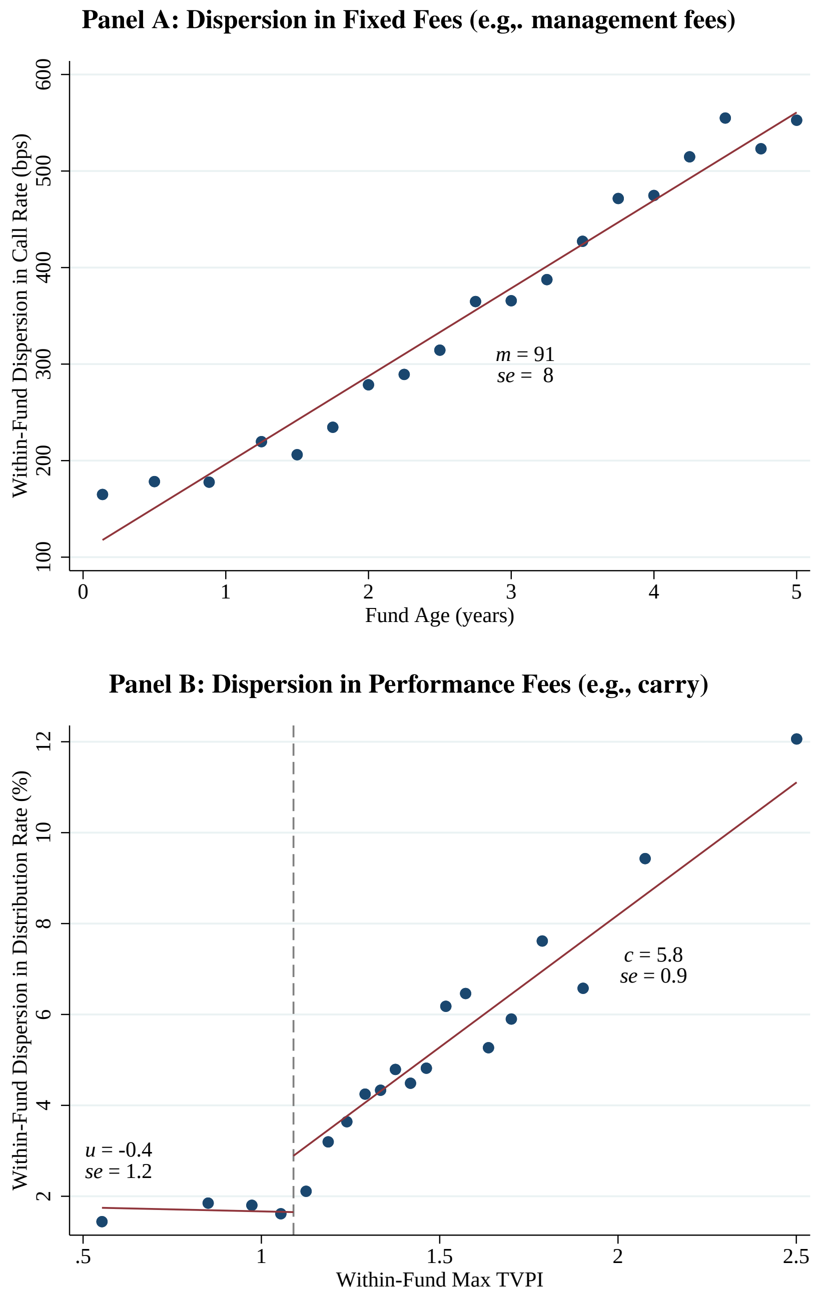


Panel B: Funds with at least five investors



Notes: This plot shows the fraction of funds with either one or two net-of-fee return clusters, where the number of clusters is determined by the Silhouette method (Rousseeuw, 1987) or the Gap statistic (Tibshirani et al., 2001). Returns are measured using DVPI. Panel A is based on all funds in the master sample and Panel B is based on those with at least five investors. See Section 2.2.2 for the master sample definition and Section 3.1 for more details on the clustering methods.

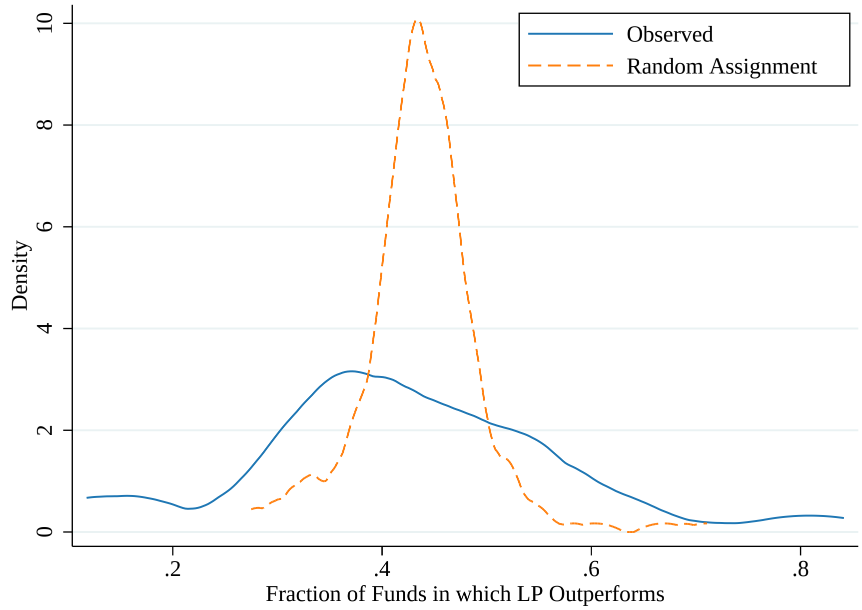
Figure 6: Estimating Within-Fund Fee Dispersion



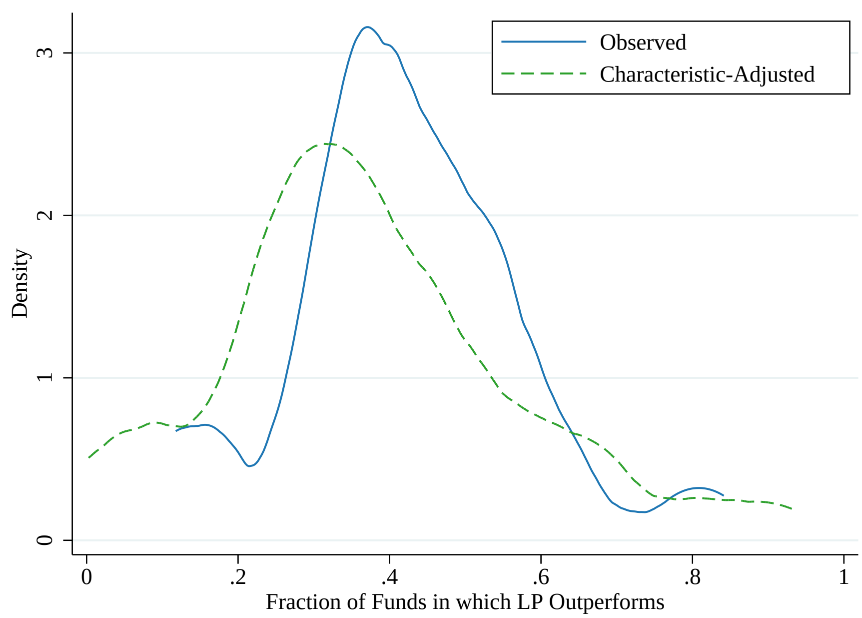
Notes: This plot depicts how we estimate dispersion in fixed fees (e.g., management, m) and performance-based fees (e.g., carry, c) within the average fund. **Panel A** shows a binned scatter plot of the within-fund standard deviation of call rates for fund f at time t against f's age at the same date. The call rate equals the fraction of committed capital that has been called for investment. Panel A is made using all data on funds whose age is at most five years old. **Panel B** is a binned scatter plot where the y-axis is the within-fund standard deviation of distribution rates for f at time t. Distribution rate is the percent of distributions relative to commitment amount. The x-axis of the plot is f's maximum TVPI at time t. Both variables are shown after partialing out vintage fixed effects. Funds must have more than five investors at time t to be included. The vertical dotted line in Panel B marks the boundary of funds with a TVPI of 1.09, a proxy for those that are profitable enough to charge carry. On either side of the boundary, we report the slope of the line of best fit and its standard error. In both panels, standard errors are clustered by fund and estimates are based on funds in the master sample (see Section 2.2.2) that were determined to have multiple investor tiers by the Silhouette score approach (Rousseeuw, 1987) described in Section 3.1. See Section 3.2 and Table 2 for more estimation details.

Figure 7: The Distribution of Pension Effects

Panel A: Observed Data vs Random Assignment Model

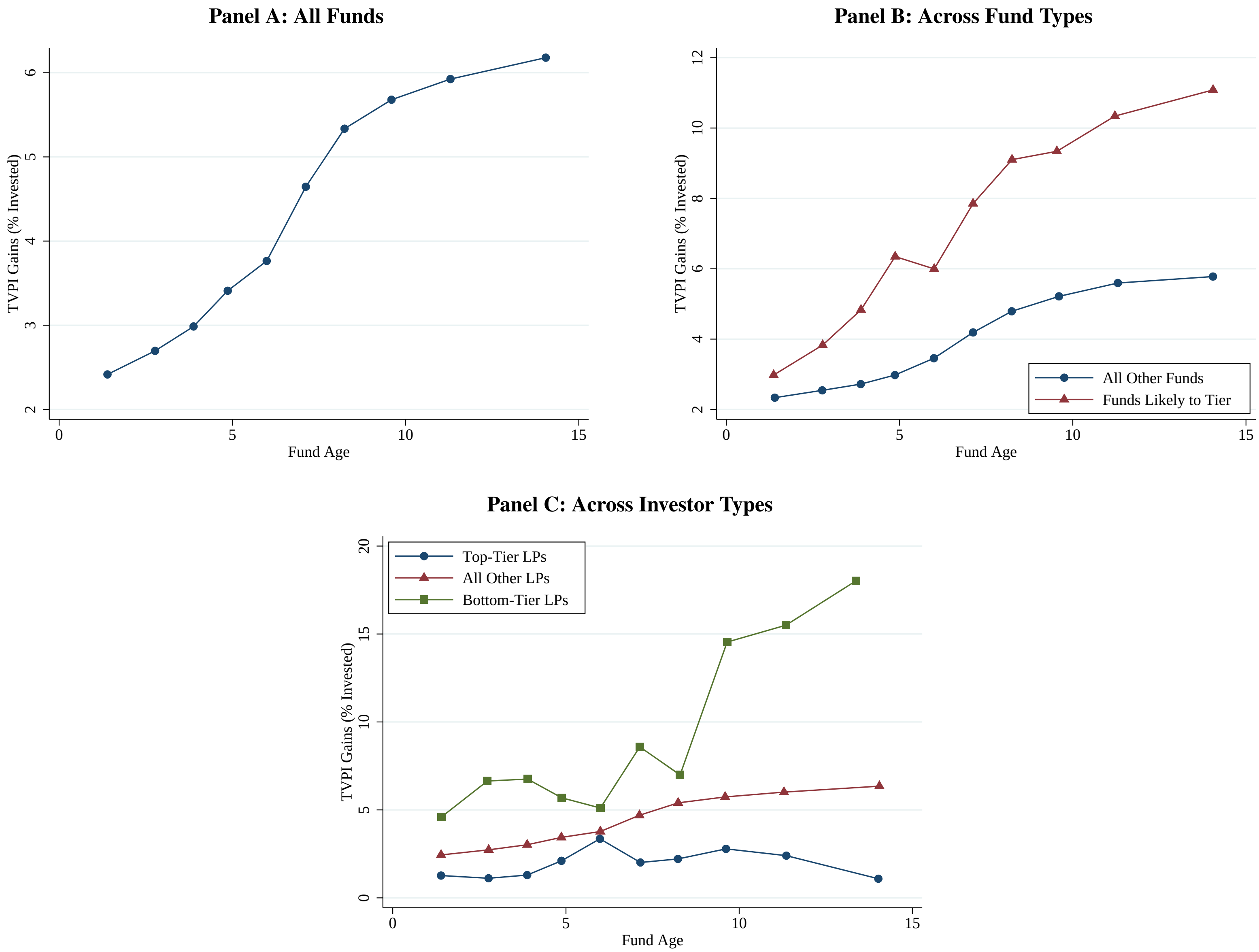


Panel B: Before and After Characteristic-Adjustment



Notes: This plot shows the fraction of funds in which a pension outperforms other investors ("pension effects"). The blue line in Panel A is created by regressing an indicator for whether pension p earns above-median returns in fund f, y_{pf}, on fixed effects for pension and age vigintile. The estimated pension fixed effects are then shrunk towards their mean using an empirical Bayes estimate and shown in blue. The orange line shows simulated pension effects based on the random assignment of contracts to pensions in each fund (see Section 4.1.1). In Panel B, we evaluate how observable characteristics account for the observed pension effects. To do so, we regress returns r_{pft} of pension p in fund f at time t on a vector of characteristics and fixed effects for fund-date. We use the residuals from the regression to determine characteristic-adjusted above-median return status in each fund $\tilde{y}_{pf}$, re-estimate pension effects, and apply the empirical Bayes procedure. The resulting characteristic-adjusted pension effects are plotted in green, alongside the pension effects before adjustments. Returns are measured using TVPI. See Sections 4.1.1 and 4.2.2 for more details.

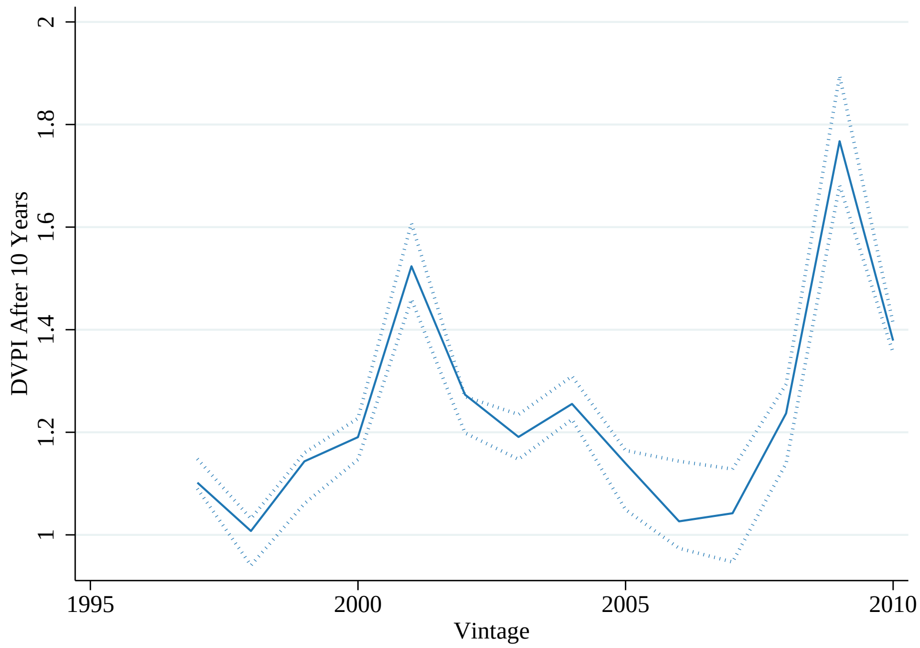
Figure 8: Forgone Return Curves



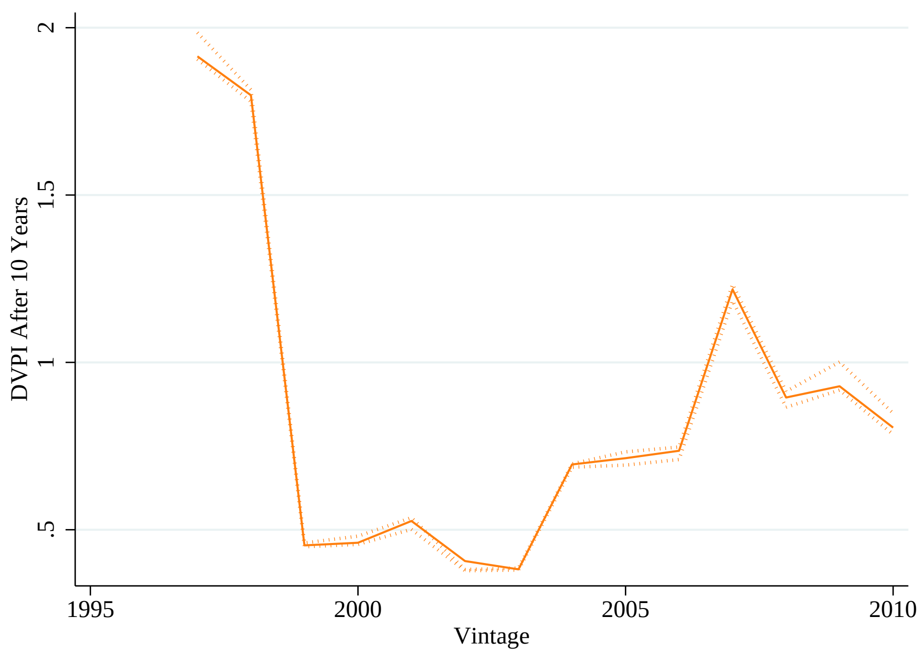
Notes: This plot shows forgone capital per dollar of contributions over the fund lifecycle. Formally, let r_{pfa} equal the TVPI of investor p in fund f at age a and define $\widetilde{r}_{fa}$ as the average return of LPs who are in the top-tier of fees. As in Section 4.1, top-tier LPs are defined as those with above-median returns for the majority of a fund's life. Forgone capital for each investor is defined as $\Gamma_{pfa} = K_{pfa} \max(\widetilde{r}_{fa} - r_{pfa}, 0)$, where K_{pfa} equals cumulative contributions. To aggregate, we group observations based on declines of fund age a, sum the total amount of forgone capital Γ in each group, and scale the total by the sum of contributions in the group. We then plot the resulting amount of forgone capital per dollar of contributions against the average fund age in each decile. Panel A is based on all funds. Panel B focuses on funds that are likely to tier, defined as non-VC funds that are undersubscribed and use a placement agent. Panel C divides investors into one of three buckets based on the fraction of funds in which they are a top-tier investor. See Section 6.1 for more details.

Figure 9: The Impact of Fees on Aggregate Performance

Panel A: Private Equity



Panel B: Venture Capital



Notes: This plot shows aggregate performance based on different reporting conventions within funds. The sample of funds used in the plot are those that appear in both our dataset and the Preqin dataset on WRDS, conditional on the fund having complete data in its ninth year of life. For each fund, we consider three different measures of returns: (i) the minimum observed return in the fund, r_f^{min}; (ii) the maximum observed return in the fund, r_f^{max}; or (iii) the return reported on WRDS, r_f^W. In all cases, we use DVPI and measure returns as of the last available quarter in each fund's ninth year of life. For each return measure and vintage year, we compute aggregate returns using a fund size-weighted average across all funds in each vintage. The solid line in the plot shows the aggregate return based on the data from WRDS. The shaded lines correspond to aggregate returns based on either the minimum and maximum observed return in each fund.