NBER WORKING PAPER SERIES

PORTFOLIO MANAGEMENT IN PRIVATE EQUITY

Gregory W. Brown
Celine Yue Fei
David T. Robinson

Working Paper 31664
http://www.nber.org/papers/w31664

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA02138
September 2023

The authors thank the Institute for Private Capital and Burgiss for support of this project, including access to the primary data set utilized in the analysis, and Wendy Hu for providing research support. We also thank William Cong, Peter Cornelius, Olivier Darmouni, Irem Demirci, Paolo Fulghieri, Will Gornall (discussant), Oleg Gredil (discussant), Arpit Gupta (discussant), Lu Han, Ulrich Hege, Gerard Hoberg, Philip Howard, Niklas Huether, Viktoria Ivashina, Arthur Korteweg (discussant), Peng (Peter) Liu, Tong Liu, Erik Loualiche, Victor Lyonnet, Song Ma, Nadya Malenko, Filippo Mezzanotti, Gordon Phillips, Andrea Rossi, Albert Sheen, Morten Sorensen (discussant), Christophe Spaenjers (discussant), Andreas Stathopoulous, Eva Steiner, Daniel Urban, Neng Wang, Michael Weisbach, Mark Westerfield (discussant), Ting Xu, Ayako Yasuda as well as participants of the 2023 American Finance Association (AFA) Annual Meeting, 2023 Financial Intermediation Research Society (FIRS) Annual Meeting, 2023 Columbia Private Equity Conference, 2023 Mitsui Finance Symposium, 2023 SAIF Summer Institute of Finance Conference, 2023 Lapland Investment Fund Summit, 2023 SKEMA Corporate Restructuring Conference, 2022 USC Finance, Organizations and Markets Conference, 2022 WashU St. Louis Annual Conference on Corporate Finance, 2022 UNC Annual Private Equity Research Consortium, 2022 Conference on Financial Economics and Accounting, 2022 Society of Quantitative Analysts, 2021 IPC Spring Research Symposium conference, seminars at Case Western Reserve University, University of St. Gallen, Institute of Financial Studies at Southwestern University of Finance and Economics, Tulane University, and Shanghai Tech University for useful comments and suggestions. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Portfolio Management in Private Equity
Gregory W. Brown, Celine Yue Fei, and David T. Robinson
NBER Working Paper No. 31664
September 2023
JEL No. G10,G11,G2,G20,G24

ABSTRACT

General Partners (GPs) in private equity face a trade-off between focusing their skills and effort on fewer investments to earn higher returns, or investing more broadly to reduce risk through diversification. Using a novel, deal-level dataset of 5,925 global investments from 1999 to 2016, we show that these portfolio considerations are important for understanding fund-level private equity returns. The largest investments in PE funds typically have the lowest returns on average, but are also the least risky. Returns and risk are both increasing in industry or geographic concentration. And while GP-specific return variation (e.g., skill) only accounts for 4%-6% of the total return variation of a typical investment, it accounts for around 40% of the return variation at the fund level. These findings show that GPs use portfolio construction, and not just deal selection, to seek risk-adjusted fund-level returns.

Gregory W. Brown
University of North Carolina at Chapel Hill
Kenan-Flagler Business School
CB 3440 Kenan Center #301
Chapel Hill, NC 27599
gregwbrown@unc.edu

Celine Yue Fei
University of North Carolina at Chapel Hill
300 Kenan Center Drive, Kenan Center
Chapel Hill, NC 27599
Celine_Fei@kenan-flagler.unc.edu

David T. Robinson
Fuqua School of Business
Duke University
100 Fuqua Drive
Durham, NC 27708
and NBER
davidr@duke.edu

1 Introduction

The private equity industry has grown over the last forty years from a relatively niche investment segment to a $6 trillion asset class. It is now a major component of the portfolios of many insurance companies, pension funds, university endowments, and sovereign wealth funds. Accordingly, many industry and academic observers have approached private equity performance measurement through the lens of the *investor's* portfolio management decision, asking, for example, how a retirement system might optimally add private equity commitments to a portfolio, including public equities, fixed income, and other asset classes (see Korteweg and Westerfield (2022) for a review and also Gredil, Liu, and Sensoy (2020), Giommetti and Sorensen (2021), Gourier, Phalippou, and Westerfield (Forthcoming)). This approach implicitly informs relative performance measures in private equity and is the subject of much practical debate in industry circles on private equity investing (Kaplan and Schoar (2005), Robinson and Sensoy (2013), Harris, Jenkinson, Kaplan, et al. (2023)).

And yet there is another, perhaps more important, portfolio management decision in private equity, one that is intrinsic to the delegated nature of the investment process itself. Specifically, general partners (GPs), who are responsible for investing the capital committed by limited partners (LPs), deploy their fund's capital sequentially, building a portfolio of individual investments. Indeed, the very term *portfolio company*, commonly used to refer to the targets of private equity investment, reflects the fact that the GP is forming a portfolio of investments when they raise a fund and deploy the capital. Nevertheless, most discussions of GP investment behavior couch their investment decision in terms of "deal selection", treating the merits of an individual investment in isolation rather than in terms of "portfolio formation."

This paper explores how GP-level portfolio formation considerations affect private equity investment behavior. To fix ideas, a useful straw man is that GPs evaluate each potential investment in isolation. Under this null hypothesis, GP investment behavior would follow the predictions of a standard search model. In particular, a GP facing randomly arriving investment opportunities would invest in all opportunities above a threshold expected return until the committed capital was fully drawn: the threshold return would embed the oppor-

tunity cost of capital. In such a world, there would be no reason for the size or riskiness of one deal to be related to the size or riskiness of other investments in the GP's portfolio. The alternative hypothesis, which guides our analysis, is that GPs evaluate potential investments as part of a broader portfolio formation strategy linked to their specialization or skill. Under the alternative, one would expect deal characteristics to be correlated across investments in a manner consistent with the incentives provided by the limited partner agreements that govern GP investment behavior and with the GPs' career concerns. We develop this analytical framework more carefully in Section 2.

To explore portfolio formation at the GP level, our analysis exploits a novel dataset of 5,925 portfolio company holdings for 467 distinct buyout funds with vintage years from 1999-2016 operated by 315 distinct private equity firms. Our investment and return data are updated through 2020.[1] The data are broad in terms of industry and geography. For example, 52% of the portfolio companies in our data are located in North America, 34% are in Western Europe, and portfolio companies operate in all 11 GICS sectors (consumer discretionary, health care, information technology, etc.) Section 3 presents our data and discusses the main variables used in the analysis.

Our analysis begins by exploring the tradeoff between generalized and specialized investment. These are presented in Section 4. Here, we find that sector- and geographically specialized funds generate higher returns. These results are consistent with the idea that GPs choose deals to capitalize on their expertise in specific geographies or industries. Yet, specialized funds are, by definition, less diversified and so are likely to bear more region-specific or industry-specific risks. In fact, we find that more specialized funds have higher overall risk. Thus, a risk-return tradeoff exists at the fund level with regard to the degree of specialization in investment focus.

To dig deeper, we next ask how GPs allocate the money inside the fund to build a portfolio of different investments. Section 5 focuses on two dimensions: deal size and deal timing.[2] Ranking holdings from largest to smallest as a percentage of fund size, we find that the largest investments have, on average, the lowest returns. This stands in sharp contrast

[1] We do not consider fund vintages after 2016 since most newer funds are in their investment period.

[2] Unlike Lopez-de-Silanes, Phalippou, and Gottschalg (2015), we focus on investments held at the fund level rather than aggregate across funds to the firm level.

to evidence from mutual funds and hedge funds, which invest in public companies whose highest returns come from outsized positions in their "best ideas" (Antón, Cohen, and Polk, 2021). In fact, returns within a typical PE fund increase monotonically as the investment size rank increases through the fifth largest investment made by a PE fund. To see whether our finding is due to the diseconomies of scale and quick flips, we control for fund size and deal duration and find similar results.[3]

Why do PE buyout funds make outsized bets on low-returning investments? One reason could be the return-risk trade-off discussed above: GPs do not want to "bet the ranch" on any one investment. Indeed, when we calculate ratios of average deal returns to average risk by size rank of the investments, we find these are essentially constant. We calculate these ratios both in terms of nominal money multiples (akin to a Sharpe ratio) and using public market equivalents (more akin to an information ratio) and find similar results. In other words, larger investments are associated with lower risk so that, on average, the risk-return profile of investments is similar regardless of investment size. This is consistent with a portfolio strategy where GPs identify good investments of various sizes but are mindful of the risks.[4] This risk management concern is consistent with the findings of Braun, Jenkinson, and Schemmerl (2020), who find that large deals (relative to fund size) are significantly more likely to be offered for co-investment. In sum, in private equity, managers take their biggest bets on their "safest ideas" instead of "best ideas."

In a similar vein, we sort deals based on their investment year within the fund to study deal timing. The "risk-management" view would suggest that managers start with safer (lower-return) deals and then explore riskier (higher-return) deals as they are increasingly confident that they have locked in returns sufficient to meet the hurdle rates required to earn carried interest. Alternatively, the "exploration" view would predict that managers first bet on riskier (higher-return) deals and then invest in safer (lower-return) deals later on to hedge risks. Empirically, we find that funds start with higher-return deals. That is, the earlier deals in the fund are of higher return and higher risk. The deals made in the later

[3]Lopez-de-Silanes, Phalippou, and Gottschalg (2015) show that quick flips are associated with the highest returns. In our sample, we do find that quick flips tend to be smaller within funds.

[4]Our data do not allow us to examine whether managers apply lower leverage to larger deals to mitigate risk. This is an interesting avenue for future research to pursue.

years earn lower returns and are less volatile. But again, when we examine the ratios of returns to risk (using multiples or PMEs) by year of investment, we find that they are also essentially constant across time.

Do general partners and their investors benefit from the portfolio management decisions they make? Is portfolio management skill a source of persistent differences across GPs? To investigate this issue, Section 6 extends the variance decomposition framework of Korteweg and Sorensen (2017) to deal-level data. By comparing deal-level estimates and fund-level estimates of the components of return variance attributable to GP skill, timing, and idiosyncratic noise, we can link portfolio formation to overall manager skill. When we aggregate across investments within a fund, our performance attribution measures are close to those reported in Korteweg and Sorensen (2017): we find that about 40% of overall variation is attributable to GP skill, and about 48% is attributable to luck. At the individual deal level, however, we find that only 4% is GP skill, and upwards of 90% is luck. This comparison implies that a *significant* source of GP skill is the portfolio selection across individual investments. In some sense, this is the smoking gun indicating that portfolio-formation considerations are salient in private equity fund formation.

In spite of the centrality of the issue we study in this paper, almost no work to date has studied the GP's investment decision through the lens of portfolio construction. The closest are Lopez-de-Silanes, Phalippou, and Gottschalg (2015), which documents the decreasing returns to scale of buyout funds, and Spaenjers and Steiner (2021), which shows that specialized investors in hotels outperform generalist funds. In venture capital, Gompers, Kovner, and Lerner (2009) and Hochberg, Mazzeo, and McDevitt (2015) investigate the specialization of funds. Almost all existing work in private equity research either focuses on the performance of individual investments made by GPs or else focuses on the net-of-fee returns earned by investors in private equity funds.

In sum, our results show a high degree of interdependence between individual investments in a private equity fund. These results are consistent with the idea that GPs, motivated by career concerns and the incentives of limited partner agreements, approach fund investment through the lens of portfolio formation. Nevertheless, our findings open many new questions for future research. We discuss these in Section 7.

2 Analytical Framework

This section lays out a simple framework to guide our empirical analysis. The goal is to motivate a series of empirical tests that allow us to explore whether we should think of GP investment behavior as being driven by portfolio formation considerations or whether each deal should be viewed as an isolated individual investment decision.

A natural starting point is to consider what investment behavior would look like if GPs had no concerns at all about portfolio formation. If they simply evaluated the merits of each investment in isolation, their investment behavior would be well described by the predictions of a classic dynamic search model along the Lippman and McCall (1976) in which by exerting effort to search for deals, they would face investment opportunities arriving at some Poisson arrival rate, where each deal would have an expected return R_i that conformed to some known distribution $F(R)$ with some known variance. In this world, the optimal investment policy would be to reject all deals below some threshold return R^*. Because capital in a fund is limited, and because the GPs' time and attention are limited, the optimal value of R^* would be chosen in a manner that captured the opportunity cost of foregone investment and would generally be a function of the cost of search and the variance of returns, as is standard in real options analysis.

If this simple framework described GP behavior, there would be no specific relationship between the size of a particular deal and its risk characteristics or between the size/risk of one deal and that of a later deal in the portfolio. The set of investments would essentially look like a left-truncated sample of the underlying distribution from which the GP was selecting returns.

Of course, GPs have career concerns and operate under limited partner agreements that create incentives for them *not* to view each investment in isolation. Limited partner agreements specify management fees as well as carried interest–a portion (typically 20%) of the net return that they generate on exited investments above some hurdle rate. Metrick and Yasuda (2010) show that a substantial portion of the lifetime compensation for general partners comes from the stream of management fees they generate by raising subsequent funds.

Preserving the continuation value of a management fee stream can create incentives to

play it safe if management fees are a large enough component of total lifetime compensation. To see how, consider the following stylized choice between two risky investments. Investment A has a 1/3 chance of a 6x return and a 2/3 chance of a total loss. Investment B has a 2/3 chance of a 1.5x return and a 1/3 chance of a zero net return. A risk-neutral manager maximizing carried interest in the current fund would always choose Investment A. But now, suppose that the fund manager can only raise a follow-on fund if they exceed a return of 1.2x; otherwise, they exit the industry and face an outside wage of zero. Because Investment A involves a 2/3 chance of not raising a follow-on fund, the manager has incentives to choose Investment B for many empirically plausible management fee streams.

More generally, in the closely related setting of hedge fund management, Lan, Wang, and Yang (2013) show how a risk-neutral fund manager can endogenously behave in a risk-avoiding or risk-seeking manner. Risk-neutral managers behave in a risk-averse manner when the risk of fund closure is especially salient; they have an incentive to play it safe so as to preserve the fund's going concern value in order to continue collecting management fees in the future along the lines of the illustration above. The greater the liquidation risk/costs, the stronger this incentive is, and the more prudently the manager behaves. And yet, a counter-balancing force comes from the fact that risk-neutral managers who possess skills may leverage their strategy to exploit their ability to generate alpha, effectively behaving in a risk-seeking manner.

The analysis of Lan, Wang, and Yang (2013) illustrates a tradeoff between exploiting specialized skills and minimizing downside risk. In the context of PE investing, this suggests trading off specialization, which likely requires a smaller, more focused portfolio that takes bigger bets on a smaller number of investments while minimizing the idiosyncratic risk they face. Diversification can be achieved either by making a greater number of smaller investments, spreading investments across industries and/or regions to benefit from lower correlations between portfolio companies, or investing in less risky companies or structuring their investments so they are less risky.

These considerations imply that we should expect more concentrated funds to earn higher returns but also face greater risk. In contrast, funds that are more diversified would be expected to earn lower returns on average but have lower volatility.

Timing provisions in limited partner agreements surrounding how carried interest is earned also create dynamic incentives as the fund's capital is being drawn down. In particular, most buyout funds have 'whole-fund' carry provisions that require the GP to exceed a hurdle rate on all previous investments before a manager can draw carried interest on a specific deal.[5] Given that the hurdle rate typically applies to the entire fund's committed capital, this could incentivize the manager to play it safe with larger investments and take more risks with smaller ones.

Agency considerations are at the heart of the design of limited partnership agreements in private equity Robinson and Sensoy (2013). Indeed, Ewens, Jones, and Rhodes-Kropf (2013) argue that requiring GPs to bear idiosyncratic risk is part of the solution to the agency conflict between LPs and GPs. Agency considerations suggest that managers who experience early successes may have an incentive to play it safe in later investments, both as a result of 'quiet life' considerations as well as from the desire to preserve carried interest that has already been earned in prior deals. (Later losses could result in clawbacks of the earlier carry.)

To test for the tradeoff between specialization and diversification, Section 4 examines the relation between fund-level performance measures and measures of concentration. To explore dynamic considerations in private equity funds, the remaining sections tie deal-level performance to deal size and timing within the life of a fund.

3 Data

Our analysis makes use of a new dataset of private equity buyout holdings (i.e., investment "deals") and returns information from Burgiss, a global provider of data management services for the limited partner community. Additional information on the Burgiss holding-level dataset is provided in Brown, Harris, Hu, et al. (2020). Sourced directly from LPs, Burgiss provides a set of detailed, verified, and cross-checked portfolio company information for a large sample of institutional-quality private equity funds. In this analysis, we specifically

[5]Exact details vary significantly from one partnership agreement to the next. See Robinson and Sensoy (2013) or Hüther, Robinson, Sievers, et al. (2020) for more.

examine buyout funds because the portfolio characteristics of buyout funds are especially under-researched. While this paper is among the first to use the holdings-level dataset from Burgiss, fund-level data from Burgiss have been used extensively in recent academic work.[6]

Because including funds with incomplete holdings information may significantly affect the precision of our concentration measures, we require a nearly complete history of holdings for a fund to be included in our sample. As a result, we focus on funds with vintages from 1999 to 2016, where the vintage year is defined as the year of the first portfolio company investment. We only include vintage years through 2016 because more recent funds are still in their investment period. In addition, we require the following three conditions: that all the holdings in the fund have deal size information (no missing values); that the fund made at least three, but fewer than 50 investments; and that the total deal size is at least 25% and not more than 200% of the total amount of committed capital to the fund. These restrictions result in a sample of 468 buyout funds investing in 5,925 portfolio companies.

Variable definitions are provided in Table 1. Our main performance measure is the Kaplan and Schoar (2005) public market equivalent (PME). This measure is the ratio of the discounted sum of cash flows received to the discounted sum of capital invested.[7] These measures can be computed at the deal level or at the fund level. For example, a deal-level PME of 1.15 would indicate that an investor ended up with 15% more in present value than they put into the investment. A fund-level PME of 0.9 would indicate that the present value of distributions from the fund was only 90% of the present value of capital paid into the fund. Because it is customary to use the realized returns from a broad market index as the discount rate, the PME is often interpreted as an excess return over what would have been earned if the capital invested in private equity had instead been invested in a public market index. We use gross PMEs that represent the experience of GPs in individual deals (i.e., that exclude fund-level fees and carried interest) because of data availability.

[INSERT Table 1 AROUND HERE]

[6]See, for example, Brown, Gredil, and Kaplan (2019) and Brown, Harris, Hu, et al. (2021); Harris, Jenkinson, and Kaplan (2014) and Harris, Jenkinson, Kaplan, et al. (2018, 2023).

[7]In robustness checks, we also use total fund value to paid-in capital (TVPI), which is defined as the ratio of the (undiscounted) sum of current fund net asset value and total distributions to the total amount invested. See the appendix tables.

As discussed in length in Harris, Jenkinson, and Kaplan (2014), the calculation and interpretation of PME depends on the choice of the public market benchmark index. In general, the findings in Harris, Jenkinson, and Kaplan (2014) suggest that the average PME is fairly robust to a range of different public market benchmarks.[8] Following Braun, Jenkinson, and Stoff (2017), we choose the benchmark index based on the holding region. We use Russell 3000 for holdings in North America, the Asia and Pacific MSCI performance index for Asian and Pacific holdings, the Europe MSCI performance index for European holdings, and the MSCI World performance index for other holdings. All indices are in USD as we utilize data in Burgiss that has all been converted to USD.[9]

[INSERT Table 2 AROUND HERE]

Table 2 reports the summary statistics for the primary variables used in our analysis. Panel A reports the summary statistics of deal-level characteristics. Our deal sample has a mean (median) PME of 1.63 (1.27) and an interquartile range of 0.62 to 2.17. Panel B reports the summary statistics of fund-level variables. The online appendix contains correlation matrices for deal- and fund-level variables.

4 Concentration, Specialization and Fund Performance

In this section, we investigate the tradeoff between diversification and specialization by examining the relationship between fund performance and alternative measures of investment concentration. First, we examine the relationship between portfolio investment size concentration and fund returns–that is, we ask how funds that concentrate their capital in a small number of investments compare with funds that spread their capital across a larger number of deals. Second, we examine how industry and geographic specialization relate to fund performance.

[8] For example, the S&P 500, NASDAQ Composite, Russell 3000, Russell 2000, and Fama-French 8th, 6th, 4th, and 2nd size deciles.

[9] At this time, the Burgiss holding database only provides the investment entry and exit year (instead of the exact date). We assume July is the entry month and June is the exit month.

4.1 Measures

In our data, the distribution of deal size within a fund follows a power law distribution.[10] As such, we use granularity (Gabaix, 2011) as a fund-level measure of portfolio concentration. Granularity is the parameter of the Power Law distribution, i.e., α in $e^{-\alpha \text{rank}}$ of the within-fund size rank distribution.[11] The larger the *Granularity*, the more highly concentrated a fund's investments in large-size deals. As reported in Panel B of Table 2, the median buyout fund has a *Granularity* degree of 0.79. The *Granularity* ranges from -0.02 to 2.46 for the sample under study. The wide range of values indicates significant cross-sectional variations in terms of the degree of portfolio concentration.

To measure fund specialization, we compute Herfindahl-Hirschman indices (HHI) at the industry and geography level within a fund. These are calculated as the sum of squared within-fund portfolio weights for each of the 11 GICS industry sectors and each of the 11 geographic regions covered by Burgiss, where each weight is the share of total deal size in each sector.[12] By construction, the HHI is always in the range of [0,1]. The larger the value of the index, the more specialized the fund is in a particular industry or geographic region.

Panel B of Table 2 tabulates summary statistics for industry and geographic HHIs. The mean (median) industry HHI is 0.26 (0.22) and the mean (median) geographic HHI is 0.65 (0.66). The large 5%-to-95% ranges for both variables indicate considerable cross-sectional variation in industry and geographic specialization.

[INSERT Figure 1 AROUND HERE]

Figure 1 plots the industry (panel (a)) and geographic (panel(b)) HHI by fund vintage years. The industry HHI is stable over time – the median and aggregate industry HHI is around 0.2 to 0.3 for all years. The geographic HHI is decreasing over time, yet in most years, the median and aggregate geographic HHI is around 0.6 to 0.8. Interestingly, the median (solid black line) is always above the aggregate (dashed red line) level. This indicates that the

[10] See Internet Appendix Figure B1.

[11] We thank Arpit Gupta for suggesting this measure. Details on variable construction in Table 1.

[12] Appendix Table A1 presents the detailed composition of sectors for our sample of deals in this study. We also include a "other" sector when the holding is not classified as one of the 11 GICS sectors. This "other" sector represents only 38 holdings which is less than 1% of the sample.

median fund generally focuses on a subset of regions relative to the universe of all available deals when making investment decisions.

Panels (c) and (d) plot the frequency distribution of funds of different industry (c) and geographic (d) HHI intervals. Consistent with panels (a) and (b), most funds have an industry HHI of 0.1 to 0.4, suggesting fair diversification across industries for buyout funds. Regarding geographic specialization, it becomes much more specialized – most funds have a geographic HHI from 0.5 to 1.

4.2 Regression Results

We estimate two sets of fund-level regressions. Our measure of fund-level returns is the weighted average, deal-level, gross-of-fee PME. As a proxy for fund-level risk, we use the standard deviation of deal-level PMEs within a fund. Table 3 reports results from regressions with fund concentration, specialization, and other control variables as the independent variables. Columns (1) to (4) present the results for fund returns as the dependent variable. Column (1) shows a positive relation between the *Granularity* and *value-weighted PME*, indicating a higher return for more granular funds. One increase in the granularity degree of the within-fund deal size distribution is related to a 0.24 increase in *value-weighted PME*. Results are significant at the 95% level.

[INSERT Table 3 AROUND HERE]

Column (2) in Table 3 presents results on the relation between fund-level industry specialization and return. We find a positive and statistically significant relation between *HHI sector* and *PME*, indicating higher returns for funds with more industry specialization. Column (3) presents the results on the fund-level relationship between performance and geographic specialization of the portfolio companies held by the fund. Similar to the results on industry specialization, funds with a higher geographic specialization level have a higher return *ceteris paribus*. Both results are statistically significant at the 95% confidence level. Moving from perfectly diversified (0) to perfectly concentrated (1) for the *HHI sector* (*HHI region*) index is associated with a 0.41 (0.40) increase in *PME*. In column (4), we include all

three measures: the *Granularity*, the *HHI sector*, and the *HHI region*. We continue to find a positive and significant relationship between concentration/specialization and fund returns.

To see how deal risk is related to concentration and specialization, the dependent variable in columns (5) to (8) is the standard deviation of PMEs of deals within the fund. Column (5) shows that a higher level of portfolio concentration is related to a higher standard deviation of PMEs. The implication is that even if large deals are less risky (as shown in the next section), the overall effect of portfolio concentration is to increase the volatility of returns. A one standard deviation increase in granularity is associated with a 0.23 increase in the *standard deviation of PME* of deals in the fund. This result is highly statistically significant.

Columns (6) and (7) focus on sector and geographic specialization, respectively, and show a statistically significant positive relation between specialization and our risk measure. Funds with greater regional or industry specialization are associated with higher overall volatility in performance. In the specification that includes all variables and vintage fixed effects (column (8)), an increase from fully diversified to fully concentrated in the *HHI sector* (*HHI region*) index is associated with a 0.39 (0.53) increase in the standard deviation of deal PMEs. The coefficient for the *HHI sector* variable is significant at the 90% confidence level, and the coefficient for the *HHI region* variable is significant at the 95% level.

Overall, our results suggest that concentration in a few large deals generates higher fund-level returns at the cost of higher risk. Similarly, industry and geography specialization is also associated with higher returns and risk.

We provide a battery of robustness checks in the appendix. These include using total value to paid-in capital as an alternative performance measure, using equally weighted instead of value-weighted returns, examining downside risk, and using "jackknife" resampling for return and risk. The results are qualitatively unchanged using these alternative measures of risk and return.

5 Dynamic Investment Considerations

5.1 Are Bigger Deals Better Deals?

First, we explore whether PE funds take the biggest bets on the best deals, as has been shown in the mutual fund setting (Antón, Cohen, and Polk, 2021), or whether hurdle-rate considerations, as described in Section 2 lead them to play it safe with bigger deals. To capture variation in deal size, we rank deals within a given fund according to the dollar value of the investment. For example, a *Deal-Size Rank* of one means that the deal is the largest investment made by a fund, a *Deal-Size Rank* of two represents the second largest deal in the fund, and so on. We group deals ranked tenth-largest or smaller into a single rank category.[13] The largest deal is often exceptionally large – more than $1/7^{\text{th}}$ of total fund value. This concentration falls off quickly so that the fifth largest deal is generally about $1/12^{\text{th}}$ of total fund value.

We start by plotting summary statistics by *Deal-Size Rank* in Figure 2. Panel (a) shows average deal PMEs as the solid blue line and indicates that a fund's largest deal has the lowest return on average with a mean PME of 1.26. As the *Deal-Size Rank* increases, i.e. as deals get smaller, the mean return increases monotonically for the first five *Deal-Size Rank* groups and then becomes relatively stable around 1.7. The results are similar across different percentiles except for the very bottom of the distribution where PMEs are zero regardless of rank and for the maximum PME, which exhibits no pattern.[14] This suggests that the relation between *Deal-Size Rank* and PMEs is not driven by outliers but instead a fairly stable feature across funds.

[INSERT Figure 2 AROUND HERE]

The fact that larger deals have lower returns stands in contrast to findings from the mutual fund and hedge fund literature. As noted above, risk management considerations on the part of the GP are likely an important element of this result. Because larger deals in PE funds are typically a much larger percentage of fund value than the largest position

[13] Figure B1 Panel (a) shows the distribution of deal sizes (dollar value) for different ranks.
[14] Appendix Table A4 provides a tabulation of the returns by percentiles for each *Deal-Size Rank*.

in a mutual fund, larger PE deals pose a greater risk to overall fund performance. Thus, in keeping with the predictions of Lan, Wang, and Yang (2013), GPs may take measures to minimize the risk of the deal because its size threatens the continuation value of the fund franchise, even if the GPs themselves are inherently risk-neutral. We consider the standard deviation of PMEs across funds as a proxy for the average risk of each *Deal-Size Rank* group and plot these values in panel (a) of Figure 2 as the dashed red line.[15] The results indicate that variation in PMEs is lower for relatively large deals. The largest deal group has a standard deviation of PMEs of 1.24, and values trend up to about 1.6 for the fifth largest deal in a fund, and there is no noticeable trend for higher ranks.

These results are consistent with the lower average performance for relatively large investments being a deliberate choice by GPs. Specifically, larger deals within a fund also have lower risk, consistent with the notion that GPs do not want to "bet the ranch" on relatively large deals.

In robustness tests, we generate similar plots using TVPI as the performance measure. We observe a very similar pattern consistent with larger deals having lower returns but also lower riskiness. In this case, the ratio of return to risk is perhaps more akin to a Sharpe ratio (since these are in absolute terms) and is again almost constant across *Deal-Size Rank*.

[INSERT Figure 3 AROUND HERE]

5.2 Are Early Bets Safer Bets?

The second dimension we study is how the deal sequence relates to deal return (PME) and riskiness. We sort deals by their investment sequence within the fund (i.e., by *Deal Sequence*) from the early deals (sequence one) to late deals (sequence five and greater). Plot (a) in Figure 3 shows average deal PMEs as the solid blue line.[16] The results indicate that, on average, deal returns in the first few years of a fund's life have higher returns with a mean PME of 1.68. Mean (and median) PMEs decrease for deals done later in the investment

[15] Appendix Table A4 tabulates these values.

[16] Panel A of Table A5 reports summary statistics of deal PMEs for each *Deal Sequence*.

period (i.e., for *Deal Sequence* 4 and 5). The dashed red line in Plot (a) of Figure 3 shows the riskiness of the deals made in different years of the fund life (as measured by the standard deviation of PMEs) and that risk follows the same tendency to decline as funds age. Plot (b) of Figure 3 shows that the ratio of mean to standard deviation of PMEs is quite stable around 1.00 for all *Deal Sequence* groups. In addition, we plot the patterns using TVPI as the performance measure in the Appendix and find similar results. The pseudo-Sharpe ratio (using TVPI) is even more stable across deal sequences than the pseudo-information ratio (using PME).

5.3 Regression Analysis

Taken together, the results above provide graphical evidence that GPs systematically balance the risk and return of their portfolios by managing the relative size and timing of investments. Next, we explore regressions to quantify the magnitude and statistical significance of these findings.

[INSERT Table 4 AROUND HERE]

Table 4 reports the results. As in the univariate analysis, the regression results indicate that larger deals tend to have lower returns. To ease the interpretation of size results, we use the negative of the rank (i.e., -1 for rank 1) in the regression, so a larger number implies a larger relative size. Columns (1) to (3) of Table 4 report results without GP fixed effects. In column (1), we regress deal PME on *Deal-Size Rank* and *Deal Sequence*. A one-notch smaller position in *Deal-Size Rank* is associated with an average increase in PME of 0.05, similar to what we inferred from averages shown in Table A4. A one-year increase in deal sequence is associated with an average decrease in PME of 0.05. Column (2) includes the absolute deal size to capture the effect of absolute (versus relative) deal size. The effect of absolute deal size is negative and marginally significant, but the coefficients on *Deal-Size Rank* and *Deal Sequence* are largely unchanged.

Column (3) shows the results with the full set of control variables and location and industry fixed effects. We find that the negative and statistically significant relationship

between PME and *Deal-Size Rank* persists after controlling for other variables. However, the magnitude of the relation is somewhat smaller. For example, in the specification (3), a one-notch smaller position in rank is associated with a 0.03 increase in PME. The negative and statistically significant relation between *Deal Sequence* and PME also persists in the presence of other control variables, and the magnitude does not change. The coefficient on the dummy variable indicating fully exited deals is positive and significant, consistent with anecdotal evidence that better deals realize exits more quickly (and counter to the hypothesis that value creation is still occurring or interim valuations are conservative). Similarly, longer-duration deals have lower PMEs. Controlling for deal exit dummy and deal duration also mitigates the concerns that our results on deal sequence merely reflect the mechanical fact that earlier deals are more likely to exit up to a given period or earlier deals may have a shorter duration.

Columns (4) – (6) report the specifications with GP fixed effects. We observe the same relationship that relatively larger deals within the fund have lower returns. We also observe the same negative relation between deal sequence and deal PME, implying that later deals have lower returns. All coefficients on *Deal-Size Rank* and *Deal Sequence* remain negative and significant at the 99% confidence level. This implies that the incentive to balance between different deals within the fund is not GP-specific. The magnitude of the coefficients on *Deal-Size Rank* and *Deal Sequence* is even slightly larger after controlling for GP fixed effects, suggesting that the portfolio management pattern is stronger after accounting for any differences in GP styles.

Taken together, our findings suggest that PE managers actively manage risk-return trade-offs in deals by allocating equity investments in portfolio companies with various risk-return profiles. While the data do not allow us to accurately determine whether this is done through deal selection, capital structure, or a combination, our findings are consistent with a deliberate portfolio management process, not a random selection of investments. Speaking to relative deal size, our finding stands in contrast to the "best ideas" findings in mutual funds and hedge funds, where relatively large positions have higher average returns. Instead, we find that managers balance deals with different sizes to control for the overall riskiness of the portfolio. Regarding timing, we find evidence that managers take on more risk and invest in

higher-return deals earlier in fund life and then lean toward deals with lower risk and return later in the investment period.

5.4 GP Heterogeneity

A natural question in light of the findings above is whether the dynamic patterns we see in the data are present for all funds or are more pronounced for certain types of GPs. In other words, are timing considerations more pronounced for better GPs or worse GPs?

To test for this, we extract the GP fixed effects from the estimation presented in Column (6) of Table 4 and sort them into terciles. Because these are fixed effects from a regression of performance on deal characteristics, this effectively allocates GPs into one of three average performance groupings. Then we repeat the regressions from Table 4 for each tercile.

[INSERT Table 5 AROUND HERE]

Table 5 reports the results. The strength of the results varies by GP performance. Specifically, the results are strongest for the top-performing GPs, about the same as for the full sample results for the average-performing GPs, and only marginally significant for the low-performing GPs. In addition, Internet Appendix Table C15 shows results from interactions of the *Deal-Size Rank* and *Deal Seq.* with estimated GP fixed effects. The coefficients for the interaction terms are negative and statistically significant at 5% or 1%. These results indicate that the deal sizing and sequencing considerations are associated with stronger average performance.

5.5 Does Early Success Spur Stronger Performance?

The results in the previous subsection suggest that the dynamic portfolio formation considerations we have identified are not necessarily a direct reflection of agency costs associated with delegated portfolio management, even if agency considerations are fundamental to the design of limited partner agreements.

To further explore these issues, we next ask how early performance affects later performance. Here, we can distinguish between two alternative hypotheses. One is that stronger

performance in early deals is indicative of stronger performance overall, in which case we would expect stronger performance early in the life of a fund to predict stronger performance later in a fund's life. This would especially be true if stronger early performance facilitated greater risk-taking on the part of GPs, as in Lan, Wang, and Yang (2013). In contrast, stronger performance early in the life of a fund could instead induce managers to take less risk, either because "quiet life" considerations were more salient or because the "whole fund" carried interest features of the limited partner agreement encouraged managers not to risk making losses that would eat away at carry that had already been earned. (Given the complexities of carried interest calculations, this need not necessarily operate through a behavioral channel, such as the disposition effect.)

[INSERT Table 6 AROUND HERE]

Table 6 presents the results of this analysis. The dependent variable is the deal-level PME. The regression includes a dummy variable *Early Success* equaling one if the fund made at least one deal with a PME greater than or equal to three within the first two years of the fund's life, a dummy variable *After Year 2* that equals one if the observation is a deal made after the first two years of a fund's life and the interaction of these two variables. The insignificant loading on *Early Success* indicates that funds with stronger early performance do not, in general, have average deal-level performance that exceeds funds without early home runs. The reason can be seen through the loading on the interaction term. The negative loading indicates that funds with high-performing deals early in the fund's life generally experience lower returns in later deals. This is true across a variety of specifications: Column (1) includes the two variables and their interaction; Column (2) introduces controls for fund size, deal size, whether the deal was exited, and the holding period of the deal; Columns (3) and (4) introduce GP, sector, and geography fixed effects. In all specifications, there is a highly statistically significant negative loading on the interaction term between early deal success and later deal success. This is consistent with two scenarios, one in which the terms of limited partner agreements create incentives to "crystallize carried interest" and one in which the underlying agency conflict at the heart of the limited partner agreement itself

gives rise to less effort in later deals; effectively encouraging GPs to rest on their laurels. Regardless of the underlying mechanism, the results point to a pronounced interdependence across deals within a given fund, which speaks against the idea that deals are evaluated in isolation.

5.6 Robustness

One potential concern regarding our analysis is that some funds have not fully invested all the capital committed by limited partners by the end of the sample period. This might affect our results if there is a trend in deal size over the investment period of recent funds. For example, suppose funds invest larger and higher returning deals later in the investment period. In that case, our sample could be missing these realizations and thus overestimate the negative relationship between deal size and return. In addition, evidence indicates that NAVs suffer from smoothing and systematic misvaluation (Barber and Yasuda, 2017, Brown, Gredil, and Kaplan, 2019). However, due to data availability, we have little choice but to use reported NAVs at our ending date for unrealized deals. Of course, cash flows are not subject to the same biases, so we have more confidence in results based only (or primarily) on realized deals. Given the high proportion of unrealized deals for more recent vintages, the accuracy of the valuation information is important for proper return measurement.

To address these concerns, we examine the subsample of deals held by funds with a *Fraction Invested* ratio above 80%. We also allow for a time gap of at least four years for investments to be realized (our sample ends in 2016, while the data is collected in the third quarter of 2020). Results are robust to this sample restriction (and presented in Internet Appendix Table C7). We also report results including *Deal-Size Rank* (Table C8) and *Deal Sequence* (Table C9) separately; do robustness checks using the full rank number and sequence number of deals (i.e., we do not group deals with a rank higher than ten together; similarly, we do not group deals with a sequence later than five together). Results are reported in Table C10. Table C11 reports results on TVPI. Our main findings are essentially unchanged.

6 Portfolio Formation and Performance Persistence

The results thus far illustrate the importance of investment timing considerations within a fund, as well as the tradeoff between specialization and diversification in investment, suggesting that portfolio formation considerations are important for understanding private equity investment performance. Given that these results vary across GPs in a manner that is related to their overall average performance, in this section, we extend the analysis further by exploring how portfolio formation considerations affect performance persistence. More generally, we ask how portfolio formation considerations affect how we attribute overall variation in performance to differences in GP skill, timing of investment, and luck.

To do this, we implement return variance decomposition techniques pioneered by Korteweg and Sorensen (2017). Their original analysis operates at the *fund* level, whereas our data allow us to operate at both the fund and deal levels. This allows us to compare the results we obtain from the fund- and deal-level analysis to establish how our understanding of persistent differences in GP skill is affected by their portfolio formation behavior.

6.1 Fund-level Decomposition

The model of Korteweg and Sorensen (2017) decomposes the total variance in returns into three components (conditional on appropriate covariates, i.e., controls): a GP-specific random effect, a GP-time random effect, and a fund-specific random effect. Formally, the model can be written as follows:

$$r_{iu} = X'_{iu}\beta + \Sigma_{\tau=t_{iu}}^{t_{iu}+N_u}(\gamma_i + \delta_{i\tau}) + \epsilon_{iu} \tag{1}$$

where i represents the GP firms, u represents funds, X_{iu} are the observed covariates (following Korteweg and Sorensen (2017), we control for fund vintage fixed effects in the estimation), t_{iu} denotes the fund's first year of operation (vintage year), N_u denotes the fund life, γ_i is the GP-specific random effect, $\delta_{i\tau}$ is the GP-time random effect, and ϵ_{iu} is the fund-specific random effect. Fund return is measured using the natural logarithm of TVPI. (Using the PME measure is inappropriate in the estimation because the time-specific variation is directly

modeled.) As PE funds do not have a clear exit time, we set the fund life to ten for all funds (as in the original Korteweg and Sorensen (2017) model).

The GP random effects, γ_i, are intended to capture the time-invariant average differences in GP quality. The original purpose of the GP-year (GP-time) random effects, $\delta_{i\tau}$, in Korteweg and Sorensen (2017), is to test whether a large part of return persistence in return is due to the overlap of consecutive funds that are managed by the same PE firm because overlapping funds are exposed to identical market conditions during their period of overlap. The final, fund-specific effect, ϵ_{iu}, captures fund-specific idiosyncratic performance shocks, or in essence, luck. The total variance in fund returns is then the sum of the variances of the three sets of random effects:

$$\sigma^2(r_f) = N^2\sigma^2(\gamma_i) + N\sigma^2(\delta_{i\tau}) + \sigma^2(\epsilon_{iu}). \tag{2}$$

Dividing this equation by the left-hand side expresses the total variation in terms of the proportion explained by GP skill, the proportion explained by overlap, and the proportion explained by luck.

$$\text{Signal to Noise (Skill)}\% = \frac{N^2\sigma^2(\gamma)}{\sigma^2(r_f)}, \text{Overlap Effect}\% = \frac{N\sigma^2(\delta)}{\sigma^2(r_f)}, \text{Noise}\% = \frac{\sigma^2(\epsilon)}{\sigma^2(r_f)}. \tag{3}$$

[INSERT Table 7 AROUND HERE]

Our estimates of the fund-level variance decomposition are presented in Table 7. As a comparison, columns (1) and (2) are the results from Korteweg and Sorensen (2017), which is based on the sample of buyout funds in Preqin from 1969 to 2001. Columns (3) and (4) are results estimated using our sample of buyout funds in Burgiss from 1999 to 2016. Panel A shows the magnitudes of the three random effects as measured by their standard deviations $\sigma(\gamma)$, $\sigma(\delta)$, and $\sigma(\epsilon)$. In the specifications in the even columns, we also include fund vintage years as covariates. Panel B reports the variance decomposition, which is easier to interpret. The top part of Panel B shows the magnitude of the decomposed variance of the three random effects ($N^2\sigma^2(\gamma)$, $N\sigma^2(\delta)$, and $\sigma^2(\epsilon)$) and the total variance of the fund returns.

The bottom part of Panel B shows the percentages of each random effect in explaining the total variance, which are the skill, overlap effect, and luck (as defined in Equation 3).

Column (3) shows that in our sample, 36.22% of the total variance in fund returns is due to skill, 13.14% is due to overlap effect, and 50.96% is due to luck. In column (4), we include vintage year fixed effects as covariates; the importance of skill in explaining return variation increases slightly, and the importance of the overlap effects and luck decreases: 38.99% of the total variance in returns is due to skill, 13.00% is due to overlap effect, and 48.01% is due to luck. This is consistent with the fact that part of the noise in returns (luck) is due to market conditions in different years.

Our estimates of $\sigma(\gamma)$ and $\sigma(\delta)$ are close to the estimates in Korteweg and Sorensen (2017). For example, in specifications with vintage year fixed effects as covariates, our estimate of $\sigma(GPRE)$ is 0.033 while their estimate of $\sigma(\gamma)$ is 0.029 (0.4 basis points smaller than ours). Our estimate of $\sigma(GPYearRE)$ is 0.059, while their estimate of $\sigma(\delta)$ is 0.071 (1.2 basis points larger than ours). The gap in the estimates of $\sigma^2(\epsilon)$ is larger. Our estimate is 0.363, while their estimate is 0.481 (11.8 basis points larger than ours). This results in the finding that skills explain 38.99% of the fund-level return in our sample, 15.68% larger in Korteweg and Sorensen (2017) sample. Our estimates of the importance of the overlap effect in explaining fund-level return (13.00%) are very similar to the Korteweg and Sorensen (2017) estimate of (13.82%). Our estimates of the importance of noise in explaining fund-level return is 48.01%, which is 15.13% lower than Korteweg and Sorensen (2017)'s estimates.

6.2 Deal-level Decomposition

One may or may not learn more about GP skills using the deal-level returns than estimated using the fund-level returns.[17] On the one hand, more granular return data partially solves the mechanical overlapping issue that gives rise to the importance of the GP-year fixed effects that capture persistence due to overlapping investment periods. This implies that the variance in the GP-year effect is likely to be smaller at the deal level than at the fund level, and thus GP skill may explain more of the return variation at the deal level. On the other hand, return is noisier at the deal level than at the fund level. As a result, the variance in

[17] See footnote 19 in Korteweg and Sorensen (2017).

the error term can be larger at the deal level, leaving a smaller proportion of the variance explained by GP skill. Of course, GPs can mitigate this idiosyncratic risk exposure through the selection of imperfectly correlated investments.

To operationalize our GP-Deal level model, we estimate Equation 1 on deal-level data as follows:

$$\widehat{DealReturn}_{ij} = X_{ij}'\beta + \Sigma_{\tau=t_{ij}}^{t_{ij}+N_{ij}}(GPRE_i + GPYearRE_{i\tau}) + \nu_{ij} \tag{4}$$

where i represents the GP firm, j stands for deal, t_{ij} denotes the deal's first year (entry year), N_{ij} denotes the deal life which is the number of years from the deal entry to the deal exit, $GPRE_i$ is the GP-specific random effect, $GPYearRE_{i\tau}$ is the GP-year random effect, and ν_{ij} is the deal-specific random effects. Deal return is measured using the natural logarithm of TVPI. The results are presented in Table 8.

The total variance in $DealReturn_{iu}$ is the sum of the variances of the above-discussed three random effects,

$$\sigma^2(DealReturn) = N^2\sigma^2(GPRE) + N\sigma^2(GPYearRE) + \sigma^2(\epsilon) \tag{5}$$

and we further define the percentage of skill (i.e., signal-to-noise ratio), overlap effect, and noise as

$$\text{Signal to Noise (Skill)}\% = \frac{N^2\sigma^2(GPRE)}{\sigma^2(DealReturn)}, \tag{6a}$$

$$\text{Overlap Effect}\% = \frac{N\sigma^2(GPYearRE)}{\sigma^2(DealReturn)}, \tag{6b}$$

$$\text{Noise}\% = \frac{\sigma^2(\epsilon)}{\sigma^2(DealReturn)}. \tag{6c}$$

[INSERT Table 8 AROUND HERE]

Table 8 reports the estimates of the GP-Deal model. Panel A shows the magnitudes of the GP-specific, GP-year-specific, and error-term random effects. Specifications also differ in whether deal entry year and exit year fixed effects are controlled for, as indicated in the

table. Panel B reports the variance decomposition results.

Comparing fund-level and deal-level estimates, we see both a reduction in the overlapping effect and an increase in the noise component. Take the specification in column (4) as an example, where both deal entry and exit effects are controlled for, the overlapping effect at the deal level reduces 5.80%, 7.20% smaller than the fund-level overlapping effect. In contrast, the noise component increases to 90.20%, which is larger than the fund-level noise component by 42.19%. Overall, the increase in the noise component dominates the reduction in the overlapping effect and results in skill only accounting for 4.01% of the return variance on the deal level. In other words, by examining deal-level returns, investors learn much less about GP skills than by looking at the fund-level returns. Fund-level returns, which inherently reflect the portfolio formation considerations we focus on in this paper, demonstrate a much higher degree of GP-specific skill and a much lower degree of idiosyncratic noise than deals examined in isolation. In some sense, this evidence is the "smoking gun" illustrating the importance of portfolio formation for GPs.

7 Conclusion

In this paper, we use novel, investment-level data to explore how portfolio formation considerations drive the investment behavior of general partners in private equity. We focus both on the tradeoff that general partners in private equity face between specialization and diversification, as well as timing considerations within the life of a fund. Studying the investment decisions of private equity general partners through the lens of portfolio formation allows us to learn more about how the fund-level risk and return measures are influenced by the interdependencies between individual fund-level investment decisions.

In keeping with the idea that an important source of value creation in private equity is focused, context-specific knowledge about a sector or a market, we find that focused funds earn higher returns than funds that pursue a diversified investment strategy. This focus, however, comes at the cost of higher volatility. On balance, the ratio of market-adjusted performance to overall volatility is similar across strategies, suggesting that there is an equilibrium trade-off at the fund level between focus and diversification.

Digging more deeply into the sequence and size of individual investments, we find that general partners do not place the biggest bets on their "best" (i.e., highest return) ideas; instead, their biggest bets are their safest and in keeping with a risk-return trade-off, also their lowest returning investments on average. In fact, we show that within a fund, ranking individual investments in increasing order of size would rank them in decreasing order of riskiness up through the fifth largest investment. At the same time, we find evidence of *exploration* early in the life of the fund, as smaller and riskier investments tend to occur earlier, while larger, safer investments occur later.

The findings in our paper are consistent with a model in which rational, risk-neutral GPs who are interested in maximizing lifetime income arising from fees and carried interest income balance the desire to generate stronger performance in a fund against the potential losses associated with taking greater risk in a fund and potentially sacrificing the going concern value of future funds. The findings are also consistent with behaviorally motivated theories of GP behavior. In addition, our findings raise interesting questions in organizational economics about the optimal formation of investment teams inside private equity funds, as well as how the use of leverage in an individual buyout transaction is influenced not just by market conditions but by risk-mitigation concerns that play through the other deals in the fund's portfolio. We leave these questions for future work.

References

Antón, M., Cohen, R. B., & Polk, C. (2021). Best ideas. *Available at SSRN 1364827*.

Barber, B. M., & Yasuda, A. (2017). Interim fund performance and fundraising in private equity. *Journal of Financial Economics*, *124*(1), 172–194.

Braun, R., Jenkinson, T., & Schemmerl, C. (2020). Adverse selection and the performance of private equity co-investments. *Journal of Financial Economics*, *136*(1), 44–62.

Braun, R., Jenkinson, T., & Stoff, I. (2017). How persistent is private equity performance? evidence from deal-level data. *Journal of Financial Economics*, *123*(2), 273–291.

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

Brown, G. W., Harris, R., Hu, W., Jenkinson, T., Kaplan, S. N., & Robinson, D. T. (2021). Can investors time their exposure to private equity? *Journal of Financial Economics*, *139*(2), 561–577.

Brown, G. W., Harris, R. S., Hu, W., Jenkinson, T., Kaplan, S. N., & Robinson, D. T. (2020). Private equity portfolio companies: A first look at burgiss holdings data. *Available at SSRN 3532444*.

Cavagnaro, D. R., Sensoy, B. A., Wang, Y., & Weisbach, M. S. (2019). Measuring institutional investors' skill at making private equity investments. *The Journal of Finance*, *74*(6), 3089–3134.

Ewens, M., Jones, C. M., & Rhodes-Kropf, M. (2013). The price of diversifiable risk in venture capital and private equity. *The Review of Financial Studies*, *26*(8), 1854–1889.

Gabaix, X. (2011). The granular origins of aggregate fluctuations. *Econometrica*, *79*(3), 733–772.

Giommetti, N., & Sorensen, M. (2021). Optimal allocation to private equity. *Available at SSRN 3761243*.

Gompers, P., Kovner, A., & Lerner, J. (2009). Specialization and success: Evidence from venture capital. *Journal of Economics & Management Strategy*, *18*(3), 817–844.

Gourier, E., Phalippou, L., & Westerfield, M. M. (Forthcoming). Capital commitment. *Journal of Finance*.

Gredil, O., Liu, Y., & Sensoy, B. A. (2020). Diversifying private equity. *Available at SSRN 3535677*.

Harris, R. S., Jenkinson, T., & Kaplan, S. N. (2014). Private equity performance: What do we know? *The Journal of Finance*, *69*(5), 1851–1882.

Harris, R. S., Jenkinson, T., Kaplan, S. N., & Stucke, R. (2018). Financial intermediation in private equity: How well do funds of funds perform? *Journal of Financial Economics*, *129*(2), 287–305.

Harris, R. S., Jenkinson, T., Kaplan, S. N., & Stucke, R. (2023). Has persistence persisted in private equity? evidence from buyout and venture capital funds. *Journal of Corporate Finance*, 102361.

Hochberg, Y. V., Mazzeo, M. J., & McDevitt, R. C. (2015). Specialization and competition in the venture capital industry. *Review of Industrial Organization*, *46*(4), 323–347.

Hüther, N., Robinson, D. T., Sievers, S., & Hartmann-Wendels, T. (2020). Paying for performance in private equity: Evidence from venture capital partnerships. *Management Science*, *66*(4), 1756–1782.

Kaplan, S. N., & Schoar, A. (2005). Private equity performance: Returns, persistence, and capital flows. *The Journal of Finance*, *60*(4), 1791–1823.

Korteweg, A. G. (2011). Markov chain monte carlo methods in corporate finance. *Available at SSRN 1964923*.

Korteweg, A. G., & Sorensen, M. (2017). Skill and luck in private equity performance. *Journal of Financial Economics*, *124*(3), 535–562.

Korteweg, A. G., & Westerfield, M. M. (2022). Asset allocation with private equity. *Available at SSRN 4017858*.

Lan, Y., Wang, N., & Yang, J. (2013). The economics of hedge funds. *Journal of Financial Economics*, *110*(2), 300–323.

Lippman, S. A., & McCall, J. J. (1976). The economics of job search: A survey. *Economic inquiry*, *14*(2), 155–189.

Lopez-de-Silanes, F., Phalippou, L., & Gottschalg, O. (2015). Giants at the gate: Investment returns and diseconomies of scale in private equity. *Journal of Financial and Quantitative Analysis*, *50*(3), 377–411.

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

Robinson, D. T., & Sensoy, B. A. (2013). Do private equity fund managers earn their fees? compensation, ownership, and cash flow performance. *The Review of Financial Studies*, *26*(11), 2760–2797.

Spaenjers, C., & Steiner, E. (2021). The value of specialization in private equity: Evidence from the hotel industry. *HEC Paris Research Paper No. FIN-2020-1410*.

Figures

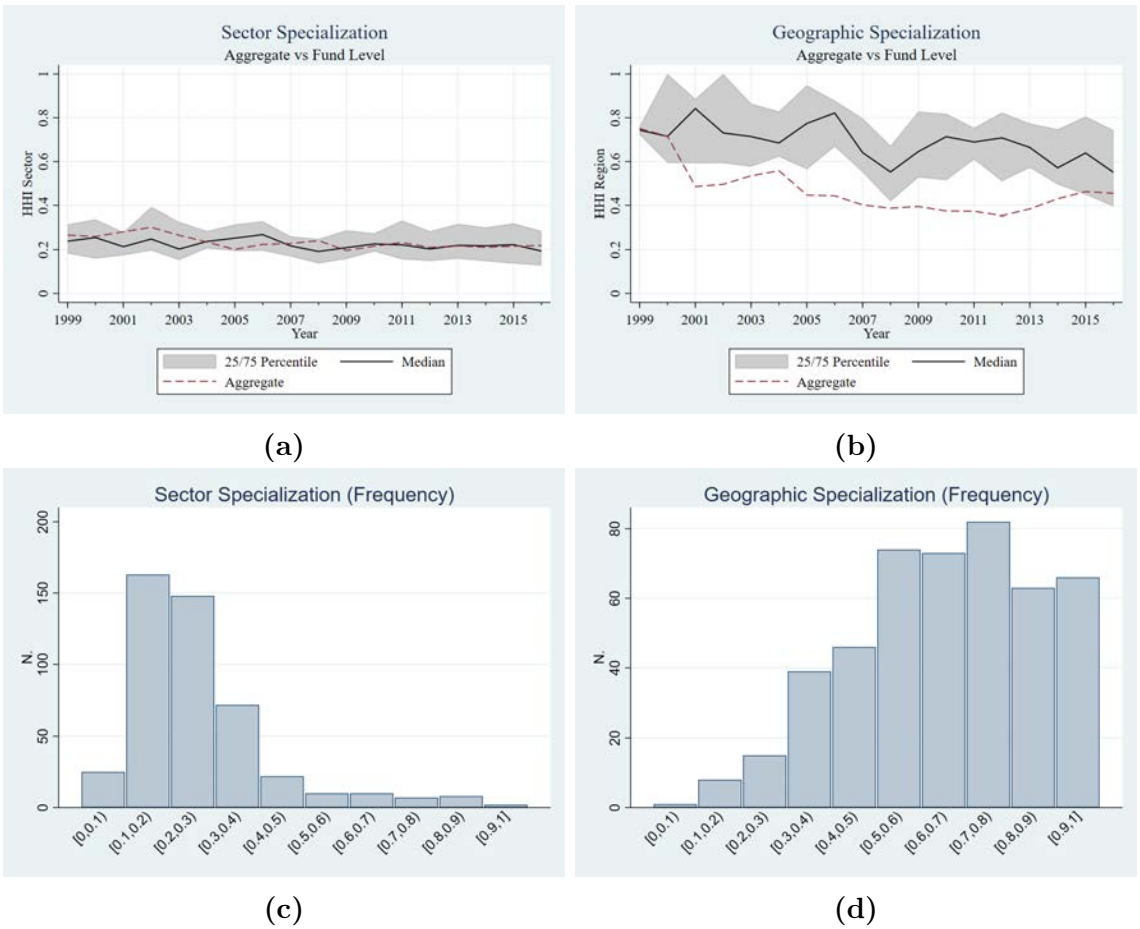
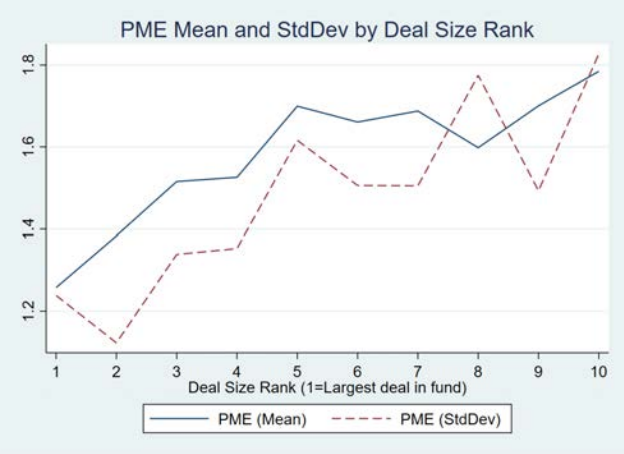
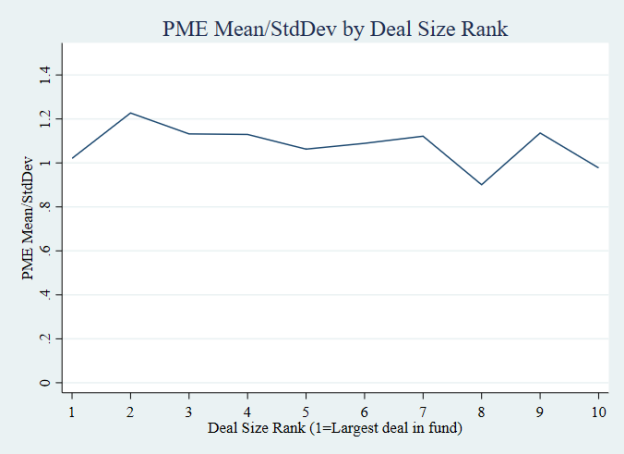


Figure 1: Industry/Geographic Specialization

Note: This figure plots the basic information of specialization indices. The top panels show the industry HHI (Panel (a)) and the region HII (Panel (b)) in each deal entry year/fund vintage. The solid black line represents the median fund-level HHI, and the grey shade ranges from the 25th to 75th percentile fund-level HII. The dashed red line is for the aggregate-level HHI (over the entire deal sample). The bottom panels show the frequency distribution of funds of different industry HHI (Panel (c)) and the region HHI (Panel (d)). The y-axis is the number of funds, and the x-axis is the range of HHI.



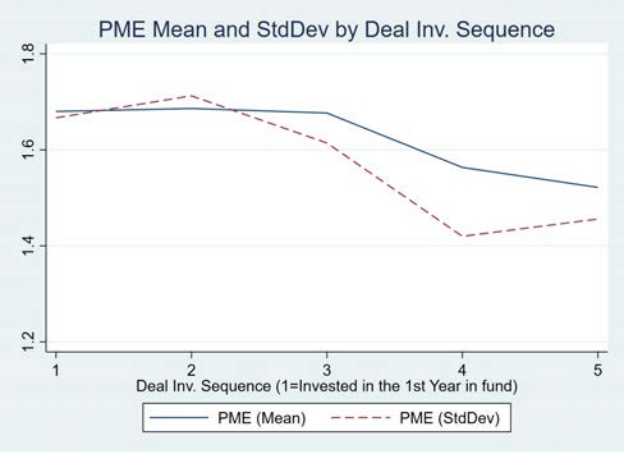
(a)



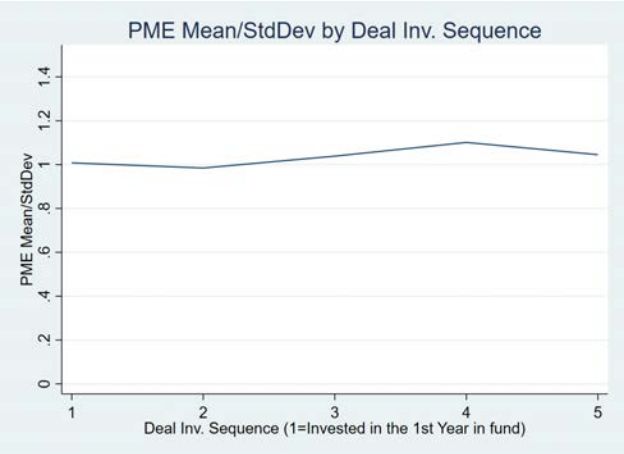
(b)

Figure 2: Deal Return by Deal Size Rank in the Fund

Note: This figure plots the mean and standard deviation of deal return in each deal-size rank category. The horizontal axis is the rank of the deal (*Deal-Size-Rank*). In Panel (a), the vertical axis is the mean (Mean, blue solid line) or standard deviation (StdDev, red dashed line) of the deal PME in a given rank category. In Panel (b), the vertical axis is the ratio between the mean of the deal PME and the standard deviation of the deal PMEs. *Deal-Size-Rank* is the rank of deals within the fund, sorted by size. 1 to 10 refers to the largest to the smallest. E.g., rank 1 means that the deal is the largest in the fund it belongs to. Rank 2 means that the deal is the second largest in the fund. We group deals with ranks larger than ten into one category: rank 10.



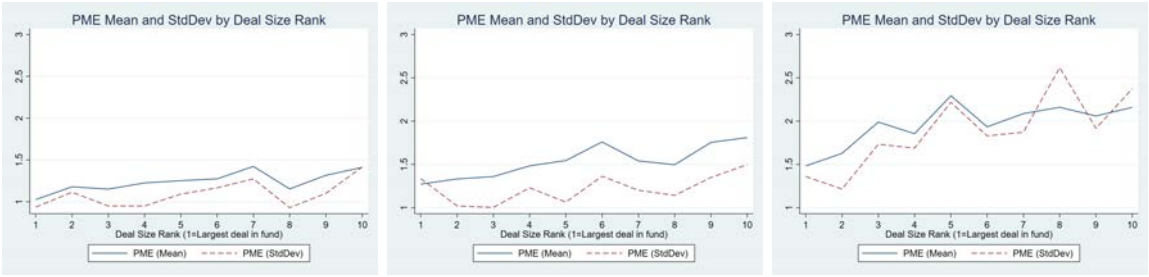
(a)



(b)

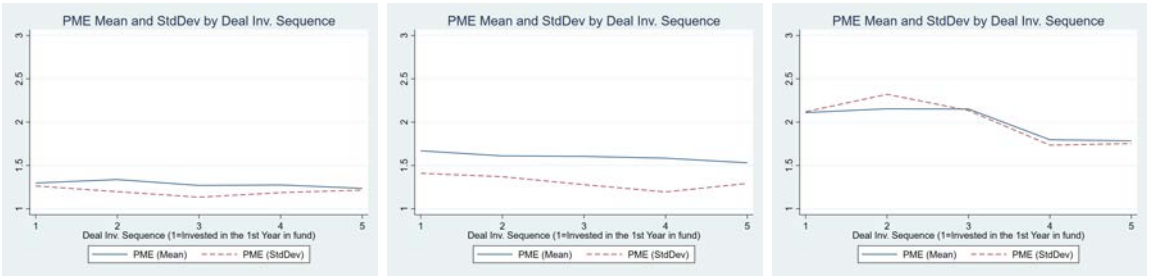
Figure 3: Deal Return by Deal Timing Sequence in the Fund

Note: This figure plots the mean and standard deviation of deal return in each deal timing sequence category. The horizontal axis is the sequence of the deal (*Deal Inv. Sequence*). In Panel (a), the vertical axis is the mean (Mean, blue solid line) or standard deviation (StdDev, red dashed line) of the deal PME in a given sequence category. In Panel (b), the vertical axis is the ratio between the mean of the deal PME and the standard deviation of the deal PMEs. *Deal Inv. Sequence* is the timing sequences of deals within the fund, sorted by investment year. One to five refers to the earliest to the latest. E.g. sequence one means that the deal is made in the first year when the fund begins to make investments. Sequence two means that the deal is made in the second year of the fund's investment period. We sequence deals to the year level due to data availability. We group deals with sequences larger than five into one category: rank five.



(a) Below 33% **(b) 33% - 67%** **(c) Above 67%**

By Deal Size Ranks



(d) Below 33% **(e) 33% - 67%** **(f) Above 67%**

By Deal Timing Sequences

Figure 4: Deal PME and Deal Size/Timing, Tertile Analysis by $\widehat{GPFE}$

Note: This figure plots figures in Figure 2 and Figure 3 for subsamples divided by tertiles of $\widehat{GPFE}$ where $\widehat{GPFE}$ are the predicted value for GP fixed effects in Table 4. In Panel (a) - (c), the horizontal axis is the rank of the deal (*Deal-Size-Rank*, the rank of the deal within the fund, sorted by size). In Panel (d) - (f), the horizontal axis is the sequence of the deal (*Deal Inv. Sequence*,the timing sequences of the deal within the fund, sorted by investment year). The vertical axis is the mean (Mean, blue solid line) or standard deviation (StdDev, red dashed line) of the deal PMEs in a given size rank/timing sequence category.

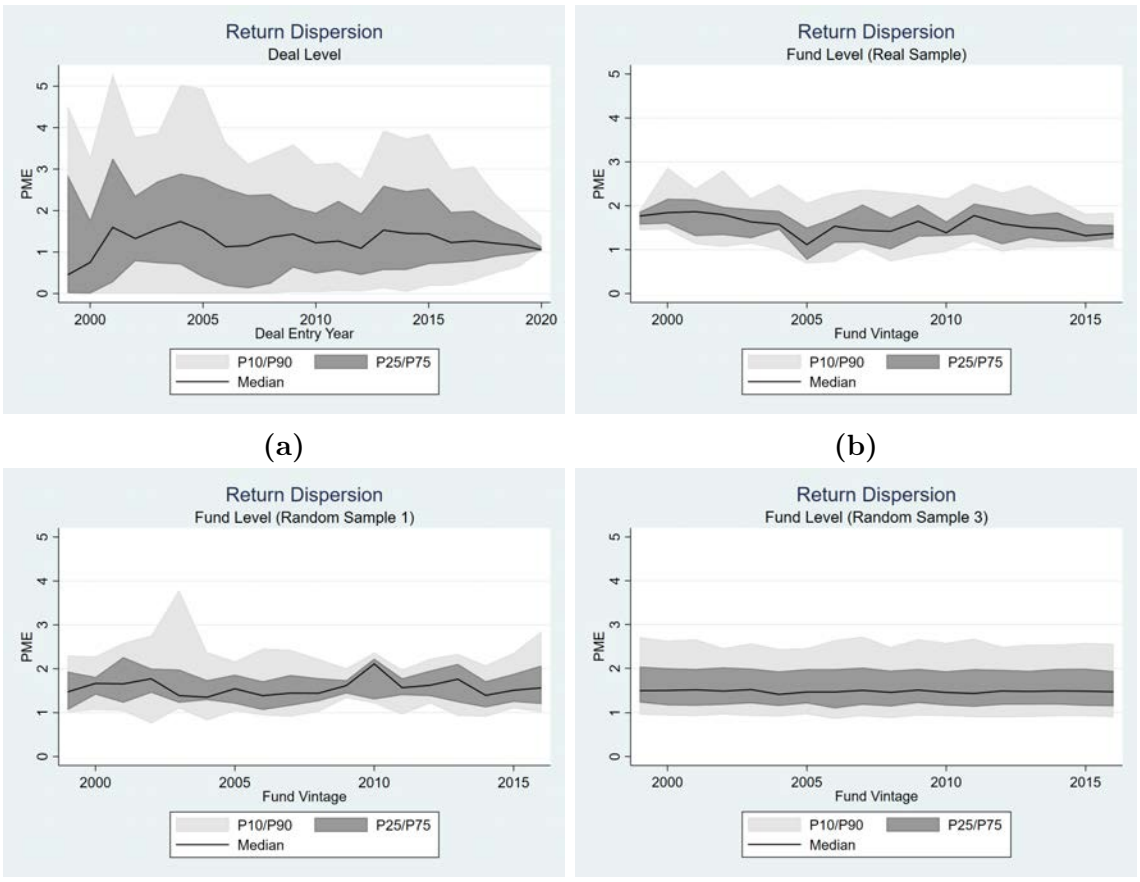


Figure 5: Return Dispersion: Fund-Level v.s. Deal-Level

Note: This figure shows the deal- (Panel (a)) and fund- (Panel (b)) level PME dispersion by deal entry year/fund vintage. Panels (c) and (d) plot the PME dispersion for randomly generated samples (as in Table A4). In Panel (c), the random sample is generated by drawing PMEs randomly (using uniform distribution) from the data without repetition. We keep the fund structure the same as in the data (i.e. how many deals one fund has). In Panel (d), the random sample is generated by drawing PMEs randomly from the data with repetition. The fund structure is also randomly generated (using uniform distribution). The dispersion is represented by the gap between the 90th and 10th percentiles and the 75th and 25th percentiles of PME.

Table 1: Variable Definitions

Variable	Definition
After Year 2	A dummy that equals one if the observation is a deal made after the first two years of a fund's life.
Deal Count	The total number of deals made by the fund.
Deal Duration	For exited deals: the number of years between the exited year and investment year. For active deals: the number of years between the investment year and the end of our sample period (2020).
Deal PME	Kaplan-Schoar Public Market Equivalent, the ratio of total value (market value + total proceeds) to total investment amount, each benchmarked by a public return. The total value is the actual realized value (including escrow) for realized deals and the latest reported market value for unrealized deals. The benchmark indices for PME are chosen in the following way: Russell 3000 for holdings in North America, the Asia & Pacific MSCI performance index for Asian & Pacific holdings, the Europe MSCI performance index for European holdings, and MSCI World performance index for other holdings. All indices are in USD.
Deal-Size Rank	The size rank of deals within a fund, with rank 1 denoting the largest deal. Rank 10 includes all deals ranked ten or higher.
Deal Sequence	The deal's investment year-$min_{j \in f}$investment year$_j + 1$. Deals made in year five and later are grouped together.
Early Success	A dummy that equals one if the fund made at least one deal with a PME greater than or equal to 3 within the first two years of the fund's life.
Exit Dummy	A dummy equaling one when the deal is exited and zero is the deal status is active.
Fraction Invested	The sum of deal sizes by the fund divided by fund size.
Fund Duration	Deal-size weighted sum of deal duration.
Fund PME	The deal-size weighted sum of deal-level PMEs held by the fund.
Fund Size	The total amount of committed capital in the fund (USD Billions).
Granularity	The Power Law parameter. We assume that the size distribution of the deals within a fund follows the Power Law ($e^{-a \times Rank}$). We calculate the a parameter for each deal in the fund and then average them across all deals in the fund.
HHI Sector	The sum of squared weights of deals in each sector s held by the fund where weights are deal size relative to fund size.
HHI Region	The sum of squared weights of deals in each region r held by the fund where weights are deal size relative to fund size.
North America	A dummy that equals to one if the fund invests all deals in North America.

Table 2: Summary Statistics for Key Variables

This table provides summary statistics for the primary variables in our analysis. Panel A reports the summary statistics for deal-level variables and Panel B for fund-level variables. Variable definitions are provided in Table 1.

Panel A: Deal-Level Summary Statistics								
Variables	N	Mean	S.D.	Min	5th per.	Median	95th per.	Max
After Year 2	5,925	0.58	0.49	0.00	0.00	1.00	1.00	1.00
Deal Duration	5,925	5.79	2.81	1.00	2.00	5.00	11.00	20.00
Deal Entry Year	5,925	2012	5.06	1999	2003	2014	2019	2020
Deal Exit Year	3,199	2014	4.32	2001	2006	2015	2020	2020
Deal PME	5,925	1.63	1.59	0.00	0.00	1.27	4.38	15.45
Deal Sequence	5,925	2.96	1.41	1.00	1.00	3.00	5.00	5.00
Deal Seq. (Non-Truncate)	5,925	3.09	1.66	1.00	1.00	3.00	6.00	13.00
Deal Size (billion, USD)	5,925	0.08	0.13	0.00	0.01	0.04	0.30	2.06
Deal-Size Rank	5,925	6.52	3.19	1.00	1.00	7.00	10.00	10.00
Deal-Size Rank (Non-Truncate)	5,925	8.32	6.47	1.00	1.00	7.00	21.00	49.00
Early Success	5,925	0.52	0.50	0.00	0.00	1.00	1.00	1.00
Exit Dummy	5,925	0.54	0.50	0.00	0.00	1.00	1.00	1.00
Panel B: Fund-Level Summary Statistics								
Variables	N	Mean	S.D.	Min	5th per.	Median	95th per.	Max
Deal Count	467	12.69	6.12	4.00	6.00	11.00	25.00	49.00
Fraction Invested	467	0.86	0.15	0.34	0.59	0.87	1.11	1.30
Fund Duration	467	5.23	2.18	0.82	2.06	5.04	9.18	12.73
Fund PME (Value-Weighted)	467	1.54	0.54	0.21	0.77	1.46	2.51	4.61
Fund PME Volatility	467	1.35	0.74	0.19	0.50	1.19	2.92	6.05
Fund Size (billion, USD)	467	1.20	1.77	0.04	0.12	0.60	4.54	17.52
Fund Vintage	467	2,010	4.65	1,999	2,001	2,012	2,016	2,016
Granularity	467	0.81	0.36	-0.02	0.26	0.79	1.47	2.46
HHI Sector	467	0.26	0.15	0.04	0.10	0.22	0.59	1.00
HHI Region	467	0.65	0.21	0.08	0.29	0.66	1.00	1.00
I(North American)	467	0.32	0.47	0.00	0.00	0.00	1.00	1.00

Table 3: Performance, Concentration, and Specialization

This table reports point estimates from OLS regressions connecting risk and return to concentration, industry specialization, and geographic specialization. The sample period spans from 1999 to 2016. In columns (1) to (4), the dependent variable is the weighted average of the Kaplan-Schoar PME of all deals in a given fund. Weights are the dollar value invested in the deal. In columns (5) to (8), the dependent variable is the standard deviation of deal-level PMEs in a given fund. The key independent variables are *Granularity*, *HHI Sector*, and *HHI Region*. Variable definitions are in Table 1. Fund vintage fixed effects are indicated in each column. Standard errors are clustered at the GP level. t statistics are reported in parentheses. *, **, and *** denote significance at the 90%, 95%, and 99% confidence levels, respectively.

	Fund PME (Value-Weighted)				Fund S.D. of Deal PME			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Granularity	0.24**			0.19*	0.64***			0.59***
	(2.21)			(1.75)	(3.79)			(3.34)
HHI Sector		0.41**		0.37**		0.52**		0.39*
		(2.44)		(2.16)		(2.48)		(1.82)
HHI Region			0.40**	0.36**			0.64***	0.53**
			(2.55)	(2.36)			(2.85)	(2.40)
Deal Count	-0.00	-0.00	-0.01	-0.00	0.02***	0.01	0.01	0.02***
	(-0.42)	(-1.22)	(-1.38)	(-0.04)	(2.97)	(1.56)	(1.47)	(3.29)
Fund Size	0.02	0.00	0.01	0.03	0.02	-0.05**	-0.04*	0.02
	(1.19)	(0.14)	(0.47)	(1.31)	(0.67)	(-2.21)	(-1.87)	(0.87)
Fund Duration	1.12***	0.74***	0.42	0.32	0.76**	0.12	-0.50	-0.31
	(4.04)	(2.67)	(1.36)	(0.96)	(2.06)	(0.32)	(-1.09)	(-0.66)
Fraction Invested	-0.19***	-0.19***	-0.19***	-0.19***	-0.09**	-0.09**	-0.08**	-0.09**
	(-6.44)	(-6.43)	(-6.36)	(-6.41)	(-2.52)	(-2.46)	(-2.32)	(-2.42)
North America	-0.09	-0.11*	-0.12*	-0.15**	-0.13	-0.14*	-0.18**	-0.20**
	(-1.42)	(-1.74)	(-1.95)	(-2.32)	(-1.59)	(-1.76)	(-2.05)	(-2.36)
Fund Vintage FE	Y	Y	Y	Y	Y	Y	Y	Y
Observations	467	467	467	467	467	467	467	467
Adjusted R^2	0.139	0.137	0.137	0.150	0.156	0.122	0.127	0.166

Table 4: Deal Level: Within-Fund Deal Size Rank, Timing Sequence and PME

This table reports point estimates from OLS regressions relating performance to deal size rank and deal sequence. The sample period spans from 1999 to 2016. The dependent variable is the deal-level PME. The first key independent variable is *Deal-Size Rank*, which minus one times the deal-size rank. (Flipping the sign of the rank allows for easier interpretation of the point estimate.) We group deals with ranks larger than ten into the tenth rank. The second key independent variable is *Deal Seq.*, which is the sequence of deals within the fund, in ascending order of investment year: i.e., *Deal Seq.*=1 means that the deal is invested in the first year of the fund, *Deal Seq.*=2 means that the deal is invested in the second year of the fund, and so forth. All deals after and including year 5 are grouped together. Variable definitions are in Table 1. Fixed effects for GP, industry and geographic location are indicated in each column. Standard errors are clustered at the GP level. t statistics are reported in parentheses. *, **, and *** denote significance at the 90%, 95%, and 99% confidence levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
Deal-Size Rank	-0.05***	-0.04***	-0.03***	-0.05***	-0.05***	-0.04***
	(-6.18)	(-5.61)	(-4.07)	(-6.82)	(-5.83)	(-4.61)
Deal Sequence	-0.05***	-0.05***	-0.05***	-0.05***	-0.05***	-0.06***
	(-2.63)	(-2.62)	(-3.02)	(-2.70)	(-2.70)	(-3.29)
Deal Size		-0.41*	0.06		-0.22	0.37
		(-1.77)	(0.19)		(-0.85)	(1.10)
Exit Dummy			0.56***			0.57***
			(11.91)			(9.55)
Fund Size			-0.02			-0.06
			(-1.06)			(-1.33)
Deal Duration			-0.09***			-0.09***
			(-11.89)			(-9.54)
GP FE	N	N	N	Y	Y	Y
Industry FE	N	N	Y	N	N	Y
Geography FE	N	N	Y	N	N	Y
Observations	5925	5925	5925	5925	5925	5925
Adjusted R^2	0.010	0.011	0.077	0.040	0.039	0.094

Table 5: Deal Level: Within-Fund Deal Size Rank, Timing Sequence and PME – Tercile Analysis by Estimates of GP Skill

This table reports parameter estimates from OLS regressions relating performance to deal size rank and deal sequence for subsamples divided by estimated GP fixed effects taken from Table 4. Subsamples are divided by terciles of GP fixed effects. The sample period spans from 1999 to 2016. The dependent variable is *Deal PME*. The key independent variables are *Deal-Size Rank* and *Deal Sequence*. Columns 3 and 6 include the same control variables as in Table 4. Columns (1)-(3) omit GP fixed effects, while columns (4)-(6) include them. Standard errors are clustered at the GP level. Variable definitions are in Table 1. t statistics are reported in parentheses. *, **, and *** denote significance at the 90%, 95%, and 99% confidence levels, respectively.

Panel A: Lowest tercile of GP Fixed Effects

	(1)	(2)	(3)	(4)	(5)	(6)
Deal-Size Rank	-0.03***	-0.03**	-0.02	-0.02*	-0.02*	-0.01
	(-2.62)	(-2.48)	(-1.37)	(-1.93)	(-1.74)	(-0.45)
Deal Sequence	-0.02	-0.02	-0.01	-0.02	-0.02	-0.02
	(-0.80)	(-0.78)	(-0.59)	(-0.64)	(-0.64)	(-0.52)
Deal Size		-0.18	0.42		0.06	0.48
		(-0.60)	(0.97)		(0.17)	(1.01)
Observations	1961	1961	1961	1961	1961	1961
Adjusted R^2	0.007	0.007	0.086	0.009	0.009	0.067

Panel B: Middle tercile of GP Fixed Effects

	(1)	(2)	(3)	(4)	(5)	(6)
Deal-Size Rank	-0.06***	-0.06***	-0.05***	-0.06***	-0.06***	-0.05***
	(-5.63)	(-5.42)	(-4.05)	(-5.80)	(-4.43)	(-3.66)
Deal Sequence	-0.04	-0.04	-0.05**	-0.04	-0.04	-0.07**
	(-1.59)	(-1.59)	(-2.10)	(-1.39)	(-1.39)	(-2.26)
Deal Size		0.02	-0.07		-0.24	0.06
		(0.09)	(-0.17)		(-0.48)	(0.13)
Observations	2013	2013	2013	2013	2013	2013
Adjusted R^2	0.019	0.019	0.118	0.003	0.002	0.085

Panel C: Highest tercile of GP Fixed Effects

	(1)	(2)	(3)	(4)	(5)	(6)
Deal-Size Rank	-0.06***	-0.05***	-0.05***	-0.08***	-0.07***	-0.07***
	(-3.31)	(-2.96)	(-3.25)	(-4.02)	(-3.66)	(-3.62)
Deal Sequence	-0.10***	-0.10***	-0.11***	-0.09**	-0.09**	-0.12***
	(-3.10)	(-3.19)	(-3.54)	(-2.51)	(-2.54)	(-3.56)
Deal Size		-0.82**	-0.04		-0.42	0.78
		(-2.61)	(-0.09)		(-0.91)	(1.14)
Observations	1951	1951	1951	1951	1951	1951
Adjusted R^2	0.012	0.014	0.076	0.012	0.012	0.062

Table 6: Deal-level Performance Following Early Success

This table reports point estimates from OLS regressions of deal-level PMEs on indicators for "early success" funds and later investments. The sample period spans from 1999 to 2016. The early success dummy variable equals one if the fund made at least one deal with a PME greater than or equal to 3 within the first two years of the fund's life. The dummy variable *After Year 2* equals one if the observation is a deal made after the first two years of a fund's life. Variable definitions are in Table 1. Fixed effects for GP, industry, and geographic location are indicated in each column. Standard errors are clustered at the GP level. t statistics are reported in parentheses. *, **, and *** denote significance at the 90%, 95%, and 99% confidence levels, respectively.

	(1)	(2)	(3)	(4)
Early Success Dummy	0.13*	0.08	0.05	0.01
	(1.95)	(1.27)	(0.45)	(0.11)
After Year 2 Dummy	0.37***	0.28***	0.35***	0.26***
	(6.61)	(4.77)	(5.94)	(4.17)
Early Success × After Year 2	-0.79***	-0.69***	-0.83***	-0.75***
	(-7.71)	(-6.76)	(-8.01)	(-7.08)
Exit Dummy		0.53***		0.56***
		(11.69)		(9.78)
Deal Size		-0.34		-0.38
		(-1.17)		(-1.23)
Fund Size		0.00		-0.01
		(0.37)		(-0.32)
Deal Duration		-0.09***		-0.09***
		(-11.75)		(-9.57)
GP FE	No	No	Yes	Yes
Industry FE	No	Yes	No	Yes
Geography FE	No	Yes	No	Yes
Observations	5925	5925	5925	5925
Adjusted R^2	0.035	0.097	0.050	0.105

Table 7: Fund-Level Return Variance Decomposition

This table reports posterior means and standard deviations of parameters of GP-Fund model (Equation 1). Panel A shows the parameter estimates, and Panel B reports the variance decomposition estimates. Specifications where we also include fund vintage fixed effects as covariates are indicated in each column. The model is estimated by Markov chain Monte Carlo (MCMC) using ten thousand burn-in cycles followed by 100,000 samples, saving every tenth draw. Signal to noise (skill%), overlap effect%, and error% are defined in Equation 3. The fund life N is fixed to ten. Return is measured in the natural logarithm TVPI. Variable definitions are in Table 1. Posterior standard deviations (Bayesian standard errors) are in brackets.

	(1)	(2)	(3)	(4)
Panel A: Parameter Estimates				
	KS Sample (1969 to 2001, Preqin)		Our Sample (1999 to 2016, Burgiss)	
$\sigma(GPRE)$	0.028	0.029	0.033	0.033
	(0.002)	(0.002)	(0.003)	(0.003)
$\sigma(GPYearRE)$	0.080	0.071	0.063	0.059
	(0.010)	(0.010)	(0.012)	(0.010)
$\sigma(\epsilon)$	0.499	0.481	0.397	0.363
	(0.030)	(0.031)	(0.034)	(0.032)
Vintage Year FE	N	Y	N	Y
N. of GPs	285	285	315	315
N. of Funds	624	624	467	467
Panel B: Variance Decomposition				
$N^2\sigma^2(GPRE)$	0.077	0.086	0.113	0.108
	(0.011)	(0.012)	(0.021)	(0.020)
$N\sigma^2(GPYearRE)$	0.066	0.051	0.041	0.036
	(0.015)	(0.014)	(0.016)	(0.012)
$\sigma^2(\epsilon)$	0.250	0.233	0.159	0.133
	(0.031)	(0.031)	(0.027)	(0.024)
$\sigma^2(FundReturn)$	0.393	0.369	0.312	0.277
	(0.030)	(0.030)	(0.032)	(0.027)
signal to noise (skill%)	19.59%	23.31%	36.22%	38.99%
overlap effect%	16.79%	13.82%	13.14%	13.00%
noise%	63.61%	63.14%	50.96%	48.01%

Table 8: Deal-Level Return Variance Decomposition

This table reports posterior means and standard deviations of parameters of GP-Deal model (Equation 4). Panel A shows the parameter estimates, and Panel B reports the variance decomposition estimates. Specifications where we also include deal entry year and deal exit year fixed effects as covariates are indicated in each column. The model is estimated by Markov chain Monte Carlo (MCMC) using ten thousand burn-in cycles followed by 100,000 samples, saving every tenth draw. Signal to noise (skill%), overlap effect%, and error% are defined in Equation 6. The deal life N is the deal life which is calculated as deal exit year-deal entry year+1. Return is measured in the natural logarithm TVPI. More details on the variable construction and the sample restriction are in Appendix Table 1. Posterior standard deviations (Bayesian standard errors) are in brackets.

	(1)	(2)	(3)	(4)
Panel A: Parameter Estimates				
$\sigma(GPRE)$	0.034	0.037	0.037	0.033
	(0.003)	(0.004)	(0.004)	(0.003)
$\sigma(GPYearRE)$	0.194	0.170	0.166	0.107
	(0.016)	(0.016)	(0.015)	(0.015)
$\sigma(\epsilon)$	1.023	1.030	1.024	1.023
	(0.013)	(0.013)	(0.013)	(0.012)
Deal Entry Year FE	N	Y	N	Y
Deal Exit Year FE	N	N	Y	Y
N. of GPs	315	315	315	315
N. of Deals	5925	5925	5925	5925
Panel B: Variance Decomposition				
$N^2\sigma^2(GPRE)$	0.048	0.057	0.058	0.047
	(0.009)	(0.012)	(0.012)	(0.009)
$N\sigma^2(GPYearRE)$	0.221	0.169	0.161	0.067
	(0.037)	(0.032)	(0.028)	(0.019)
$\sigma^2(\epsilon)$	1.047	1.062	1.049	1.046
	(0.027)	(0.027)	(0.026)	(0.025)
$\sigma^2(DealReturn)$	1.316	1.287	1.269	1.160
	(0.040)	(0.036)	(0.033)	(0.028)
signal to noise (skill%)	3.66%	4.41%	4.60%	4.01%
overlap effect%	16.72%	13.06%	12.70%	5.80%
noise%	79.62%	82.53%	82.71%	90.20%

Internet Appendices

Internet Appendix A: Additional Descriptive Tables

Table A1: Sector and Region Distributions

Sector	Freq.
Communication Services	318
Consumer Discretionary	1158
Consumer Staples	445
Energy	112
Financials	352
Health Care	794
Industrials	1282
Information Technology	987
Materials	350
Other	38
Real Estate	45
Utilities	44
Region	**Freq.**
Africa	21
Asia	228
Caribbean	18
Central America	4
Eastern Europe	66
Middle East	57
North America	3085
South America	105
South Pacific	172
Southeast Asia	138
Western Europe	2031

Table A2: Deal-level Correlations

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
(1) After Year 2	1.000									
(2) Deal Duration	-0.189***	1.000								
(3) Deal PME	-0.029**	-0.144***	1.000							
(4) Deal Seq.	0.861***	-0.206***	-0.039***	1.000						
(5) Deal Seq. (Non-Truncate)	0.798***	-0.188***	-0.036***	0.963***	1.000					
(6) Deal Size	0.008	0.101***	-0.052***	-0.004	-0.012	1.000				
(7) Deal-Size Rank	-0.006	-0.093***	0.094***	0.021	0.031**	-0.234***	1.000			
(8) Deal-Size Rank (Non-Truncate)	-0.032**	-0.064***	0.064***	-0.005	0.005	-0.202***	0.797***	1.000		
(9) Exit Dummy	-0.131***	0.172***	0.162***	-0.139***	-0.115***	-0.063***	0.133***	0.155***	1.000	
(10) Early Success	-0.090***	0.031**	0.144***	-0.088***	-0.092***	-0.052***	0.091***	0.125***	0.130***	1.000

*** p<0.01, ** p<0.05, * p<0.1

Table A3: Fund-level Correlations

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
(1) Deal Count	1.000										
(2) Fund Duration	0.294***	1.000									
(3) Fraction Invested	0.207***	0.677***	1.000								
(4) Fund PME	-0.007	-0.020	-0.071	1.000							
(5) Fund PME Volatility	0.079*	0.084*	0.153***	0.576***	1.000						
(6) Fund Size	0.426***	0.138***	0.100**	-0.038	-0.081*	1.000					
(7) Granularity	-0.581***	-0.337***	-0.212***	0.103**	0.138***	-0.655***	1.000				
(8) HHI Sector	-0.082*	0.429***	0.212***	0.045	0.058	-0.079*	0.057	1.000			
(9) HHI Region	0.073	0.734***	0.445***	0.064	0.118**	-0.086*	-0.076*	0.437***	1.000		
(10) Fund Vintage	-0.210***	-0.372***	-0.785***	-0.109**	-0.264***	0.000	0.103**	-0.036	-0.245***	1.000	
(11) I(North American)	-0.246***	-0.041	-0.044	-0.044	-0.052	-0.232***	0.243***	0.226***	0.265***	0.038	1.000

*** p<0.01, ** p<0.05, * p<0.1

Table A4: Deal Level: Summary Statistics of PME by Deal-Size Rank

This table shows the summary statistics of the deal-level *PME* ranked by deal size within a fund, where rank 1 is the largest deal in the fund, rank 2 is the second largest, and so forth. The smallest deals for funds with more than ten deals are grouped into rank 10. Variable definitions are in Table 1.

Panel A: Distribution of deal-level PMEs by size rank

					Percentiles						
Rank	N	Mean	S.D.	Min	5th	25th	Median	75th	95th	Max	mean/sd
1	467	1.26	1.24	0.00	0.00	0.39	1.01	1.74	3.55	11.55	1.01
2	467	1.39	1.13	0.00	0.00	0.47	1.26	1.94	3.47	5.81	1.23
3	467	1.51	1.33	0.00	0.00	0.61	1.17	2.10	3.86	7.51	1.14
4	467	1.52	1.35	0.00	0.00	0.61	1.27	2.07	3.95	9.26	1.13
5	464	1.68	1.58	0.00	0.00	0.62	1.32	2.23	4.20	12.63	1.06
6	456	1.68	1.54	0.00	0.00	0.72	1.34	2.21	4.08	15.43	1.09
7	440	1.68	1.51	0.00	0.00	0.75	1.31	2.29	4.81	9.93	1.12
8	402	1.60	1.77	0.00	0.00	0.57	1.18	2.09	4.06	15.45	0.90
9	360	1.71	1.51	0.00	0.01	0.77	1.40	2.22	4.06	13.56	1.13
10+	1935	1.78	1.82	0.00	0.00	0.64	1.35	2.30	5.06	14.41	0.98

Table A5: Deal Level Performance by Order of Deal

This table shows the summary statistics of the deal-level *PME* from earliest deal to latest deal. The group "5+" includes the fifth and all subsequent deals. Variable definitions are in Table 1.

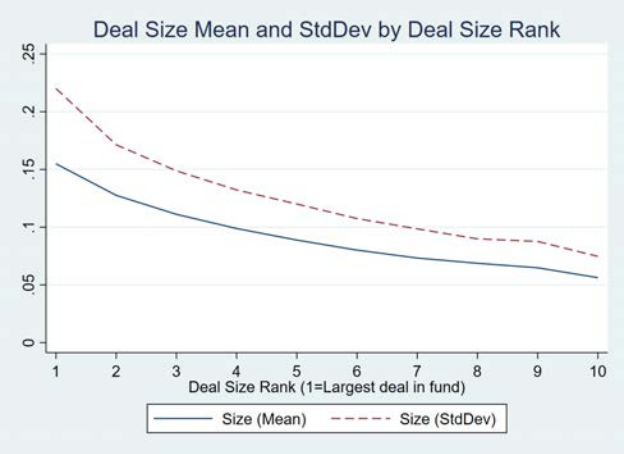
Panel A: Deal PME

					Percentiles						
Rank	N	Mean	S.D.	Min	5th	25th	Median	75th	95th	Max	mean/sd
1	1157	1.68	1.67	0.00	0.00	0.53	1.33	2.33	4.52	15.45	1.01
2	1324	1.69	1.71	0.00	0.00	0.54	1.31	2.25	4.93	15.43	0.98
3	1226	1.68	1.61	0.00	0.00	0.65	1.31	2.23	4.31	14.41	1.04
4	1030	1.56	1.42	0.00	0.00	0.68	1.26	2.02	4.20	13.93	1.10
5+	1188	1.52	1.46	0.00	0.00	0.67	1.18	1.93	4.01	12.77	1.05

Panel B: Deal Size (USD, Billions)

					Percentiles						
Rank	N	Mean	S.D.	Min	5th	25th	Median	75th	95th	Max	mean/sd
1	1157	0.08	0.14	0.00	0.00	0.02	0.04	0.09	0.27	2.06	0.58
2	1324	0.08	0.12	0.00	0.01	0.02	0.04	0.09	0.30	0.95	0.71
3	1226	0.09	0.14	0.00	0.01	0.02	0.04	0.10	0.31	1.75	0.66
4	1030	0.08	0.12	0.00	0.01	0.02	0.04	0.09	0.32	1.05	0.73
5+	1188	0.08	0.11	0.00	0.01	0.02	0.04	0.09	0.27	1.16	0.72

Internet Appendix B: Additional Figures

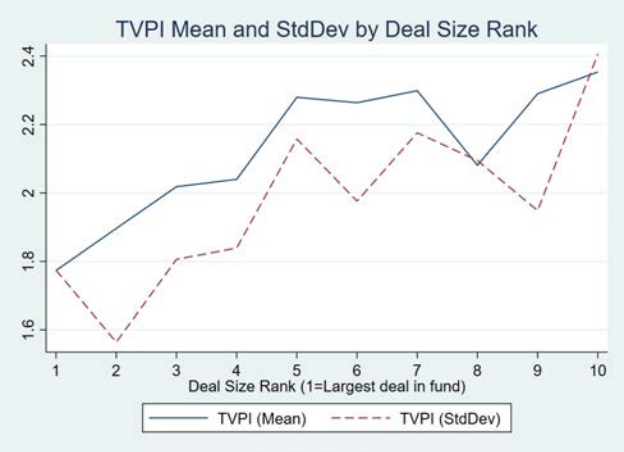


(a)

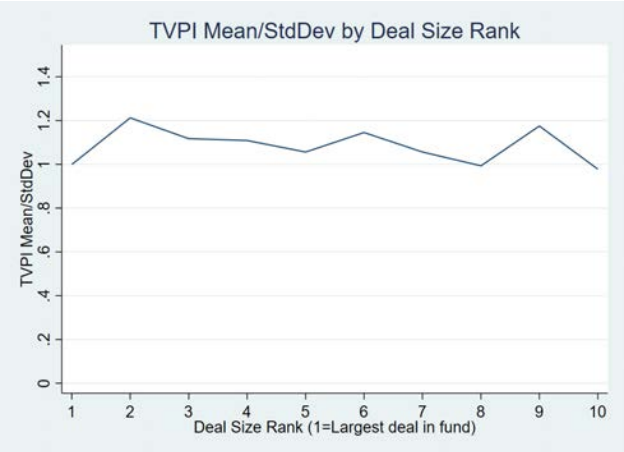
(b)

Figure B1: Deal Size by Deal-Size Rank/Timing Sequence in the Fund

Note: This figure plots the mean and standard deviation of deal sizes in each rank/sequence category. In Panel (a), the horizontal axis is the rank of the deal (*Deal-Size-Rank*). The vertical axis is the mean (Mean, blue solid line) or standard deviation (StdDev, red dashed line) of the deal size in a given rank category. Deal size is measured in USD billions. *Deal-Size-Rank* is the rank of the deal within the fund, sorted by size. 1 to 10 refers to the largest to the smallest. E.g., rank 1 means that the deal is the largest in the fund it belongs to. Rank 2 means that the deal is the second largest in the fund. We group deals with ranks larger than ten into one category: rank 10. In Panel (b), the horizontal axis is the sequence of the deal (*Deal Inv. Sequence*). The vertical axis is the mean (Mean, blue solid line) or standard deviation (StdDev, red dashed line) of the deal size in a given rank category. Deal size is measured in USD billions. *Deal Inv. Sequence* is the timing sequences of deals within the fund, sorted by investment year. One to five refers to the earliest to the latest. E.g. sequence one means that the deal is made in the first year when the fund begins to make investments. Sequence two means that the deal is made in the second year of the fund's investment period. We sequence deals to the year level due to data availability. We group deals with sequences larger than five into one category: rank five.



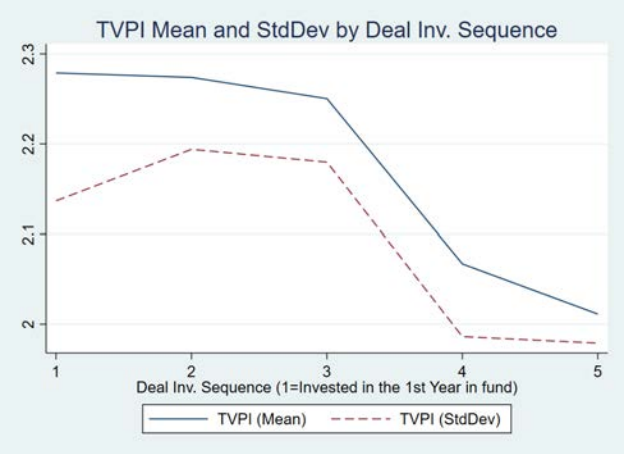
(a)



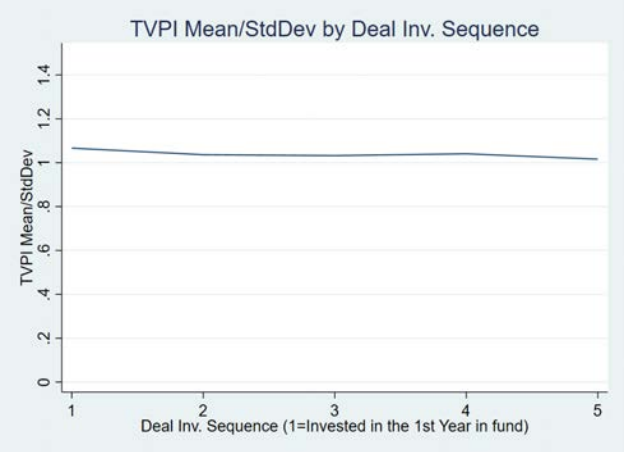
(b)

Figure B2: Deal Return (TVPI) by Deal Size Rank in the Fund

Note: This figure plots the mean and standard deviation of deal TVPI in each deal-size rank category. The horizontal axis is the rank of the deal (*Deal-Size-Rank*). In Panel (a), the vertical axis is the mean (Mean, blue solid line) or standard deviation (StdDev, red dashed line) of the deal TVPI in a given rank category. In Panel (b), the vertical axis is the ratio between the mean of the deal TVPI and the standard deviation of the deal PMEs. *Deal-Size-Rank* is the rank of deals within the fund, sorted by size. 1 to 10 refers to the largest to the smallest. E.g., rank 1 means that the deal is the largest in the fund it belongs to. Rank 2 means that the deal is the second largest in the fund. We group deals with ranks larger than ten into one category: rank 10.



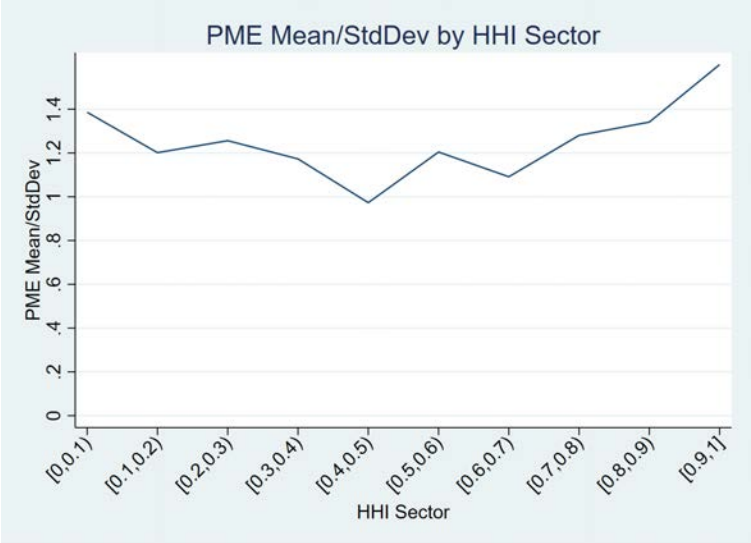
(a)



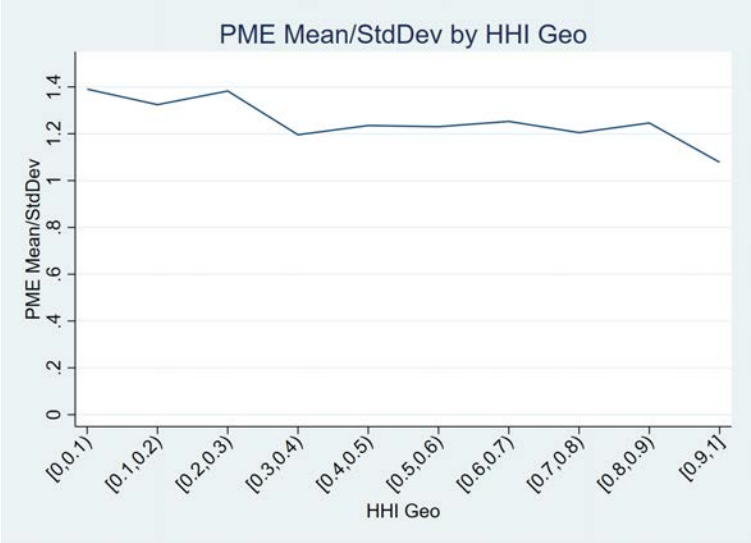
(b)

Figure B3: Deal Return (TVPI) by Deal Timing Sequence in the Fund

Note: This figure plots the mean and standard deviation of deal TVPI in each deal timing sequence category. The horizontal axis is the sequence of the deal (*Deal Inv. Sequence*). In Panel (a), the vertical axis is the mean (Mean, blue solid line) or standard deviation (StdDev, red dashed line) of the deal TVPI in a given sequence category. In Panel (b), the vertical axis is the ratio between the mean of the deal TVPI and the standard deviation of the deal PMEs. *Deal Inv. Sequence* is the timing sequences of deals within the fund, sorted by investment year. One to five refers to the earliest to the latest. E.g. sequence one means that the deal is made in the first year when the fund begins to make investments. Sequence two means that the deal is made in the second year of the fund's investment period. We sequence deals to the year level due to data availability. We group deals with sequences larger than five into one category: rank five.



(a)



(b)

Figure B4: Deal Mean-to-sd Ratio by Fund HHI

Note: This figure plots the mean to standard deviation ratio of deal PME within the same fund for each HHI category. Panel (a) plots the HHI sector, and Panel (b) plots the HHI geography. The horizontal axis is the ten HHI category intervals from 0 to 1, with 0.1 increment for each interval. The vertical axis is the average across all funds in the HHI category of the mean-to-sd ratio. More specifically, the ratio is a ratio between the mean of the deal PMEs and the standard deviation of the deal PMEs within the fund.

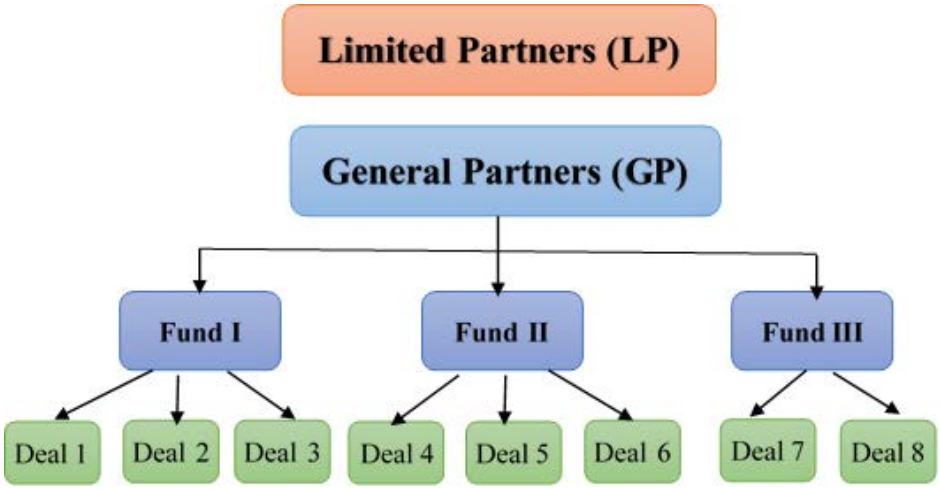


Figure B5: Hierarchical Structural of the Private Equity Industry

Note: This figure gives a graphic illustration of the hierarchical structure of the PE industry, with General Partners (GPs) at the top, managing funds in the middle that invest in deals at the bottom.

Internet Appendix C: Additional Analysis Tables

Table C1: Fund Level: Concentration, Specialization and TVPI

This table reports the cross-sectional regression results for the relationship between concentration, industry specialization, geographic specialization, and TVPI at the fund level. The sample period spans from 1999 to 2016 (both included). In columns (1) to (4), the dependent variable is *value-weighted TVPI* which is the weighted average of the total value to paid-in capital (multiples) performance measure of all deals in a given fund. Weights are the dollar value invested in the deal. In columns (5) to (8), the dependent variable is *S.D. of Deal TVPI* which is the standard deviation of deal TVPIs in a given fund. The key independent variables are *Granularity*, *HHI Sector*, and *HHI Region*. More details on the variable construction and the sample restrictions are in Appendix Table 1. Fund vintage fixed effects are indicated in each column. Standard errors are clustered at the GP level. t statistics are reported in parentheses. *, **, and *** denote significance at the 90%, 95%, and 99% confidence levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Value-Weighted TVPI				S.D. of Deal TVPI			
Granularity	0.35**			0.30*	0.84***			0.76***
	(2.15)			(1.80)	(3.91)			(3.41)
HHI Sector		0.62**		0.56**		0.82***		0.65**
		(2.54)		(2.18)		(2.84)		(2.20)
HHI Region			0.27	0.21			0.77**	0.62**
			(1.18)	(0.94)			(2.54)	(2.10)
N. of Deals	0.00	-0.00	-0.00	0.00	0.03***	0.01*	0.01	0.03***
	(0.21)	(-0.52)	(-0.81)	(0.52)	(3.02)	(1.79)	(1.53)	(3.34)
Fund Size	0.05	0.01	0.02	0.05	0.03	-0.05**	-0.04*	0.04
	(1.55)	(0.62)	(0.69)	(1.58)	(1.03)	(-2.21)	(-1.88)	(1.21)
Fund Duration	1.44***	0.87**	0.89**	0.74	1.06**	0.14	-0.47	-0.33
	(3.74)	(2.26)	(2.03)	(1.55)	(2.27)	(0.30)	(-0.81)	(-0.55)
Fraction Invested	-0.17***	-0.17***	-0.16***	-0.17***	-0.04	-0.04	-0.03	-0.03
	(-4.51)	(-4.49)	(-4.44)	(-4.47)	(-0.87)	(-0.81)	(-0.65)	(-0.74)
I(North American)	0.26***	0.23***	0.24***	0.20**	0.11	0.08	0.06	0.01
	(2.96)	(2.65)	(2.76)	(2.34)	(1.09)	(0.79)	(0.53)	(0.11)
Fund Vintage FE	Y	Y	Y	Y	Y	Y	Y	Y
Observations	467	467	467	467	467	467	467	467
Adjusted R^2	0.223	0.222	0.213	0.229	0.221	0.188	0.188	0.233

Table C2: Fund-Level: Concentration, Specialization and Return (Equal-Weighted)

This table reports the cross-sectional regressions on the relationship between deal concentration and industry/geographic specialization degree and returns at the fund level. The sample period spans from 1999 to 2016 (both included). n columns (1) to (4), the dependent variable is equal-weighted PME, and in columns (5) to (8), the dependent variable is equal-weighted TVPI. The key independent variables are *Granularity*, *HHI Sector*, and *HHI Region*. More details on the variable calculations and the sample restrictions are in Appendix Table 1. Fund vintage fixed effects are indicated in each column. Standard errors are clustered at the GP level. t statistics are reported in parentheses. *, **, and *** denote significance at the 90%, 95%, and 99% confidence levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Fund PME (Equal-Weighted)				Fund TVPI (Equal-Weighted)			
Granularity	0.34***			0.28**	0.49***			0.41**
	(2.74)			(2.15)	(2.86)			(2.31)
HHI Sector		0.50***		0.44**		0.81***		0.72***
		(2.85)		(2.46)		(3.16)		(2.69)
HHI Region			0.61***	0.55***			0.54**	0.46**
			(3.64)	(3.41)			(2.30)	(2.02)
N. of Deals	0.00	-0.00	-0.00	0.00	0.01	0.00	-0.00	0.01
	(0.53)	(-0.47)	(-0.57)	(1.04)	(1.00)	(0.08)	(-0.22)	(1.44)
Fund Size	0.02	-0.01	-0.00	0.03	0.04	-0.00	0.00	0.05*
	(1.14)	(-0.63)	(-0.07)	(1.36)	(1.64)	(-0.05)	(0.24)	(1.73)
Fund Duration	1.20***	0.72**	0.14	0.06	1.67***	0.92**	0.64	0.47
	(3.75)	(2.24)	(0.41)	(0.16)	(4.15)	(2.25)	(1.40)	(0.95)
Fraction Invested	-0.19***	-0.19***	-0.19***	-0.19***	-0.18***	-0.18***	-0.17***	-0.17***
	(-5.66)	(-5.62)	(-5.55)	(-5.61)	(-4.68)	(-4.64)	(-4.56)	(-4.66)
I(North American)	-0.12*	-0.14**	-0.17**	-0.20***	0.25***	0.21**	0.21**	0.16*
	(-1.69)	(-2.02)	(-2.36)	(-2.74)	(2.72)	(2.36)	(2.23)	(1.75)
Fund Vintage FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	467	467	467	467	467	467	467	467
Adjusted R^2	0.135	0.129	0.136	0.157	0.234	0.231	0.222	0.250

Table C3: Fund Level: Concentration, Specialization and Downside Risk

This table reports the cross-sectional regression results for the relationship between concentration, industry specialization, geographic specialization, and downside risks at the fund level. We measure downside risks by semi-standard deviation, i.e., only the below-mean return variations. The sample period spans from 1999 to 2016 (both included). In columns (1) to (4), the dependent variable is *Semi-S.D. of Deal PME* which is the semi-standard deviation of deal PMEs in a given fund. Weights are the dollar value invested in the deal. In columns (5) to (8), the dependent variable is *Semi-S.D. of Deal TVPI*, which is the semi-standard deviation of deal TVPIs (the total value to paid-in capital) in a given fund. The key independent variables are *Granularity*, *HHI Sector*, and *HHI Region*. More details on the variable construction and the sample restrictions are in Appendix Table 1. Fund vintage fixed effects are indicated in each column. Standard errors are clustered at the GP level. t statistics are reported in parentheses. *, **, and *** denote significance at the 90%, 95%, and 99% confidence levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Semi-S.D. of Deal PME				Semi-S.D. of Deal TVPI			
Granularity	0.32***			0.28**	0.42***			0.38**
	(2.94)			(2.47)	(2.88)			(2.46)
HHI Sector		0.31**		0.25		0.37		0.29
		(2.03)		(1.54)		(1.62)		(1.16)
HHI Region			0.37***	0.31**			0.35*	0.28
			(2.86)	(2.45)			(1.88)	(1.47)
N. of Deals	0.01**	0.00	0.00	0.01**	0.01*	0.00	0.00	0.01**
	(1.98)	(0.65)	(0.57)	(2.40)	(1.91)	(0.72)	(0.58)	(2.17)
Fund Size	0.01	-0.02**	-0.02*	0.01	0.02	-0.02*	-0.02*	0.02
	(0.48)	(-2.04)	(-1.71)	(0.68)	(0.82)	(-1.95)	(-1.66)	(0.96)
Fund Duration	0.53**	0.18	-0.16	-0.11	0.80***	0.36	0.08	0.18
	(2.26)	(0.79)	(-0.61)	(-0.38)	(2.69)	(1.30)	(0.21)	(0.43)
Fraction Invested	-0.07***	-0.07***	-0.07***	-0.07***	-0.05**	-0.05**	-0.05*	-0.05*
	(-3.20)	(-3.16)	(-3.00)	(-3.08)	(-2.06)	(-2.02)	(-1.88)	(-1.96)
I(North American)	-0.07	-0.09*	-0.10**	-0.12**	0.12*	0.11*	0.09	0.07
	(-1.48)	(-1.78)	(-1.99)	(-2.40)	(1.89)	(1.78)	(1.49)	(1.23)
Fund Vintage FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	467	467	467	467	467	467	467	467
Adjusted R^2	0.214	0.193	0.197	0.224	0.308	0.286	0.286	0.313

Table C4: Fund Level: Concentration, Specialization and "Jackknife" PME

This table reports the cross-sectional regression results for the relationship between concentration, industry specialization, geographic specialization, and "jackknife" (leave one out) returns at the fund level. "Jackknife" returns are calculated as the observation deal return minus the average return of all deals made in the same year (leaving one out). The sample period spans from 1999 to 2016 (both included). In columns (1) to (4), the dependent variable is *Mean of Deal "Jackknife" PME*. Weights are the dollar value invested in the deal. In columns (5) to (8), the dependent variable is *S.D. of Deal "Jackknife" PME*. The key independent variables are *Granularity*, *HHI Sector*, and *HHI Region*. More details on the variable construction and the sample restrictions are in Appendix Table 1. Fund vintage fixed effects are indicated in each column. Standard errors are clustered at the GP level. t statistics are reported in parentheses. *, **, and *** denote significance at the 90%, 95%, and 99% confidence levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Mean of Deal "Jackknife" PME				S.D. of Deal "Jackknife" PME			
Granularity	0.22**			0.18*	0.64***			0.58***
	(2.08)			(1.65)	(3.76)			(3.32)
HHI Sector		0.38**		0.34**		0.52**		0.39*
		(2.32)		(2.04)		(2.50)		(1.82)
HHI Region			0.36**	0.32**			0.61***	0.50**
			(2.28)	(2.10)			(2.70)	(2.25)
N. of Deals	-0.00	-0.00	-0.00	0.00	0.02***	0.01	0.01	0.02***
	(-0.07)	(-0.81)	(-0.95)	(0.30)	(3.00)	(1.58)	(1.49)	(3.32)
Fund Size	0.02	0.00	0.01	0.03	0.02	-0.05**	-0.04*	0.02
	(1.18)	(0.18)	(0.48)	(1.29)	(0.64)	(-2.23)	(-1.91)	(0.83)
Fund Duration	1.06***	0.71**	0.42	0.34	0.75**	0.11	-0.46	-0.28
	(3.82)	(2.55)	(1.40)	(1.01)	(2.03)	(0.29)	(-1.03)	(-0.59)
Fraction Invested	-0.20***	-0.20***	-0.20***	-0.20***	-0.09**	-0.08**	-0.08**	-0.08**
	(-6.67)	(-6.67)	(-6.63)	(-6.66)	(-2.33)	(-2.26)	(-2.13)	(-2.22)
I(North American)	-0.07	-0.09	-0.10	-0.12*	-0.12	-0.14*	-0.17*	-0.19**
	(-1.15)	(-1.44)	(-1.62)	(-1.96)	(-1.50)	(-1.68)	(-1.93)	(-2.24)
Fund Vintage FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	467	467	467	467	467	467	467	467
Adjusted R^2	0.130	0.129	0.128	0.139	0.155	0.121	0.125	0.164

Table C5: Fund Level: Concentration, Specialization and "Jackknife" TVPI

This table reports the cross-sectional regression results for the relationship between concentration, industry specialization, geographic specialization, and "jackknife" (leave one out) returns at the fund level. "Jackknife" returns are calculated as the observation deal return minus the average return of all deals made in the same year (leaving one out). The sample period spans from 1999 to 2016 (both included). In columns (1) to (4), the dependent variable is *Mean of Deal "Jackknife" TVPI*. Weights are the dollar value invested in the deal. In columns (5) to (8), the dependent variable is *S.D. of Deal "Jackknife" TVPI*. The key independent variables are *Granularity*, *HHI Sector*, and *HHI Region*. More details on the variable construction and the sample restrictions are in Appendix Table 1. Fund vintage fixed effects are indicated in each column. Standard errors are clustered at the GP level. t statistics are reported in parentheses. *, **, and *** denote significance at the 90%, 95%, and 99% confidence levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Mean of Deal "Jackknife" TVPI				S.D. of Deal "Jackknife" TVPI			
Granularity	0.34**			0.29*	0.83***			0.75***
	(2.21)			(1.89)	(3.91)			(3.40)
HHI Sector		0.54**		0.48**		0.83***		0.66**
		(2.45)		(2.07)		(2.97)		(2.30)
HHI Region			0.28	0.22			0.75**	0.60**
			(1.25)	(1.00)			(2.50)	(2.04)
N. of Deals 0.00	-0.00	-0.00	0.01	0.03***	0.01*	0.01	0.03***	
	(0.55)	(-0.16)	(-0.39)	(0.82)	(3.07)	(1.84)	(1.56)	(3.40)
Fund Size	0.05	0.02	0.02	0.05*	0.03	-0.05**	-0.04*	0.03
	(1.65)	(0.70)	(0.80)	(1.68)	(0.95)	(-2.28)	(-1.95)	(1.13)
Fund Duration	1.40***	0.89**	0.83**	0.74	1.00**	0.07	-0.50	-0.37
	(3.77)	(2.35)	(1.97)	(1.58)	(2.17)	(0.16)	(-0.87)	(-0.62)
Fraction Invested	-0.19***	-0.19***	-0.19***	-0.19***	-0.03	-0.02	-0.02	-0.02
	(-5.01)	(-5.00)	(-4.98)	(-4.97)	(-0.62)	(-0.57)	(-0.40)	(-0.49)
I(North American)	0.29***	0.27***	0.27***	0.24***	0.11	0.08	0.06	0.01
	(3.56)	(3.23)	(3.30)	(2.88)	(1.10)	(0.79)	(0.56)	(0.13)
Fund Vintage FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	467	467	467	467	467	467	467	467
Adjusted R^2	0.104	0.101	0.094	0.109	0.222	0.190	0.188	0.234

Table C6: Fund Level: Concentration, Specialization and Downside Risk ("Jackknife" Returns)

This table reports the cross-sectional regression results for the relationship between concentration, industry specialization, geographic specialization, and downside risks ("jackknife" returns) at the fund level. We measure downside risks ("jackknife" returns) by semi-standard deviation (similar to Table C3). The sample period spans from 1999 to 2016 (both included). "Jackknife" returns are calculated as the observation deal return minus the average return of all deals made in the same year (leaving one out). In columns (1) to (4), the dependent variable is *Semi-S.D. of Deal " Jackknife" PME*. Weights are the dollar value invested in the deal. In columns (5) to (8), the dependent variable is *Semi-S.D. of Deal "Jackknife" TVPI*. The key independent variables are *Granularity*, *HHI Sector*, and *HHI Region*. More details on the variable construction and the sample restrictions are in Appendix Table 1. Fund vintage fixed effects are indicated in each column. Standard errors are clustered at the GP level. t statistics are reported in parentheses. *, **, and *** denote significance at the 90%, 95%, and 99% confidence levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Semi-S.D. of Deal "Jackknife" PME				Semi-S.D. of Deal "Jackknife" TVPI			
Granularity	0.28***			0.25**	0.37***			0.33**
	(2.64)			(2.19)	(2.60)			(2.18)
HHI Sector		0.31**		0.25		0.38*		0.31
		(2.09)		(1.65)		(1.81)		(1.36)
HHI Region			0.33**	0.28**			0.34*	0.27
			(2.57)	(2.20)			(1.86)	(1.48)
N. of Deals	0.01*	0.00	0.00	0.01**	0.01*	0.00	0.00	0.01**
	(1.91)	(0.77)	(0.65)	(2.31)	(1.94)	(0.92)	(0.76)	(2.22)
Fund Size	0.00	-0.02**	-0.02*	0.01	0.01	-0.03**	-0.02*	0.01
	(0.26)	(-2.03)	(-1.73)	(0.44)	(0.37)	(-2.20)	(-1.94)	(0.51)
Fund Duration	0.51**	0.18	-0.10	-0.08	0.57*	0.15	-0.10	-0.05
	(2.22)	(0.82)	(-0.39)	(-0.28)	(1.94)	(0.56)	(-0.28)	(-0.13)
Fraction Invested	-0.07***	-0.07***	-0.07***	-0.07***	-0.04	-0.04	-0.03	-0.04
	(-3.24)	(-3.21)	(-3.07)	(-3.13)	(-1.46)	(-1.43)	(-1.30)	(-1.37)
I(North American)	-0.06	-0.07	-0.09*	-0.10**	0.11*	0.09	0.08	0.06
	(-1.26)	(-1.58)	(-1.73)	(-2.14)	(1.71)	(1.57)	(1.33)	(1.04)
Fund Vintage FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	467	467	467	467	467	467	467	467
Adjusted R^2	0.197	0.182	0.184	0.206	0.279	0.264	0.262	0.285

Table C7: Deal Level: Within-Fund Deal Size Rank, Timing Sequence and PME (Robustness –Fraction Invested $\geq$ 80%)

This table reports the robustness check results where we restrict to deals held by funds that has a Fraction Invested ratio $\geq$ 80%. The results are for the cross-sectional regressions about within-fund deal size rank and PME and within-fund deal sequence and PME. The sample period spans from 1999 to 2016 (both included). The dependent variable is Deal PME which is the Kaplan-Schoar Public Market Equivalence performance measure at the deal level. The first key independent variable is Deal-Size Rank which is the rank of deal within the fund, sorted by size. 1 to 10 refers to the largest to the smallest. E.g., rank 1 means that the deal is the largest in the fund that it belongs to. Rank 2 means that the deal is the second largest in the fund. We group deals with ranks larger than 10 into one single category: rank 10. To make the results be interpreted more easily, we use the negative of the rank (i.e., -1 for rank 1) in the regression. The second key independent variable is Deal Seq., which is the sequence of deal within the fund, sorted by investment year. One to five refers to the earliest to the latest. E.g., Sequence one means that the deal is invested in the first year in the fund that it belongs to. Sequence two means that the deal is invested in the second year in the fund. We group deals with sequences larger than five into one single category: sequence five. More details on the variable calculation as well as the sample restrictions are in Appendix Table 1. Fixed effects for GP, industry, and geographic location are indicated in each column. Standard errors are clustered at GP level. t statistics are reported in parentheses. *, **, and *** denote significance at the 90%, 95%, and 99% confidence levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
			Deal PME			
Deal-Size Rank	-0.04***	-0.04***	-0.03***	-0.05***	-0.04***	-0.04***
	(-4.98)	(-4.45)	(-3.09)	(-5.44)	(-4.51)	(-3.51)
Deal Seq.	-0.04*	-0.04*	-0.04*	-0.03	-0.03	-0.05**
	(-1.68)	(-1.67)	(-1.93)	(-1.42)	(-1.42)	(-2.14)
Absolute Deal Size		-0.30	0.21		-0.19	0.41
		(-1.32)	(0.66)		(-0.75)	(1.23)
Exit Dummy			0.50***			0.45***
			(9.91)			(6.47)
Fund Size			-0.02			-0.08*
			(-1.32)			(-1.82)
Deal Duration			-0.09***			-0.09***
			(-11.11)			(-8.20)
GP FE	No	No	No	Yes	Yes	Yes
Industry FE	No	No	Yes	No	No	Yes
Geography FE	No	No	Yes	No	No	Yes
Observations	4484	4484	4484	4484	4484	4484
Adjusted R^2	0.008	0.009	0.072	0.038	0.038	0.088

Table C8: Deal-Level: Deal Size Position and PME

This table reports the results for the cross-sectional regressions about within-fund deal size rank and PME. The sample period spans from 1999 to 2016 (both included). Panel A covers all funds. In Panel B, we restrict to deals held by funds that has a Fraction Invested ratio $\geq$ 80%. The dependent variable is Deal PME which is the Kaplan-Schoar Public Market Equivalence performance measure at the deal level. The key independent variable is Deal-Size Rank which is the rank of deal within the fund, sorted by size. 1 to 10 refers to the largest to the smallest. E.g., rank 1 means that the deal is the largest in the fund that it belongs to. Rank 2 means that the deal is the second largest in the fund. We group deals with ranks larger than 10 into one single category: rank 10. To make the results be interpreted more easily, we use the negative of the rank (i.e., -1 for rank 1) in the regression. More details on the variable calculation as well as the sample restrictions are in Appendix Table 1. Fixed effects for GP, industry, and geographic location are indicated in each column. Standard errors are clustered at GP level. t statistics are reported in parentheses. *, **, and *** denote significance at the 90%, 95%, and 99% confidence levels, respectively.

Panel A: All Funds						
	(1)	(2)	(3)	(4)	(5)	(6)
			Deal PME			
Deal-Size Rank	-0.05***	-0.04***	-0.03***	-0.05***	-0.05***	-0.04***
	(-6.14)	(-5.55)	(-4.11)	(-6.84)	(-5.84)	(-4.74)
Absolute Deal Size		-0.41*	0.10		-0.21	0.42
		(-1.74)	(0.33)		(-0.82)	(1.24)
Exit Dummy			0.58***			0.61***
			(12.14)			(10.42)
Fund Size			-0.02			-0.06
			(-1.24)			(-1.36)
Deal Duration			-0.09***			-0.08***
			(-11.54)			(-9.15)
GP FE	No	No	No	Yes	Yes	Yes
Industry FE	No	No	Yes	No	No	Yes
Geography FE	No	No	Yes	No	No	Yes
Observations	5925	5925	5925	5925	5925	5925
Adjusted R^2	0.009	0.010	0.075	0.038	0.038	0.091
Panel B: Fraction Invested ratio $\geq$ 80%						
	(1)	(2)	(3)	(4)	(5)	(6)
			Deal PME			
Deal-Size Rank	-0.04***	-0.04***	-0.03***	-0.05***	-0.04***	-0.04***
	(-4.98)	(-4.45)	(-3.19)	(-5.48)	(-4.55)	(-3.67)
Absolute Deal Size		-0.30	0.25		-0.18	0.47
		(-1.30)	(0.78)		(-0.71)	(1.37)
Exit Dummy			0.52***			0.48***
			(10.17)			(7.32)
Fund Size			-0.03			-0.08*
			(-1.43)			(-1.80)
Deal Duration			-0.09***			-0.09***
			(-10.87)			(-8.11)
GP FE	No	No	No	Yes	Yes	Yes
Industry FE	No	No	Yes	No	No	Yes
Geography FE	No	No	Yes	No	No	Yes
Observations	4484	4484	4484	4484	4484	4484
Adjusted R^2	0.007	0.008	0.071	0.038	0.037	0.086

Table C9: Deal-Level: Deal Timing Sequence and PME

This table reports the results for the cross-sectional regressions about within-fund deal timing sequence and PME. The sample period spans from 1999 to 2016 (both included). Panel A covers all funds. In Panel B, we restrict to deals held by funds that has a Fraction Invested ratio $\geq 80\%$. The dependent variable is Deal PME which is the Kaplan-Schoar Public Market Equivalence performance measure at the deal level. The key independent variable is Deal Seq., which is the sequence of deal within the fund, sorted by investment year. One to five refers to the earliest to the latest. E.g., Sequence one means that the deal is invested in the first year in the fund that it belongs to. Sequence two means that the deal is invested in the second year in the fund. We group deals with sequences larger than five into one single category: sequence five. More details on the variable calculation as well as the sample restrictions are in Appendix Table 1. Fixed effects for GP, industry, and geographic location are indicated in each column. Standard errors are clustered at GP level. t statistics are reported in parentheses. *, **, and *** denote significance at the 90%, 95%, and 99% confidence levels, respectively.

Panel A: All Funds						
	(1)	(2)	(3)	(4)	(5)	(6)
			Deal PME			
Deal Seq.	-0.04**	-0.04**	-0.06***	-0.05***	-0.05***	-0.07***
	(-2.54)	(-2.53)	(-3.08)	(-2.69)	(-2.70)	(-3.45)
Absolute Deal Size		-0.66***	-0.48*		-1.02***	-0.44
		(-2.72)	(-1.70)		(-3.35)	(-1.52)
Exit Dummy			0.59***			0.58***
			(12.34)			(9.75)
Fund Size			0.01			-0.01
			(0.63)			(-0.32)
Deal Duration			-0.10***			-0.10***
			(-12.15)			(-10.08)
Industry FE	No	No	Yes	No	No	Yes
Geography FE	No	No	Yes	No	No	Yes
Observations	5925	5925	5925	5925	5925	5925
Adjusted R^2	0.001	0.004	0.074	0.029	0.032	0.090
Panel B: Fraction Invested ratio $\geq 80\%$						
	(1)	(2)	(3)	(4)	(5)	(6)
			Deal PME			
Deal Seq.	-0.04*	-0.04*	-0.04**	-0.03	-0.03	-0.05**
	(-1.69)	(-1.68)	(-2.04)	(-1.47)	(-1.48)	(-2.30)
Absolute Deal Size		-0.54**	-0.23		-0.85***	-0.25
		(-2.26)	(-0.88)		(-3.09)	(-0.90)
Exit Dummy			0.52***			0.46***
			(10.11)			(6.54)
Fund Size			-0.00			-0.05
			(-0.11)			(-1.10)
Deal Duration			-0.09***			-0.10***
			(-11.36)			(-8.75)
Industry FE	No	No	Yes	No	No	Yes
Geography FE	No	No	Yes	No	No	Yes
Observations	4484	4484	4484	4484	4484	4484
Adjusted R^2	0.001	0.003	0.070	0.030	0.032	0.085

Table C10: Deal Level: Within-Fund Deal Size Rank, Timing Sequence and PME
(Robustness – Not Grouping Rank and Sequence Category)

This table reports the robustness check results where we use the original sequence number and rank number of the deals (i.e., we do not group deals with rank higher than ten together, similarly, we do not group deals with a sequence later than five together). To make the results be interpreted more easily, we use the negative of the rank (i.e., -1 for rank 1) in the regression. We conduct the cross-sectional regressions about within-fund size rank, sequence and PME at the deal level. The sample period spans from 1999 to 2016 (both included). Panel A covers all funds. In Panel B, we restrict to deals held by funds that has a Fraction Invested ratio $\geq$ 80%. More details on the variable calculation as well as the sample restrictions are in Appendix Table 1. Fixed effects for GP, industry, and geographic location are indicated in each column. Standard errors are clustered at GP level. t statistics are reported in parentheses. *, **, and *** denote significance at the 90%, 95%, and 99% confidence levels, respectively.

Panel A: All Funds						
	(1)	(2)	(3)	(4)	(5)	(6)
			Deal PME			
Deal-Size Rank	-0.02***	-0.02***	-0.01**	-0.03***	-0.03***	-0.02***
	(-3.93)	(-3.48)	(-2.14)	(-5.18)	(-4.50)	(-2.91)
Deal Seq.	-0.04**	-0.04**	-0.04***	-0.04**	-0.04**	-0.05***
	(-2.32)	(-2.34)	(-2.77)	(-2.22)	(-2.23)	(-2.92)
Absolute Deal Size		-0.47**	-0.03		-0.23	0.28
		(-1.98)	(-0.10)		(-0.93)	(0.81)
Exit Dummy			0.57***			0.57***
			(12.40)			(9.71)
Fund Size			-0.02			-0.06
			(-0.87)			(-1.16)
Deal Duration			-0.09***			-0.09***
			(-12.07)			(-9.69)
GP FE	No	No	No	Yes	Yes	Yes
Industry FE	No	No	Yes	No	No	Yes
Geography FE	No	No	Yes	No	No	Yes
Observations	5925	5925	5925	5925	5925	5925
Adjusted R^2	0.008	0.009	0.075	0.038	0.038	0.092
Panel B: Fraction Invested ratio $\geq$ 80%						
	(1)	(2)	(3)	(4)	(5)	(6)
			Deal PME			
Deal-Size Rank	-0.02***	-0.02**	-0.01	-0.03***	-0.02***	-0.01*
	(-2.83)	(-2.42)	(-1.23)	(-3.79)	(-2.97)	(-1.67)
Deal Seq.	-0.03*	-0.03*	-0.04**	-0.02	-0.02	-0.04**
	(-1.68)	(-1.70)	(-1.98)	(-1.32)	(-1.34)	(-2.12)
Absolute Deal Size		-0.38	0.05		-0.28	0.20
		(-1.61)	(0.14)		(-1.11)	(0.56)
Exit Dummy			0.51***			0.45***
			(10.38)			(6.53)
Fund Size			-0.02			-0.07
			(-0.90)			(-1.47)
Deal Duration			-0.09***			-0.09***
			(-11.45)			(-8.56)
GP FE	No	No	No	Yes	Yes	Yes
Industry FE	No	No	Yes	No	No	Yes
Geography FE	No	No	Yes	No	No	Yes
Observations	4484	4484	4484	4484	4484	4484
Adjusted R^2	0.005	0.006	0.070	0.036	0.036	0.086

Table C11: Deal Level: Within-Fund Deal Size Rank, Timing Sequence and TVPI

This table reports the cross-sectional regression results for the relationship between within-fund deal size rank, within-fund deal sequence, and TVPI. The sample period spans from 1999 to 2016 (both included). All are the same as in Table 4 except that the dependent variable is *Deal TVPI*, which is the total value to paid-in capital (multiples) performance measure at the deal level. The first key independent variable is *Deal-Size Rank* which is the negative of the rank of the deal within the fund, sorted by size. 1 to 10 refers to the largest to the smallest. E.g., rank 1 means that the deal is the largest in the fund it belongs to. Rank 2 means that the deal is the second largest in the fund. We group deals with ranks larger than ten into one category: rank 10. To interpret the results more easily, we use the negative of the rank (i.e., -1 for rank 1) in the regression. The second key independent variable is *Deal Seq.*, which is the sequence of deals within the fund, sorted by investment year. One to five refers to the earliest to the latest. E.g., Sequence one means that the deal is invested in the first year in the fund it belongs to. Sequence two means that the deal is invested in the second year in the fund. We group deals with sequences larger than five into one category: sequence five. More details on the variable calculation and the sample restrictions are in Appendix Table 1. Fixed effects for GP, industry and geographic location are indicated in each column. Standard errors are clustered at the GP level. t statistics are reported in parentheses. *, **, and *** denote significance at the 90%, 95%, and 99% confidence levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
			Deal TVPI			
Deal-Size Rank	-0.05***	-0.05***	-0.04***	-0.06***	-0.05***	-0.06***
	(-5.21)	(-4.83)	(-3.92)	(-5.57)	(-4.54)	(-4.82)
Deal Seq.	-0.08***	-0.08***	-0.04	-0.10***	-0.10***	-0.05**
	(-3.09)	(-3.07)	(-1.64)	(-3.94)	(-3.92)	(-2.06)
Absolute Deal Size		-0.42	0.08		-0.46	0.55
		(-1.36)	(0.18)		(-1.20)	(1.11)
Exit Dummy			0.79***			0.82***
			(12.02)			(9.95)
Fund Size			-0.03			-0.10
			(-1.26)			(-1.64)
Deal Duration			-0.01			-0.00
			(-0.65)			(-0.21)
GP FE	No	No	No	Yes	Yes	Yes
Industry FE	No	No	Yes	No	No	Yes
Geography FE	No	No	Yes	No	No	Yes
Observations	5925	5925	5925	5925	5925	5925
Adjusted R^2	0.009	0.009	0.058	0.044	0.044	0.079

Table C12: Deal Level: Within-Fund Deal Size Rank, Timing Sequence and "Jackknife" PME

This table reports the cross-sectional regression results for the relationship between within-fund deal size rank, within-fund deal sequence, and TVPI. The sample period spans from 1999 to 2016 (both included). All are the same as in Table 4 except that the dependent variable is *Deal "Jackknife" PME*. "Jackknife" PME is calculated as the observation deal PME minus the average PME of all deals made in the same year (leaving one out). The first key independent variable is *Deal-Size Rank* which is the negative of the rank of the deal within the fund, sorted by size. 1 to 10 refers to the largest to the smallest. E.g., rank 1 means that the deal is the largest in the fund it belongs to. Rank 2 means that the deal is the second largest in the fund. We group deals with ranks larger than ten into one category: rank 10. To interpret the results more easily, we use the negative of the rank (i.e., -1 for rank 1) in the regression. The second key independent variable is *Deal Seq.*, which is the sequence of deals within the fund, sorted by investment year. One to five refers to the earliest to the latest. E.g., Sequence one means that the deal is invested in the first year in the fund it belongs to. Sequence two means that the deal is invested in the second year in the fund. We group deals with sequences larger than five into one category: sequence five. More details on the variable calculation and the sample restrictions are in Appendix Table 1. Fixed effects for GP, industry and geographic location are indicated in each column. Standard errors are clustered at the GP level. t statistics are reported in parentheses. *, **, and *** denote significance at the 90%, 95%, and 99% confidence levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
			Deal "Jackknife" PME			
Deal-Size Rank	-0.05***	-0.04***	-0.03***	-0.05***	-0.05***	-0.04***
	(-6.24)	(-5.71)	(-4.08)	(-6.81)	(-6.03)	(-4.46)
Deal Seq.	-0.02	-0.02	-0.04**	-0.02	-0.02	-0.04**
	(-1.21)	(-1.21)	(-2.07)	(-1.19)	(-1.19)	(-2.32)
Absolute Deal Size		-0.37*	0.09		-0.13	0.35
		(-1.68)	(0.32)		(-0.53)	(1.05)
Exit Dummy			0.51***			0.53***
			(11.03)			(9.14)
Fund Size			-0.02			-0.04
			(-1.22)			(-1.06)
Deal Duration			-0.10***			-0.10***
			(-13.40)			(-10.52)
GP FE	No	No	No	Yes	Yes	Yes
Industry FE	No	No	Yes	No	No	Yes
Geography FE	No	No	Yes	No	No	Yes
Observations	5925	5925	5925	5925	5925	5925
Adjusted R^2	0.009	0.009	0.078	0.037	0.037	0.092

Table C13: Deal Level: Within-Fund Deal Size Rank, Timing Sequence and "Jackknife" TVPI

This table reports the cross-sectional regression results for the relationship between within-fund deal size rank, within-fund deal sequence, and TVPI. The sample period spans from 1999 to 2016 (both included). All are the same as in Table 4 except that the dependent variable is *Deal "Jackknife" TVPI*. "Jackknife" TVPI is calculated as the observation deal TVPI minus the average TVPI of all deals made in the same year (leaving one out). The first key independent variable is *Deal-Size Rank* which is the negative of the rank of the deal within the fund, sorted by size. 1 to 10 refers to the largest to the smallest. E.g., rank 1 means that the deal is the largest in the fund it belongs to. Rank 2 means that the deal is the second largest in the fund. We group deals with ranks larger than ten into one category: rank 10. To interpret the results more easily, we use the negative of the rank (i.e., -1 for rank 1) in the regression. The second key independent variable is *Deal Seq.*, which is the sequence of deals within the fund, sorted by investment year. One to five refers to the earliest to the latest. E.g., Sequence one means that the deal is invested in the first year in the fund it belongs to. Sequence two means that the deal is invested in the second year in the fund. We group deals with sequences larger than five into one category: sequence five. More details on the variable calculation and the sample restrictions are in Appendix Table 1. Fixed effects for GP, industry and geographic location are indicated in each column. Standard errors are clustered at the GP level. t statistics are reported in parentheses. *, **, and *** denote significance at the 90%, 95%, and 99% confidence levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
			Deal "Jackknife" TVPI			
Deal-Size Rank	-0.05***	-0.04***	-0.04***	-0.05***	-0.05***	-0.05***
	(-4.78)	(-4.43)	(-3.62)	(-5.33)	(-4.76)	(-4.43)
Deal Seq.	-0.03	-0.03	-0.02	-0.03	-0.03	-0.02
	(-1.10)	(-1.10)	(-1.08)	(-1.34)	(-1.34)	(-0.79)
Absolute Deal Size		-0.36	0.19		-0.11	0.58
		(-1.14)	(0.44)		(-0.32)	(1.25)
Exit Dummy			0.47***			0.57***
			(7.50)			(7.01)
Fund Size			-0.03			-0.07
			(-1.42)			(-1.45)
Deal Duration			-0.06***			-0.04***
			(-4.68)			(-2.92)
GP FE	No	No	No	Yes	Yes	Yes
Industry FE	No	No	Yes	No	No	Yes
Geography FE	No	No	Yes	No	No	Yes
Observations	5925	5925	5925	5925	5925	5925
Adjusted R^2	0.005	0.005	0.040	0.031	0.031	0.057

Table C14: Deal Level: Within-Fund Deal Size Rank, Timing Sequence and PME – Non-Linear

This table reports the cross-sectional regression results for the relationship between within-fund deal size rank, within-fund deal sequence, and PME. The sample period spans from 1999 to 2016 (both included). The dependent variable is *Deal PME*, which is the Kaplan-Schoar Public Market Equivalence performance measure at the deal level. The first key independent variable is *Deal-Size Rank* which is the negative of the rank of the deal within the fund, sorted by size. 1 to 10 refers to the largest to the smallest. We group deals with ranks larger than ten into one category: rank 10. To interpret the results more easily, we use the negative of the rank (i.e., -1 for rank 1) in the regression. The second key independent variable is *Deal Seq.*, which is the sequence of deals within the fund, sorted by investment year. One to five refers to the earliest to the latest. We group deals with sequences larger than five into one category: sequence five. More details on the variable calculation and the sample restrictions are in Appendix Table 1. Fixed effects for GP, industry, and geographic location are indicated in each column. Standard errors are clustered at the GP level. t statistics are reported in parentheses. *, **, and *** denote significance at the 90%, 95%, and 99% confidence levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
			Deal PME			
Deal-Size Rank	-0.11***	-0.10***	-0.08**	-0.10***	-0.10***	-0.08***
	(-3.52)	(-3.24)	(-2.52)	(-3.33)	(-3.16)	(-2.60)
Deal-Size Rank^2	-0.01*	-0.00*	-0.00	-0.00	-0.00	-0.00
	(-1.87)	(-1.74)	(-1.42)	(-1.48)	(-1.48)	(-1.24)
Deal Seq.	0.06	0.06	0.04	0.03	0.03	0.01
	(0.80)	(0.83)	(0.56)	(0.34)	(0.34)	(0.11)
Deal Seq.^{2}	-0.02	-0.02	-0.02	-0.01	-0.01	-0.01
	(-1.46)	(-1.49)	(-1.33)	(-0.95)	(-0.95)	(-0.91)
Absolute Deal Size		-0.39*	0.05		-0.23	0.36
		(-1.72)	(0.16)		(-0.86)	(1.08)
Exit Dummy			0.57***			0.57***
			(12.05)			(9.58)
Fund Size			-0.02			-0.06
			(-0.99)			(-1.31)
Deal Duration			-0.09***			-0.09***
			(-11.84)			(-9.55)
GP FE	N	N	N	Y	Y	Y
Industry FE	N	N	Y	N	N	Y
Geography FE	N	N	Y	N	N	Y
Observations	5925	5925	5925	5925	5925	5925
Adjusted R^2	0.011	0.012	0.077	0.040	0.040	0.094

Table C15: Deal Level: Within-Fund Deal Size Rank, Timing Sequence and PME – Interactions with Estimated GP Fixed Effects (GP Skills)

This table reports the results where we interact *Deal-Size Rank* and *Deal Seq.* with estimated GP fixed effects $\widehat{GPFE}$. $\widehat{GPFE}$ is the predicted value for GP fixed effects in Table 4. To interpret the results more easily, we use the negative of the rank (i.e., -1 for rank 1) in the regression. We conduct the cross-sectional regressions about within-fund size rank, sequence, and PME at the deal level. The sample period spans from 1999 to 2016 (both included). More details on the variable calculation and the sample restrictions are in Appendix Table 1. Fixed effects for industry and geographic location are indicated in each column. Standard errors are clustered at the GP level. t statistics are reported in parentheses. *, **, and *** denote significance at the 90%, 95%, and 99% confidence levels, respectively.

	(1)	(2)	(3)
		Deal PME	
Deal-Size Rank	-0.05***	-0.05***	-0.04***
	(-6.86)	(-6.44)	(-5.62)
Deal-Size Rank×$\widehat{GPFE}$	-0.08***	-0.08***	-0.09***
	(3.66)	(3.71)	(4.49)
Deal Seq.	-0.06***	-0.05***	-0.06***
	(-3.67)	(-3.68)	(-3.88)
Deal Seq.× $\widehat{GPFE}$	-0.13**	-0.14**	-0.15***
	(-2.39)	(-2.43)	(-2.64)
$\widehat{GPFE}$	0.88***	0.87***	0.89***
	(5.37)	(5.36)	(5.93)
Absolute Deal Size	-0.36**	0.26	
	(-2.21)	(1.13)	
Exit Dummy		0.63***	
		(14.96)	
Fund Size		-0.03**	
		(-2.54)	
Deal Duration		-0.09***	
		(-11.70)	
Industry FE	No	No	Yes
Geography FE	No	No	Yes
Observations	5925	5925	5925
Adjusted R^2	0.076	0.077	0.144

Table C16: Deal Level: Within-Fund Deal Size Rank, Timing Sequence and PME – Interaction With GP Experience

This table reports the results of the cross-sectional regression for the relationship between within-fund deal size rank, within-fund deal sequence, and PME (Table 4), interacting with *GP Experience*. *GP Experience* is the number of deals done by the GP before the observation deal. The sample period spans from 1999 to 2016 (both included). The dependent variable is *Deal PME*. The key independent variables are *Deal-Size Rank* and *Deal Seq.* The variable construction and the sample restrictions are the same as in Table 4, and more details are provided in Appendix Table 1. Fixed effects for GP, industry and geographic location are indicated in each column. Standard errors are clustered at the GP level. t statistics are reported in parentheses. *, **, and *** denote significance at the 90%, 95%, and 99% confidence levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
			Deal PME			
Deal-Size Rank	-0.05***	-0.05***	-0.04***	-0.06***	-0.06***	-0.04***
	(-5.62)	(-5.30)	(-3.68)	(-5.98)	(-5.54)	(-3.69)
Deal Seq.	-0.04*	-0.05**	-0.06***	-0.05**	-0.05**	-0.07***
	(-1.96)	(-2.02)	(-2.94)	(-2.04)	(-2.06)	(-3.00)
GP Experience	0.01	0.02	-0.01	-0.06*	-0.06*	-0.06*
	(0.30)	(0.60)	(-0.41)	(-1.88)	(-1.78)	(-1.71)
Deal-Size Rank × GP Experience	0.00	0.00	0.00	0.00	0.00	-0.00
	(1.02)	(1.42)	(0.51)	(0.78)	(0.83)	(-0.50)
Deal Seq. × GP Experience	-0.00	0.00	0.01	0.01	0.01	0.01
	(-0.00)	(0.02)	(0.64)	(1.23)	(1.23)	(0.85)
Absolute Deal Size		-0.42*	0.06		-0.09	0.37
		(-1.87)	(0.19)		(-0.32)	(1.07)
Exit Dummy			0.56***			0.53***
			(11.83)			(8.11)
Fund Size			-0.02			-0.04
			(-1.01)			(-0.98)
Deal Duration			-0.09***			-0.10***
			(-11.70)			(-8.57)
GP FE	No	No	Yes	No	No	Yes
Industry FE	No	Yes	Yes	No	Yes	Yes
Geography FE	No	Yes	Yes	No	Yes	Yes
Observations	5925	5925	5925	5925	5925	5925
Adjusted R^2	0.010	0.011	0.077	0.041	0.041	0.094

Table C17: Number of Deals by Fund Sequences

This table summarizes the number of deals by fund sequences. The fund sequence information is from Tim Jenkinson. 5,534 of the 5,925 deals in our sample have the fund sequence information.

Sequence	Freq.	Percent
1	514	9.29%
2	930	16.81%
3	1,169	21.12%
4	1,102	19.91%
5	574	10.37%
6	441	7.97%
7	342	6.18%
8	265	4.79%
9	71	1.28%
10	74	1.34%
11	19	0.34%
12	33	0.60%
Total	5,534	100.00

Table C18: Deal Level: Within-Fund Deal Size Rank, Timing Sequence and PME – Interaction With Fund Sequences

This table reports the results of the cross-sectional regression for the relationship between within-fund deal size rank, within-fund deal sequence, and PME (Table 4), interacting with I(Early Fund Seq.). I(Early Fund Seq.) is a dummy variable for funds with sequences one and two. The fund sequence information is from Tim Jenkinson. 5,534 of the 5,925 deals in our sample have the fund sequence information. The sample period spans from 1999 to 2016 (both included). The dependent variable is *Deal PME*. The key independent variables are *Deal-Size Rank* and *Deal Seq.* The variable construction and the sample restrictions are the same as in Table 4, and more details are provided in Appendix Table 1. Fixed effects for GP, industry and geographic location are indicated in each column. Standard errors are clustered at the GP level. t statistics are reported in parentheses. *, **, and *** denote significance at the 90%, 95%, and 99% confidence levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
			Deal PME			
Deal-Size Rank	-0.05***	-0.04***	-0.03***	-0.05***	-0.05***	-0.04***
	(-4.61)	(-4.16)	(-2.98)	(-5.10)	(-4.38)	(-3.49)
Deal Seq.	-0.05***	-0.05***	-0.06***	-0.06***	-0.06***	-0.07***
	(-2.65)	(-2.64)	(-3.16)	(-2.92)	(-2.92)	(-3.45)
I(Early Fund Seq.)	-0.02	-0.06	-0.05	0.13	0.11	0.05
	(-0.10)	(-0.35)	(-0.34)	(0.60)	(0.52)	(0.21)
Deal-Size Rank × I(Early Fund Seq.)	-0.00	-0.01	0.00	-0.00	-0.00	0.00
	(-0.21)	(-0.36)	(0.02)	(-0.00)	(-0.08)	(0.19)
Deal Seq. × I(Early Fund Seq.)	0.02	0.02	0.03	0.03	0.03	0.03
	(0.55)	(0.56)	(0.89)	(0.71)	(0.71)	(0.64)
Absolute Deal Size		-0.36	0.06		-0.15	0.34
		(-1.59)	(0.19)		(-0.51)	(0.97)
I(Exit)			0.56***			0.56***
			(11.22)			(8.61)
Fund Size			-0.01			-0.05
			(-0.84)			(-1.02)
Deal Duration			-0.09***			-0.10***
			(-11.24)			(-8.59)
GP FE	No	No	No	Yes	Yes	Yes
Industry FE	No	No	Yes	No	No	Yes
Geography FE	No	No	Yes	No	No	Yes
Observations	5534	5534	5534	5534	5534	5534
Adjusted R^2	0.009	0.010	0.075	0.038	0.038	0.092

Table C19: Number of Deals and Number of Funds as Early Successes

This table summarizes the number of deals and the number of funds that have returns (PME/TVPI) larger than a particular cutoff within the first N years. This is what we call "early successes". For example, Performance Measure: PME, Cutoff =2, N. of Years =2, N. of Deals =784, N. of Funds =356 (line 1) means that 784 deals made in the first two years of the fund investment period have a PME ≥ 2 and 356 funds have at least one deal with PME ≥ 2 made in the first two years.

Performance Measure	Cutoff	N. of Years	N. of Deals	N. of Funds
PME	2	2	784	356
PME	2	3	1,153	406
PME	2	4	1,413	420
PME	2	5	1,563	426
PME	3	2	364	222
PME	3	3	523	279
PME	3	4	648	311
PME	3	5	713	321
PME	4	2	187	137
PME	4	3	255	177
PME	4	4	317	199
PME	4	5	352	211
TVPI	2	2	1,166	415
TVPI	2	3	1,724	444
TVPI	2	4	2,118	452
TVPI	2	5	2,357	455
TVPI	3	2	667	322
TVPI	3	3	964	371
TVPI	3	4	1,185	388
TVPI	3	5	1,299	394
TVPI	4	2	353	226
TVPI	4	3	526	282
TVPI	4	4	655	306
TVPI	4	5	726	318
TVPI	5	2	219	153
TVPI	5	3	315	205
TVPI	5	4	389	226
TVPI	5	5	425	240

Table C20: Deal-Level: Early Success and Returns (Robustness Check 1)

This table reports point estimates from OLS regressions of deal-level PMEs on indicators for "early success" funds and later investments. The sample period spans from 1999 to 2016. The Early Success$_{PME=3,t=3}$ dummy variable equals one if the fund made at least one deal with a PME greater than or equal to 3 within the first three years of the fund's life. The dummy variable *After Year 3* equals one if the observation is a deal made after the first three years of a fund's life. Variable definitions are in Table 1. Fixed effects for GP, industry, and geographic location are indicated in each column. Standard errors are clustered at the GP level. t statistics are reported in parentheses. *, **, and *** denote significance at the 90%, 95%, and 99% confidence levels, respectively.

	(1)	(2)	(3)	(4)
Early Success$_{PME=3,t=3}$	0.15**	0.11	-0.01	-0.06
	(2.02)	(1.54)	(-0.11)	(-0.51)
After Year 3	0.26***	0.17***	0.27***	0.17**
	(4.09)	(2.72)	(3.95)	(2.38)
Early Success$_{PME=3,t=3}$ × After Year 3	-0.59***	-0.51***	-0.63***	-0.56***
	(-6.63)	(-5.95)	(-6.73)	(-5.93)
Exit Dummy		0.53***		0.56***
		(11.48)		(9.58)
Absolute Deal Size		-0.31		-0.39
		(-1.05)		(-1.28)
Fund Size		0.00		-0.01
		(0.08)		(-0.20)
Deal Duration		-0.10***		-0.10***
		(-12.39)		(-9.90)
Industry FE	No	Yes	No	Yes
Geography FE	No	Yes	No	Yes
Observations	5925	5925	5925	5925
Adjusted R^2	0.032	0.095	0.041	0.098

Table C21: Deal-Level: Early Success and Returns (Robustness Check 2)

This table reports point estimates from OLS regressions of deal-level PMEs on indicators for "early success" funds and later investments. The sample period spans from 1999 to 2016. The Early Success$_{PME=4,t=2}$ dummy variable equals one if the fund made at least one deal with a PME greater than or equal to 4 within the first two years of the fund's life. The dummy variable *After Year 2* equals one if the observation is a deal made after the first two years of a fund's life. Variable definitions are in Table 1. Fixed effects for GP, industry, and geographic location are indicated in each column. Standard errors are clustered at the GP level. t statistics are reported in parentheses. *, **, and *** denote significance at the 90%, 95%, and 99% confidence levels, respectively.

	(1)	(2)	(3)	(4)
Early Success$_{PME=4,t=2}$	0.09	0.06	-0.01	-0.03
	(1.30)	(1.00)	(-0.08)	(-0.36)
After Year 2	0.25***	0.17***	0.23***	0.16***
	(5.26)	(3.54)	(4.50)	(2.83)
Early Success$_{PME=4,t=2}$ × After Year 2	-0.87***	-0.77***	-0.92***	-0.83***
	(-6.82)	(-6.20)	(-7.37)	(-6.79)
Exit Dummy		0.54***		0.57***
		(11.72)		(10.07)
Absolute Deal Size		-0.29		-0.35
		(-1.07)		(-1.16)
Fund Size		0.00		-0.01
		(0.12)		(-0.13)
Deal Duration		-0.09***		-0.09***
		(-11.79)		(-9.64)
Industry FE	No	Yes	No	Yes
Geography FE	No	Yes	No	Yes
Observations	5925	5925	5925	5925
Adjusted R^2	0.037	0.100	0.053	0.108

Table C22: Deal-Level: Early Success and Returns

This table reports the cross-sectional regressions for the deal-level return for "early-success" funds. The sample period spans from 1999 to 2016 (both included). Panel A reports results that compare deals in "early-success" funds and deals in "non-early-success" funds. The key independent variable is Early Success$_{PME=3,t=2}$, a dummy variable that equals one if the fund where the deal sits made at least one deal with $PME \geq 3$ within the first two years. Columns (1) to (4) report the results for the subsample of deals made in the "early period" (first two years here). Columns (5) to (8) report the results for the subsample of deals made after the "early period". Panel B reports results comparing deals made during the "early period" and afterward. The key independent variable is *After Year 2*, a dummy variable that equals one if the deal is made after the first two years. Columns (1) to (4) report the results for the subsample of deals made by "early-success" funds. Columns (5) to (8) report the results for the subsample of deals made by "non-early-success" funds. In both panels, in columns (1), (2), (5), and (6), the dependent variable is deal PME, and in columns (3), (4), (7), and (8), the dependent variable is deal TVPI. More details on the variable calculation and the sample restrictions are in Appendix Table 1. Fixed effects for GP, industry, and geographic location are indicated in each column. Standard errors are clustered at the GP level. t statistics are reported in parentheses. *, **, and *** denote significance at the 90%, 95%, and 99% confidence levels, respectively.

Panel A: Early Success Funds v.s. Non-Early Success Funds								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Early Deals				Follow-on Deals			
	Deal PME		Deal TVPI		Deal PME		Deal TVPI	
Early Success$_{PME=3,t=2}$	0.82***	0.86***	0.91***	0.96***	0.05	-0.04	0.01	-0.12
	(11.71)	(7.71)	(9.29)	(5.65)	(0.87)	(-0.31)	(0.17)	(-0.76)
Exit Dummy	0.35***	0.43***	0.45***	0.61***	0.65***	0.66***	0.93***	0.94***
	(4.90)	(4.72)	(4.41)	(4.64)	(11.48)	(9.29)	(11.42)	(9.72)
Deal Size	0.08	-0.12	0.17	-0.12	-0.67**	-0.66**	-0.96**	-0.85**
	(0.18)	(-0.30)	(0.26)	(-0.20)	(-2.11)	(-1.99)	(-2.25)	(-2.02)
Fund Size	-0.01	-0.04	-0.01	-0.06	0.01	-0.02	0.01	-0.05
	(-0.28)	(-0.91)	(-0.35)	(-1.11)	(0.66)	(-0.26)	(0.37)	(-0.75)
Deal Duration	-0.11***	-0.12***	-0.06***	-0.06**	-0.09***	-0.09***	0.02	0.03
	(-9.47)	(-6.88)	(-3.47)	(-2.58)	(-7.86)	(-6.18)	(0.94)	(1.06)
GP FE	No	Yes	No	Yes	No	Yes	No	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geography FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2481	2481	2481	2481	3444	3444	3444	3444
Adjusted R^2	0.128	0.144	0.083	0.118	0.072	0.086	0.069	0.081
Panel B: Early Deals v.s. Follow-on Deals								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Early Success Funds				Non-Early Success Funds			
	Deal PME		Deal TVPI		Deal PME		Deal TVPI	
After Year 2	-0.42***	-0.48***	-0.42***	-0.53***	0.30***	0.25***	0.41***	0.33***
	(-4.76)	(-5.33)	(-3.62)	(-4.78)	(5.15)	(4.05)	(4.90)	(3.62)
Exit Dummy	0.63***	0.71***	0.82***	0.97***	0.42***	0.40***	0.69***	0.66***
	(7.87)	(6.46)	(7.48)	(6.63)	(9.24)	(6.69)	(10.23)	(7.45)
Absolute Deal Size	-0.83	-1.06*	-0.94	-1.46*	-0.12	-0.06	-0.31	-0.12
	(-1.43)	(-1.90)	(-1.12)	(-1.91)	(-0.48)	(-0.25)	(-0.86)	(-0.37)
Fund Size	0.03	-0.02	0.03	-0.06	-0.01	-0.01	-0.01	-0.01
	(1.00)	(-0.51)	(0.59)	(-1.36)	(-0.38)	(-0.15)	(-0.43)	(-0.17)
Deal Duration	-0.10***	-0.10***	-0.02	-0.00	-0.08***	-0.09***	0.00	-0.00
	(-8.51)	(-6.73)	(-0.86)	(-0.05)	(-8.78)	(-6.57)	(0.36)	(-0.09)
GP FE	No	Yes	No	Yes	No	Yes	No	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geography FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3075	3075	3075	3075	2850	2850	2850	2850
Adjusted R^2	0.078	0.095	0.053	0.081	0.092	0.113	0.076	0.097

Table C23: Deal Level: Early Success and Industry/Geographic Specialization

This table reports the cross-sectional regressions on whether follow-on deals are more likely to be the same as in the industries/regions of early successes. The sample period spans from 1999 to 2016 (both included). We restrict to the subsample of deals made after the "early period" (first two years here), i.e., follow-on deals. The key independent variable is Early Success$_{PME=3,t=2}$, a dummy variable that equals one if the fund where the deal sits made at least one deal with $PME \geq 3$ within the first two years. In columns (1) to (3), the dependent variable is I(Sector) which is a dummy variable that equals to one if the deal is in the same sector as one of the early deals (first two years here) by the same fund. In columns (1) to (3), the dependent variable is I(Region) which is a dummy variable that equals to one if the deal is in the same region as one of the early deals (first two years here) by the same fund. More details on the variable calculation and the sample restrictions are in Appendix Table 1. Fixed effects for GP, industry, and geographic location are indicated in each column. Standard errors are clustered at the GP level. t statistics are reported in parentheses. *, **, and *** denote significance at the 90%, 95%, and 99% confidence levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
		I(Sector)			I(Geographic)	
Early Success$_{PME=3,t=2}$	0.13***	0.12***	0.06	0.03**	0.02*	-0.02
	(4.96)	(5.12)	(1.57)	(2.09)	(1.73)	(-1.58)
Exit Dummy		0.00	-0.01		-0.01	0.01
		(0.13)	(-0.56)		(-1.60)	(0.69)
Absolute Deal Size		-0.06	0.07		0.10	0.12**
		(-0.49)	(0.63)		(1.54)	(2.02)
Fund Size		0.00	-0.02*		-0.01***	-0.00
		(0.18)	(-1.70)		(-3.71)	(-0.11)
Deal Duration		-0.00	0.00		-0.00*	0.00
		(-0.58)	(0.07)		(-1.81)	(0.15)
GP FE	No	No	Yes	No	No	Yes
Industry FE	No	Yes	Yes	No	Yes	Yes
Geography FE	No	Yes	Yes	No	Yes	Yes
Observations	3444	3444	3444	3444	3444	3444
Adjusted R^2	0.017	0.079	0.210	0.004	0.167	0.312

Table C24: Deal Level: Industry/Geographic Specialization and Return

This table reports the cross-sectional regressions on the relationship between industry/geographic specialization and return. The sample period spans from 1999 to 2016 (both included). Panel A examines whether deals in a fund's specialized industry/region have higher returns. The dependent variable is *Deal PME*, which is the Kaplan-Schoar Public Market Equivalence performance measure. In columns (1) to (3), the key independent variable is Sector%, the percentage of deals in a fund's specialized industry. In columns (4) to (6), the key independent variable is Region%, the percentage of deals in a fund's specialized region. Panel B examines whether investments outside a fund's specialized industry/region have lower returns. The dependent variable is *Deal PME*, which is the Kaplan-Schoar Public Market Equivalence performance measure. In columns (1) to (3), the key independent variable is I(Non-Specialized Sector), a dummy variable equal to one if the deal is outside a fund's specialized industry (i.e., the industry with less than 50% of deals). In columns (4) to (6), the key independent variable is I(Non-Specialized Region), a dummy variable equal to one if the deal is outside a fund's specialized region (i.e., the region with less than 50% of deals). More details on the variable calculation and the sample restrictions are in Appendix Table 1. Fixed effects for GP, industry, and geographic location are indicated in each column. Standard errors are clustered at the GP level. t statistics are reported in parentheses. *, **, and *** denote significance at the 90%, 95%, and 99% confidence levels, respectively.

Panel A: Specialization Weights and PME						
	(1)	(2)	(3)	(4)	(5)	(6)
			Deal PME			
Sector%	0.10	0.06	0.06			
	(0.94)	(0.59)	(0.46)			
Region%				0.21	0.12	0.22
				(1.62)	(1.00)	(1.47)
Exit Dummy		0.61***	0.63***		0.61***	0.62***
		(12.57)	(10.68)		(12.62)	(10.67)
Absolute Deal Size		-0.44	-0.42		-0.44	-0.44
		(-1.54)	(-1.40)		(-1.55)	(-1.46)
Fund Size		0.01	-0.01		0.01	-0.01
		(0.43)	(-0.30)		(0.65)	(-0.27)
Deal Duration		-0.09***	-0.09***		-0.09***	-0.09***
		(-11.79)	(-9.69)		(-11.74)	(-9.66)
GP FE	No	No	Yes	No	No	Yes
Industry FE	No	Yes	Yes	No	Yes	Yes
Geography FE	No	Yes	Yes	No	Yes	Yes
Observations	5925	5925	5925	5925	5925	5925
Adjusted R^2	0.000	0.072	0.087	0.001	0.072	0.087

Panel B: Out of Specialization and PME						
	(1)	(2)	(3)	(4)	(5)	(6)
			Deal PME			
I(Non-Specialized Sector)	-0.02	-0.01	-0.02			
	(-0.36)	(-0.15)	(-0.34)			
I(Non-Specialized Region)				-0.06	-0.05	-0.08
				(-0.58)	(-0.49)	(-0.73)
Exit Dummy		0.61***	0.63***		0.61***	0.63***
		(12.51)	(10.69)		(12.59)	(10.67)
Absolute Deal Size		-0.44	-0.42		-0.44	-0.43
		(-1.54)	(-1.40)		(-1.54)	(-1.42)
Fund Size		0.01	-0.01		0.01	-0.01
		(0.40)	(-0.31)		(0.46)	(-0.31)
Deal Duration		-0.09***	-0.09***		-0.09***	-0.09***
		(-11.82)	(-9.69)		(-11.83)	(-9.68)
GP FE	No	No	Yes	No	No	Yes
Industry FE	No	Yes	Yes	No	Yes	Yes
Geography FE	No	Yes	Yes	No	Yes	Yes
Observations	5925	5925	5925	5925	5925	5925
Adjusted R^2	-0.000	0.072	0.087	-0.000	0.072	0.087

Table C25: Fund Level: Return Dispersion – Real Sample v.s. Random Samples

This table reports the return dispersion for the real sample and random samples across vintage years. Columns (1) to (4) report the difference between the 90th and 10th percentiles of returns, columns (5) to (8) report the difference between the 75th and 25th percentiles of returns, and columns (9) to (12) report the standard deviation of returns. "Actual data" is our sample data. "Random Sample 1" is generated by drawing PMEs randomly (using uniform distribution) from the data without replacement, keeping the number of deals done by a fund the same as found in the data. "Random Sample 2" is the same, except that PMEs are drawn randomly with replacement. "Random Sample 3" is generated by drawing PMEs randomly from the data with replacements but allowing the number of investments per fund also to vary randomly.

Panel A: PME												
Vintage	90th − 10th Percentile				75th − 25th Percentile				Std. Dev.			
	Actual	R1	R2	R3	Sample	R1	R2	R3	Actual	R1	R2	R3
1999	0.46	1.12	1.08	1.50	0.29	0.54	0.52	0.73	0.18	0.42	0.40	0.67
2000	1.41	1.25	1.42	1.50	0.56	0.56	0.63	0.74	0.47	0.42	0.48	0.68
2001	1.27	1.04	1.07	1.50	0.84	0.52	0.54	0.74	0.55	0.46	0.50	0.68
2002	1.74	1.34	1.38	1.50	0.64	0.74	0.72	0.75	0.56	0.48	0.48	0.67
2003	1.03	1.26	1.22	1.53	0.66	0.55	0.51	0.76	0.39	0.42	0.41	0.69
2004	1.49	1.17	1.27	1.50	0.43	0.61	0.62	0.75	0.68	0.53	0.56	0.67
2005	1.38	1.24	1.24	1.52	0.74	0.70	0.69	0.74	0.54	0.53	0.54	0.67
2006	1.56	1.42	1.45	1.51	0.57	0.63	0.62	0.75	0.76	0.56	0.56	0.69
2007	1.34	1.31	1.43	1.51	0.87	0.61	0.65	0.75	0.65	0.52	0.55	0.68
2008	1.59	1.25	1.28	1.51	0.72	0.63	0.64	0.75	0.64	0.55	0.56	0.68
2009	1.39	1.28	1.31	1.49	0.73	0.57	0.58	0.74	0.48	0.48	0.51	0.68
2010	1.21	1.23	1.29	1.49	0.33	0.65	0.61	0.74	0.59	0.54	0.58	0.67
2011	1.32	1.31	1.33	1.51	0.70	0.66	0.64	0.75	0.52	0.51	0.54	0.69
2012	1.35	1.35	1.35	1.50	0.81	0.64	0.64	0.75	0.55	0.54	0.55	0.67
2013	1.43	1.40	1.39	1.53	0.53	0.67	0.66	0.76	0.58	0.55	0.59	0.69
2014	1.09	1.26	1.31	1.52	0.67	0.65	0.67	0.76	0.59	0.56	0.59	0.69
2015	0.75	1.26	1.29	1.52	0.39	0.64	0.65	0.75	0.32	0.53	0.56	0.69
2016	0.81	1.35	1.39	1.53	0.31	0.66	0.67	0.76	0.30	0.55	0.58	0.69
Panel B: TVPI												
Vintage	90th − 10th Percentile				75th − 25th Percentile				Std. Dev.			
	Sample	R1	R2	R3	Sample	R1	R2	R3	Sample	R1	R2	R3
1999	1.82	1.49	1.58	2.07	1.01	0.71	0.76	1.01	0.71	0.56	0.59	0.91
2000	1.78	1.75	1.78	2.08	1.03	0.79	0.78	1.02	0.64	0.59	0.60	0.90
2001	1.84	1.57	1.49	2.04	1.49	0.76	0.72	1.00	0.79	0.68	0.67	0.89
2002	2.69	1.68	1.89	2.02	1.12	0.91	0.94	1.00	0.86	0.59	0.66	0.88
2003	1.67	1.74	1.70	2.02	0.50	0.69	0.72	1.00	0.51	0.58	0.57	0.89
2004	2.25	1.60	1.60	2.05	0.95	0.79	0.82	1.01	0.97	0.69	0.73	0.90
2005	1.55	1.71	1.59	2.05	0.96	0.91	0.88	1.01	0.66	0.71	0.71	0.91
2006	2.25	1.87	1.89	2.07	0.95	0.85	0.85	1.02	0.88	0.72	0.74	0.92
2007	1.48	1.74	1.84	2.07	0.63	0.83	0.86	1.01	0.64	0.69	0.74	0.92
2008	2.19	1.71	1.69	2.05	1.35	0.84	0.84	1.01	0.90	0.75	0.75	0.91
2009	1.76	1.72	1.80	2.04	0.89	0.77	0.82	1.00	0.69	0.65	0.71	0.91
2010	1.75	1.65	1.64	2.06	0.59	0.86	0.88	1.01	0.77	0.72	0.75	0.89
2011	2.11	1.72	1.75	2.07	0.98	0.86	0.89	1.02	0.82	0.67	0.72	0.92
2012	2.32	1.75	1.84	2.05	0.96	0.83	0.85	1.02	0.91	0.71	0.74	0.91
2013	1.72	1.89	1.85	2.09	0.90	0.92	0.88	1.03	0.78	0.74	0.76	0.93
2014	1.45	1.73	1.77	2.06	0.65	0.91	0.89	1.02	0.75	0.73	0.77	0.92
2015	1.21	1.71	1.70	2.08	0.48	0.85	0.92	1.02	0.46	0.70	0.73	0.92
2016	1.12	1.85	1.89	2.06	0.52	0.92	0.92	1.02	0.39	0.76	0.78	0.92

Internet Appendix D: Hierarchical Linear Model

D.1 GP-Fund-Deal Model a la Cavagnaro, Sensoy, Wang, et al. (2019)

In this section, we present a three-layer GP-Fund-Deal hierarchical model a la Cavagnaro, Sensoy, Wang, et al. (2019). Using the GP-Fund-Deal model, we can quantify the variation across GPs in their intrinsic value in terms of portfolio management. In the model, idiosyncratic risks are captured by deal-specific random effects. Because we are interested in comparing the GP-specific effects with and without idiosyncratic risks, we need to estimate the GP-Fund layer using the estimation of deal-specific random effects from the Fund-Deal layer. Therefore, we estimate the model in two stages. That is to say, in order to distinguish between GP skills in deal selection and portfolio management, we need to estimate GP skills in two different models, as presented below.

First Stage (Fund-Deal Level) In the first stage, we estimate a hierarchical linear model at the fund-deal level using the deal return as the left-hand side variable.

$$DealReturn_{iuj} = X'\beta + \Sigma_{\tau=t_{iuj}}^{t_{iuj}+DealLife_{iuj}}(FundRE_u + FundYearRE_{u\tau}) + \epsilon_{uj} \tag{D.1}$$

where i represents the GP firms, u represents funds, j represents deals, X are the observed covariates (in the estimation, we control for deal entry and exit year fixed effects), t_{iuj} denotes the deal's first year (entry year), $DealLife_{iuj}$ denotes the deal life which is the number of years from the deal entry to the deal exit, $FundRE_u$ is the fund-specific random effect, $FundYearRE_{u\tau}$ is the fund-time random effect, and ϵ_{uj} is the fund-level deal-specific random effect. Deal return is measured using TVPI.

Second Stage (GP-fund level) In the second stage, we estimate a hierarchical linear model at the GP-fund level using the *adjusted* deal return as the left-hand side variable.

$$\widehat{DealReturn}_{iuj} = \Sigma_{\tau=t_{iuj}}^{t_{iuj}+DealLife_{iuj}}(GPRE_i + GPYearRE_{i\tau}) + \nu_{ij} \tag{D.2}$$

where i represents the GP firms, u represents funds, j stands for deals, t_{iuj} denotes the deal's first year (entry year), $DealLife_{iuj}$ denotes the deal life which is the number of years from the deal entry to the deal exit, $GPRE_i$ is the GP-specific random effect, $GPYearRE_{i\tau}$ is the GP-time random effect, and ν_{ij} is the GP-level deal-specific random effects. Deal return is measured using TVPI.

We examine two models that differ in how the *adjusted* deal return is calculated. The estimated GP-specific effect ($GPRE_i$) from Model 1 tells how skillful GPs are at achieving high-return deals. The estimated GP-specific effect ($GPRE_i$) from Model 2 tells how skillful GPs are at managing funds.

In Model 1, the adjusted return for each deal is after subtracting estimated fund-specific effect and fund-year-specific effect from the *actual* deal returns (Equation D.3). The variation left in the *adjusted* deal returns reflects GP's skills to achieve high return deals (the deal-specific effect ϵ_{uj}).

$$\widehat{DealReturn}_{iuj} = DealReturn_{iuj} - X'\hat{\beta} - \Sigma_{\tau=t_{iuj}}^{t_{iuj}+DealLife_{iuj}}(\widehat{FundRE}_u + \widehat{FundYearRE}_{u\tau}) \tag{D.3}$$

Where $\widehat{FundRE}_{iu\tau}$ is the estimated fund random effect, $\widehat{FundYearRE}_{u\tau}$) is the estimated fund-year random effect.

In Model 2, when computing the *adjusted* return for each deal, we subtract the estimated fund-year-specific effect and the fund-level deal-specific effect from the *actual* deal return (Equation D.3′). The variation left in the adjusted deal returns reflects the GP's value-added in fund management (the fund-specific effect $FundRE_u$).

$$\widehat{DealReturn}_{iuj} = DealReturn_{iuj} - X'\hat{\beta} - \Sigma_{\tau=t_{iuj}}^{t_{iuj}+DealLife_{iuj}}\widehat{FundYearRE}_{u\tau} - \widehat{\epsilon}_{uj} \tag{D.3′}$$

Where $\widehat{FundYearRE}_{iu\tau}$ is the estimated fund-time random effect, and $\widehat{\epsilon}_{uj}$ is the estimated fund-level deal-specific random effect.

The assumptions on the parameter distribution are the same in the GP-Fund level model (KS model) and discussed in the Internet Appendix. The variance decomposition in the first

and second stages is similar.

D.2 Bayesian Estimator

Following KS and CSWW, we estimate our model using a Bayesian estimator, techniques of Markov Chain Monte Carlo (MCMC), and Gibbs sampling.[1] While the model could also be estimated by classical maximum likelihood estimation (MLE), Bayesian estimation is better for estimating parameters with non-negativity constraints (e.g., variance), small sample inferences (the mean number of funds managed by GPs is 1.74, and the mean number of deals in a fund is 11.78 in our sample), and allowing for non-normal distributions, which suit the skewed return pattern in PE.

Each MCMC cycle g in the algorithm consists of two steps. Following the approach outlined in sections A1 to A5 of the appendix in Korteweg and Sorensen (2017), the first step is to obtain a random draw of each parameter in the KS model. We employ the priors and initial values described in section A7 of the Korteweg and Sorensen (2017) appendix. The priors are sufficiently diffuse for the results to be determined by the data instead of by the prior assumptions. In KS model, the random effects ϵ_{uj} are redefined so that their mean is the fund effect $FundRE_u$. We instead leave them as mean zero to ease interpretation of the second step of our estimation.

At the end of the first stage, we adjust the total return of each fund according to equation Equation D.3 or Equation D.3′ in order to account for the fund-time random effects (or plus deal-specific random effects). Then, based on the current cycle's adjusted returns, we obtain a draw of each parameter in Equation Equation D.2 and their variances. As in the first stage, the priors in the second step are also diffuse so that the data, rather than prior assumptions, drive the outcomes.

D.3 Results

[INSERT Table D1 AROUND HERE]

[1] For more details on the estimation procedure and why the Bayesian estimator fits the PE setting, see Korteweg and Sorensen (2017) and Korteweg (2011).

Table D1 shows the results estimated from the first stage of the model (Equation D.1), which decomposes the variance into fund-specific random effect, fund-year-specific random effect, and deal-specific random effect (conditional on the fund). The estimates are close to those in the GP-deal model, as reported in Internet Appendix Table B9. The first-stage analysis provides several findings. First, we find that noise accounts for 80% to 90% (depending on specifications) of the deal-level total return variation. This implies that idiosyncratic risk at the deal level accounts for a large part of return variation. Second, overlapping effects account for 16.14% without controlling for deal year fixed effects and account for 4.62% after controlling for deal entry and exit year fixed effects. The difference between the overlapping effects, which are random effects, and the deal year fixed effects is that the former captures the time effects only common to all deals in the same fund and the latter captures the time effects common to all deals. Therefore, our findings suggest that 71.38% ($\frac{16.14-4.62}{16.14}$) of the overlapping effects at the deal level come from factors affecting all deals such as common market risk exposure, and the rest 28.62% come from fund-level common strategies that impose a time-variant effect to deal return. Finally, fund-specific effects account for 4.52% of the deal returns after controlling for deal entry and exit year fixed effects.

[INSERT Table D2 AROUND HERE]

Table D2 reports the results of second-stage estimates. Columns (1) to (4) report the results of Model 1 where we subtract fund-specific fund-year-specific effects when adjusting the deal return entering the second-stage estimation (Equation D.3). The results tell us how GP skill affects deal returns if we take into account both fund-specific effects (i.e., how good GPs are at portfolio management) and deal-specific effects (i.e., how good GPs are at selecting deals conditional on funds). We find that GP-skill only accounts for 6.43% of the total variance of returns. Noise accounts for more than 87.06% of the return variation. This implies that GP-specific effects do not play an important role in achieving high-return deals. Put differently, the variance in how much one GP can systematically differ from another GP in the deal idiosyncratic component (luck) is small.

Columns (5) to (8) report the results of Model 2 where we subtract both fund-year-specific and deal-specific effects when adjusting the deal returns entering the second-stage estimation (Equation D.3′). This case measures how GP skills in the fund-specific component (i.e., GP's portfolio management skills) affect returns. We find that skill accounts for a much larger part of the return variation, 45.69%, in this case. This implies that GPs differ to a large degree in their portfolio management skills. The noise component only accounts for 8.53% of the return variation. The difference in the results from Model 1 and Model 2 sheds light on which dimension lies GP skills. Our findings suggest that GPs mainly differ in their portfolio management skills. Specific deal-picking skills are limited given the high riskiness of the PE industry.

To clarify, we mean deal-picking in a broad sense that includes both ex-ante selection and ex-post monitoring. Also, skills that can lead to an overall high return for all deals in the fund are included in the fund-specific component (and thus in portfolio management). However, our findings suggest that GPs do not differ much in finding particular high-return deals. This is consistent with our findings in the regression analysis that GPs do not chase extreme alpha in the scarification of bringing additional riskiness to the whole fund.

There is also a difference in the overlapping effect between Models 1 and 2. In Model 1, GP-year random effects are around half of the GP random effects while the GP-year random effects are close to GP random effects in Model 2. Examples of the overlapping effect at the GP level are the GP common strategies (e.g., management-team-year specific effect) and risk exposures to the market conditions that are common for funds managed by the same GP. The contrast in the results in the GP-year effects between Models 1 and 2 implies that GP styles are around twice as important as GP-year strategies for return with idiosyncratic risks, but as important as GP-year strategies for return deducting idiosyncratic risks.

Table D1: Bayesian Estimation of GP-Fund-Deal Three-Layer Hierarchical Linear Model (First Stage)

This table reports posterior means of parameters of Bayesian models described in section 6. The model specification is the GP-fund-deal three-layer model and we report the first stage results in this table. Panel A shows the parameter estimates, and Panel B reports the variance decomposition estimates for the same parameters. Specifications where we also include deal entry year and deal exit year fixed effects as covariates are indicated in each column. The model is estimated by Markov chain Monte Carlo (MCMC) using ten thousand burn-in cycles followed by 100,000 samples, saving every tenth draw. signal to noise (skill)%, overlap effect%, and error% are defined in Equation 6. The fund life N is the deal life which is calculated as deal exit year-deal entry year+1. Return is measured in TVPI. More details on the variable calculations and the sample restrictions are in Appendix Table 1. Posterior standard deviations (Bayesian standard errors) are in brackets.

	(1)	(2)	(3)	(4)
Panel A: Parameter estimates				
$\sigma(FundRE)$	0.033	0.037	0.041	0.035
	(0.003)	(0.004)	(0.004)	(0.003)
$\sigma(FundYearRE)$	0.189	0.161	0.151	0.094
	(0.016)	(0.017)	(0.016)	(0.016)
$\sigma(\epsilon)$	1.017	1.026	1.019	1.020
	(0.013)	(0.013)	(0.013)	(0.012)
Deal Entry Year FE	N	Y	N	Y
Deal Exit Year FE	N	N	Y	Y
N. of Funds	467	467	467	467
N. of Deals	5925	5925	5925	5925
Panel B: Variance decomposition				
$N^2\sigma^2(FundRE)$	0.046	0.056	0.069	0.052
	(0.009)	(0.012)	(0.015)	(0.010)
$N\sigma^2(FundYearRE)$	0.209	0.152	0.134	0.053
	(0.036)	(0.032)	(0.029)	(0.018)
$\sigma^2(\epsilon)$	1.034	1.052	1.038	1.041
	(0.027)	(0.027)	(0.026)	(0.025)
$\sigma^2(DealReturn)$	1.289	1.260	1.242	1.146
	(0.037)	(0.034)	(0.031)	(0.027)
signal to noise (skill%)	3.55%	4.44%	5.59%	4.52%
	(0.007)	(0.009)	(0.012)	(0.009)
overlap effect%	16.14%	12.05%	10.80%	4.62%
	(0.025)	(0.024)	(0.022)	(0.015)
noise%	80.30%	83.51%	83.62%	90.87%
	(0.024)	(0.022)	(0.019)	(0.014)

Table D2: Bayesian Estimation of GP-Fund-Deal Three-Layer Hierarchical Linear Model (Second Stage)

This table reports posterior means of parameters of Bayesian models described in section 6. for the second stage of the GP-fund-deal three-layer. Columns (1) to (4) report the result of Model 1 where the adjusted returns are computed by subtracting fund random effects fund-year random effects. Columns (5) to (8) report the result of Model 2 where the adjusted returns are computed by subtracting fund-year random effects and deal-specific random effects. Panel A shows the parameter estimates, and Panel B reports the variance decomposition estimates for the same parameters. The model is estimated by Markov chain Monte Carlo (MCMC) using ten thousand burn-in cycles followed by 100,000 samples, saving every tenth draw. signal to noise (skill%), overlap effect%, and error% are defined in Equation 6. The fund life N is the deal life which is calculated as deal exit year-deal entry year+1. Return is measured in TVPI. More details on the variable calculations and the sample restrictions are in Appendix Table 1. Posterior standard deviations (Bayesian standard errors) are in brackets.

	Model 1				Model 2			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Parameter estimates								
$\sigma(GPRE)$	0.026	0.027	0.027	0.031	0.021	0.028	0.028	0.032
	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.004)	(0.004)	(0.006)
$\sigma(GPYearRE)$ 0.054	0.073	0.078	0.114	0.077	0.105	0.113	0.117	
	(0.008)	(0.012)	(0.013)	(0.015)	(0.008)	(0.015)	(0.015)	(0.022)
$\sigma(\epsilon)$	1.017	1.024	1.013	1.014	0.070	0.102	0.108	0.123
	(0.012)	(0.012)	(0.012)	(0.012)	(0.006)	(0.025)	(0.024)	(0.045)
N. of GPs	315	315	315	315	315	315	315	315
N. of Deals	5925	5925	5925	5925	5925	5925	5925	5925
Panel B: Variance decomposition								
$N^2\sigma^2(GPRE)$	0.052	0.057	0.058	0.076	0.034	0.062	0.065	0.083
	(0.008)	(0.010)	(0.010)	(0.015)	(0.005)	(0.020)	(0.019)	(0.037)
$N\sigma^2(GPYearRE)$	0.017	0.031	0.036	0.077	0.035	0.065	0.075	0.082
	(0.005)	(0.010)	(0.012)	(0.020)	(0.007)	(0.019)	(0.020)	(0.033)
$\sigma^2(\epsilon)$	1.034	1.049	1.027	1.028	0.005	0.011	0.012	0.017
	(0.025)	(0.025)	(0.024)	(0.025)	(0.001)	(0.007)	(0.006)	(0.015)
$\sigma^2(DealReturn)$	1.103	1.138	1.121	1.181	0.074	0.138	0.152	0.182
	(0.027)	(0.030)	(0.029)	(0.039)	(0.011)	(0.042)	(0.041)	(0.080)
signal to noise (skill%)	4.68%	5.03%	5.15%	6.43%	46.39%	44.87%	42.58%	45.69%
	(0.007)	(0.008)	(0.008)	(0.012)	(0.039)	(0.044)	(0.045)	(0.044)
overlap effect%	1.58%	2.74%	3.21%	6.51%	46.92%	47.40%	49.58%	45.79%
	(0.005)	(0.009)	(0.010)	(0.016)	(0.037)	(0.043)	(0.042)	(0.045)
noise%	93.74%	92.22%	91.64%	87.06%	6.69%	7.73%	7.85%	8.53%
	(0.009)	(0.013)	(0.014)	(0.021)	(0.006)	(0.020)	(0.018)	(0.033)