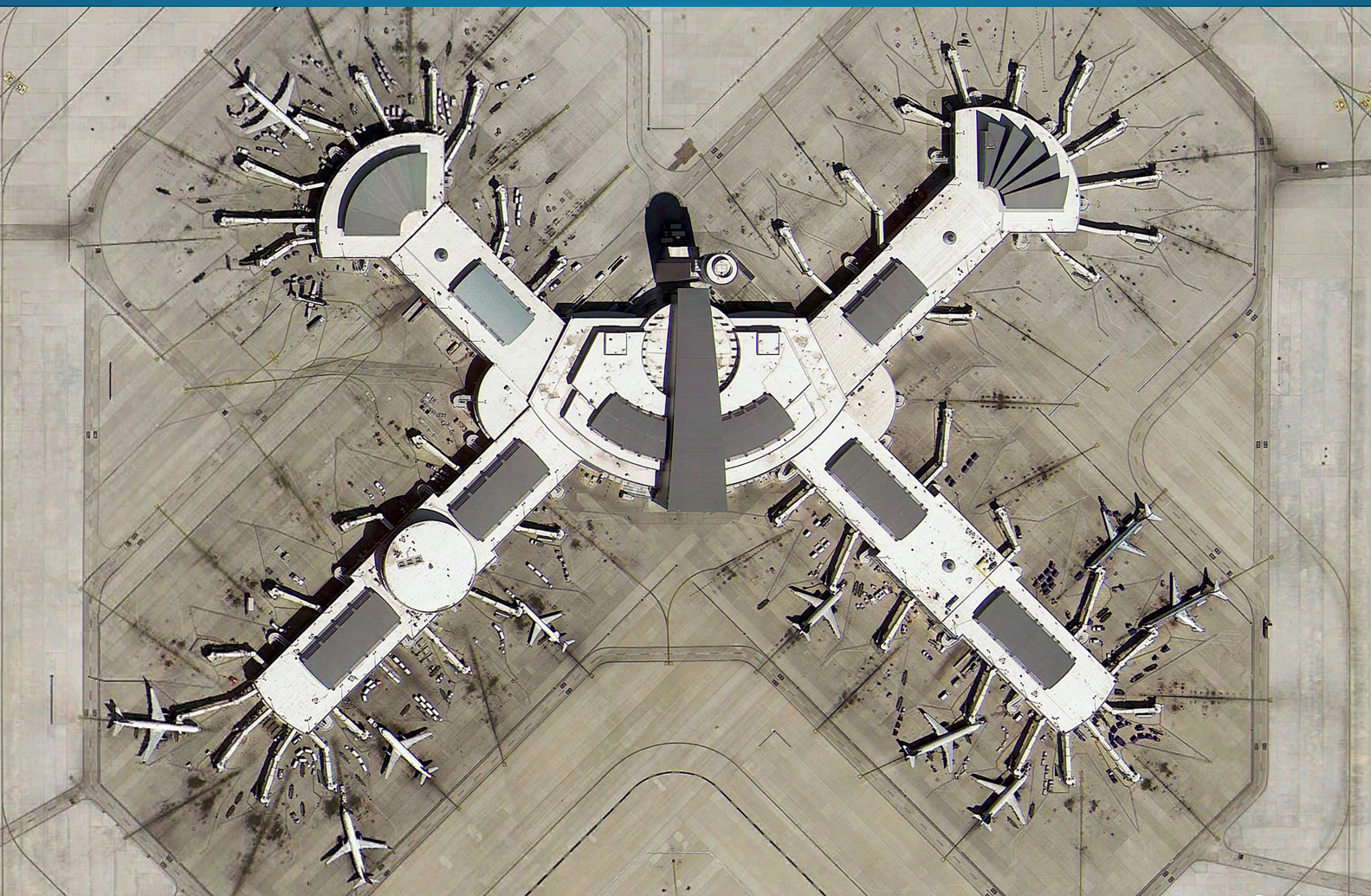


Robust Benchmarks for investors in private infrastructure funds

infraMetrics® White Papers



infraMetrics®
by  **QEDHEC** *infra*

November 2021

Executive Summary

The infraMetrics fund strategy analyser enables the gross and net performance of unlisted infrastructure funds to be benchmarked using robust IRR and multiple quartiles that are not biased or skewed by the limitation of manager contributed data. This tool makes thousands of observations of the typical performance of infrastructure funds available, in hundreds of segments, along with dozens of geographies and 20 years of vintages, all updated quarterly with no lag. Simulated results are both congruent with contributed market data at the aggregate level over a long period, and more robust and precise at the vintage year or sub-segment level. With this tool, infrastructure manager selection and fund monitoring are no longer hindered by unreliable and biased reported fund performance data.

Scant contributed data leaves investors and managers guessing about fund performance

Ranking and selecting managers based on quartiles is not just a matter of sorting funds by IRR and picking the top of the list. The notion of quartile implies an underlying statistical distribution of returns and a relative ranking i.e. ranking funds or managers by quartile is a basic form of performance benchmarking. Using quartiles to rank observations requires either knowing the underlying distribution of returns is or observing a sufficiently large number of realised performance metrics in order to estimate the quartiles of that distribution empirically with reasonable accuracy.

Contributed datasets typically include one or two dozen new observations per vintage. In this paper, we show that using 10-50 data points per year from a well-defined distribution leads to completely different estimates of the

quartiles boundaries from one year to the next. In other words, **when only 10 to 50 observations of IRRs or multiples are available in any given vintage year, the confidence with which investors can estimate quartile boundaries is so low that the resulting rankings of managers and funds are meaningless.** The same fund may find itself switching from the top quartiles to the bottom quartiles simply because of a difference in contributed performance data in two consecutive years.

In effect, using IRR and multiple quartiles to rank and select private fund managers is practically impossible until at least 1,000 (ideally 10k) observations are available to investors or consultants. Clearly, there are not enough infrastructure funds in the world to report that much performance data, let alone in the same vintage year.

We show that aggregating contributed data over several consecutive years, as is sometimes done, does not improve precision much: with five years of data (c.100 observations), the 95% confidence interval of the quartiles boundaries is still very large for the purpose of ranking annual rates of returns (between 700 and 1,200 basis points).

With such sparse contributed fund performance data, investors cannot know with a reasonable degree of confidence what the quartile of the net IRR or net multiple distributions really are. Since quartile rankings are an important part of the private fund investment decision and monitoring process, in the case of infrastructure funds, investors (LPs) and managers (GPs) do not have any useable point of reference to evaluate performance.

Moreover, data paucity gets worse as one tries to drill down to the subsegments of the infras-

structure universe, making any comparisons or assessments between styles and fund risk profiles impossible as well.

A large dataset of simulated fund achieves much better results if using market valuations and cash flows

With a large number of observations of fund performance, reliable estimates of the quartile boundaries can be obtained. In the absence of enough observable data, such a high number of observations can only be achieved through simulation. Simulating private fund returns has been suggested before in industry research. The solution is to 'bootstrap' the estimation of returns and multiples quartile boundaries by creating a large sample of possible investment outcomes with at least 1,000 data points per vintage or segment.

For these results to be relevant to investors two important conditions need to be met:

- The valuations of the assets purchased and sold by the simulated funds need to represent the market values of the underlying assets at the time, as well as the distributions made by each invested company in any given year;
- The simulated behaviour of the funds should correspond to the typical path followed by such investment vehicles.

We build our approach using the market value of hundreds of infrastructure investments and their dividend cash flows, obtained from the infraMetrics® database. This data presents a number of advantages that make it the best available input for a simulation exercise:

- The infraMetrics database of 700+ tracked unlisted infrastructure firms is designed to track a broader investible universe of 7,000+ firms in 25 countries in a representative manner. Hence, it does not suffer from reporting and survivorship biases.

- infraMetrics valuations are re-calibrated each quarter to match the most current level of expected returns, including the unlisted infrastructure equity risk premium and the relevant yield curve (see Blanc-Brude and Gupta, 2021, for a summary).

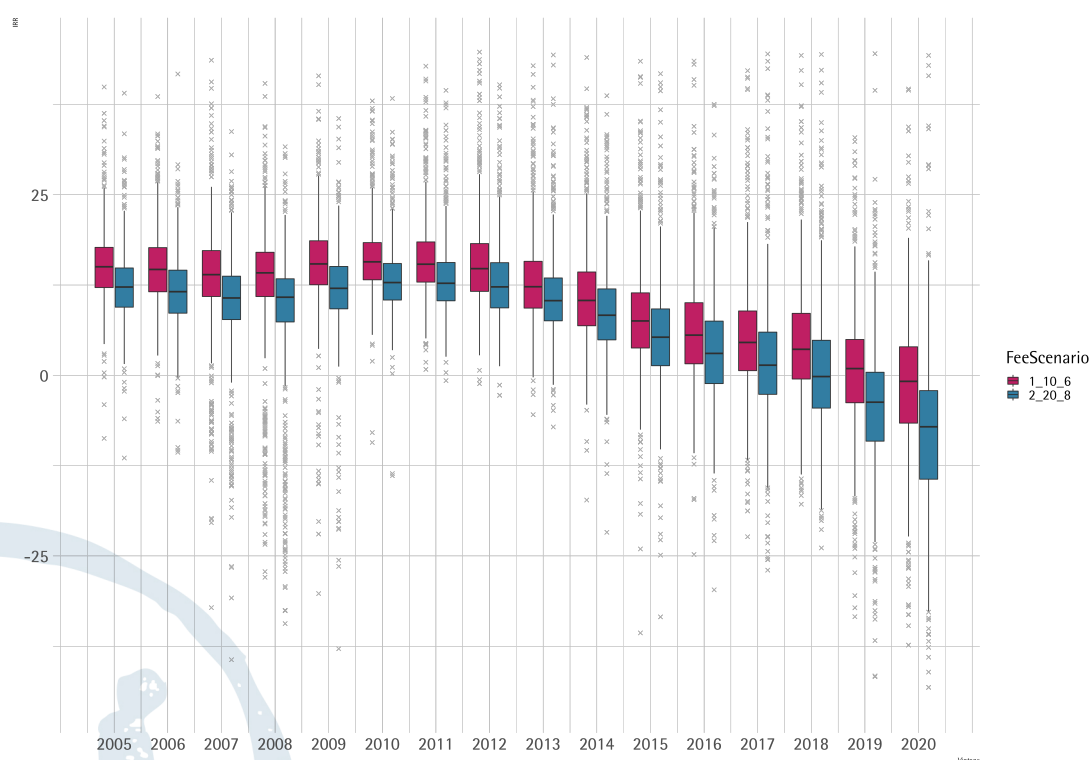
Next, our fund simulations rely on a number of assumptions about the investment pace and size of private infrastructure funds, fee levels, ability to deploy capital, etc. (see paper and Appendix). We developed these assumptions through a combination of desk-based research and semi-structured interviews with investment managers. Market participants were also involved in the validation and 'reality check' of the simulation results (see the acknowledgments at the end of this Executive Summary).

With the infraMetrics fund strategy analyser, investors can avoid selecting the wrong managers and be able to track performance over time

In back-tests, we find that simulated results are congruent with contributed data at the aggregate level i.e., over a long period: they are close enough to suggest that the simulation is capable of producing market-like results. Alignment of the results with market data is due to the use of infraMetrics' market valuations and realised asset-level cash flows as the inputs of a bottom up simulation. Thanks to a range of fee scenarios available, the results can be used to compare performance across different fee structures as shown on figure 1. They are also more robust and precise than contributed data at the vintage year, strategy or sub-segment level. The infraMetrics fund analyser produces results for more than 100 segments including Global Core and Core+, Europe, UK or Australia and most TICC® segments of the infrastructure universe.

Our results also highlight the possibility of fund selection errors when using small samples: a fund may appear to be in the bottom/top quartile

Figure 1: Comparison of net fund IRR by fee level: 2&20 + 8% hurdle rate vs 1&10 + 6% hurdle rate



Source: infraMetrics®

when it actually is not. In the paper, we describe examples of real funds that are given a mediocre quartile rank using contributed data, whereas the iFSA benchmark places these funds in the top quartile. This is the case of Type I error (false negative). LPs making this error miss the opportunity to invest with a good manager, while managers find themselves unable to showcase their skills.

We also show examples of Type II error (false positive): against a contributed TVPI benchmark two funds can appear in the top quartile, whereas, benchmarked using iFSA data they are in the 3rd and 2nd quartiles respectively. LPs making such errors would unknowingly select a poor-performing manager.

Using robust and granular fund performance benchmarks is the only way for LPs to select managers on the basis of their relative quartile rankings. Likewise, such data is necessary for GPs to showcase their skills and compare their performance against a fair and robust benchmark.

The paper also suggests two important use cases of this data:

- **Manager selection:** The due diligence process to select the best possible manager differs across LPs and includes both qualitative and quantitative factors. Invariably, one of those factors is the past performance of the manager's funds. Using robust and granular fund performance benchmarks is the only way for LPs to select managers on the basis of their relative quartile rankings. Such data is necessary for GPs to showcase their skills and compare their performance against a fair and robust benchmark.
- **Fund monitoring:** funds with very similar ex ante characteristics (size, announced strategy, geography, etc.) could perform very differently depending on the investment decisions made by the manager, especially since individual funds are usually limited to just a few investments (typically less than a dozen), making active investment choices highly relevant to

monitor. Thank to the methodology and data used for iFSA, not only can benchmark performance metrics be computed by vintage, but they can also be computed each year for a single vintage. We describe how investors might track the transition of certain funds between quartiles from year to year thanks to granular and robust benchmarks of performance quartiles.

Acknowledgments: the authors wish to thank for their comments and contributions to this paper Dr Noël Amenc, Dr Tim Whittaker, BlackRock, Landmark Partners, Evli Fund Management, Yielco, Mercer, DWS, GMPF & Local Pension Partnership.



Contents

Executive Summary	2
1 Motivation: with contributed data, fund quartile rankings are like a lottery	8
2 Proof of Concept	10
2.1 Step 1: Fund quartiles estimated with contributed data are not reliable	10
2.2 Step 2: Simulating private funds achieves superior quartile estimation results	13
2.3 Step 3: Back-testing	15
3 Benchmark Design Methodology	20
3.1 Fund characteristics	20
3.2 Selection of underlying investments	20
3.3 Computing fund cash flows	22
4 Fund Strategy Benchmark Metrics	24
4.1 IRR	24
4.2 TVPI	26
4.3 DPI	26
4.4 RVPI	27
4.5 PME	27
5 Example Use Cases	29
5.1 Manager selection	29
5.2 Fund performance monitoring	30
A Appendix	32
A.1 Sample size and robustness of quarter estimates	32
A.2 Investment ratio assumptions	32
A.3 Number of investments made	32
A.4 Fund size assumptions	32
A.5 Deal success rate assumptions	32
References	37
EDHEC<i>infra</i> Publications (2018-2021)	38

Frédéric Blanc-Brude is the Director of EDHEC Infrastructure Institute, a dedicated research unit developing a unique body of applied research on infrastructure investment from the perspective of large asset owners. He is also the CEO of Scientific Infra, a provider of unlisted infrastructure equity and debt index data and analytics. He is a member of the editorial board of the *Journal of Alternative Investments*. He holds a PhD in Finance (King's College London) and degrees from the London School of Economics, the Sorbonne, and Sciences Po Paris. He is a regular contributor to the G20 working group on long-term infrastructure investment, has advised the European Insurance and Occupational Pensions Authority (EIOPA) on the prudential treatment of infrastructure investments and also represents EDHEC on the Advisory Board of the Global Infrastructure Facility of the World Bank. EDHEC*infra* was founded in 2016.

Abhishek Gupta is an Associate Director at the EDHEC Infrastructure Institute and the Head of infraMetrics Product Development. He has more than 10 years of experience in asset management and alternative investments including stints at Goldman Sachs and Partners Group. He holds a Masters of Science in Financial Engineering from Nanyang Business School and a Bachelor of Technology from the Indian Institute of Technology.

1. Motivation: with contributed data, fund quartile rankings are like a lottery

Today, 80% of institutional investors exposed to unlisted infrastructure equity are invested via managed private investment funds. As a result, fund manager selection and performance monitoring are key aspects of the investment process in infrastructure. Indeed, most individual infrastructure portfolios are concentrated in a limited number of investments reflecting active manager choices.

To select skilled managers, investors and fund of fund managers typically rely on rankings by quartiles of net IRR and multiples and aim to work with asset managers that are consistently in the top quartiles. Likewise, to monitor performance, investors need to compare the reported performance of the funds they are invested in with that of comparable funds and, again, hope to achieve top quartile results.

However, this process is hindered by the limited availability of infrastructure fund performance data. There are at least five reasons why such data is sufficiently scarce and biased to make both manager selection and monitoring very challenging. Indeed, it is more akin to a lottery than a rigorous, criteria-based selection process:

- first, available sample sizes are small (usually less than 30 data points) and, as we demonstrate below, estimating quartile boundaries reliably is impossible with so little data;
- second, the data contributed by managers can suffer from several biases (reporting, selection and survivorship biases), making the estimation of quartiles of manager and fund performance unreliable;
- third, in the case of some strategies and geographies, too few funds may exist in the first place to achieve any robust estimate of the quartiles of returns and multiples even if all the available data could be collected;
- fourth, because this data is contributed and processed by humans, it is sometimes wrong: either the exact investment year or the performance data itself can sometimes be inaccurate. Such human errors are compounded by the limited number of data points available: with 30 data points per vintage or fewer to rely on, there is no law of large numbers to “wash out” human errors and a single inaccurate data point can create large deviation in reported quartiles;
- fifth, the same is true of outliers: if reported data includes one or two very high or very low IRRs, with a small sample, estimated quartile boundaries are not robust. As far as we know, there is no outlier treatment in existing commercial datasets used to rank funds and managers.

Finally, contributed fund data is also typically stale i.e. available with a lag of one to three years, depending on the age of the fund. New funds usually do not report any performance data for the first two or three years, and more mature funds tend to report with a lag of up to four quarters.

Moreover, since most funds also arbitrarily set a fixed hurdle rate at 7 or 8%, in the absence of robust performance quartile data, there typically is no relative benchmark against which infrastructure funds and managers can be assessed.

Such a paucity of performance data for infrastructure funds means that asset managers (GPs) can struggle to demonstrate whether they are performing adequately or not, while investors (LPs) are left none the wiser about the skills or performance persistence of their asset managers.

Hence, fund quartile ranking can be a lottery as we show in this paper, top quartile funds can be ranked in lower, even the lowest quartile, and vice-versa. This is of course problematic for LPs who wish to select good managers, but also for these same good managers who may not be able to showcase their skills if they cannot be benchmarks fairly. We show such examples using real-world fund data in section 5.

EDHECinfra has developed an industrial-grade solution to this endemic data paucity problem in the private infrastructure fund space with the Fund Strategy Analyser component of its infraMetrics platform: thanks to its access to the market valuations and distributions of hundreds of individual infrastructure equity investments in 25 countries, over 20 years and in dozens of market segments, the infraMetrics Fund Strategy Analyser (iFSA) provides unbiased, robust and consistent quartile estimates of the performance of unlisted infrastructure investment funds.

iFSA uses the infraMetrics database to mimic the typical behaviour of private infrastructure investment funds and produce robust estimates of the IRR, multiples and PME quartiles that would be reported if thousands of funds existed in the market and faithfully reported their performance data in each segment and each vintage, every quarter.

This white paper shows that unless investors have large datasets, they cannot rely on quartile ranking and that creating such datasets is possible for infrastructure funds thanks to the infraMetrics technology.



2. Proof of Concept

In this section, we show that using the contributed infrastructure fund performance datasets typically available to investors to estimate performance quartiles leads to random and even misleading results. We show that a much larger amount of data is needed to achieve reliable quartile estimates of the relative performance of funds and fund managers, and that much better results can be achieved by simulating 'typical' funds, given access to enough data for the funds underlying assets, their valuations and cash flows.

In what follows, we first demonstrate the lack of robustness of quartile rankings based on small samples, then discuss the greater robustness of simulated fund performance. We also provide back-testing results confirming the superiority of simulated results compared with contributed data, especially when it comes to evaluating individual vintages or strategies.

2.1 Step 1: Fund quartiles estimated with contributed data are not reliable

Quartiles of net IRRs and multiples have been used to select and assess private investment funds for decades, and unlisted infrastructure funds are no exception. Ranking and selecting managers based on quartiles is not just a matter of sorting funds by IRR and picking the top of the list. The notion of quartiles implies an underlying statistical distribution of returns and a relative ranking; i.e. ranking funds or managers by quartile is a basic form of performance benchmarking.

Using quartiles to rank observations requires either knowing what the underlying distribution of returns is or to observing a sufficiently large number of realised performance metrics to estimate the quartiles of that distribution empirically with reasonable accuracy.

Unfortunately, the distribution of private infrastructure fund returns in any given year is unknown and unobservable. Also, as we show below, using contributed performance data to estimate quartiles boundaries leads to unreliable results due to the paucity of available data.

For example, looking at the Preqin dataset of unlisted infrastructure fund performance metrics (see www.preqin.com), recent vintage years typically exhibit 10-20 contributors for net IRRs and 15-35 contributors for net multiples.

As of Q3 2021, the full Preqin dataset includes 228 observations of infrastructure fund IRRs going back to 1993 (one observation) and includes at least 10 observations per vintage from 2006 onwards. Thereafter, the number of unique fund observations per vintage ranges between 8 in 2009 to 24 in 2016 with an average of 15 observations per vintage year.

What are the consequences of using such small samples to describe the empirical quartiles of a the underlying distribution of returns?

In a given vintage year, the true distribution of returns includes all the possible outcomes that a large number of hypothetical funds could have experienced. Assume for simplicity that the underlying distribution of infrastructure fund returns follows a normal law that does not change from one vintage year to another. Each data point in a manager-contributed dataset would correspond to an individual draw from this distribution, which is used to estimate the quartile boundaries of fund performance in that vintage year.

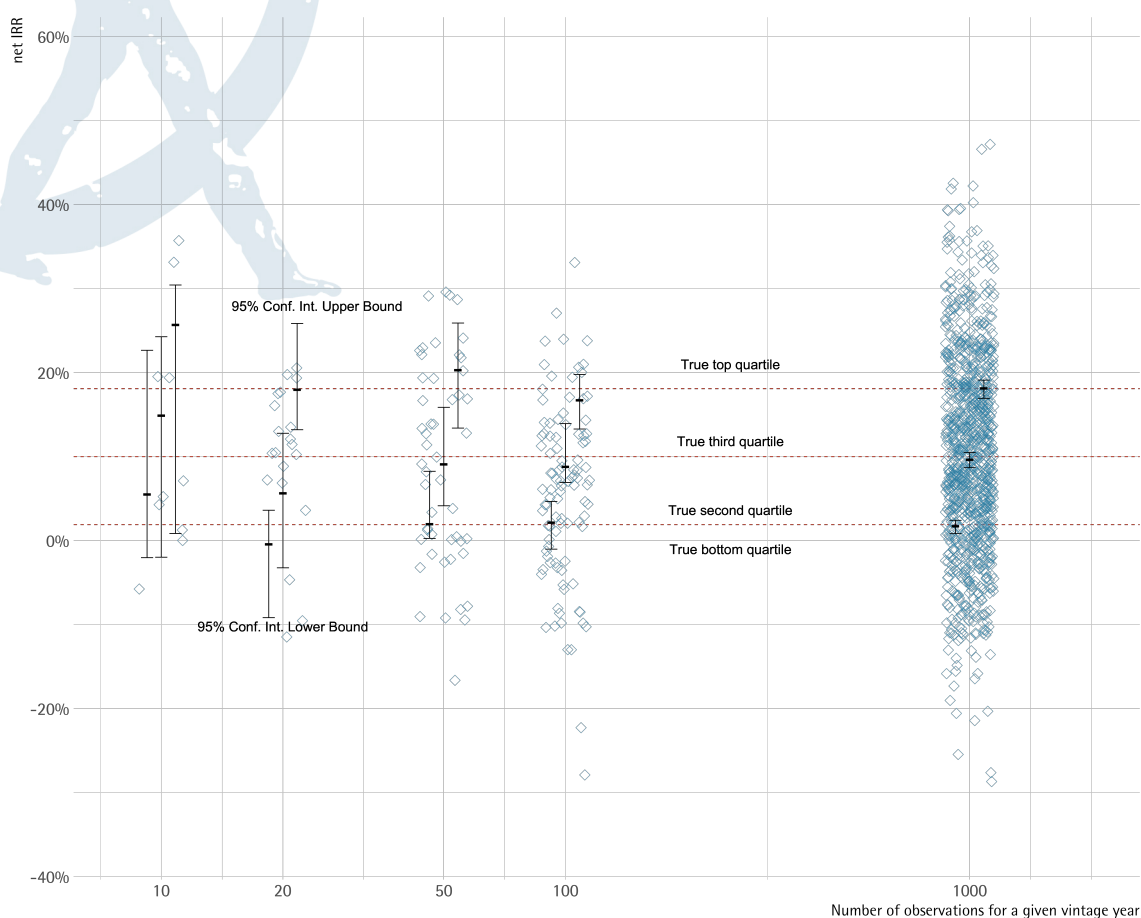
Table 1 shows the empirical quartiles for five rounds of drawing 10 observations from a well-

Table 1: Precision of IRR quarter estimates for five consecutive rounds: with 10 observations

True Value of bottom quartile: 1.9%, 2nd quartile: 10%, top quartile: 18.1%

Vintage	n	Quantile	Estimate	Error	Lower c.i.	Upper c.i.	95% Conf. range
Bottom quartile boundary estimates, five consecutive vintage years							
1	10	25%	9.6%	770 bp	-6.5%	18.5%	2500 bp
2	10	25%	-1.1%	-300 bp	-2.6%	6.3%	890 bp
3	10	25%	7.6%	570 bp	-5%	19.7%	2480 bp
4	10	25%	5.1%	320 bp	2.7%	15.8%	1310 bp
5	10	25%	3.9%	200 bp	-10%	13.5%	2360 bp
3rd/2nd quartile boundary estimates, five consecutive vintage years							
1	10	50%	16.9%	690 bp	-3.1%	25.6%	2870 bp
2	10	50%	15.7%	570 bp	6.8%	20.4%	1360 bp
3	10	50%	12.7%	270 bp	9.3%	29.6%	2030 bp
4	10	50%	6.1%	-390 bp	-1.5%	28.5%	3000 bp
5	10	50%	14.5%	450 bp	-1%	34.5%	3560 bp
Top quartile boundary estimates, five consecutive vintage years							
1	10	75%	11.8%	-630 bp	7.4%	21%	1360 bp
2	10	75%	20.8%	270 bp	12.9%	28.2%	1530 bp
3	10	75%	17.8%	-30 bp	10.2%	24.6%	1440 bp
4	10	75%	21.9%	380 bp	13.2%	34.4%	2120 bp
5	10	75%	17.6%	-50 bp	6.6%	24%	1740 bp

Figure 2: Illustrative example of estimating IRR quartile boundaries and 95% confidence intervals with small and larger datasets given a known underlying distribution of returns



Source:infraMetrics

defined normal law, that is, observing 10 IRR data points for five consecutive vintage years when we already know what the quartiles are. Table A2-A8 in the Appendix shows the same results for five consecutive rounds of making 20, 50, 100, 1000 and 10k observations. We see that obtaining 10 or 20 data points per year from the same well-defined distribution leads to completely different estimates of the quartiles boundaries from one year to the next.

With 10 observations per year, quarter estimates are off by several hundred basis points relative to their true value and their 95% confidence interval has a range of 1,000-3,000 basis points. With 20 or 50 observations of net IRR per year, the estimation error and the size of the confidence intervals are still so large that fund rankings can be considered random. For example, with 20 observations (table 9), the typical sample size available in the Preqin dataset, a fund with a 12% IRR which should be ranked in the second best quartile (since the true median is 10%) could still be considered bottom quartile because available observations can lead to a 300-400 basis point error in the estimation of the median IRR. Likewise, a fund with a 15% IRR could be considered a top quartile fund when the true top quartile boundary is 18%, but with too few observations the estimation error is easily large enough to miscategorise such funds.

Tables 12 and 13 show that increasing the number of observations can lead to considerably higher confidence in the estimation of quartile boundaries of net IRR. With 10,000 observations estimation errors are in the 0-30 basis points range and 95% confidence intervals have a range about 60 basis points. At this level of precision and robustness, ranking and comparing funds by quartile based on reported net IRRs becomes feasible and meaningful.

In other words, **when only 10 to 30 observations of IRRs or multiples are available in any given vintage year, the confidence with**

which investors can estimate quartile boundaries is so low that the resulting rankings of managers and funds are meaningless. The same fund may find itself switching from the top quartiles to the bottom quartiles simply because of a difference in contributed performance data in two consecutive years.

Figure 2 shows this result graphically: for a given return distribution, the three horizontal red lines indicate the true value of IRR quartile boundaries. The chart shows empirically derived quartile boundaries and their 95% confidence intervals for sample sizes going from 10 to 1,000.

Figure 2 shows that up to 50 observations, the estimation of the true quartile boundaries is still so imprecise that it is not possible to distinguish between quartiles with a reasonable level of confidence as the 95% confidence intervals of the quartile boundaries frequently overlap. From 100 observations, it is possible to distinguish between the quartile boundaries confidence intervals, but estimation errors remain large. With 1,000 observations, on the right of the chart, the true quartiles are approximated sufficiently closely and precisely to serve as a point of reference to rank investment funds.

Thus, using IRR and multiple quartiles to rank and select private fund managers is practically impossible until at least 1,000 (ideally 10k) observations are available to investors or consultants.

Clearly, there are not enough infrastructure funds in the world to report that much performance data, let alone in the same vintage year. Moreover, the true distribution of infrastructure fund returns is unlikely to follow a time-invariant normal law as in our example above. With non-normal and dynamic return distributions, even more observations may be needed to achieve robust estimates of IRR quartile boundaries.

We also see that aggregating contributed data over several consecutive years, as is sometimes

done, can only improve precision so much: with five years of data (c.100 observations), the 95% confidence interval of the quartiles boundaries is still very large for the purpose of ranking annual rates of returns (between 700 and 1,200 basis points).

Aggregating all the data available in the Preqin dataset since 1993 yields more than 200 observations. This more precise but also backward-looking and not helpful for investors to rank funds today. Indeed, the underlying distribution of returns is unlikely to be static over time. This is a well-known phenomenon for investors who have experienced the yield compression of infrastructure of the past two decades. The type of infrastructure funds available and their strategies have also changed over the past 20 years, and so has the variance (the risk) of their return distribution. Hence using a rolling average or even all previously reported returns creates a backward-looking (in this case, upward) bias in quartile boundary estimates.

Finally, beyond the central issue of sample size for quartile boundary estimation and robustness, individual outliers can have a significant impact when samples are very small, as is the case for contributed infrastructure fund performance data. If even one of the 10 or 20 reported IRR data points in a given vintage year is a significant outlier, adding or removing this single observation completely changes quantile estimation results as shown in table 2. As far as we are aware, quarter rankings currently produced using contributed data do not include any systematic outlier treatment. In fact, with so little data, it is not possible to estimate the parameters of the distribution of returns, thus it is not very clear how to define outliers either. In table 2, we remove a single data points in two vintages which exhibit one IRR above 50%. As the table show, the presence of this single data point completely changes the estimates of IRR quartiles.

In conclusion, with sparse contributed fund performance data, investors cannot know with a reasonable degree of confidence what the quartile of the net IRR or net multiple distributions really are. Since quartile rankings are an important part of the private fund investment decision and monitoring process, in the case of infrastructure funds, investors (LPs) and managers (GPs) do not have any useable point of reference to evaluate performance.

Moreover, data paucity gets worse as one tries to drill down to the subsegments of the infrastructure universe, making any comparisons or assessments between styles and fund risk profiles also impossible.

2.2 Step 2: Simulating private funds achieves superior quartile estimation results

Figure 2 above shows that with a large number of observations of fund performance, reliable estimates of the quartile boundaries can be obtained. In the absence of enough observable data, such a high number of observations can only be achieved through simulation. The solution to the issue described above is to 'bootstrap' the estimation of returns and multiples quartile boundaries by creating a large sample of possible investment outcomes with at least 1,000 data points per vintage or segment.

Simulating private fund returns has been suggested before in industry research as an approach that can palliate for the lack of observable data described above (see for example Cornel, 2017, for BlackRock's approach to private equity funds).

For these results to be relevant to investors, two important conditions must be met:

- The valuations of the assets purchased and sold by the simulated funds need to represent the market values of the underlying assets at

Table 2: Differences between empirical net IRR quartiles of contributed data before and after removing a single positive outlier, vintage years 2009 and 2018

	Vintage Year	Obs.	Bottom quartile (25% quantile)	2nd Quartile (Median)	Top Quartile (75% quantile)	Max Value
Raw data	2009	9	0.075%	7.10%	10.11%	448%
w/o outlier*		8	-4.05%	5.20%	9.04%	13.2%
Difference		1	-412.5bp	-190bp	-107bp	-4340bp
Raw data	2018	19	7.25%	9.80%	16.40%	55.2%
w/o outlier*		18	7.02%	7.44%	13.40%	23%
Difference		1	-22.5bp	-36bp	-300bp	-322bp

* removing values above 50% net IRR. Source: Preqin, EDHECinfra

the time, as well as the distributions made by each invested company in any given year;

- The simulated behaviour of the funds should correspond to the typical path followed by such investment vehicles.

We build our approach using the market value of hundreds of infrastructure investments and their dividend cash flows obtained from the infraMetrics® database. This data presents a number of advantages that make it the best available input for a simulation exercise:

- The infraMetrics database of 700+ tracked unlisted infrastructure firms is designed to track a broader investible universe of 7,000+ firms in 25 countries in a representative manner. Hence, it does not suffer from reporting and survivorship biases.
- infraMetrics valuations are re-calibrated each quarter to match the most current level of expected returns, including the unlisted infrastructure equity risk premia and the relevant yield curve (see Blanc-Brude and Gupta, 2021, for a summary).

Next, our fund simulations rely on a number of assumptions about the investment pace and size of private infrastructure funds, fee levels, ability to deploy capital, etc (see below and in the Appendix). We develop these assumptions through a combination of desk-based research and semi-structured interviews with investment managers. Market participants were also involved in the validation and 'reality check' of the

simulation results (see acknowledgments in the executive summary).

We simulate the path of thousands of potential funds using this data, in each vintage and in each segment of the market.

It should be noted that such simulations aim to represent all the standard or typical outcomes of the activity of private investment funds i.e. some extreme cases will, by definition, fall outside the scope of the simulations. For example, a fund manager may have a legal dispute with its LPs and terminate the fund. Another fund may use financial engineering to achieve extremely high returns, or just be very lucky. Such scenarios produce outliers in the distribution of fund returns and are outside the scope of the simulation. They are also irrelevant since our objective is to estimate accurate and robust quartile boundaries i.e. outliers should not make any difference to the results. As opposed to the small contributed datasets, with a large number of observations, outliers do not have any significant impact on quartile boundary estimation.

In summary, simulations allow us to produce much more robust results than simply observing contributed data:

- No selection bias: By covering a large number of possible simulated paths across the entire universe of infrastructure, this technique ensures that the fund benchmarks are free from the selection biases that is typically seen in those contributed benchmarks that rely on the reported track records of a subset of GPs

and LPs;

- No survivorship bias: With contributed data, poor-performing GPs may stop contributing, thus making the benchmarks biased towards 'survivors'. With iFSA, by design some of the simulated paths result in funds investing in poorly performing assets and/or failing to deploy capital, especially in periods of market stress. Such 'zombie' funds remain in the iFSA dataset making the results free of survivorship bias;
- Robust quantile estimates: contributed dataset are found to be very limited in individual vintages or subsegments but a bootstrapping (resampling) approach enables as much data as is needed to be generated to ensure the robustness of the results;
- Granularity of fund strategies: As a result, data is available for granular strategies, such as core funds or renewable energy funds etc, across the entire fund lifecycle (J-curve) and in multiple geographies;
- Up-to-date data: contributed benchmarks exhibit reporting lags and the timing of reporting may differ between contributors as well, resulting in blended data that is not strictly comparable. infraMetrics produces results at T+10 from the end of every quarter end, thus, ensuring that investors have access to up-to-date information.

2.3 Step 3: Back-testing

2.3.1 Comparing aggregate datasets

We first compare the infraMetrics net IRR fund simulation results and the Preqin dataset on an aggregate basis, for the period 2005–2018 (Preqin data is not available beyond the 2018 vintage).

As discussed above, this creates a backward-looking bias that precludes using such results for the purpose of benchmarking funds today,

but this bias is common to both datasets and with 200+ data points the Preqin quartile boundaries are now more accurate albeit still biased, as shown above.

Table 3 shows descriptive statistics for the two datasets and figure 3 shows a scatterplot of the data and the estimates and confidence intervals of the quartile boundaries. We see that while contributed data tends to have higher mean and quartile boundaries, the two datasets correspond to the same range of IRR values, with the exception of one large outlier in the contributed dataset. However, simulated quartile boundaries are much more precise due to the number of data points available, as shown on figure 3, where the confidence interval whiskers are very small for the simulated data.

It is worth noting that simulated results also fall within the confidence interval of contributed data points, as shown by the horizontal dashed lines on figure 3. Thus, the largest available sample of contributed data agrees with the simulation results about the overall distribution of the data taken in aggregate over 13 vintage years. This is a first validation of the ability of simulation to generate 'market-like' results.

This is also evident looking at figure 4, which shows how, in aggregate and over many years, both contributed and simulated datasets tend to tell the same story.

We see that in a few cases, contributed datapoints exhibit very high returns compared with simulated ones. This can be the result of managers using strategies that are riskier than what is implied by the simulation such as investing in private equity-style assets or using fund-level or holdCo-level leverage to boost returns. Such manager-specific decisions are not meant to be captured by the simulation, which only reflects the path of thousands of 'standard' funds, but does not, for example, include assumptions about the use of additional

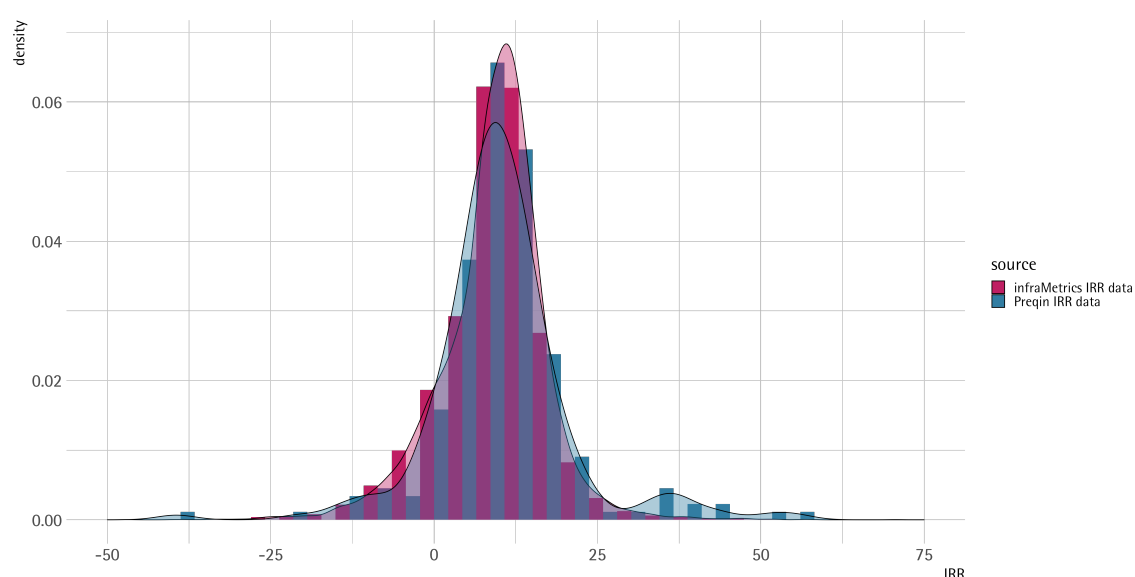
Table 3: Descriptive statistics of the infraMetrics and Preqin datasets of net infrastructure fund IRR, 2005-2018 vintages

	Mean	Min	Q1	Q2	Q3	Max	Obs.
InfraMetrics net IRR	8.84	-50.2	5.2	9.8	13.5	134.5	13,993
Preqin net IRR	12.33	-39.5	5.6	9.1	14.0	448.0	206

Figure 3: Scatter-plot of infraMetrics and Preqin net IRR datasets, quartile boundaries and confidence intervals, aggregates for 2005-2018 vintages



Figure 4: Probability density plots of infraMetrics and Preqin net IRR datasets, aggregates for 2005-2018 vintages



Source: EDHECinfra, Preqin. Note that the data on the plot has been rescaled to compare densities directly because the infraMetrics dataset includes 38 times more data, making a regular histogram difficult to read.

Figure 5: Scatter-plot of infraMetrics (1,000 obs.) and Preqin net IRR datasets, quartile boundaries and confidence intervals ('whiskers') for the 2010 and 2011 vintages



Figure 6: Scatter-plot of infraMetrics (1,000 obs.) and Preqin net IRR datasets, quartile boundaries and confidence intervals ('whiskers') for the 2015 and 2018 vintages



Source: EDHECinfra, Preqin

leverage, style drift or any other reason why individual funds might not fall within the distribution of returns of infrastructure funds.

Hence, while the aggregate distribution of net IRRs in the contributed and simulated cases are different, they are still close enough to suggest that the simulation is capable of producing market-like results.

2.3.2 Comparing individual vintage years

This comparison is, however, only valid at the aggregate level. When comparing contributed and simulated data in individual vintage years, which is what investors and asset managers really need to do, the two datasets are very different.

Figures 5 and 6 show scatter-plots and quartile boundary estimates for simulated vs contributed data in individual fund vintage years. As before in our example above, the very limited number of contributed data points within individual vintages leads to highly uncertain quartile estimates that often overlap between quartiles.

In other words, the true quartiles are unknown when using only contributed data. The same conclusion applies when looking at contributed data by strategy or sector within vintage years or even across multiple vintage years: contributed data is too scarce and quartiles are unknowns or random.

Conversely, simulated quartiles continue to be robust and precise within vintages as well as within strategies or sectors.

Figure 6 also shows examples of biases in the contributed data of more recent vintage years: simulated results include data for funds in their early development which often exhibit negative returns because of the so-called J-curve. These funds do not contribute data to the Preqin database which only reports positive net IRRs for the 2015 or 2018 vintages, leading to quartile boundary estimates that are biased upwards since

only funds that are able to generate positive returns in these vintages report their data.

Investors have to wait for a number of years for the performance of more funds to be reported as shown in figure 5 for older vintages like 2010 and 2011. Still, even in these older vintages we see that limited data availability leads to very imprecise quartile boundary estimation and meaningless fund or manager quartile rankings.

In conclusion, simulated results are both congruent with contributed data at the aggregate level over a long period, and more robust and precise at the vintage year or sub-segment level. Alignment of the results with market data is simply due to the use of market valuations and realised asset-level cash flows as the inputs of a bottom up simulation.

Next, we describe the simulation methodology used to obtain these results.

3. Benchmark Design Methodology

The methodology relies on a Monte Carlo simulation of bottom-up data on infrastructure companies along with some assumptions on the investing behaviour of a fund. Each simulation run can then be broadly divided into four sections: defining the fund characteristics, selecting the underlying investments, computing the fund cash flows and calculating the fund metrics such as IRR and TVPI. This process is explained in figure 7.

3.1 Fund characteristics

This step lays the foundation of the approach. Based on the data from actual fund prospectus and other industry consultations, we define the following characteristics of a fund:

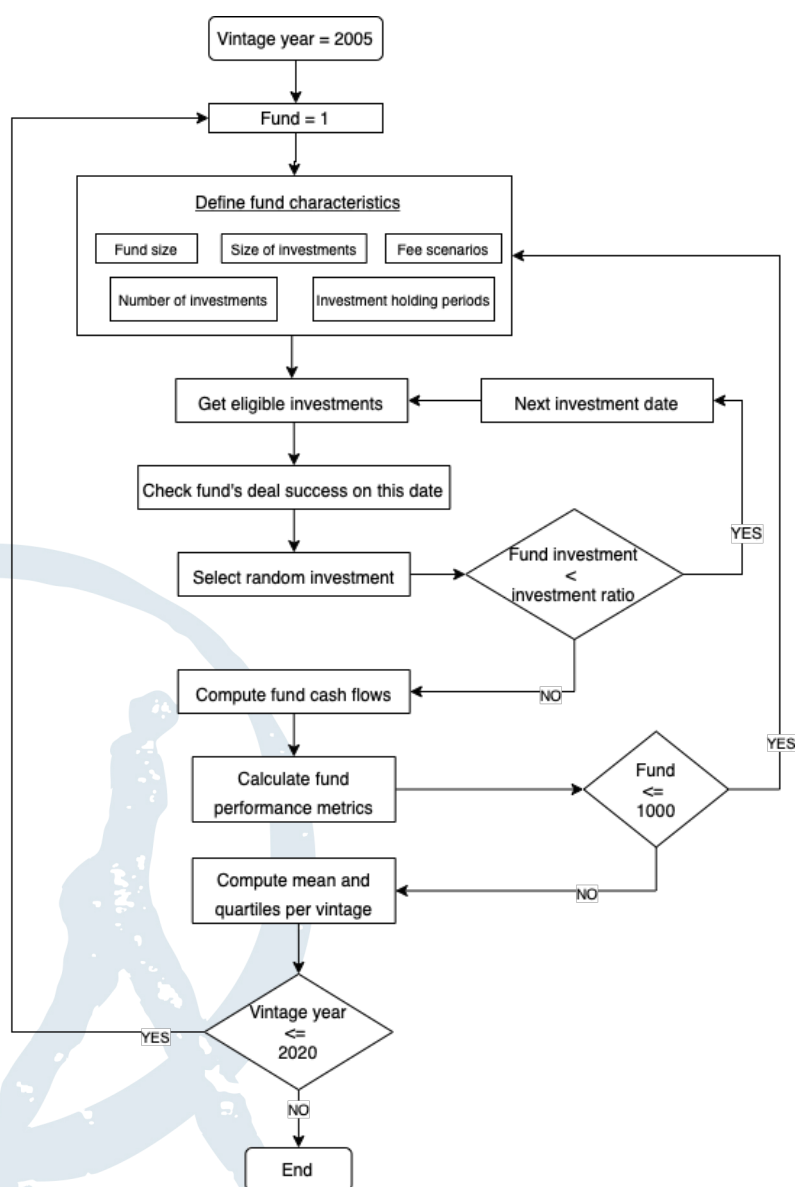
- **Fund size:** With the ever-growing investor interest in the unlisted infrastructure asset class, average fund size has ballooned from USD200m in 2000 to USD1bn in 2019. In this simulation, we have assumed fund size to be distributed between USD100m to USD2bn, with some custom probabilities, such that the average tracks this historical evolution. The probability assumptions for the fund size are detailed in the Appendix.
- **Number of investments in a fund:** A closed-end infrastructure fund typically invests in 5-20 deals, so we make an ex-ante assumption of a uniform distribution over this range. The final number of deals is also impacted by the market activity in any given investment year.
- **Holding period of an investment:** Holding periods typically depend on the market conditions, investment performance etc. However, after several rounds of consultations with industry participants, we have assumed a uniform distribution between four and eight years.

- **Deal success rate:** For any given investment year, we assume a deal success probability depending on market activity. This determines which funds are eligible to make an investment at any given time. This data is calibrated based on the historical number of deals/number of funds ratio (these probabilities are available in the appendix). It can be seen that in the crisis periods, there is less opportunity to make new deals which leads to an increase in dry powder in some funds, for e.g., the deal activity had dropped by about 90% in the 2008-09 period.
- **Investment size:** We assume that capital is equally deployed (at the price given by prevailing NAV) to all the randomly selected companies in the fund.
- **Fees:** In closed-end fund structures, fee has several components and also various computation methods. In this solution, we pre-compute the net of fees performance metrics of all the funds using the following six different scenarios and allow our users to choose as per their needs. Performance fees are computed at the fund level and includes a 100% catch-up provision for GPs. There is no second close but this will be included in a future version of the tool.

3.2 Selection of underlying investments

EDHECinfra has collected the financial accounts of more than 700 companies since 2000. These includes the crucial information of dividends paid by an operating company to its equity investors. Furthermore, our asset pricing models allow us to value each of these companies in our universe on a quarterly or monthly basis. Translating these data points into fund terminology, we have historical information of NAV and distributions of hundreds of unlisted infrastructure companies in

Figure 7: Fund simulation algorithm (e.g. Vintage 2005)



Source: infraMetrics®

Table 4: Fee scenarios used in the iFSA simulations

Management fees	Performance fees	Hurdle rate
2.00%	20%	8%
1.50%	20%	8%
1.50%	20%	6%
1.50%	15%	8%
1.25%	15%	8%
1.00%	10%	6%

the 25 most active markets in the world, going back 20 years.

With this dataset, we shortlist the companies eligible for investment given the fund's strategy, vintage and other characteristics. We also check whether the fund is eligible to make a deal on that investment date based on a deal success rate assumption. Out of the possible options, we then randomly invest in one company, which becomes unavailable for investment for the rest of the investment period. This process is followed until the fund has invested up to the investment ratio or the fund is abandoned.

3.3 Computing fund cash flows

Every time a fund makes an investment, we draw down capital from the fund corresponding to the NAV of the underlying company multiplied by the investing fund's stake. We also rescale any distributions paid back from the company and the NAV by the fund's stake. This gives us the cash-flows and NAV data for the fund's investment in any selected company.

Finally, we pool all the cash flows and NAV data at the investment level, to obtain quarterly fund cash flows and NAV gross of any fees.

$$\begin{aligned} DD_t &= \sum_n dd_t \\ D_t &= \sum_n d_t \\ NAV_t &= \sum_n nav_t \end{aligned}$$

Where, dd_t , d_t and nav_t are the drawdown, distribution and net asset value at the investment level respectively at time t , n is the number of investments in the fund, DD_t , D_t and NAV_t are the drawdown, distribution and net asset value at the fund level respectively at time t

We then apply multiple fee scenarios, as described above, to compute the fund level cash flows

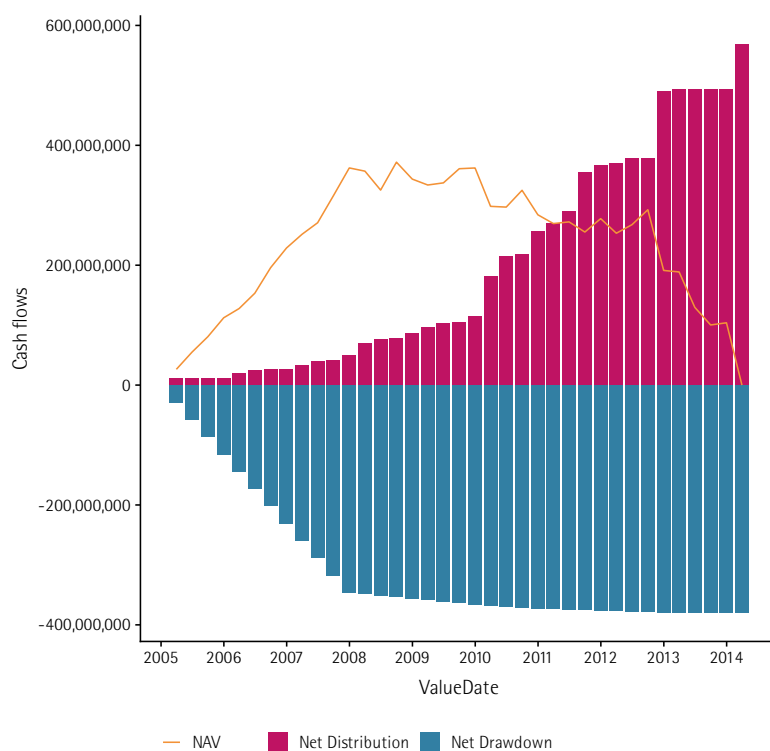
net of all fees. \widehat{DD}_t is the net drawdown from investors which includes the quarterly payment of management fees in addition to the capital drawn by underlying investments and \widehat{D}_t is the net distribution to the investors after the payment of performance fees to the manager.

We then apply multiple fee scenarios, as described above, to compute the fund level cash flows net of all fees. \widehat{DD}_t is the net drawdown from investors which includes the quarterly payment of management fees in addition to the capital drawn by underlying investments and \widehat{D}_t is the net distribution to the investors after the payment of performance fees to the manager.

Figure 8 shows the net of fees cash flows of a sample fund from the 2005 simulations. For the first few years of investment, the fund draws down capital as it continues to invest in underlying assets. During this period, NAV grows and there are limited distributions paid back to investors. Following this period, the fund starts to exit the investments resulting in increased distributions and gradual decrease in NAV, before it is fully liquidated.

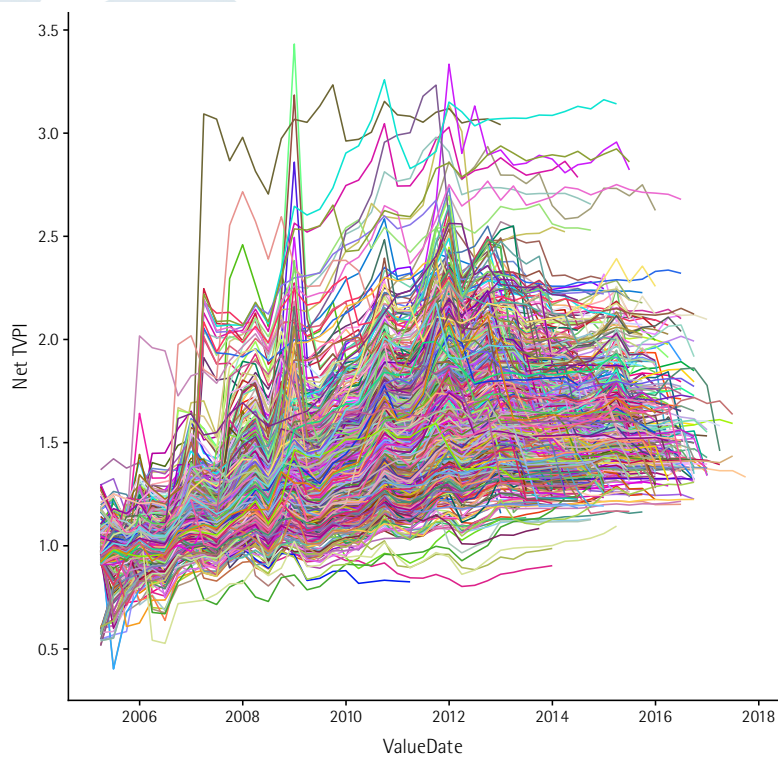
Figure 9 shows all the possible paths that different 2005 vintage funds take in the simulation. There is a wide range of final TVPI ranging anywhere from less than 1x to more than 2.5x. There are also some funds which invest in poor-performing assets and suffer up to 50% write-downs in the initial period.

Figure 8: Cash flows of a sample fund from all the vintage 2005 simulation



Source:

Figure 9: Vintage 2005 fund simulations



Source:

4. Fund Strategy Benchmark Metrics

Private market funds often have unique characteristics such as irregular timing of cash flows, size of those cash flows, etc. Due to these, return calculations and benchmarking methodologies often differ from those used in public markets. Consequently, a number of fund-style metrics have been implemented in the Fund Strategy Analyser to give infrastructure investors an intuitive understanding of the performance reported for various infrastructure strategies and segments and enable them to use these results directly against the data reported by infrastructure investment fund managers.

Analysing the performance of private funds from the point of view of LPs requires taking fees into account. iFSA returns metrics are both gross and net of fees. Indeed, fees can make a significant difference in the reported performance, particularly the fee drag of long holding periods or when portfolio distributes larger amount in earlier periods.

In this section we, we cover the calculation methodology of various performance metrics and also describe some of the results for selected fund strategies.

4.1 IRR

The internal rate of return (IRR) is one of the most popular metrics used for closed-end funds. It is the rate at which the net present value of negative cash flow equals the net present value of positive cash flow.

$$NPV = \sum_{t=1}^T \frac{(\widehat{DD}_t + \widehat{D}_t)}{(1 + IRR)^t} + \frac{NAV_T}{(1 + IRR)^T} = 0$$

It is a dollar-weighted calculation which is considered most appropriate for assessing private closed-end fund managers because it holds the manager responsible for both the amount and timing of the investment, since the limited partners have no control over when the capital is called or distributed after the initial commitment.

The most direct way to benchmark a private fund is to compare its IRR against that of other funds in the same vintage year.

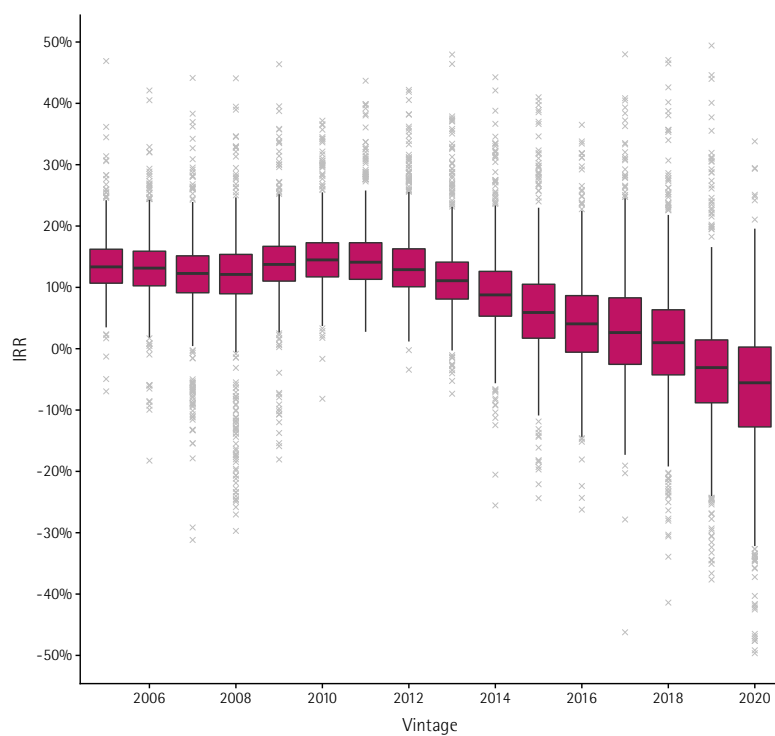
Figure 10 shows the distribution of IRR by vintage years for global infrastructure funds using the following fee scenario: management fees 1.5%, performance fees of 15% and a hurdle rate of 8%.

Note that more mature vintages (2005-2012) generally have the higher returns (mean fund return is higher than 10%). There are also some very low return funds in the vintage years just preceding the financial crisis as a lot of managers could not deploy capital or invested in infrastructure asset that performed poorly either during or after the GFC. In more recent vintages (2017-2020), lower mean returns are due to the J-curve effect (the impact of fees is greater than that of the valuation uplift) as well as the impact of Covid-19 from 2020 onwards.

We can also use quartile ranks to compare different segments of the infrastructure market: figure 11 shows a comparison between IRR between Core and Core+ strategies by vintage year. We see that Mature Core funds have an average net IRR of 10% which is approximately 5% lower than the Core+ funds, a persistent trend throughout the vintage years.

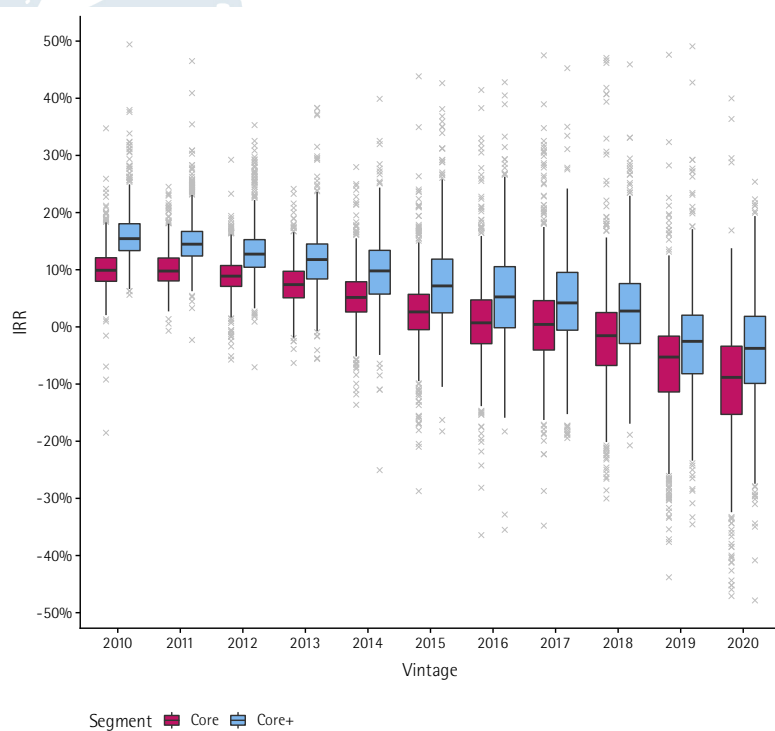
Similarly, figure 12 compares the net IRR between renewables and transport sector funds by vintage

Figure 10: Distribution of fund IRR by vintage year for global infrastructure strategy



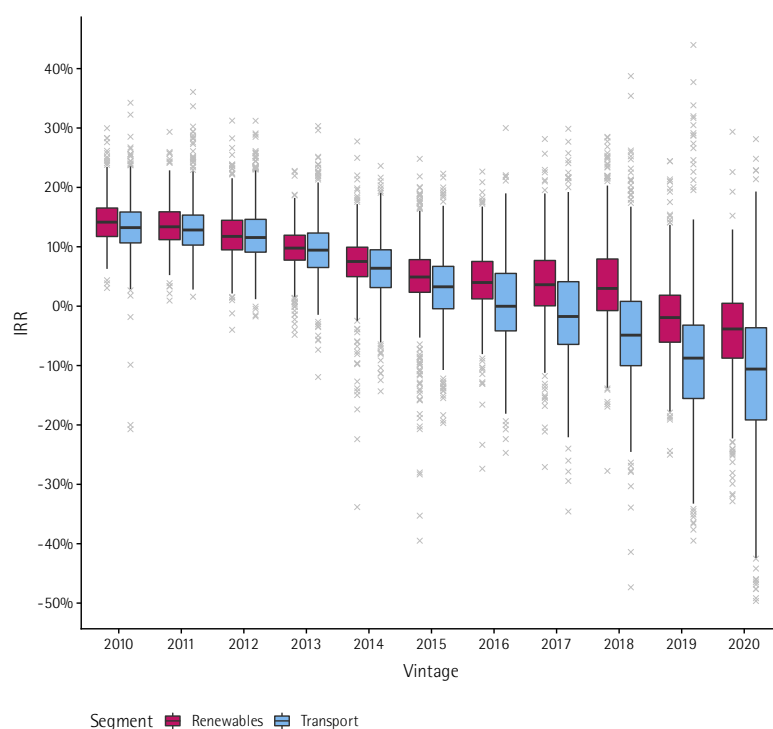
Source: infraMetrics®

Figure 11: Comparison of fund IRR by vintage year in Core and Core+ strategies



Source: infraMetrics®

Figure 12: Comparison of fund IRR by vintage year in Renewables and Transport sectors



Source: infraMetrics®

year: While the average return in mature funds was close in the two sectors, in the more recent vintages, renewable energy funds outperform transport sector funds, in part due to the Covid-19 pandemic and its impact on transport.

4.2 TVPI

Total value to paid-in (TVPI) ratio reflects the valuation multiple (realised or expected) of an investment. It is calculated by dividing the fund's cumulative net distributions to the investors (after the performance fees) and residual value by the paid-in capital (includes drawdowns for investments and management fees).

$$TVPI_T = \frac{(\sum_{t=1}^T \hat{D}_t + NAV_T)}{\sum_{t=1}^T \widehat{DD}_t}$$

Multiples ignore the time value of money and offer a quick and readily digestible means of indicating the performance of a fund. Any multiple above 1x shows that the fund has returned (or would return) more than its initial investment to its investors.

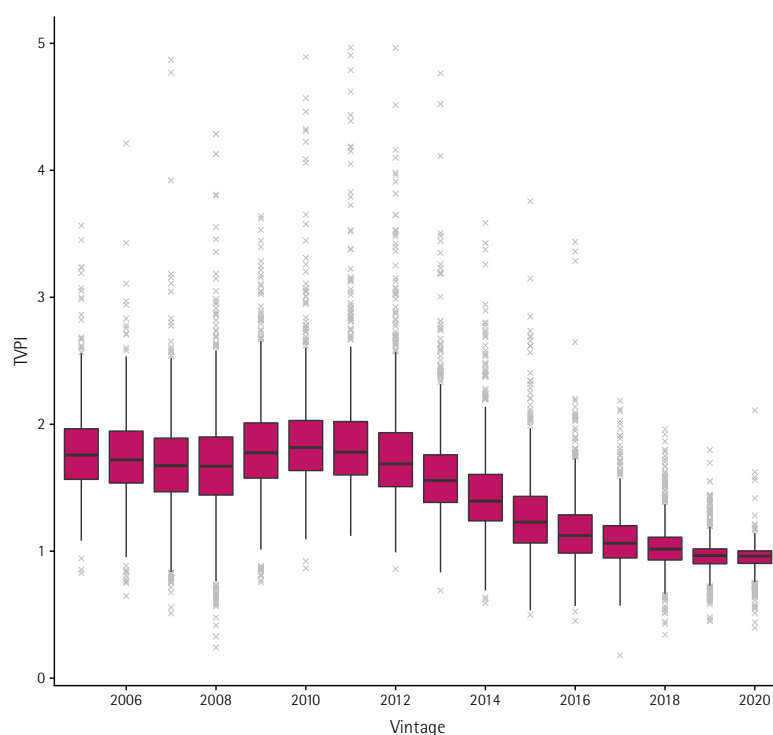
Figure 13 shows the distribution of TVPI by vintage years for global infrastructure funds. Note a similar trends as for fund IRRs, with an average TVPI of mature vintages at about 1.75x, implying that, in aggregate, funds returned 75% more than the initial investment over their full lifetime. The recent vintage year funds have an average net TVPI of a bit under 1x, indicating the effect of fees on the return and causing the J-curve effect.

4.3 DPI

Distributions to paid-in (DPI) ratio, also known as the realisation multiple, is calculated by dividing the cumulative net distributions to the investors (after the performance fees) by the paid-in capital (includes drawdowns for investments and management fees). It gives an insight into how much of the fund's return as actually been "realised" or paid out to investors.

$$DPI_T = \frac{\sum_{t=1}^T \hat{D}_t}{\sum_{t=1}^T \widehat{DD}_t}$$

Figure 13: Distribution of fund TVPI by vintage year for global infrastructure strategy



Source: infraMetrics®

4.4 RVPI

Residual value to paid-in (RVPI) ratio is calculated by dividing the market value of unrealised investments by the paid-in capital (includes drawdowns for investments and management fees). It provides a measurement of how much of the fund's return is unrealised and dependent on the market value of its investments.

$$RVPI_T = \frac{NAV_T}{\sum_{t=1}^T \widehat{DD}_t}$$

4.5 PME

Public market equivalent (PME) refers to measures of relative performance of private funds using a public index or portfolio as the benchmark. Unlisted infrastructure fund returns are not directly comparable to those of public markets, due to the asset class's illiquid nature and irregular timing of cash flows. PMEs typically consist of 'investing' the cash flows of the funds (ex post facto) into an public index in order to benchmark the performance of a fund, or a group of funds, against a public market index

while accounting for the timings of the fund cash flows. We consider two types of PMEs:

- **KS PME (Kaplan Schoar)** – The Kaplan Schoar PME (Kaplan and Schoar, 2005) measures the wealth multiple effect of investing in the private infrastructure fund versus the index. It represents the public market-adjusted equivalent to the traditional TVPI. It is calculated by compounding each fund cash flow – both capital calls and distributions – based on index performance between the date of the cash flow and the valuation date. If the KS PME is greater than 1, the private fund outperformed the public market index.

$$KS \text{ PME} = (\text{Sum of future value distributions} + NAV) / \text{Sum of future value capital calls}$$

- **Direct alpha** – This is the most recent method to compare private market returns against the public market and uses a methodology similar to that of KS PME (Gredil et al., 2014). The key difference is that direct alpha quantifies the out/underperformance of the private fund by calculating the IRR of the compounded cash flows plus fund NAV.

Table 5: Distribution of direct alpha in global infrastructure funds by vintage years

Vintage	Mean	Bottom quartile	Median	Top quartile
2005	10.9%	8.0%	10.5%	13.4%
2006	10.8%	7.4%	10.5%	13.5%
2007	8.1%	4.7%	7.9%	11.5%
2008	3.3%	1.0%	3.7%	6.9%
2009	1.8%	-0.6%	1.9%	4.5%
2010	4.2%	1.6%	4.0%	6.4%
2011	4.5%	1.6%	3.9%	6.6%
2012	3.6%	0.4%	2.8%	5.9%
2013	1.4%	-1.8%	1.1%	4.0%
2014	-1.4%	-5.2%	-1.7%	1.8%
2015	-5.7%	-10.3%	-6.2%	-1.9%
2016	-10.1%	-15.3%	-10.4%	-5.6%
2017	-13.2%	-19.5%	-13.3%	-7.8%
2018	-18.2%	-24.1%	-18.0%	-12.2%
2019	-27.4%	-33.1%	-26.9%	-21.8%
2020	-44.3%	-50.5%	-42.3%	-36.3%

In infraMetrics, we compute a MS-PME and a direct alpha measure against a broad developed equity market index (The Scientific Beta Developed Equity Index).

We also build a similar computation using a private infrastructure index: the infra300® index. While this is not a public market index, it provides a useful comparison of various fund strategies against a broad market index of unlisted infrastructure equity.

Table 5 shows the distribution of direct alpha in global infrastructure funds by vintage years measured against a public developed equity market index. We note a declining trend in alpha which, on aggregate, lasts until the vintage year 2013 funds. The funds started thereafter have, on average, underperformed the equity markets.

5. Example Use Cases

In this section, we consider two simple examples of how iFSA can be used by investors in unlisted infrastructure funds to achieve better investment outcomes.

5.1 Manager selection

Private infrastructure closed-end funds, similar to those of private equity, are created by fund managers or GPs (General Partners) that raise commitments from investors or LPs (Limited Partners). The capital is committed by LPs at the fund inception for a period of 10 to 12 years. Once fundraising is completed or the fund is 'closed', the GP typically has full discretion and responsibility for the investment decisions made during the fund's life. Indeed, such funds are sometimes called "blind-pools". Manager selection is therefore a central aspect of the investment decision taken by LPs.

The due diligence process to select the best possible manager differs across LPs and includes both qualitative and quantitative factors. Invariably, one of those factors is the past performance of the manager's funds.

In the absence of market benchmarks and given the tendency to use absolute return benchmarks, a frequent approach consists of ranking managers by the historical performance of their funds and look for signs of out-performance and persistence. Thus, a 'top quartile' manager would tend to create funds that frequently if not always find themselves in the top 25% of realised performance, consistently outperforming peers and equivalent strategies due to their skills, knowledge or access to the market for unlisted infrastructure assets.

Next, we look at the actual realised historical track record of nine funds managed by four different

managers as of 2021 June (source: Preqin). The funds were created between 2010 and 2018 in the Core, Core+ and Opportunistic strategies. The fund benchmark need to be of the same vintage as the fund's for a fair performance comparison, given the J-curve, opportunity set and market conditions at the time. In what follows for each of these nine funds, use the benchmark for their corresponding vintage year.

Tables 6 and 7 show the realised performance metrics of actual funds and their quartile ranking using the Preqin quartile ranks and the infra-Metrics quartile ranks, respectively.

Table 6 shows that:

1. The sample is very thin and inconsistent across vintages. With such a small dataset, inclusion (or exclusion) of even one fund dramatically changes quartile ranks;
2. The only available benchmark is 'all infrastructure' since there is even less data for specific fund strategies. This implies an inherent bias in the comparison exercise since a Core fund will now be compared against the same benchmark as an Opportunistic fund;
3. The data is stale (2 quarters) which can to inaccurate conclusions about the performance in period of crisis such as this one (Covid-19 pandemic).

Table 7 shows the quartile ranking using the iFSA benchmarks. Each strategy and vintage includes a sample of 1,000 simulated funds, allowing for robust and granular quartile estimates since we use strategy-specific benchmarks, making the comparison more fair for individual funds. While clearly more robust, this benchmark also

Table 6: Track record of four infrastructure fund managers as of 2021 June and their quartile ranking using Preqin, data as of 2020 December

Track record of 5 managers as of 2021 Jun						Preqin benchmark as of 2020 Dec			
Manager	Fund	Strategy	Vtg	Net IRR	Net TVPI	Benchmark	# funds	IRR Q-rank	TVPI Q-rank
Manager 1	Fund 1	Core	2010	14.1%	1.76	Infra-All	13	2	2
Manager 1	Fund 2	Core plus	2015	11.1%	1.39	Infra-All	22	2	1
Manager 2	Fund 1	Opportunistic	2012	10.5%	1.45	Infra-All	13	n/a	2
Manager 2	Fund 2	Opportunistic	2015	12.4%	1.27	Infra-All	22	2	2
Manager 3	Fund 1	Core plus	2011	9.3%	1.56	Infra-All	13	2	1
Manager 3	Fund 2	Core plus	2015	14.5%	1.43	Infra-All	22	2	1
Manager 4	Fund 1	Opportunistic	2013	17.6%	1.60	Infra-All	17	1	1
Manager 4	Fund 2	Opportunistic	2016	19.0%	1.50	Infra-All	26	1	1
Manager 4	Fund 3	Opportunistic	2018	6.0%	1.10	Infra-All	23	n/a	3

Source: Preqin, EDHECinfra

Table 7: Track record of four infrastructure fund managers as of 2021 June compared with infraMetrics® Fund Strategy Analyser (iFSA) as of 2021 June

Track record of five managers as of 2021 Jun						iFSA benchmark as of 2020 Dec			
Manager	Fund	Strategy	Vtg	Net IRR	Net TVPI	Benchmark	# funds	IRR Q-rank	TVPI Q-rank
Manager 1	Fund 1	Core	2010	14.1%	1.76	Core	1,000	1	1
Manager 1	Fund 2	Core plus	2015	11.1%	1.39	Core+	1,000	1	2
Manager 2	Fund 1	Opportunistic	2012	10.5%	1.45	Opportunistic	1,000	4	4
Manager 2	Fund 2	Opportunistic	2015	12.4%	1.27	Opportunistic	1,000	2	3
Manager 3	Fund 1	Core plus	2011	9.3%	1.56	Core+	1,000	4	3
Manager 3	Fund 2	Core plus	2015	14.5%	1.43	Core+	1,000	1	2
Manager 4	Fund 1	Opportunistic	2013	17.6%	1.60	Opportunistic	1,000	2	4
Manager 4	Fund 2	Opportunistic	2016	19.0%	1.50	Opportunistic	1,000	1	1
Manager 4	Fund 3	Opportunistic	2018	6.0%	1.10	Opportunistic	1,000	2	2

Source: Preqin, EDHECinfra

highlights the possibility of Type I (false negative) or Type II errors (false positive) with using small sample set where a fund may appear to be in the bottom/top quartile when it actually is not. Thus:

1. Manager 1: the quartile rank for two funds using Preqin's IRR data (Table 6) is '2', whereas the iFSA benchmark (Table 7) place these funds in the top quartile. This is the case of Type I error (false negative). LPs making this error miss the opportunity to invest with a good manager, while managers find themselves unable to showcase their skills;
2. Manager 3 is a case of Type II error (false positive): against the contributed TVPI data in table 6, the two funds of this manager would be in the top quartile, whereas, benchmarked iFSA data (table 7), places the two funds in the 3rd and 2nd quartiles respectively. LPs making such errors would unknowingly select a poor-performing manager.

Thus, using robust and granular fund performance benchmarks is the only way for LPs to select managers on the basis of their relative quartile rankings. Likewise, such data is necessary for GPs to showcase their skills and compare their performance against a fair and robust benchmark.

5.2 Fund performance monitoring

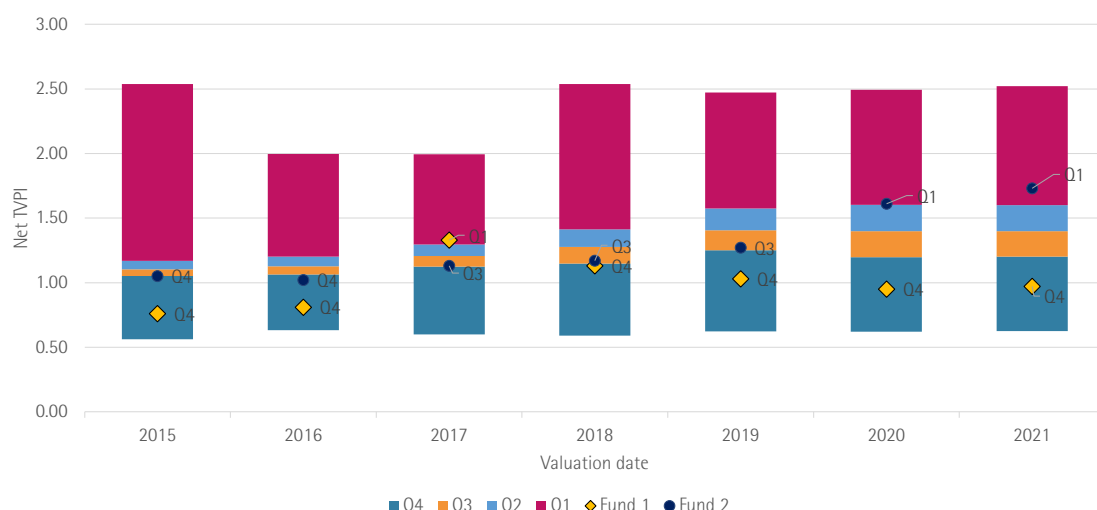
LPs need to monitor the performance of their fund investments on an ongoing basis for numerous reasons: position keeping, risk management, ongoing manager selection, etc. LPs thus need to receive useable performance metrics from fund managers, including understanding the return drivers of each fund and how the GP is performing vs the market and its peers. Indeed, funds with very similar ex ante characteristics (size, announced strategy, geography, etc.) could perform very differently depending on the investment decisions made by the manager, especially since individual funds are usually limited to just a few investments (typically

Table 8: Characteristics of two private infrastructure equity closed-end funds

	Fund 1	Fund 2
Strategy	Core+	Core+
Vintage	2014	2014
Fund size	>1bn	>2bn
Management fees	2%	2%
Performance fees	20%	20%
Hurdle rate	8%	8%

Source: Preqin, EDHECInfra

Figure 14: Annual performance comparison of two 2014 vintage funds against the annual iFSA quartiles of vintage 2014



Source: Preqin, EDHECInfra

less than a dozen), making active investment choices highly relevant to monitor.

Next, we take the example of two actual funds with very similar *prima facie* characteristics as shown in table 8.

Both launched in 2014, by Q2 2021 these two seemingly comparable funds have a net TVPI of 0.98x and 1.65x (Source: Preqin): clearly, Fund 1 made different investment choices and achieved a lower multiple, but it is not the entire story.

Thanks to the methodology and data used for iFSA, not only can benchmark performance metrics be computed by vintage, but they can also be computed each year for a single vintage.

Figure 14 shows the annual performance benchmarking of these two funds against the quartiles of 2014 vintage funds from the iFSA universe, valued every year.

Interestingly, until 2017, Fund 1 was the better performer of the two funds and was, in fact, in the top quartile of all funds, before moving to the bottom quartile the following year. Fund 2 followed a more typical J-curve pattern gradually moved to the top quartile in 2021.

It is likely that Fund 1 must have faced some issues in 2018 such as failed new investment, write-downs of existing investments, etc. which has result in this stark performance difference between the two funds by 2021.

Thus, regular monitoring against a relevant benchmark can allow LPs to better understand performance issues with their managers and understand the nature of under- or outperformance.

A. Appendix

A.1 Sample size and robustness of quarter estimates

Tables 9 to 13 show the precision of IRR quarter estimates for five consecutive rounds (e.g. vintage years) of observing 10 to 10,000 observations of the same normal return distribution with mean 10% and standard deviation 12%.

A.2 Investment ratio assumptions

Figure 16 shows the distribution of investment ratio of the simulated funds in any given vintage year. This is the result of the investment size and deal success probability assumptions.

A.3 Number of investments made

Figure 17 shows the distribution of the number of investments in a fund by vintage years. This is the result of the initial number of investment assumption and the deal success rate in any given investment year.

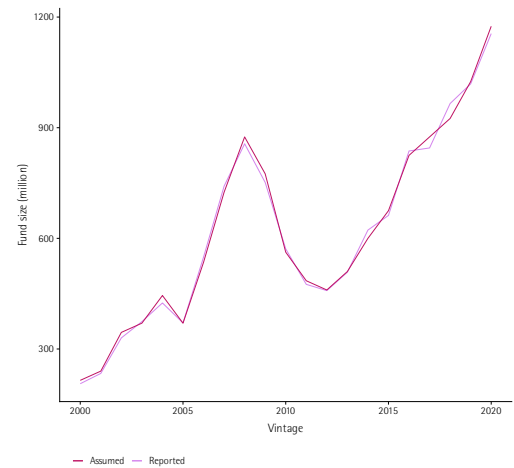
A.4 Fund size assumptions

Table 14 shows the assumed probabilities for different fund sizes across historical vintage years. It represents the average fund size in each vintage as observed in the market data (Figure A1).

A.5 Deal success rate assumptions

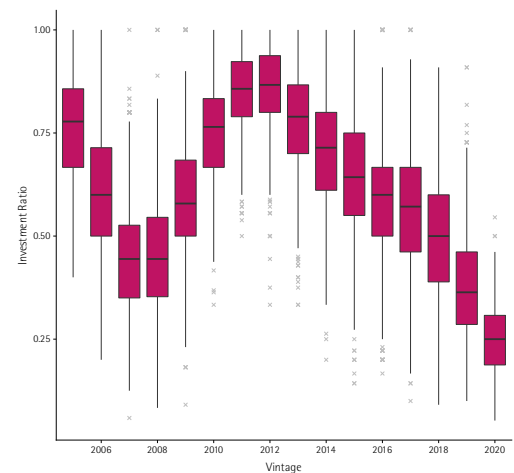
Figure 15 shows the historical probability of a deal success driven by the market activity in any given investment year.

Figure 15: Average historical fund size assumed vs observed



Source: EDHECinfra, Preqin

Figure 16: Distribution of fund investment ratio by vintage year



Source: EDHECinfra

Table 9: Precision of IRR quarter estimates for five consecutive rounds: with 20 observations

True Value of bottom quartile: 1.9%, 2nd quartile: 10%, top quartile: 18.1%

Vintage	n	Quantile	Estimate	Error	Lower c.i.	Upper c.i.	95% Conf. range
Bottom quartile boundary estimates, five consecutive vintage years							
1	20	25%	0.7%	-120 bp	-6.4%	6%	1240 bp
2	20	25%	4.9%	300 bp	0%	13.1%	1310 bp
3	20	25%	5%	310 bp	-3.3%	10.6%	1390 bp
4	20	25%	-4.6%	-650 bp	-12.6%	7%	1960 bp
5	20	25%	-1.6%	-350 bp	-10.4%	6.9%	1730 bp
3rd/2nd quartile boundary estimates, five consecutive vintage years							
1	20	50%	13.9%	390 bp	10.6%	23.8%	1320 bp
2	20	50%	9.8%	-20 bp	-1.5%	21.9%	2340 bp
3	20	50%	14.1%	410 bp	10.2%	22.7%	1240 bp
4	20	50%	10.4%	40 bp	5.2%	11.3%	620 bp
5	20	50%	9.3%	-70 bp	7.3%	17.8%	1050 bp
Top quartile boundary estimates, five consecutive vintage years							
1	20	75%	18.9%	80 bp	10%	32.7%	2280 bp
2	20	75%	18.4%	40 bp	14.7%	26.8%	1200 bp
3	20	75%	13.2%	-490 bp	7.7%	20.5%	1290 bp
4	20	75%	18.9%	80 bp	12.4%	24.2%	1180 bp
5	20	75%	20.5%	240 bp	14.6%	26.1%	1150 bp

Table 10: Precision of IRR quarter estimates for five consecutive rounds: with 50 observations

True Value of bottom quartile: 1.9%, 2nd quartile: 10%, top quartile: 18.1%

Vintage	n	Quantile	Estimate	Error	Lower c.i.	Upper c.i.	95% Conf. range
Bottom quartile boundary estimates, five consecutive vintage years							
1	50	25%	-2.7%	-460 bp	-8.8%	1%	980 bp
2	50	25%	-3%	-490 bp	-8.9%	2.8%	1160 bp
3	50	25%	5.3%	340 bp	1.9%	9.6%	770 bp
4	50	25%	-1.8%	-370 bp	-4.9%	6.2%	1110 bp
5	50	25%	1.8%	-10 bp	-1.8%	7.9%	970 bp
3rd/2nd quartile boundary estimates, five consecutive vintage years							
1	50	50%	10.2%	20 bp	5.1%	13.4%	830 bp
2	50	50%	9.6%	-40 bp	1.9%	12.9%	1100 bp
3	50	50%	8.8%	-120 bp	4.8%	16.6%	1180 bp
4	50	50%	9.5%	-50 bp	7.4%	15.5%	810 bp
5	50	50%	9.1%	-90 bp	5.6%	16.4%	1080 bp
Top quartile boundary estimates, five consecutive vintage years							
1	50	75%	20.4%	230 bp	13.7%	22.9%	920 bp
2	50	75%	17.2%	-90 bp	11.6%	22.6%	1110 bp
3	50	75%	17%	-110 bp	12.9%	22.8%	990 bp
4	50	75%	13.5%	-460 bp	11.5%	20.5%	900 bp
5	50	75%	26.5%	840 bp	16.4%	33.3%	1690 bp

Table 11: Precision of IRR quarter estimates for five consecutive rounds: with 100 observations

True Value of bottom quartile: 1.9%, 2nd quartile: 10%, top quartile: 18.1%

Vintage	n	Quantile	Estimate	Error	Lower c.i.	Upper c.i.	95% Conf. range
Bottom quartile boundary estimates, five consecutive vintage years							
1	100	25%	1.6%	-30 bp	-1.9%	5.3%	720 bp
2	100	25%	4.6%	270 bp	1.3%	7.3%	610 bp
3	100	25%	2.5%	60 bp	-0.7%	6.6%	730 bp
4	100	25%	4.3%	240 bp	0.8%	6.1%	530 bp
5	100	25%	2%	10 bp	-1.8%	6%	770 bp
3rd/2nd quartile boundary estimates, five consecutive vintage years							
1	100	50%	12%	200 bp	9%	16%	710 bp
2	100	50%	8.9%	-110 bp	6.6%	11.5%	490 bp
3	100	50%	9.9%	-10 bp	5.5%	14.2%	870 bp
4	100	50%	9.4%	-60 bp	5.2%	12%	680 bp
5	100	50%	12.1%	210 bp	9.6%	15.3%	570 bp
Top quartile boundary estimates, five consecutive vintage years							
1	100	75%	18.7%	60 bp	15%	22.4%	740 bp
2	100	75%	16.6%	-150 bp	14.9%	20.6%	570 bp
3	100	75%	17.6%	-40 bp	15.5%	19.6%	410 bp
4	100	75%	18.1%	0 bp	13.8%	21.5%	780 bp
5	100	75%	15.1%	-300 bp	14.2%	17.4%	310 bp

Table 12: Precision of IRR quarter estimates for five consecutive rounds: with 1,000 observations

True Value of bottom quartile: 1.9%, 2nd quartile: 10%, top quartile: 18.1%

Vintage	n	Quantile	Estimate	Error	Lower c.i.	Upper c.i.	95% Conf. range
Bottom quartile boundary estimates, five consecutive vintage years							
1	1,000	25%	1.2%	-70 bp	0.2%	2.4%	230 bp
2	1,000	25%	1.5%	-40 bp	0.8%	2.8%	200 bp
3	1,000	25%	1.6%	-30 bp	0.6%	2.6%	200 bp
4	1,000	25%	2.3%	30 bp	0.9%	3.2%	230 bp
5	1,000	25%	1.7%	-20 bp	0.6%	2.4%	190 bp
3rd/2nd quartile boundary estimates, five consecutive vintage years							
1	1,000	50%	10.5%	50 bp	9.6%	11.3%	170 bp
2	1,000	50%	9.2%	-80 bp	8.4%	10.1%	160 bp
3	1,000	50%	9.5%	-50 bp	8.6%	10.6%	200 bp
4	1,000	50%	9.9%	-10 bp	8.7%	10.8%	200 bp
5	1,000	50%	10.5%	50 bp	9.9%	11.3%	140 bp
Top quartile boundary estimates, five consecutive vintage years							
1	1,000	75%	17.3%	-80 bp	16.1%	18.3%	220 bp
2	1,000	75%	17.6%	-50 bp	16.4%	19.1%	270 bp
3	1,000	75%	18.2%	10 bp	17%	19.4%	240 bp
4	1,000	75%	19%	90 bp	18.4%	19.8%	140 bp
5	1,000	75%	18.1%	0 bp	17.3%	19.4%	210 bp

Table 13: Precision of IRR quarter estimates for five consecutive rounds: with 1,000 observations

True Value of bottom quartile: 1.9%, 2nd quartile: 10%, top quartile: 18.1%

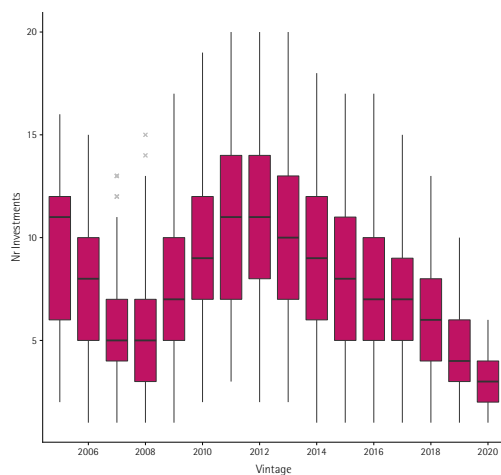
Vintage	n	Quantile	Estimate	Error	Lower c.i.	Upper c.i.	95% Conf. range
Bottom quartile boundary estimates, five consecutive vintage years							
1	10,000	25%	1.9%	0 bp	1.6%	2.3%	70 bp
2	10,000	25%	2%	10 bp	1.7%	2.3%	60 bp
3	10,000	25%	2%	10 bp	1.7%	2.3%	60 bp
4	10,000	25%	2.1%	20 bp	1.8%	2.5%	70 bp
5	10,000	25%	1.9%	0 bp	1.5%	2.2%	70 bp
3rd/2nd quartile boundary estimates, five consecutive vintage years							
1	10,000	50%	10.1%	10 bp	9.8%	10.4%	60 bp
2	10,000	50%	10.2%	20 bp	9.9%	10.5%	60 bp
3	10,000	50%	10.1%	10 bp	9.8%	10.5%	60 bp
4	10,000	50%	10%	0 bp	9.7%	10.3%	70 bp
5	10,000	50%	10%	0 bp	9.6%	10.3%	70 bp
Top quartile boundary estimates, five consecutive vintage years							
1	10,000	75%	18.3%	20 bp	18%	18.6%	70 bp
2	10,000	75%	17.9%	-20 bp	17.6%	18.3%	70 bp
3	10,000	75%	18.3%	30 bp	18%	18.7%	70 bp
4	10,000	75%	18.4%	30 bp	18.1%	18.8%	70 bp
5	10,000	75%	18.3%	20 bp	18%	18.7%	60 bp

Table 14: Distribution of infrastructure fund sizes by vintage – 2000-2020

Vintage	Very small (\$100mn)	Small (\$250mn)	Average (\$500mn)	Large (\$1000mn)	Very large (\$2000mn)
2000	40%	50%	10%	0%	0%
2001	40%	40%	20%	0%	0%
2002	20%	50%	20%	10%	0%
2003	20%	40%	30%	10%	0%
2004	20%	30%	30%	20%	0%
2005	20%	40%	30%	10%	0%
2006	10%	10%	60%	20%	0%
2007	0%	10%	60%	20%	10%
2008	0%	10%	30%	50%	10%
2009	0%	10%	50%	30%	10%
2010	0%	25%	50%	25%	0%
2011	10%	30%	40%	20%	0%
2012	10%	40%	30%	20%	0%
2013	10%	20%	50%	20%	0%
2014	0%	20%	50%	30%	0%
2015	0%	10%	50%	40%	0%
2016	0%	10%	40%	40%	10%
2017	0%	10%	30%	50%	10%
2018	0%	10%	20%	60%	10%
2019	0%	10%	20%	50%	20%
2020	0%	10%	10%	50%	30%

Source: Preqin, EDHECinfra

Figure 17: Distribution of number of investments in a fund by vintage year



Source: EDHECInfra

Table 15: Historical deal success probability by year of investment

Vintage year	Deal success probability
2005	100%
2006	100%
2007	74%
2008	42%
2009	10%
2010	35%
2011	60%
2012	100%
2013	82%
2014	86%
2015	74%
2016	63%
2017	56%
2018	55%
2019	48%
2020	54%
2021	55%

Source: Preqin, EDHECInfra

References

- Blanc-Brude, F. and A. Gupta (2021). The fair value of investments in unlisted infrastructure equity. *EDHEC Infrastructure Institute*. <https://edhecinfra.com/paper/the-fair-value-of-investments-in-unlisted-infrastructure-equity>.
- Cornel, J. (2017). Synthetic peer benchmarking for diversified private equity programs. *The Journal of Alternative Investments* 19(4), 53–66.
- Gredil, O., B. E. Griffiths, and R. Stucke (2014). Benchmarking private equity: The direct alpha method. *Available at SSRN 2403521*.
- Kaplan, S. N. and A. Schoar (2005). Private equity performance: Returns, persistence, and capital flows. *The journal of finance* 60(4), 1791–1823.



EDHEC*Infra* Publications (2018–2021)

EDHEC*Infra* Methodologies & Standards

- The Infrastructure Company Classification Standard (TICCS) - Updated March 2020
- Credit Risk Methodology - April 2020
- Infrastructure Index Methodology Standard - Updated March 2020
- Global Infrastructure Investment Data Standard - Updated March 2020
- Unlisted Infrastructure Valuation Methodology - A Modern Approach to Measuring Fair Value in Illiquid Infrastructure Investments - Updated March 2020

Selected EDHEC Publications

- Amenc, N. & F. Blanc-Brude. "The Cost of Capital of Motorway Concessions in France - A Modern Approach to Toll Regulation" (September 2020)
- F. Blanc-Brude & A. Gupta. "Unlisted Infrastructure Performance Contribution, Attribution & Benchmarking" (July 2020)
- Whittaker, T. & R. Tan. "Anatomy of a Cash Cow: An In-depth Look at the Financial Characteristics of Infrastructure Companies." (July 2020)
- Amenc, N., F. Blanc-Brude, A. Gupta, L. Lum. "Investors Should Abandon Absolute Returns Benchmarks - Lessons from the Covid-19 Lockdowns" (June 2020)
- Amenc, N., F. Blanc-Brude, A. Gupta, J-Y. Lim. "2019 Global Infrastructure Investor Survey - Benchmarking Trends and Best Practices" (April 2019)
- Whittaker, T., S. Garcia. "ESG Reporting and Financial Performance: The case of infrastructure." (March 2019)
- Blanc-Brude, F., J-L. Yim. "The Pricing of Private Infrastructure Debt - A dynamic Approach" (February 2019)
- Blanc-Brude, F., C. Tran. "Which Factors Explain Unlisted Infrastructure Asset Prices?" (January 2019)
- S. Garcia, F. Blanc-Brude, T. Whittaker. "Tome La Siguiente Salida (Take the Next Exit) - A Case Study of Road Investments Gone Wrong, Spain, 1998-2018" (March 2018)
- Amenc, N., F. Blanc-Brude "Selecting Reference Indices for the Infrastructure Asset Class" (February 2018)
- Blanc-Brude, F., T. Whittaker, and M. Hasan. "Cash Flow Dynamics of Private Infrastructure Debt" (March 2016).

For more information, please contact:

Tina Chua on +65 6438 0030

or e-mail: tina.chua@edhec.edu

EDHEC Infrastructure Institute

EDHEC Asia-Pacific

One George Street - #15-02

Singapore 049145

Tel.: +65 6438 0030

edhec.infrastructure.institute