Synthetic Peer Benchmarking for Diversified Private Equity Programs

JEROEN CORNEL

JEROEN CORNEL
is a vice president at BlackRock Private Equity Partners in Princeton, NJ.
jeroen.cornel@blackrock.com

Private equity has become a standard component in the investment portfolios of many pension plans, insurance companies, and endowments (Kaplan and Antoinette [2005]; Diller and Kaserer [2009]; Phalippou and Gottschalg [2009]; Gottschalg, Talmor, and Vasvari [2010]; Malinowski and Wittlin [2014]). In the current low yield environment, many institutional investors have further increased their allocation to private equity as an asset class (Harris, Jenkinson, and Kaplan [2014]; Gompers, Kaplan, and Mukharlyamov [2015]). It is well known that diversifying capital through several fund managers over different vintages, strategies, and regions decreases idiosyncratic or unsystematic risk (Weidig and Mathonet [2004]; Weidig, Kemmerer, and Born [2005]; Mathonet and Meyer [2007]) and optimizes risk-adjusted returns of private equity investments (Rouvinez and Kubr [2003]; Gresch and von Wyss [2011]). However, in implementing a diversified program of private equity fund investments, a number of thorny measurement issues arise, not least of which is how to do appropriate benchmarking.

One popular method of benchmarking private equity is the public market equivalent method (PME) that estimates the performance of a fictive investment in the public equity market (e.g., using the MSCI World index) while maintaining the timing and magnitude of the private equity investment (Rouvinez [2003]; Frei and Studer [2004]; Ellis, Pattni, and Tailor [2012]; Malinowski and Wittlin [2013]; Sorensen and Jagannathan [2015]). The PME method is well known, broadly applied, and valid for direct private equity investments, fund investments, and diversified programs. On the other hand, the investor might ask how similar diversified programs offered by other private equity solution providers have performed during the same time window. In order to answer this question adequately, a large universe of diversified or comingled programs would need to be available, which is currently not the case (Mathonet and Meyer [2007]; Kaserer and Diller [2011]). To address this challenge, we have developed a simulation technique that replicates a diversified private equity program based on a large universe of underlying funds and that provides a representative benchmark for a specific diversified private equity program in two key performance dimensions, i.e., internal rate of return (IRR) and multiple of invested capital. This method is a representative benchmarking method because we reconstruct underlying cash flows, subsequently aggregate these to a program level, and only then estimate the IRR. In doing so and in contrast to existing benchmarking methods, we properly address a potential mathematical pitfall, as explained in detail in the following section.

Furthermore, this technique provides probabilistic rather than deterministic results and also unravels the overall performance into a market factor, a vintage, strategy, and geographic adjustment, and an unexplained factor which, among others, may be interpreted as manager selection capabilities. The concept of unraveling or attributing performance to common factors is an often-applied framework introduced and published by coworkers at BlackRock in the case of the public market equivalent (Malinowski and Wittlin [2013]). This technique can be interpreted as a complementary technique, but for benchmarking performance against the private market and attributing this performance to private market factors.

This article presents this new method in detail and is structured as follows: the first section discusses asset diversification and explains how, depending on the manner in which underlying figures are aggregated, the performance of a diversified private equity program could be misleading. Next, we present the status quo of peer benchmarking methods for diversified private equity programs and subsequently describe in detail the proposed new method. We then discuss in depth the results of our method when applied to two examples of diversified private equity programs, followed by concluding remarks.

ASSET DIVERSIFICATION AND MODELING CAUTIONS

Often the proverb "do not put all your eggs in one basket" is used as an argument for asset diversification that can decrease to a large extent the idiosyncratic or unsystematic risk of a portfolio of investments (Weidig, Kemmerer, and Born [2005]; Mathonet and Meyer [2007]). Whereas diversification is a widely used tool to manage the risk–return characteristics of the investor also in private equity, it is often not fully understood in depth. This section elaborates further on some mathematical aspects of fund diversification that will be important for a meaningful peer benchmarking method for diversified private equity programs.

Let us consider the simplified example of a diversified program that invests only in North American private equity funds with one particular vintage year. Exhibit 1 shows the probability distributions of final multiple (total value to paid-in, or TVPI[1]) of an investment in a single fund (light gray bars) and in a diversified program consisting of 20 funds randomly selected from the same universe (dark gray bars). The corresponding cumulative distribution functions are shown as lines in the same figure and the corresponding quartiles and mean values are outlined in the caption of the figure.

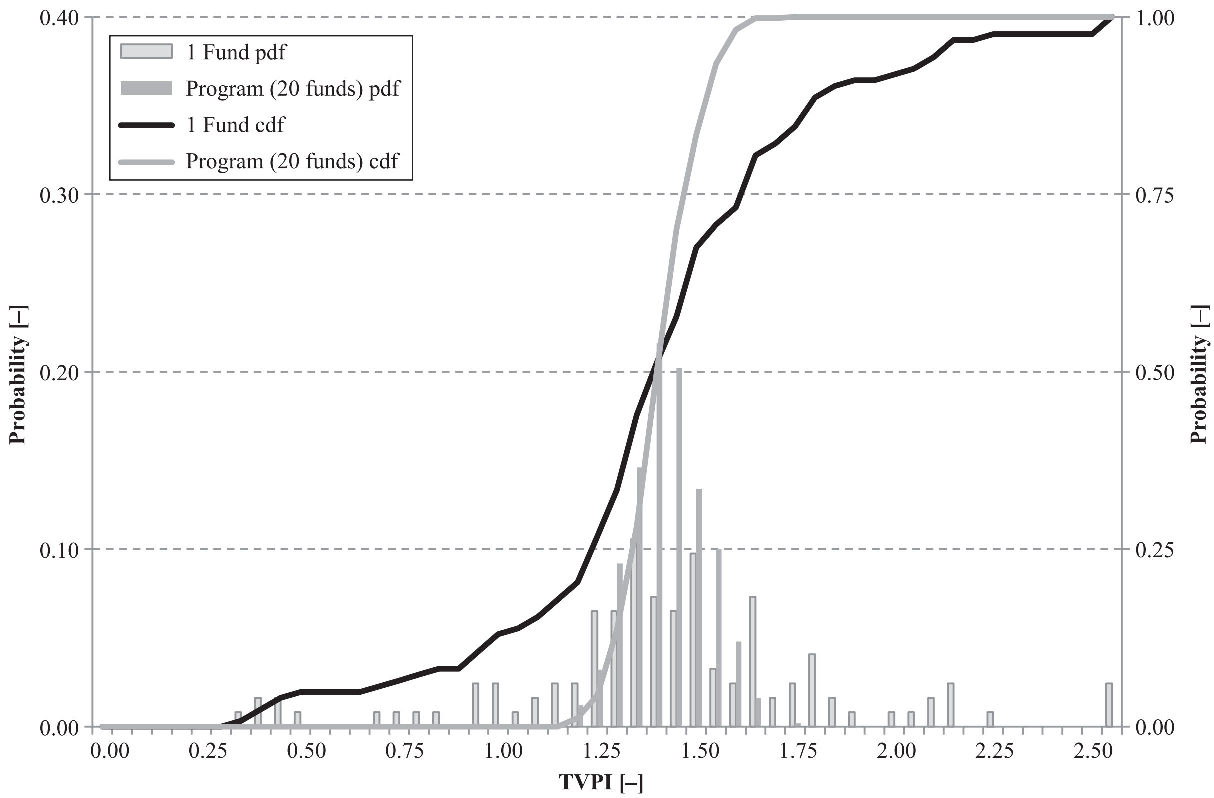
It can readily be observed that the extreme tails of the probability distribution for a fund investment are truncated for a diversified program. This means that extreme losses but also extreme gains are absent for a diversified program. One way to quantify the uncertainty around the expected TVPI is by looking at the standard deviation or volatility in each sample; these numbers are also mentioned in the caption. As expected intuitively, the diversification effect increases with the number of funds in the program, and Exhibit 2 shows the diversification benefit in terms of a multiple as a function of the number of funds in the program.

The uncertainty of outcomes decreases with each fund added to the program, and the marginal diversification benefit seems to level off at around 12 to 15 fund investments (Rouvinez and Kubr [2003]; Weidig, Kemmerer, and Born [2005]; Cornelius et al. [2013]). Clearly, the standard deviation also decreases accordingly when diversifying over a number of funds, as indicated by the numbers above the stacked bars.

Whereas the effect of asset diversification is straightforward when examining multiples (which are linear functions of the underlying cash flows), the picture becomes more complex when looking at the other key performance metric used in private equity investing: the internal rate of return.[2] With IRR values for underlying funds readily available through various commercial data vendors, it is tempting to compare IRRs of diversified programs to these values. However, because the IRR is a nonlinear function of the underlying cash flows, this approach potentially leads to material errors. Exhibit 3 shows illustrative cash flows of a diversified program that invests in three funds, USD 1 in each fund.

An investor in a diversified program would observe the aggregated or pooled series of cash flows of the underlying funds and hence would observe an IRR of 11.74%. If the investor were to calculate the simple average of fund IRRs, he would obtain an IRR of 9.87%, i.e., a difference of 187 bps. The magnitude and direction of this error depends on the size and timing of cash flows. This phenomenon, known as Jensen's inequality, states that for nonlinear functions, the function of the expectation is different from the expectation of a

Exhibit 1
Diversification Effect



Notes: This exhibit shows the diversification effect on the probability distribution of final TVPIs of an investment in a single fund (quartiles: 1.24x, 1.39x, 1.61x; mean: 1.41x; standard deviation: 0.41x) and in a program consisting of 20 randomly selected funds (quartiles: 1.34x, 1.40x, 1.47x; mean: 1.41x; standard deviation: 0.09x). All results are obtained through random sampling from a universe of 123 funds, one particular vintage year, all private equity, North America. Legend entries: pdf (probability density function), cdf (cumulative density function).

function (IRR of expected cash flows is different from the expected IRR of cash flows) (Brown [2006]; Kaserer and Diller [2011]; Ellis, Pattni, and Tailor [2012]). The investor in a diversified program should bear in mind this potentially misleading phenomenon when peer benchmarking the IRR of such a diversified program.

PEER BENCHMARKING METHODS FOR DIVERSIFIED PRIVATE EQUITY PROGRAMS

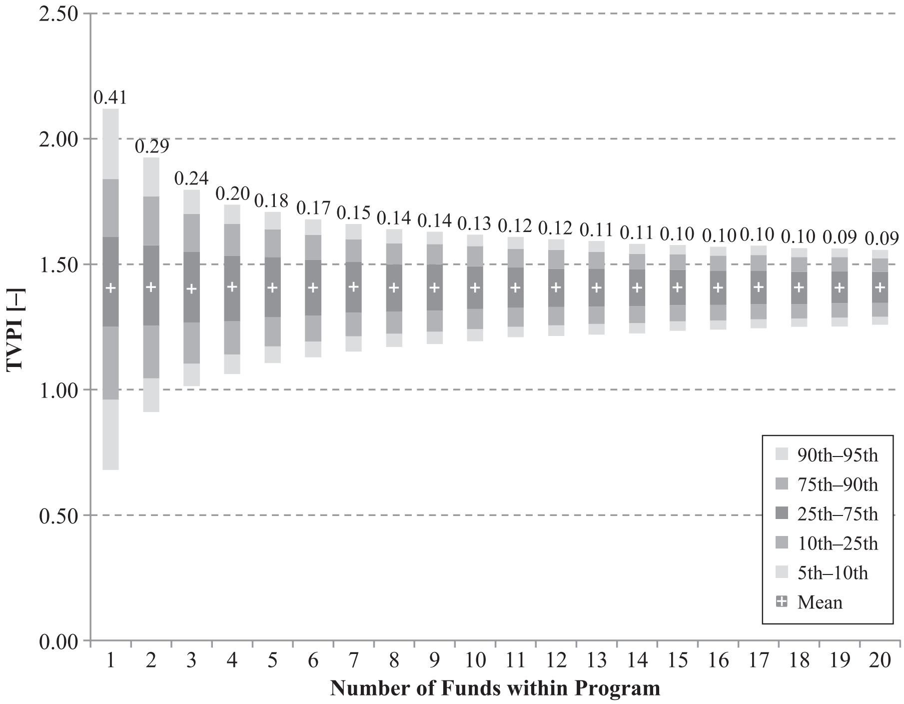
Whereas various methods exist to assess the performance of a single private equity fund investment, techniques that analyze and compare the performance of diversified programs are not readily available (Mathonet and Meyer [2007]; Day and Diller [2010]; Kaserer and Diller [2011]; Chandler, Wrigley, and Gottschalg [2015]; Demaria [2015]). A proper comparison of a diversified program's performance would require a universe of diversified programs with similar vintage year, strategy, and geographic composition. However, such universes currently available through commercial data providers are too small for a proper and meaningful comparison. To the best of the author's knowledge, three basic techniques (with variations) are used by private equity practitioners as a rough approximation to peer benchmarking of diversified programs. All of these techniques can be applied to IRR as well as to TVPI.

Status Quo

Quartile analysis of underlying funds. In this analysis, each of the underlying funds within a diversified program is benchmarked to its peer group (by vintage, strategy, and geographic focus), given a quartile ranking,

EXHIBIT 2

Probability Distribution of Final TVPI as a Function of the Number of Funds within the Diversified Program



Notes: It is important to observe that the mean remains constant, and the median is increasing slightly with each additional fund added to the program. The number above the stacked bars indicates the standard deviation. All results are obtained through random sampling from a universe of 123 funds, one particular vintage year, all private equity, North America.

EXHIBIT 3

Illustration of Jensen's Inequality in a Diversified Program Consisting of Three Funds

Time	Fund A	Fund B	Fund C	Program
1	–1.0	–	–	–1.0
2	–	–1.0	–	–1.0
3	–	–	–1.0	–1.0
4	–	0.1	–	0.1
5	0.2	1.0	–	1.2
6	2.0	–	–	2.0
7	–	–	0.3	0.3
8	–	–	1.2	1.2
IRR	17.45%	3.33%	8.83%	11.74%
		9.87%		

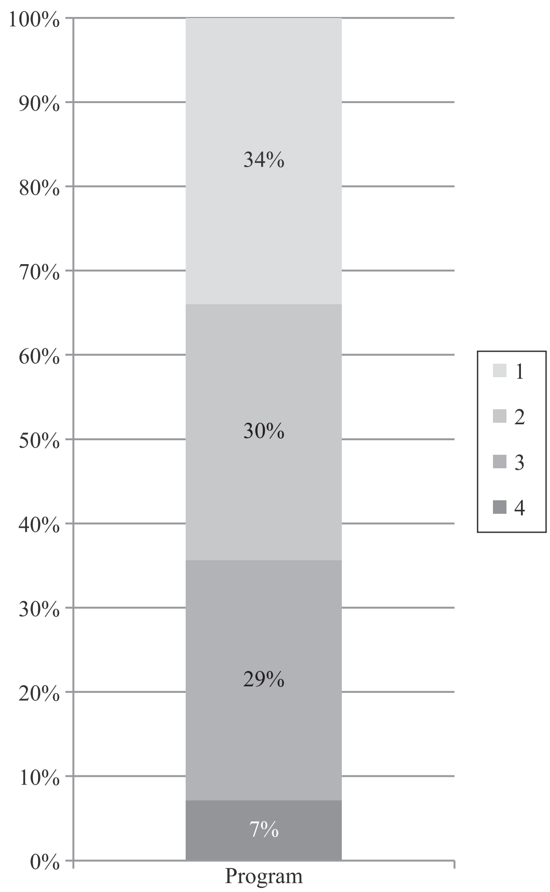
and subsequently the entire program is aggregated by either number of funds or commitments. A typical result can be seen in Exhibit 4 that shows 34% of the underlying funds are in their top quartile.[3] Although this analysis is widely used to judge fund selection capabilities, it does not quantify or properly benchmark the performance of a diversified program and provides at most a qualitative idea about its performance relative to the private equity industry.

Analysis of quartiles of underlying vintages. This analysis compares the performance of a diversified program to the quartiles of underlying vintages. Exhibits 1 and 2 already show that quartiles of single funds are fundamentally different from quartiles of diversified programs and demonstrate that little information about performance of a diversified program can be extracted from such analysis. However, to illustrate this method, let us assume a private equity program that is diversified over three consecutive vintage years. An investor might calculate quartiles for the universe of funds of each individual vintage year and compare this with the performance of a diversified program, as shown in Exhibit 5.

As can be seen, the performance of the diversified program falls in the 2nd quartile for all vintages when

EXHIBIT 4

Typical Result of a Quartile Ranking based on IRR and Aggregated by Commitment



Note: The order of the quartiles is reversed as explained in Footnote 3.

looking at IRR. Even though this analysis seems to be more quantitative than the quartile ranking, it neither takes into account that it compares a diversified program with single funds, nor takes into account the composition of the diversified program, nor properly aggregates granular underlying cash flows to calculate IRR.

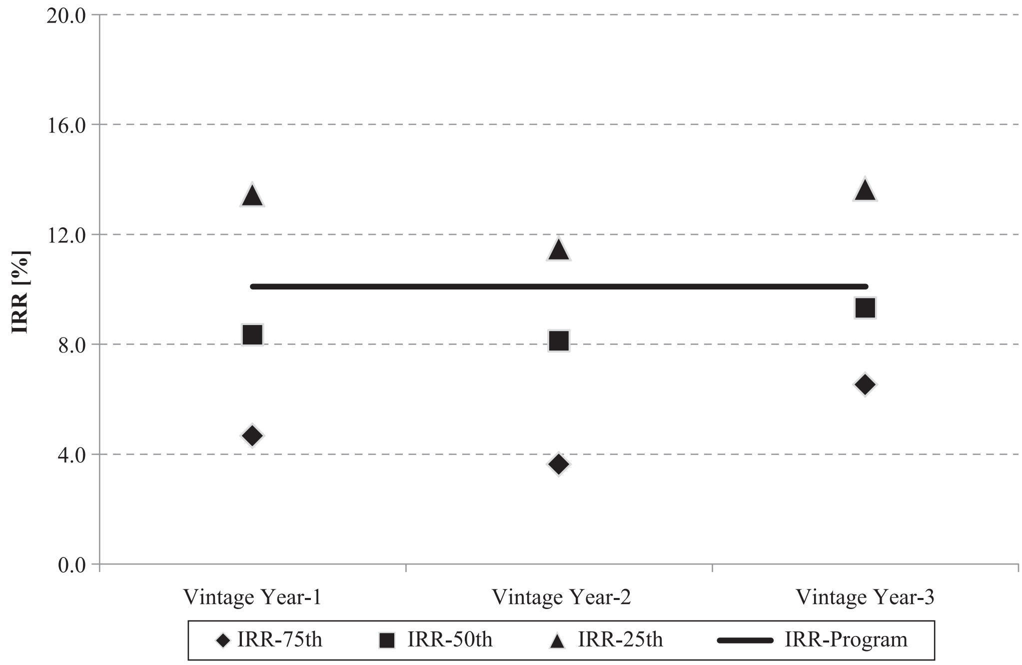
Construction of weighted benchmarks. This method is more intuitive and better approximates a true benchmarking technique for diversified programs (Rouvinez and Kubr [2003]; Weidig, Kemmerer, and Born [2005]; Mathonet and Meyer [2007, chapter 18.4.2]; Day and Diller [2010]). The technique weighs the benchmarks of the underlying peer groups (by vintage, strategy, and geographic focus) taking into account the commitments to these underlying peer groups and calculates a weighted performance metric for the diversified program. Weidig, Kemmerer, and Born [2005]; Mathonet and Meyer [2007, chapter 18.4.2]; and Day and Diller [2010] also present a probabilistic variation of this technique that randomly draws funds from a universe according to the composition of the program and aggregates TVPI and IRR to a diversified program level. Here, a technical aspect should be emphasized that will be important to understand the differences between the proposed technique and existing methods: even though the construction of weighted benchmarks might arrive at the closest approximation of a representative benchmark for a diversified program, it does not correctly aggregate cash flows when calculating IRR and hence does not address the pitfall explained in the previous section and might mislead investors.

Synthetic Peer Benchmarking

Our approach to peer benchmarking a diversified program differs in three important aspects from the aforementioned techniques: 1) it calculates a program IRR after aggregating or pooling detailed historical cash flows of underlying investments to a program level and therefore addresses Jensen's inequality as opposed to simply weighing benchmark IRRs as discussed previously; 2) it provides a probabilistic two-dimensional outcome rather than one deterministic point of reference; and 3) it attributes performance to common factors, estimates from where the over- or underperformance stemmed, and quantifies an unexplained factor which, among others, may be interpreted as manager alpha. With an ever-increasing amount of data available through various commercial data vendors, new techniques can be developed that help the investor judge how his investment in a diversified program has performed compared with the private equity industry as a whole and from where the over- or underperformance stemmed. Our technique relies on a large universe of performance data (IRR and TVPI) of individual funds coupled with a Monte Carlo simulation, and its sequence is as follows:

- We determine the composition, i.e., by vintage, strategy, and geography, taking into account the underlying investments of a diversified program measured by paid-in capital at the reporting date. Also required are the number of investments within the diversified program, the IRR and the TVPI at the reporting date (gross of manager fees,

Exhibit 5
Quartiles of Each Underlying Vintage Year



Note: Quartiles are presented in terms of IRR (symbols) compared with the gross performance of the diversified program (line).

net of underlying fees), and the fraction of the program invested into co-investments.

- For each Monte-Carlo run, we simulate a synthetic diversified program in a bottom-up fashion, and we reconstruct a synthetic history of cash flows of such program since inception—as opposed to simply weighting the benchmarks of the underlying peer groups at the reporting date. To achieve this, we first construct the time vector on a quarterly frequency from the first quarter of the oldest vintage year until the reporting date. For instance, if the oldest vintage is 2000 and the reporting date is September 30, 2014, we simulate 59 quarters.[4]
- We then draw funds from the available universe taking into account the total number of investments and the composition of the diversified program in terms of vintage, strategy, and geographic focus. For instance, when the program is composed of 40% 2007 vintages, 50% 2008 vintages, and 10% 2009 vintages, the probability of drawing a 2007, 2008, and 2009 fund in the simulation is 40%, 50%, and 10%, respectively. Each peer group on its own has a different number of funds, and this number is given, of course, by the commercial data vendor. In each Monte Carlo simulation, an underlying fund can only be drawn once. For each randomly drawn fund, we obtain one TVPI value and one IRR value—historical cash flows are not available at an individual fund level. In case the program has an allocation to co-investments, we add a spread, which has been estimated internally, to these performance values only for that fraction of the portfolio allocated to co-investments. This higher performance of co-investments can be partially attributed to a lighter fee load as compared with funds. As an important difference to existing benchmarking methods and to address the previously explained pitfall when calculating IRRs, we do not simply weigh performance metrics of underlying funds; rather, we reconstruct a series of historical cash flows for each drawn fund and subsequently aggregate those to a program level before calculating the performance metrics. This series of cash flows for each fund consists of only two cash flows, one call and one distribution. The magnitude of the distribution is given by the

TVPI, and the IRR dictates the holding period. Such simplification and reconstruction of a series of historical cash flows at an individual fund level and subsequent aggregation to a program level has previously been discussed in the literature (Rouvinez [2004], Day and Diller [2010]), and in a recent in-depth and thorough work, Renkema, van den Goorbergh, and Garcia Rivas [2015] demonstrate both theoretically and empirically that the resulting IRR is an accurate estimator for program IRR values.

- After each Monte Carlo run, we aggregate the cash flows of the underlying funds and calculate IRR and TVPI at a diversified program level. In doing so, we properly address Jensen's inequality when calculating the program IRR—an important difference to existing methods. When repeating this procedure many thousands of times, a distribution of program IRRs and program TVPIs can be obtained, and the actual performance of the diversified private equity program can be compared with the results of this Monte Carlo simulation. Each Monte Carlo run and resulting performance can be thought of as a diversified program of the same composition and number of investments as the actual program, but with the underlying investments themselves being selected blindly. The percentile in which the diversified program falls can easily be obtained numerically both for IRR and TVPI.
- The sequence described earlier enables representative synthetic peer benchmarking of a diversified private equity program, and in the following we describe the second major novelty of this technique, i.e., performance attribution and quantification of potential manager alpha. In order to attribute the performance of a diversified private equity program, we repeat the Monte Carlo simulation a number of times, each time with a different composition and only evaluate the mean of the simulation. By taking into account the composition of the broad private equity universe the program is exposed to, we obtain the market factor that intuitively explains the largest part of the performance. Subsequently, we adapt the diversified program's composition, looking separately at vintage, strategy, and geographic focus (in case they are different from the broader market universe) and determine the contributions of these characteristics to the total performance. Let us take the example of the vintage year contribution for a program investing in three consecutive vintage years: when the broad market over three underlying vintage years is composed as 35%, 35%, 30% and the program is composed 40%, 40%, 20% to the same vintage years, both compositions will result in a different performance; the former is the broad market, and the latter is the actual program adjusted for a tilt in vintage year composition. The difference between the two is the vintage year contribution or adjustment. Later on, when discussing the results of an actual program, we will elaborate further on this aspect. The sum of the vintage, strategy, and geography adjustments plus the market factor makes up that part of the performance that can be explained by factors commonly known in private equity investing.
- The unexplained factor is literally the unexplained performance between the sum of the factors described previously and the actual reported performance of the diversified program. This factor could result from various items, e.g., other factors not considered in this analysis or manager selection capabilities. This logic of attributing performance to factors is analogous to the method published by coworkers at BlackRock on the public market equivalent (Malinowski and Wittlin [2013]) and can be seen as a complementary technique.

In summary, our peer benchmarking technique differs from the commonly known methods in three important aspects:

1. Cash flows at the level of the underlying fund investments are not readily available from commercial data providers, and hence they are reconstructed from provided fund TVPI and IRR and are pooled to simulate the program's cash flows, and therefore we simulate true program IRRs. This is an important difference from other existing methods and addresses Jensen's inequality.
2. Results are probabilistic in both performance dimensions rather than providing one deterministic point of reference.
3. Performance is attributed to factors such as the market factor and adjusted for vintage year,

strategy, and geographic composition that are different from the general universe. The unexplained part of performance might, among others, be attributed to unique manager selection capabilities. To the best of the author's knowledge, this attribution to common factors and quantification of potential manager selection capabilities has not been published in the literature to date.

Let us briefly touch upon the manager's ability to generate above-average returns (alpha), as this is a key differentiator of the proposed technique and an important and ever-recurring topic in examining private equity performance (Kaserer and Diller [2011]). In this analysis, the unexplained part might be interpreted, among others, as the ability to select the best fund managers among a broad universe of fund managers. Furthermore, the vintage, strategy, and geographic adjustment can also be thought of as intentionally over- or underweighting the program with respect to the broader market. For instance, the manager might overweight his allocation to Europe in case he expects strong return potential in this region. In case these adjustments are positive contributors, they can also be interpreted as manager alpha.

We would like to emphasize again that this method compares the performance of diversified private equity programs gross of manager fees and net of underlying fees. In principle, this method could also be used to benchmark net-net performance of diversified programs, but then additional and to some extent arbitrary and disputable assumptions have to be made to estimate average general manager fee on top of the underlying fees. Most importantly, in any benchmarking method, it should be fully disclosed at what level performance is compared such that a proper "apples-to-apples" comparison can be made and legitimate conclusions can be drawn.

It is well known that exchange rates can heavily influence private equity performance, and we insulate this effect by downloading the performance of underlying funds in the reporting currency of the diversified program.

This article does not address performance of secondaries, and both examples that are discussed in depth do not contain a significant allocation to secondaries. However, this framework could be extended to include also secondaries when making assumptions about secondary market pricing.

RESOURCES

The internal data used in this work are all actual performance data audited at December 31, 2014, and the corresponding names of the comingled programs have been anonymized. We rely on a leading global provider of investment decision support tools for the private capital markets as our external data provider for the private equity industry as a whole. In this work, we consider vintages 2000 to 2014, strategies of private equity (buyout, mezzanine, distressed, special situations) and venture capital (balanced, early stage, late stage), and a geographic focus on North America and Europe. This subuniverse consists of 2,468 funds and amounts to a market capitalization of USD 2.19 trillion. The entire Monte Carlo simulation framework is implemented in MATLAB Version 7.12 by The Mathworks on a standard Hewlett-Packard workstation. Typically, we run 10,000 simulations in order to arrive at numerically stable modeling results.

RESULTS AND DISCUSSION

In this section, we discuss in depth the results of our proposed method and emphasize the important differences versus other methods. The first example is a relatively simple diversified program which focuses on North American buyout funds. The second example is a broader and more complex diversified program, which also includes an allocation to co-investments. In both examples, the results are presented in two parts: the first part provides peer benchmarking in terms of both IRR and TVPI, and the second part discusses the attribution of performance to well-known factors.

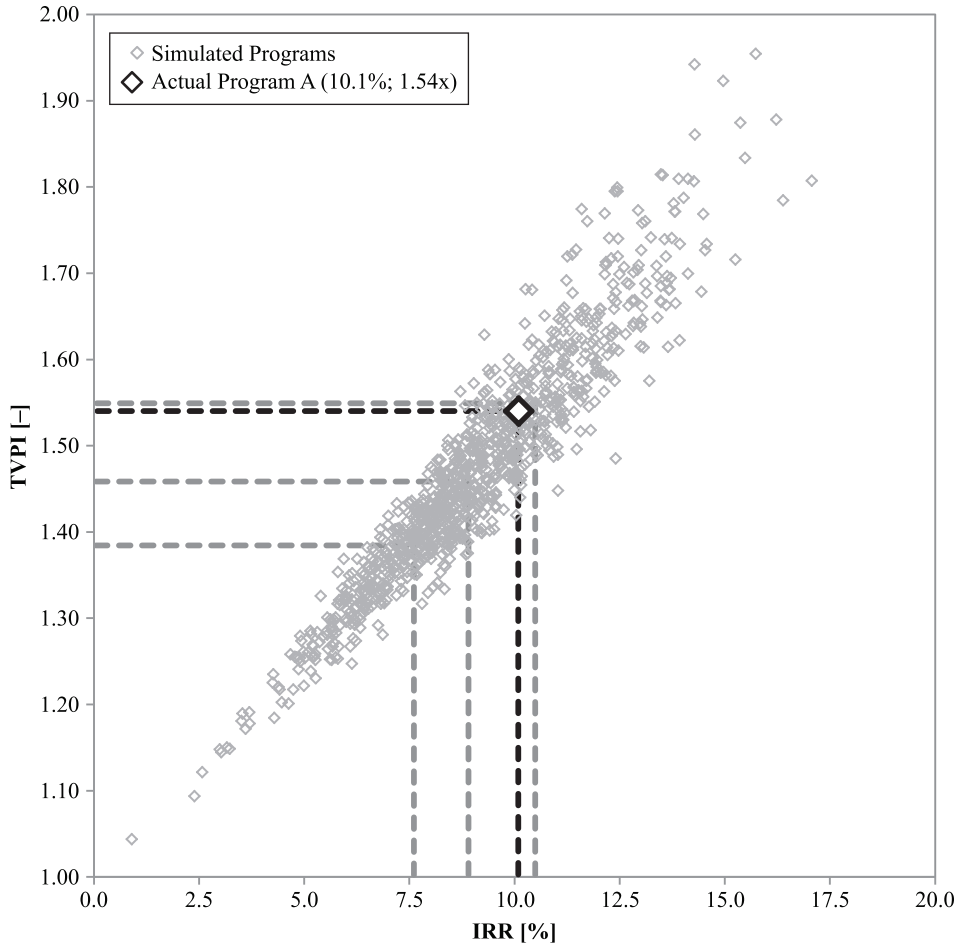
Program A

This program focuses on North America and consists of 15 buyout funds diversified over vintages 2005, 2006, 2007 and has a gross IRR and TVPI on December 31, 2014, of 10.10% and 1.54x, respectively.

Exhibit 6 shows the results of our new method and presents synthetic peer benchmarking for Program A in two dimensions. The open symbols represent the simulated IRRs and TVPIs at the reporting date obtained when taking into account the exact composition of the program. In each simulation, we draw randomly from the available universe of fund investments, reconstruct

EXHIBIT 6

A Two-Dimensional Benchmarking of Program A (IRR quartiles: 7.60%, 8.90%, and 10.49%; TVPI quartiles: 1.38x, 1.46x, and 1.55x)

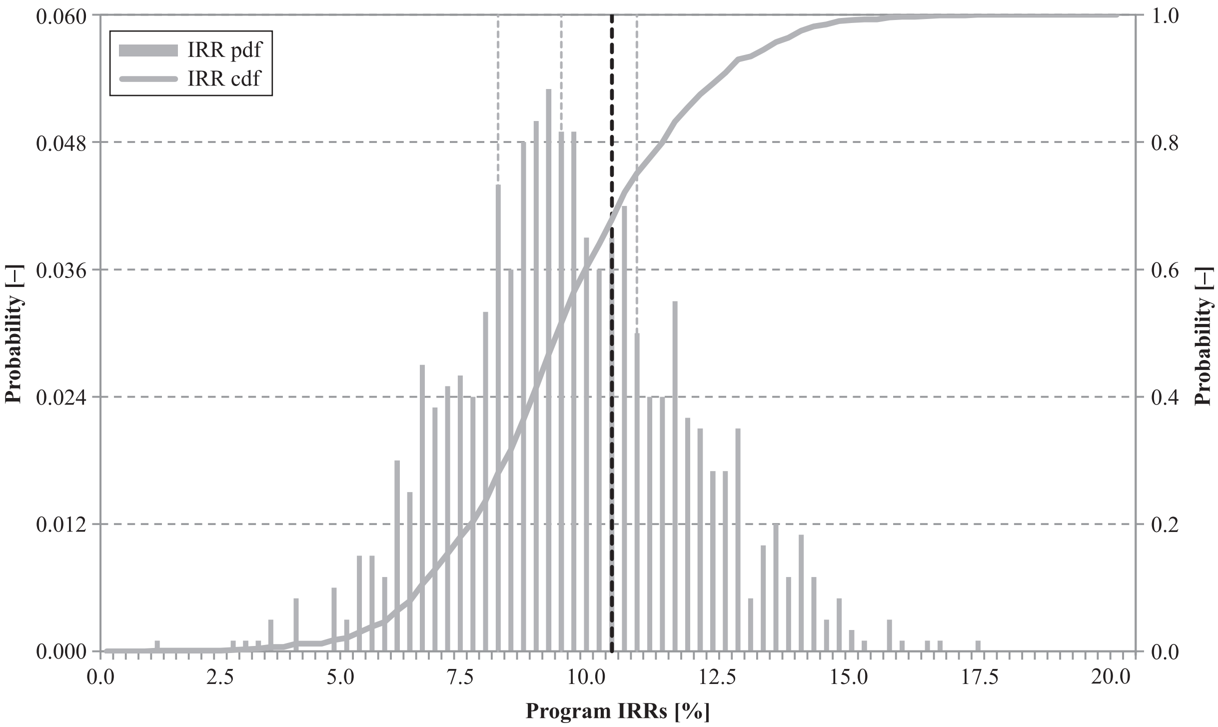


Notes: Actual program IRR and TVPI equal 10.1% and 1.54x, respectively. This analysis is based on 334 underlying funds and their performance.

the cash flows from the drawn IRR and TVPI, aggregate these cash flows to a program level, and subsequently calculate program IRR and TVPI. The cloud of symbols clearly shows the dispersion in both IRR and TVPI dimensions and, as expected, a TVPI of 1x and 2x corresponds to an IRR of approximately 0% and 20%, respectively. The dashed gray lines are the quartiles both in IRR (vertical lines) and TVPI (horizontal lines) dimensions, and the actual values are given in the caption of the figure. As mentioned before, each of these simulations and resulting symbols can be thought of as a diversified program with the same composition as Program A, but with the underlying funds being selected blindly from the entire relevant private equity universe. We then compare these simulations with the actual performance of Program A (shown as the larger black symbol) which falls in the 2nd quartile for both IRR and TVPI. Moreover, this method indicates that Program A is a 31st percentile and 27th percentile program when looking at IRR and TVPI, respectively. When projecting all these symbols along the vertical axis into the horizontal axis, the distribution of IRRs can be obtained, as shown in Exhibit 7. Also in this figure, the quartiles and actual performance are depicted. Exhibit 8 shows the same analysis, but for TVPI.

In a second step, we attribute performance to general factors such as the broad market, vintage, strategy,

Exhibit 7
Probability Distribution of Simulated Program IRRs



Notes: The gray and black dashed lines indicate the quartiles of the simulation and actual program IRR (10.1%), respectively. Legend entries: pdf (probability density function), cdf (cumulative density function).

and geography adjustments. In the case of this program, we consider only the vintage year adjustment as the program focuses solely on North American buyout funds. Exhibit 9 shows these results summarized in separate stacked bar charts, both for IRR and TVPI. The left-hand side of the figure shows that when randomly selecting funds in the North American buyout market in those vintages, the investor would have observed an IRR of 9.40%. When adjusting for the actual tilt in the vintage year composition of this particular program, which is slightly different from the broad market universe, we see that the IRR shrinks to 9.02%, meaning the tilt in vintage years contributed negative 38 bps to the overall performance. We observe, however, a gross IRR of 10.10% for Program A, and hence 108 bps make up the unexplained part of the performance. This unexplained part might be interpreted as manager selection capabilities and other factors that might not be considered in the present analysis. For TVPI, the same logic applies, and we observe an unexplained multiple of 0.07x.

Program B

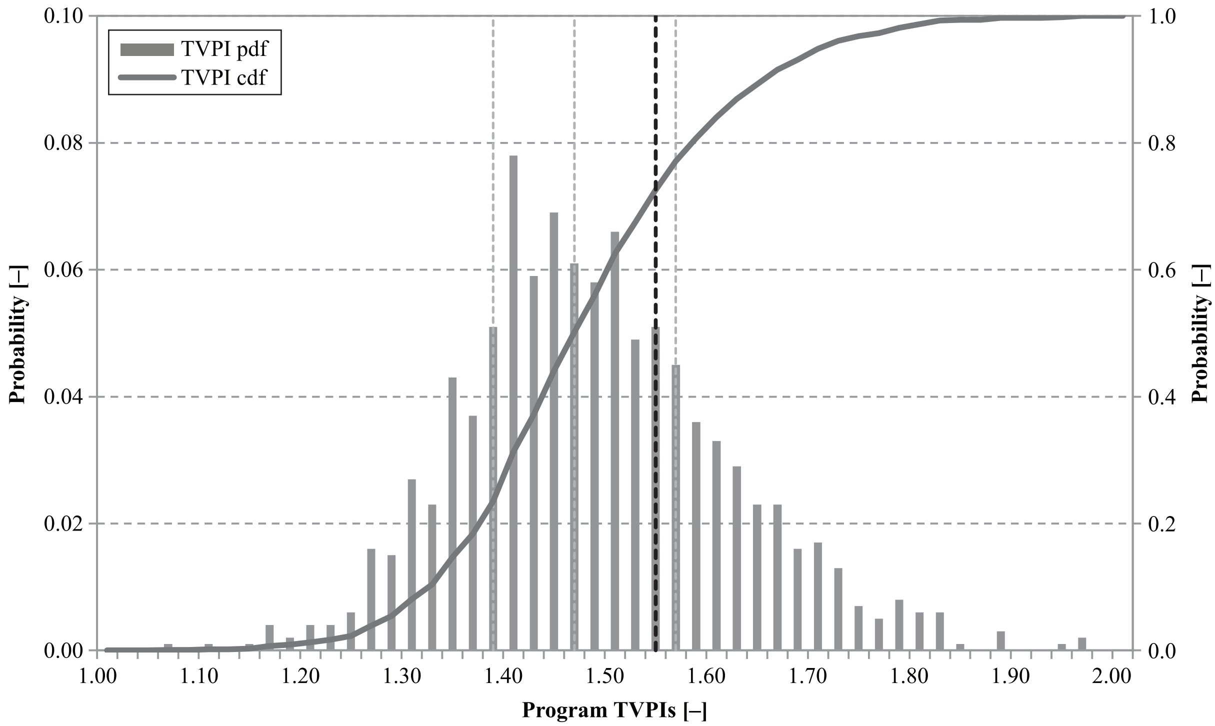
This program focuses on private equity and venture capital globally diversified over vintages 2000–2005; it consists of 61 investments including a 26% allocation to co-investments. As of December 31, 2014, Program B showed a gross IRR and TVPI of 19.04% and 1.98x, respectively.

Analogous to the previously discussed program, Exhibit 10 shows the two-dimensional peer benchmarking results for Program B. As can be seen, this program is clearly a top performing diversified program and falls within the top quartile both for IRR and TVPI. More precisely, we estimate this program to be a 10th and 9th percentile program when looking at IRR and TVPI, respectively.

Exhibit 11 shows the attribution of the performance of Program B. It can be seen that all factors contributed positively to the performance of this program. The left-hand side of the figure shows that when randomly

Exhibit 8

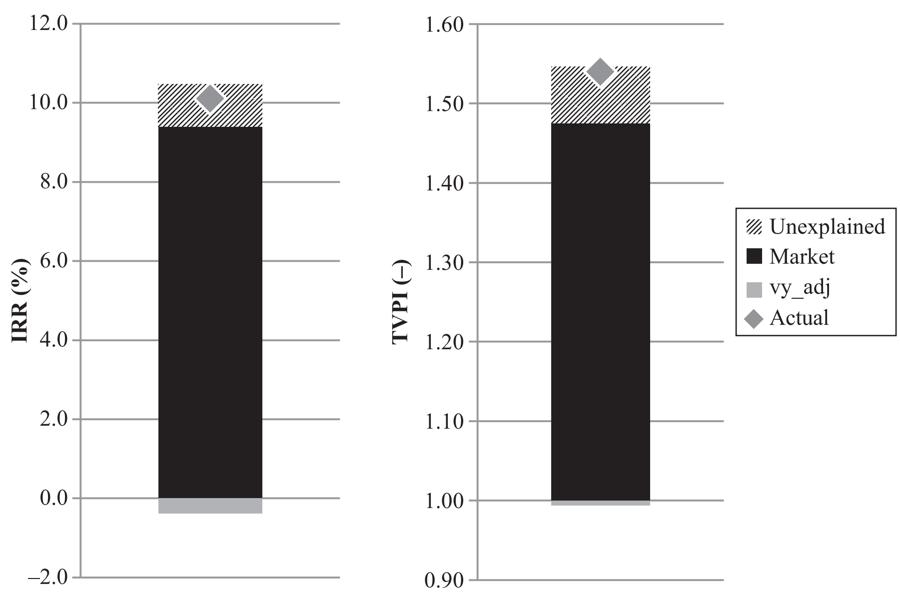
Probability Distribution of Simulated Program TVPIs



Notes: The gray and black dashed lines indicate the quartiles of the simulation and actual program TVPI (1.54x), respectively. Legend entries: pdf (probability density function), cdf (cumulative density function).

Exhibit 9

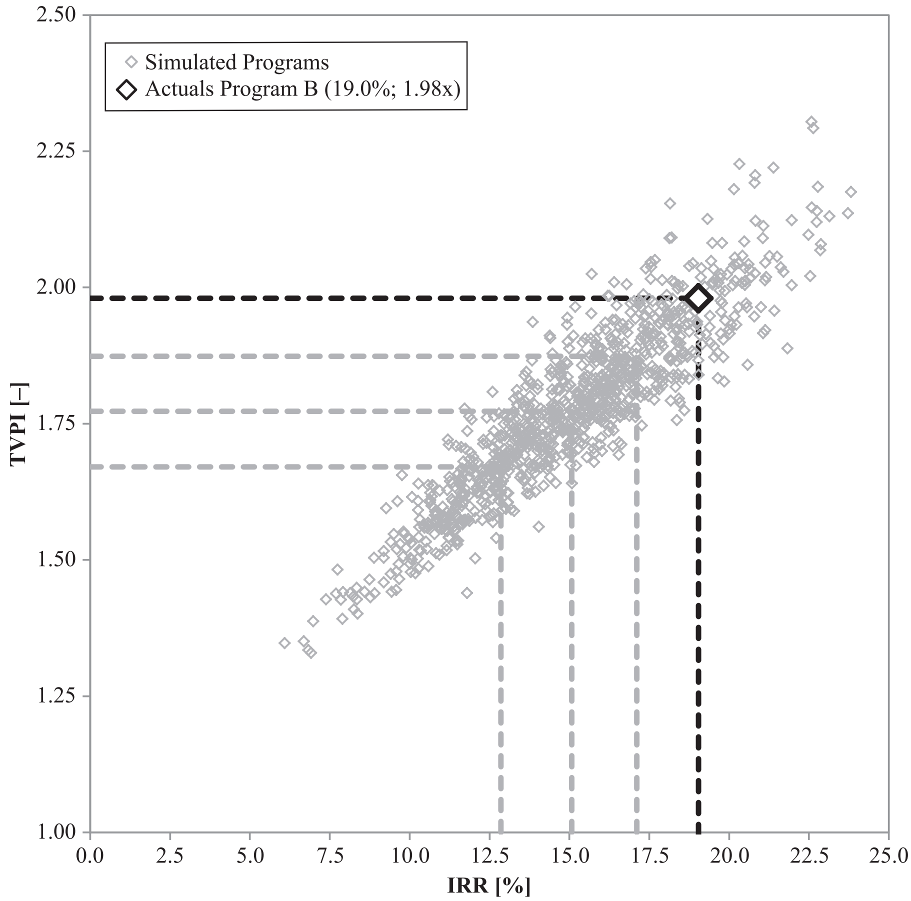
Performance Attribution of Program A



Notes: Performance attribution results are presented both in terms of IRR (left-hand side) and TVPI (right-hand side). Legend entry: vy_adj (vintage year adjustment).

EXHIBIT 10

A True Two-Dimensional Benchmarking of Program B (IRR quartiles: 12.85%, 15.07%, and 17.11%; TVPI quartiles: 1.67x, 1.77x, and 1.87x)



Notes: Actual program IRR and TVPI equal 19.0% and 1.98x, respectively. This analysis is based on 840 underlying funds and their performance.

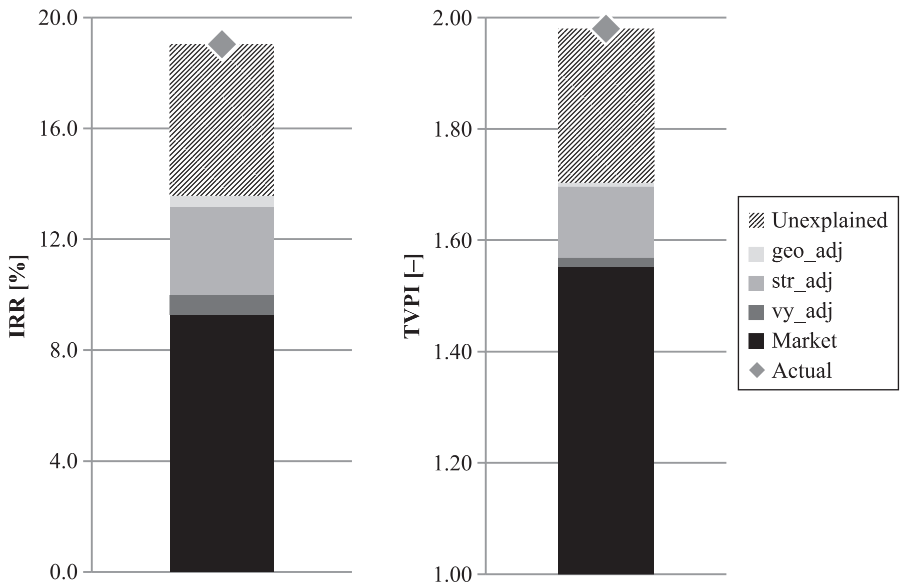
selecting funds from a universe that corresponds to those vintages, strategies, and geographies, the investor would have observed an IRR of 9.28%. When adjusting for the actual tilt in the vintage year composition, which is slightly different from the broad market universe, we see that the IRR increases to 9.98%, meaning the tilt in vintage years contributed a positive 70 bps to the overall IRR. When adjusting for the actual tilts in strategy and geographic focus, we obtain positive contributions of 319 bps and 40 bps, respectively. The large positive contribution of the strategy tilt can be explained by the fact that, when compared with the broad market universe, this program is underweighted to venture capital, which in those vintage years is known to have performed below expectations. When adding all those factors, we can explain 13.57% of the performance, leaving 547 bps as unexplained. For TVPI, the same logic is valid, and we observe an unexplained multiple of 0.28x.

CONCLUDING REMARKS

Whereas various methods exist to assess the performance of an investment in a single private equity fund, techniques that analyze and compare the performance of diversified private equity programs are not readily available. This article presents a new peer benchmarking technique for diversified programs that overcomes three key challenges that have prevented peer benchmarking of diversified programs until now: 1) a lack of comparable data—mimicking its composition and number of holdings we simulate a diversified program and additionally provide probabilistic outcomes in both performance

Exhibit 11

Performance Attribution of Program B



Notes: The results are presented both in terms of IRR (left-hand side) and TVPI (right-hand side). Legend entries: geo_adj (geography adjustment), str_adj (strategy adjustment), vy_adj (vintage year adjustment).

dimensions rather than providing one deterministic point of reference; 2) a misleading practice of averaging underlying performances of individual funds—we reconstruct underlying cash flows, aggregate to a program level, and then calculate performance; 3) no visibility on the drivers of out- (or under-) performance—we unravel performance into common factors and estimate an unexplained part of performance which might, among others, be attributed to unique manager selection capabilities.

This work focuses on private equity investments, but it should be emphasized that this methodology works as well for other illiquid alternative asset classes such as real estate and infrastructure, given a sufficient universe of empirical data.

ENDNOTES

The author is grateful to Nathalie von Niederhaeusern, Julia Wittlin, Elliot Asarnow, and Christophe Rouvinez for their helpful suggestions and fruitful discussions.

[1]The TVPI is the total value (cumulative distributions plus combined net asset value) divided by paid-in capital contributions.

[2]The IRR is an unknown rate that discounts the cash flows such that the sum of their present values equates to zero. When looking at more than two cash flows, there is no closed form solution to this problem, and the IRR has to be found through an iterative search. In some cases, there might even exist multiple solutions to this nonlinear problem.

[3]Strictly mathematically, the top performing funds fall in the 4th quartile and the underperforming funds in the 1st quartile. As a convention and in order to be consistent with most private equity professionals, we reverse this order and define top performing funds in the 1st (top) quartile and underperforming in the 4th (bottom) quartile. The same holds for the percentiles when looking more granularly at performance later on in this work.

[4]Important to note is that we randomize the starting point of a fund investment to better reflect real private equity investing behavior, e.g., a 2007 fund might start calling capital in Q3 2007 in the 1st Monte Carlo run, in Q1 2007 in the 2nd Monte Carlo run, in Q4 2007 in the 3rd Monte Carlo run, and so on and so forth.

REFERENCES

Brown, R.J. "Sins of the IRR." *Journal of Real Estate Portfolio Management*, Vol. 12, No. 9 (2006), pp. 195-199.

Chandler, A., H. Wrigley, and O. Gottschalg. "Comparing Apples to Apples with PME Benchmarks." Bloomberg, 2015.

Cornelius, P., C. Diller, D. Guennoc, and T. Meyer, *Mastering Illiquidity*. West Sussex, U.K.: John Wiley & Sons, 2013.

Day, S., and C. Diller. *Risk Management in Private Equity*. London: PEI Media, 2010.

Demaria, C. "Fee Levels, Performance, and Alignment of Interests in Private Equity." *The Journal of Alternative Investments*, Vol. 17, No. 4 (2015), pp. 95–135.

Diller, C., and C. Kaserer. "What Drives Private Equity Returns? Fund Inflows, Skilled GPs, and/or Risk?" *European Financial Management*, Vol. 15, No. 3 (2009), pp. 643–675.

Ellis, C., S. Pattni, and D. Tailor. "Measuring Private Equity Returns and Benchmarking against Public Markets." BVCA Research & Economics (2012), pp. 1–26.

Frei, A., and M. Studer. "Practitioner's Guide to Private Equity Benchmarking." *The Journal of Alternative Investments*, Vol. 7, No. 3 (2004), pp. 81–94.

Gompers, P.A., S.N. Kaplan, and V. Mukharlyamov. "What Do Private Equity Firms Say They Do?" Working Paper 15-081, Harvard Business School, 2015.

Gottschalg, O., E. Talmor, and F. Vasvari. "Private Equity Fund Level Return Attribution: Evidence from U.K. Based Buyout Funds." BVCA Research & Economics (2010), pp. 1–24.

Gresch, N., and R. von Wyss. "Private Equity Funds of Funds vs. Funds: A Performance Comparison." *The Journal of Private Equity*, Vol. 14, No. 2 (2011), pp. 43–58.

Harris, R.S., T. Jenkinson, and S.N. Kaplan. "Private Equity Performance: What Do We Know?" *The Journal of Finance*, Vol. 69, No. 5 (2014), pp. 1851–1882.

Kaplan, S.N., and S. Antoinette. "Private Equity Performance: Returns, Persistence, and Capital Flows." *The Journal of Finance*, Vol. 60, No. 4 (2005), pp. 1791–1823.

Kaserer, C., and C. Diller. *Performance Measurement and Benchmarking in Private Equity*. London: PEI—Private Equity International, 2011.

Malinowski, J., and J. Wittlin. "Shining a Brighter Light on Private Equity." BlackRock (2013), pp. 1–4.

——. "Do Buyout Funds Outperform?" BlackRock (2014), pp. 1–7.

Mathonet, P.Y., and T. Meyer. *J Curve Exposure: Managing a Portfolio of Venture Capital and Private Equity Funds*. West Sussex, U.K.: John Wiley & Sons, 2007.

Phalippou, L., and O. Gottschalg. "The Performance of Private Equity Funds." *The Review of Financial Studies*, Vol. 22, No. 4 (2009), pp. 1747–1776.

Renkema, J., R. van den Goorbergh, and C. Garcia Rivas. "RATZ IRR: Performance Measurement in the Absence of Cash Flow Data" Working paper, APG Asset Management, 2015.

Rouvinez, C. "Private Equity Benchmarking with PME+." Venture Capital International (2003), pp. 34–38.

——. "Benchmarking Private against Private." Private Equity International (2004), pp. 36–39.

Rouvinez, C., and T. Kubr. "Enhancing Private Equity Returns via Diversification." Private Equity International (2003), pp. 19–22.

Sorensen, M., and R. Jagannathan. "The Public Market Equivalent and Private Equity Performance." *Financial Analysts Journal*, Vol. 71, No. 4 (2015), pp. 43–50.

Weidig, T., A. Kemmerer, and B. Born. "The Risk Profiles of Private Equity Funds of Funds." *The Journal of Alternative Investments*, Vol. 7, No. 4 (2005), pp. 33–41.

Weidig, T., and P.Y. Mathonet. "The Risk Profiles of Private Equity." Working paper, European Investment Fund, 2004.