



A data-driven approach to venture fund portfolio building.

Graham **Schwikkard** | CEO at SyndicateRoom

Creating consistent returns from venture investing.

The venture industry has historically been characterised by oracle-style investors. These are investors who raise capital on the basis of their ability to find and pick the top performing early stage companies, investing in just a few per cycle, then doubling down on the winners. Much like the pre-ETF listed equities space, these actively managed funds dominate the landscape - despite their highly variable results that could be the result of luck. VCs only pick winners in 2.5% of their investments.¹

In this article, I'll be exploring an alternative approach to venture investing, one which attempts to consistently capture the growth of the entire market.

This approach is built on four important elements that we have derived through various analyses, including a statistical Monte Carlo approach:

1. The venture market has shown consistent growth year on year.
2. A portfolio of 50 investments a year can replicate this growth, minimising variation.
3. Access to the top 20% of the market has a marked impact on portfolio growth.
4. On average, fixed ticket investments show better returns than variable tickets.

When most venture firms talk about using data in their decision making, it typically alludes to them gathering data on a single, individual firm and attempting to find signals that the specific company is (A) on the growth trajectory and (B) has potential to grow its current valuation 10X. However, there is no existing database large and diverse enough to inform these decisions. Startup companies are simply not like listed equities, and so no firm can gather sufficient data to build a causative model that proves certain signals are definitive of growth for any one company.

Additionally, by definition the startups are trying to break existing business models and so it is easy to miss the right signals. In this way, the oracle investors are actually better suited to picking up the composite set of new data that an investment provides them. The only hope for these company-level data-driven approaches to work, is to truly narrow down the type and industry of startup that you work with, in the hope that a large enough data set can be built and used.

I raise this approach simply because it is the first model that most investors consider when they hear data-driven venture investing. But, as I've explained, it is the wrong approach due to a lack of a coherent data model. Rather, one should focus on the market level, not the company level. By focusing on the market level, it is possible to come up with a broad approach to venture investing to capture market growth, much in the same way that tracking funds in public markets attempts to capture broad market growth. It won't be exactly the same, as the data and

share purchasing mechanisms are vastly different, but the idea of capturing market growth can be replicated.

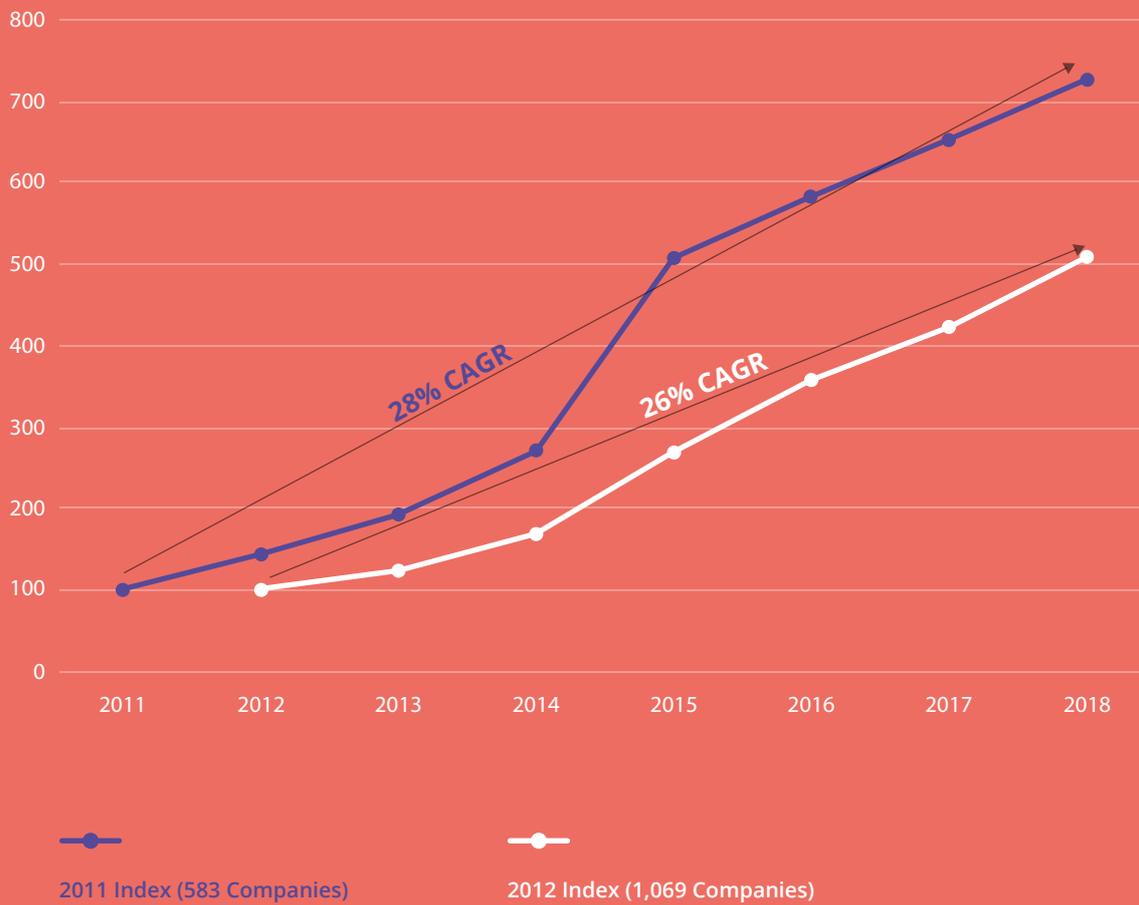
The venture market has shown consistent growth year on year

Historically, venture market growth has been difficult to analyse as the companies are private and so they are not required to give up their information willingly. However, in the UK this is not the case as every company must submit structured filings to Companies House for every fundraise they complete. These filings include the number of shares issued and the share price that they raised at, and the data is made public and published on the Companies House website. Scraping and aggregating this basic set of data means that you can build a database of every company which raised capital, how much they raised, when they raised, and at what valuation. This simple model can then be used to map the starting valuation of a particular cohort, and then track it year on year and see how the market valuation grows.

There is a slight caveat to the data unfortunately, as some companies do fill in the forms poorly and so those valuations must be discarded, but the majority of it is intact and usable. SyndicateRoom partnered with Beauhurst to pull this data together, and they validated the valuations from press releases too. Below is the total growth seen for the 2011 and 2012 cohorts:

Indexed valuation of UK startups

that raised equity finance¹ in 2011 and 2012



Sources: SR Analysis, SyndicateRoom report - Rise of the Growth Hunters, SyndicateRoom report - Early Stage Equities: A Long Term Study, Beauhurst

Notes: ¹Cohort limited to companies in seed or venture stage at start of index and where an accurate valuation could be calculated. Where the company ceased trading it was valued at zero

What is immediately striking about this data is the consistent growth. We found similar patterns for the 2013 and 2014 cohorts, which showed 24% CAGR and 23% CAGR respectively. In a high risk market where 40% of the companies are dead by year 5, it seems counterintuitive to find the returns represented by a near straight line and not a rollercoaster, or even a staircase. In part this is because the companies are not actively traded, and so aren't subject to the day to day speculators. Even though they can be over or undervalued at any one fundraising event, this is relatively rare. Our findings were supported by a British Business Bank commissioned study on the global VC market that study found when you pooled together all the venture funds across the globe, they showed similar growth of 18% between 1970-2016.²

However, there is a lot of churn in which VC's provide the best returns. Cambridge Associates found that for the last 10 years, 40%–70% of total gains were claimed by new and emerging managers.³

A recent publication by Morgan Stanley also found that Venture Capital returns are high, but highly variable (over 20% annual returns but with 20% standard deviation).⁴ So individual VC investors may show variable returns, but when aggregated it's a much more consistent picture, which, based on our data looking at all deals in the market, is what you would expect.

From our own analysis of UK VCs we found that the average growth of their portfolios is 18%, yet most target 20%. Therefore, one must ask the question: why are they not able to beat the market growth?

The next question is, how do you capture this growth? The startups are not listed, so you can't just go out and invest in every one, and the later stage growth businesses (already demonstrably performing) are very hard to get access to. As a starting point, we set out to find the right kind of portfolio to build. How many companies should you invest in? How should it be spread out? Do you invest more in higher valued companies? Should you follow your money?

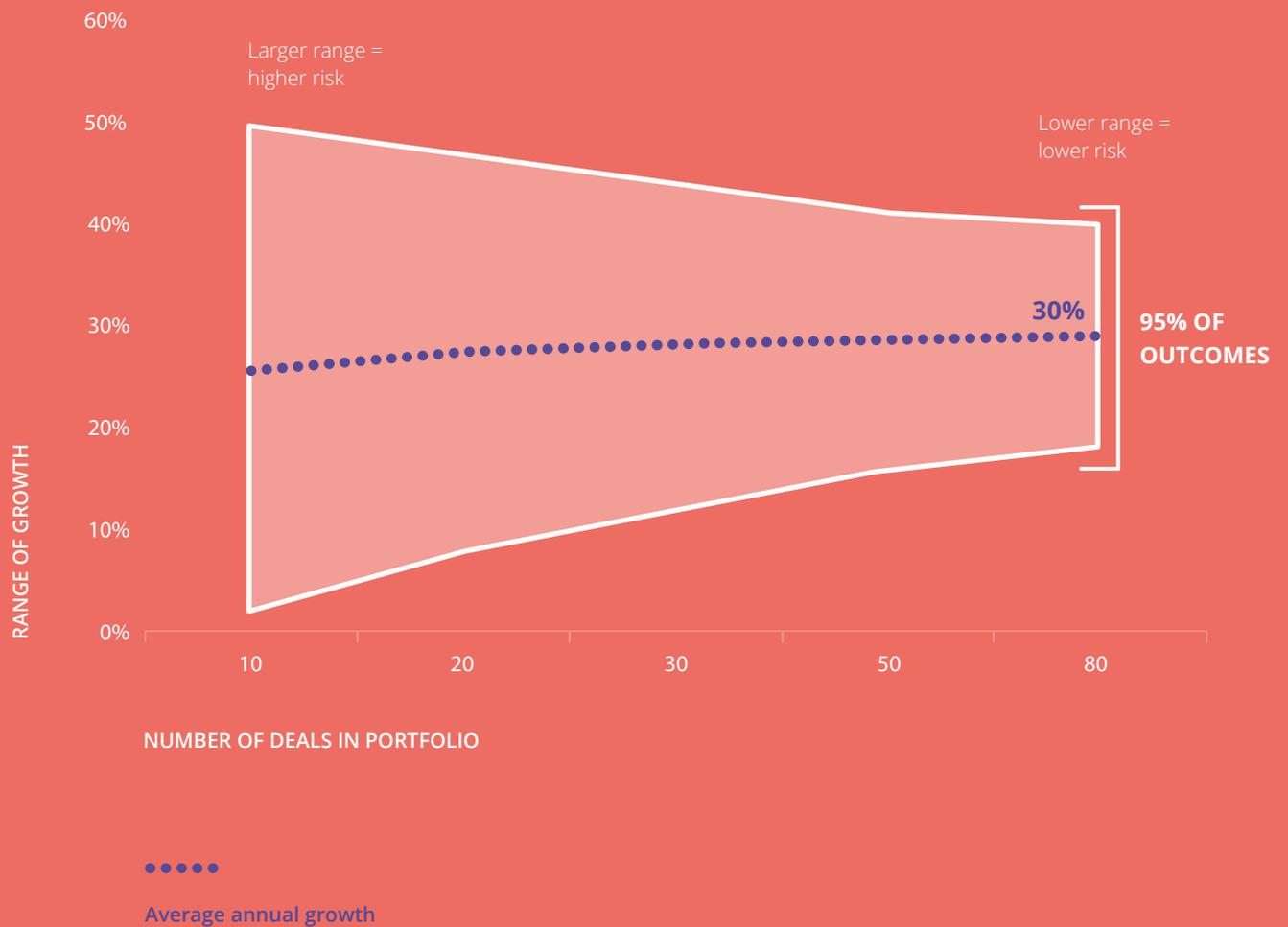
A portfolio of 50 investments a year can replicate market growth and minimise variance

Our next step in analysing this historical data set was to run various portfolio-building strategies. In order to set up the simulations, the cohorts were kept separate to simulate a fund which deployed over a single year. SyndicateRoom has historically deployed investor funds across a single year so that tax claims could be applied to a single tax year. Therefore a single cohort was analysed, although repeated across the years (i.e. 2011, 2012, and 2013) to see if the strategy remained effective. Later, the investments from the adjacent cohorts were joined together to simulate investing across the years, and this data is the basis for SyndicateRoom's institutional fund that has a three year deployment.

Each company within the cohort was assigned a simple number, starting at one and running up to the total number of companies which raised capital in that year. We then created a script to create a set of random numbers which would represent a set of investments by a venture capital fund. The size of the portfolio being built was the independent variable for this set of investment simulations - the script would create 10, 20, 30, 40, 50, 60, 70, or 80 random investments.³ The 10-80 random companies would be pulled together into a single portfolio, with the fund investing a fixed amount into each company at the time of their fundraiser for that year. It was then just a case of calculating the growth of the portfolio. However, performing this analysis for a single random selection would only represent a random, lucky (or unlucky) draw of companies. Thus, the selection was repeated 10,000 times for each portfolio size, and the subsequent mean returns and variation of returns measured.

The results are plotted below, showing the mean returns of the simulated portfolios.

Average and range of growth as the number of deals¹ in a representative portfolio increase



Sources: SR analysis on the entire cohort of early stage deals in 2011-2014, data provided by Beauhurst and Companies House

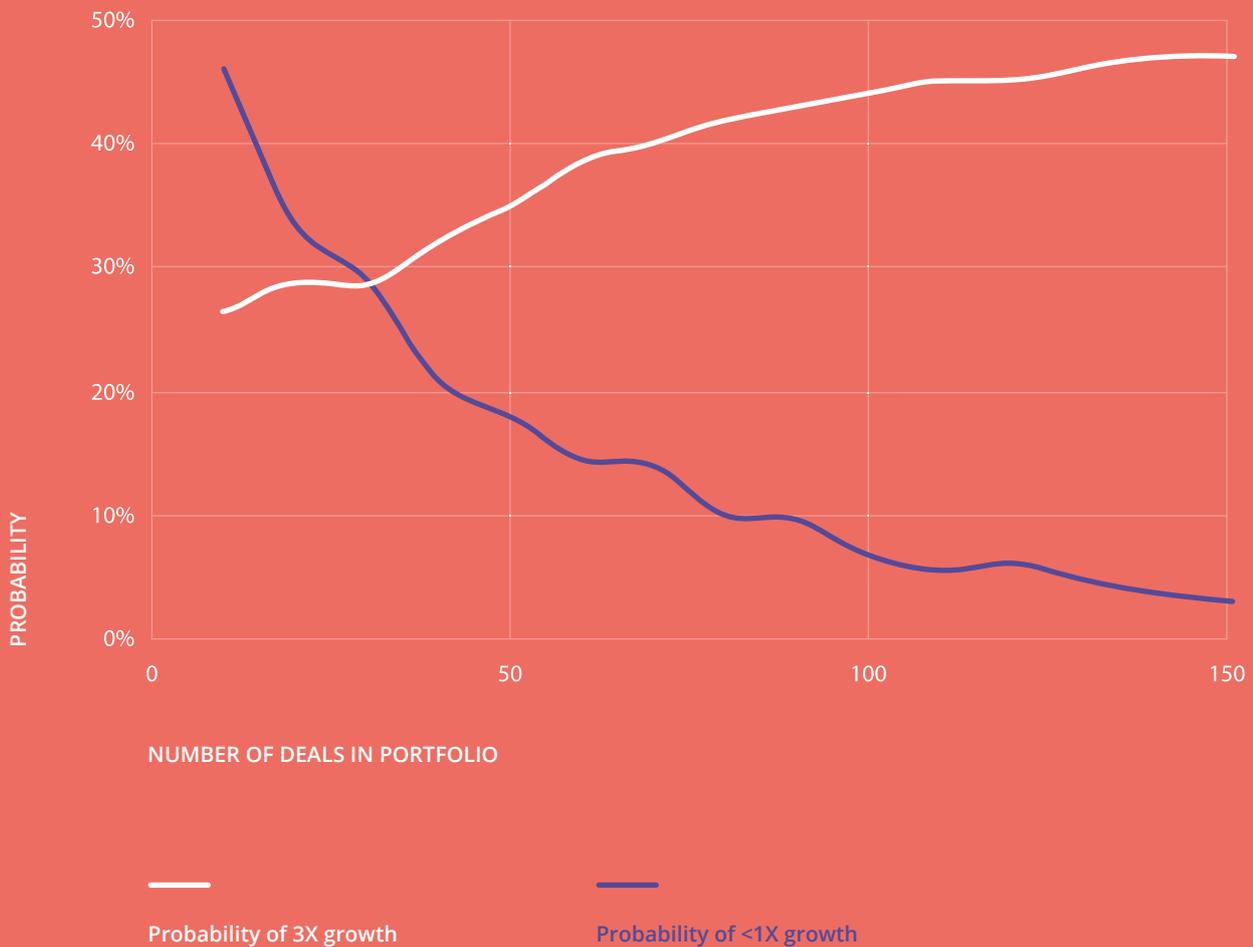
Notes: ¹Based on random selection of deals of all historical U.K startups fundraising from 2011-2014, repeating the selection 1,000 times and comparing the resultant growth. This is the entire cohort of deals, not just focussed on selection from super angels investments

The analysis shows that as the number of deals in the portfolio increases, the mean growth approaches the mean of the cohort. This is simply a product of central limit theorem - that the sample mean will approach the population mean as the size of the sample increases. Similarly, the variance of portfolio growth decreases as the portfolio size (sample size) increases. Interestingly, the variance is high enough at 10 deals that in some instances you will have greater returns than would be likely for a larger portfolio size, but overall your average growth will be lower. This type of analysis has been done previously^{6,7}, where the authors modelled the typical, theoretical spread of 0X, 1X, 5X, 10X deals within the market. The result of their analysis was the critical importance of building a portfolio of 100-150 deals. Our analysis is the first completed on actual historical data to our knowledge.

Reflecting on the average VC returns, it would suggest that a random selection for their portfolios could result in returns, on average, greater than we actually see. For SyndicateRoom's own fund, we selected 50 deals a year as a starting point for building our portfolio - it represented a balance between how many deals we could initially complete whilst reducing the variation of growth significantly. When these 50 deals are combined across three years to build a portfolio of 150 deals, we find that the probability of losing capital is significantly minimised (but will never be 0), whilst the odds of returning 3X of cash invested is higher than 50%.

Probability of growth

as portfolio size increases¹



Sources: SR analysis on the entire cohort of early stage deals in 2011-2014, data provided by Beauhurst and Companies House

Notes: ¹Based on random selection of deals of all historical U.K startups fundraising from 2011-2014, repeating the selection 1,000 times and comparing the resultant growth. This is the entire cohort of deals, not just focussed on selection from super angels investments

There is a critical point to consider in this simulation: it is not always possible to access the full set of deals in any one cohort. They are not publicly listed and so you cannot simply log on to a trading platform and randomly select your own portfolio. Moreover, the set of returns in the startup markets follows a power law market ⁸² - much of the growth is concentrated and not bell curved.

Therefore, the next part of our analysis was to look at what would happen if you could not get into the top deciles of the population.

Access to the top 20% of the market has a marked impact on portfolio growth

It is generally well accepted that startups follow a power law distribution: 40% of startups fail within 5 years, and the >10X returns are concentrated within the top 10% of companies. Our own analysis supported this power law too, with 7.5% of the population showing 10X or greater growth.

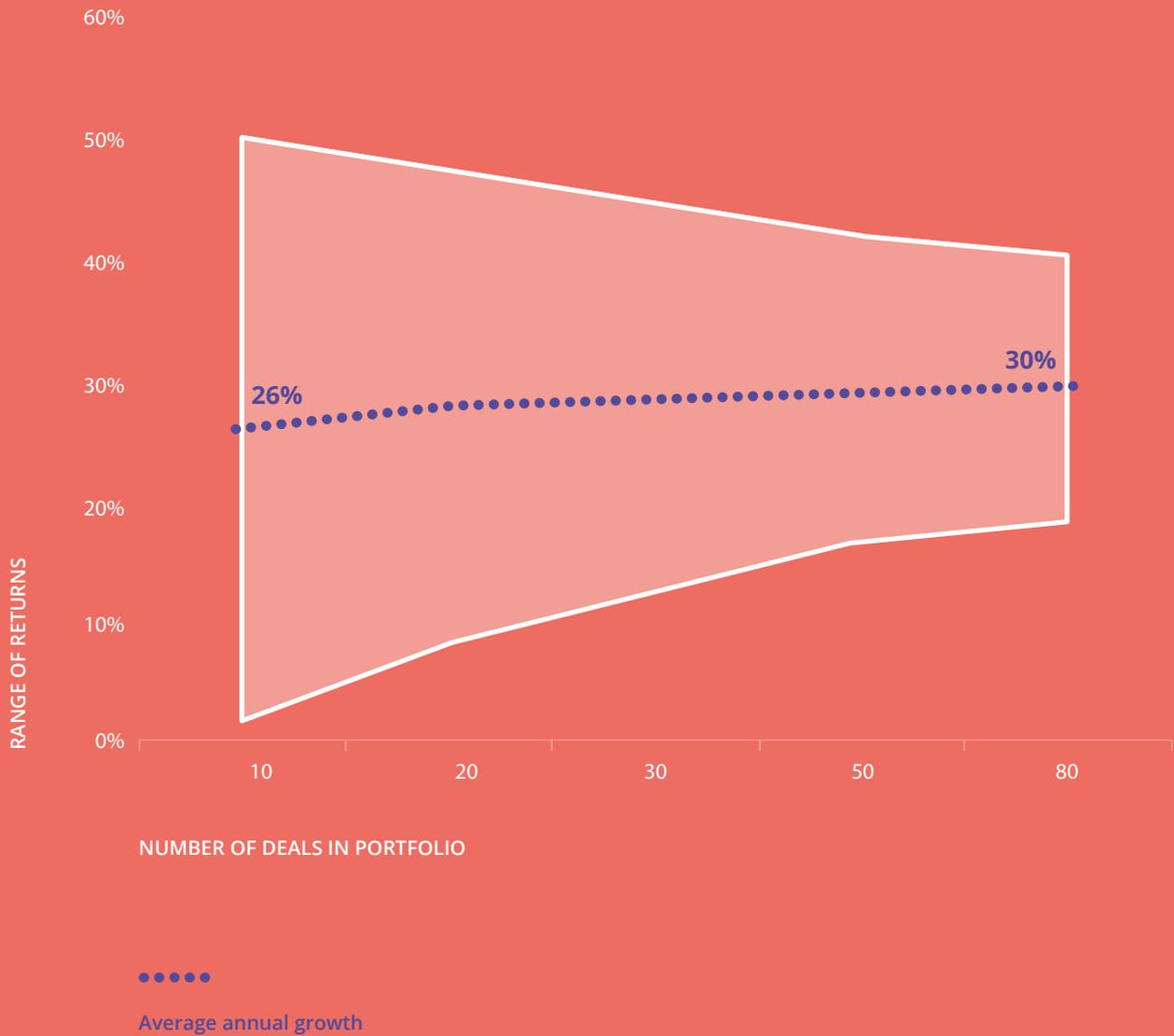
This distribution is what drives most venture capitalists to focus their capital on a few deals a year, and then double down on the startups which are showing growth. As the previous analysis has shown, when you are only completing 10 deals a year, there is a lot of pressure to pick the winners. Therefore it seems logical to perform detailed due diligence on each investee company, and make predictions based on this analysis to determine if the company has potential

to increase its valuation 10X or more. I argued earlier that it is difficult to do this based on a structured data method, and so it relies primarily on the VCs intuition and logic. A venture capitalist will not invest in a company unless its projections show a hockey stick like growth, it just isn't worth the risk otherwise because they need to bank on every one of their 10 investments being able to provide the returns of the whole portfolio. One could explore this further and examine whether it leads to irrational, overly ambitious decisions by the investee companies, but that is not a question for this paper. What we wanted to know is: how critical is access to the top 10% of deals when building a diversified portfolio?

Below are the results of the previous variance analysis at increasing portfolio sizes when we remove the top 10% of the population.

Annual growth¹

with access to all deals in the market

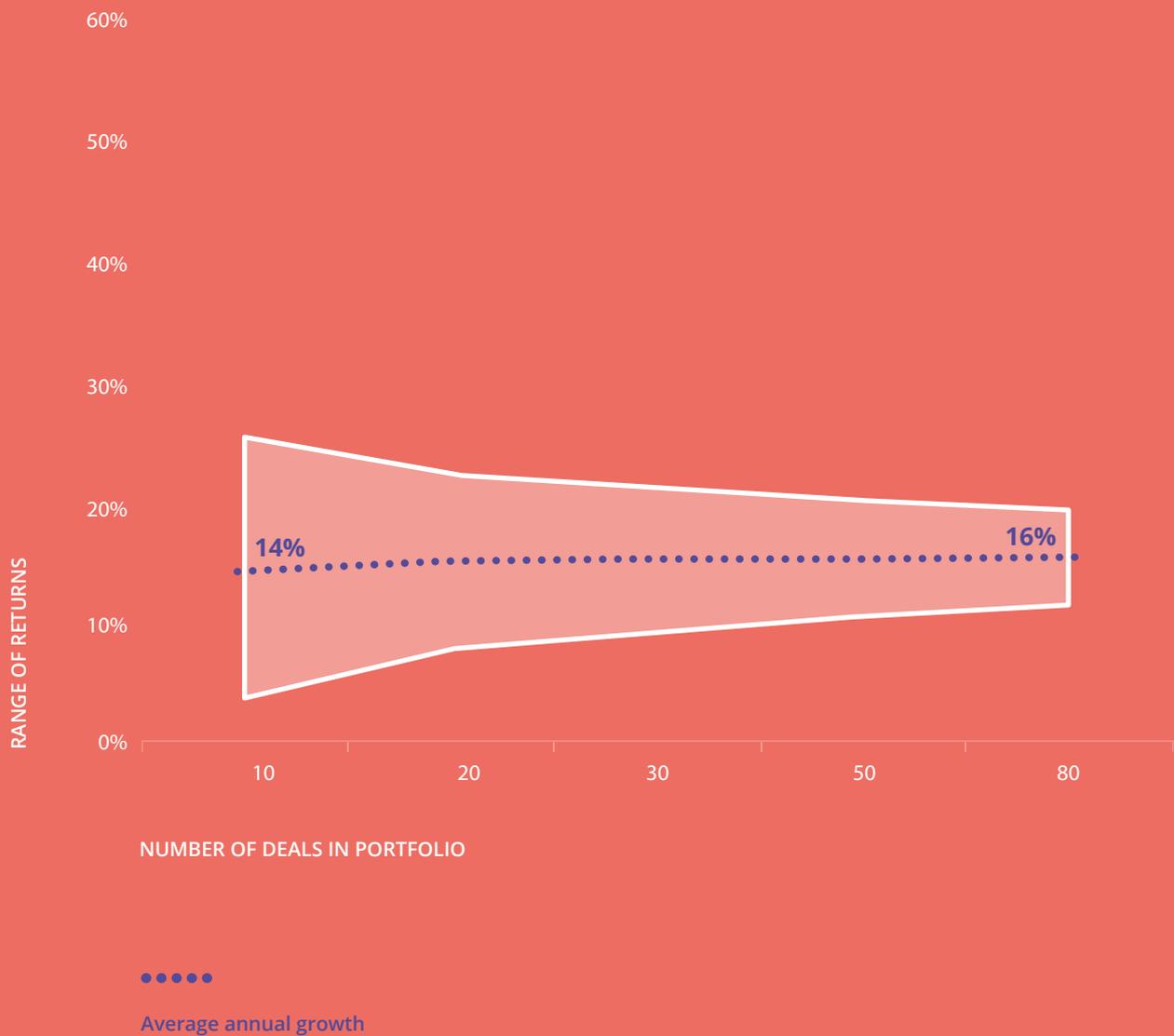


Sources: SR analysis on the entire cohort of early stage deals in 2011, data provided by Beauhurst and Companies House

Notes: ¹Confidence interval based on 2 standard deviations from the mean. Using 2011 cohort

Annual growth¹

when missing top 10% of deals in the market



Sources: SR analysis on the entire cohort of early stage deals in 2011, data provided by Beauhurst and Companies House

Notes: ¹Confidence interval based on 2 standard deviations from the mean. Using 2011 cohort

Once again, the central limit theorem comes into play and the mean growth of the population decreases, so the sample portfolio growth decreases appropriately. However, the central limit theorem also makes the variance of returns once again decrease as the portfolio increases (so bigger portfolios are still better, even if you lack access). In the population where 10% of the population has been removed, the overall growth rate is still 16%. This is still respectable, especially in comparison to most UK VCs.

A large portfolio can clearly still improve returns, but it is nevertheless critical that any fund has a way to access the whole population, and in particular the top deals. The good firms succeed because they are able to weigh the flow of deals in their favour, and the bad firms fail because they never get to see the best companies. Access is not equally distributed. You need to build the machine that regularly enables you to see and get into the best startups. If you want to build the right machine, you need the following:

- 1. Avoid adverse selection from companies.** Remove any selection methodology which could exclude the top deals. This would seem intuitive, but putting companies through a negative experience and not being friendly to founders will put off the next set of entrepreneurs from working with that fund. This applies to fees as well. Our own research has found that the average upfront fee from UK VCs is 3.9%, with a 1.1% annual management fee. This excludes any 'director fees' or 'expenses' that are charged to the company. High fees can be off-putting to companies,

possibly excluding them from your selection pool (and some of the top 20% of companies could be among them).

- 2. Plug into the right networks.**

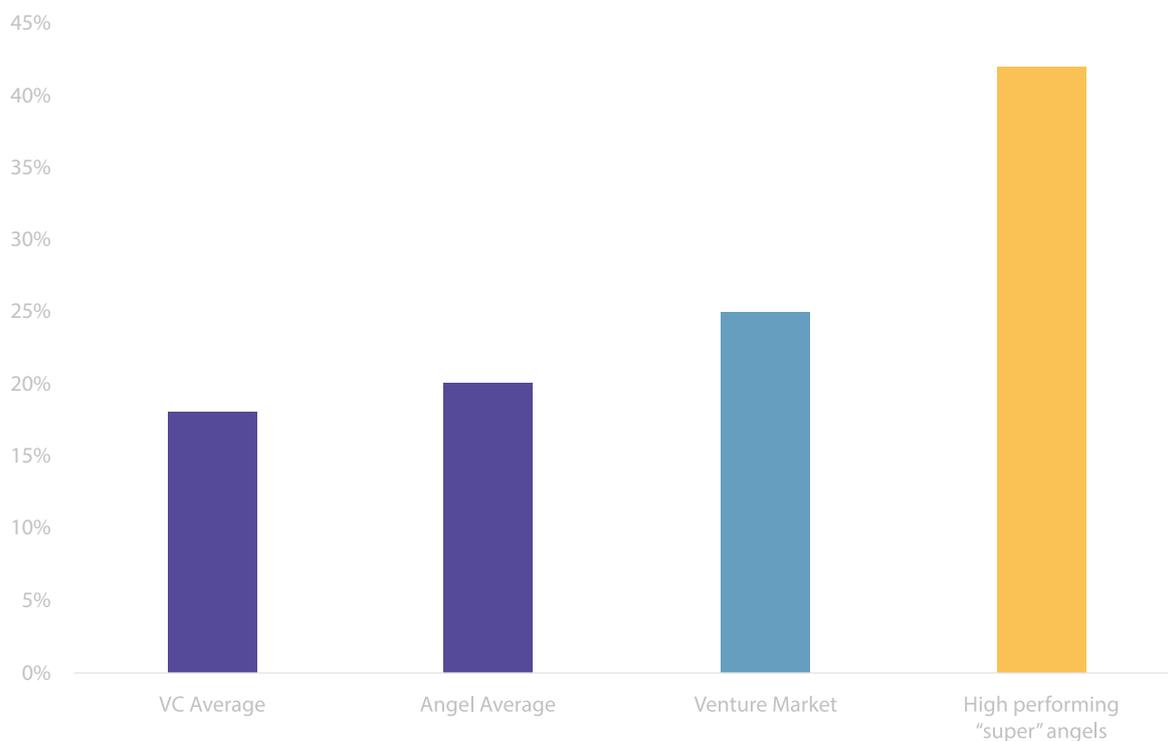
The venture market is simply a network of nodes like any other network, where early stage companies connect to an original investor, but are then connected to further investor nodes for either their original or later rounds. The companies are nodes too, connecting to other investors connected to other companies. There are multiple ways to solve this problem. Most VCs solve it by sending an analyst team to events and finding the company nodes. Some VCs rely on their profile to attract companies, either through cold introductions or finding some network connection to introduce them. Overall I believe it is a network problem, and how you find those nodes with access to the top end of the deals.

- 3. VCs must have a compelling value proposition to companies.**

In a competitive funding market, a funding offer needs to align with the jobs that need to be done by startups when raising capital. Most VCs offer their expertise and ability to provide invaluable advice, or even just their brand to be attached to the company. Our research with companies in their early rounds (typically before they have taken on an institutional investor) has found that they really just want to get on with running their business, so we have taken the route of straightforward capital.

Our compelling value proposition SyndicateRoom has solved the network problem with another set of analyses, which I plan on answering with a separate white paper. In short, we have used a large dataset to build the track records for thousands of angel investors to identify those with access to the best deal flow. We now partner with those angel investors to get into their deal flow. They are our machine. For companies, we offer fast and cheap capital, simply going in if the angel investor is going into the round. That is our compelling value proposition.

Average portfolio growth of venture investors



14

Fixed ticket sizes create the best returns

There is still one other important question that needs to be answered when structuring a fund which aims to capture market returns from a diversified portfolio: how much to put into each deal?

On average, fixed ticket investments show better returns than variable tickets.

Determining the right spread of capital through a potential portfolio was a relatively straightforward analysis. We used the previous Monte Carlo simulation that had been set up when looking at portfolio sizes and simply adjusted the initial investment into each company to either be:

1. A fixed ticket size. Meaning that each investment would be an equal percentage of the portfolio but the percentage of each round would be variable, as would the percentage acquired of the company's shares.
2. A fixed percentage of the round, meaning the percentage that each startup would make for the whole portfolio would be variable.
3. A fixed percentage of the valuation, also meaning a mixed portfolio and variable percentage of the round. But returns might be more predictable if you expect returns to grow on a specific distribution.

Strategy	% of portfolio	% of round	% of company
Fixed ticket	Equally weighted	Variable	Variable
Fixed % of round	Variable	Equally weighted	Variable
Fixed % of valuation	Variable	Variable	Equally weighted

Running this analysis again on the entire population, but then again on 10,000 simulations of a portfolio of 50 companies showed that a fixed ticket sizes created the best returns across the various cohorts.

This result may seem counterintuitive at first, as we had assumed an approach which adjusted the ticket size based on the round size or valuation would mean you took a larger position on companies which had more traction (as they were raising a larger round). Instead, what happens is that taking larger positions in these later stage rounds leaves less room for growth and so even though the returns are not bad, they are lower on average because you miss out on the early rounds of high potential companies. This is possibly more important in a fund which is maximising for consistent returns by diversification across a lot of smaller, early rounds.

	Fixed investment (ticket) size	Investing as % of round (or other variable)
Average CAGR ¹	30% £5m - £31m in 7 years	24% £5m - £22m in 7 years
Cash returns multiple	2.5X	2X
Model findings	Mixed exposures to larger and smaller deals providing diverse returns	Overexposure to larger deals with less runway of high growth

In conclusion: putting the research into action with Access EIS

Our research informed a new approach to investing in startups which aims to capture the steady growth that can be observed in the UK startup market as a whole, which led us to create our flagship fund, Access EIS. To summarise, it does this in the following way.

- 1. Diversification to minimise risk, and replicate overall market growth.** Building a portfolio of 50+ companies, instead of trying to pick a small number of winners.
- 2. Access to the best deals, and making use of existing expertise.** Co-investing with angel investors who have a proven track record of backing successful companies, along with access to the top deals.
- 3. On average, fixed ticket investments show better returns than variable tickets.**

Closing thoughts and areas for future research

Portfolio building approaches have always been a subject of discussion for the venture industry, but it is only more recently that data-driven approaches are becoming more prominent in conversation. Notable examples include SignalFire (give [this podcast](#) a listen from [@thefullratchet](#)) and their Beacon database which tracks multiple data sources to inform their venture model. Nnamdi Iregbulem wrote about [building an index fund](#) for venture capital, and the barriers to creating one (which I believe our fund has addressed somewhat). Lastly, I would recommend reading posts by [Steve Crossan](#), as well as any of the sources referenced in this paper.

I hope that this white paper contributes to the discussion and provides insight into one approach to portfolio building. This is just the start of our analysis and future plans, I can see potential for a number of improvements, namely:

- Follow on decisions. We already have an analysis to support this, and it's why the Access fund doesn't follow its money (as it is, on average, better to invest in a new company for the stage of the market we work in). However, we are working on a model which could inform a structured approach to follow ons. On this note, Craig Thomas wrote a useful analysis about [fund reserves here](#).
- Sector analysis. We already have the data to determine which sectors have historically created returns, but whether or not that is predictive needs further analysis.
- Valuation analysis. Are there signs around historical valuations that provide insight for making better odds on decisions to invest?
- Big data analysis. Right now we have gathered not enough data to warrant true machine learning applications, but I can see how to expand the data set over time and companies to provide a significant dataset which may identify new correlations.

If you would like to discuss any elements of this, please [reach out to me on twitter](#), or email me at graham@syndicatoroom.com. If you like the sound of this approach, check out our website for [more information on investing in our Access fund](#).

Citations & footnotes

1, 3: Cambridge Associates - ["Venture Capital Disrupts Itself: Breaking the Concentration Curse"](#)

2: British Business Bank - ["The Future of DC Pensions: Enabling Access to Venture Capital and Growth Equity"](#)

4: Morgan Stanley - ["Public to Private Equity in the United States: A Long-Term Look"](#)

5: Obviously, VC investments are not random in nature, our goal of this analysis was to demonstrate the impact of diversification and determine at what point you get the greatest impact. Moreover, it may not be possible to execute 50 deals a year (although SyndicateRoom's Access Fund has found a way to do this both efficiently and effectively).

6: Kevin Dick - ["Simulating Angel Investment: Kevin's Remix"](#)

7: Steve Crossan - ["Modelling suggests rational venture investors should have bigger portfolios"](#)

8: Abe Othman - ["What AngelList Data Says About Power-Law Returns In Venture Capital"](#)

9: Clint Korver - ["Picking winners is a myth, but the PowerLaw is not"](#)

Disclaimer & Risk Warning

Investing in startups is high risk and while our investment strategy addresses many of the risks, you may not get back the full amount you invest. It is important to seek advice before making an investment decision.

These materials are written and provided for general information purposes only. It is worth remembering that while a significant study for the industry as a whole, this report nevertheless uses a small sample size which limits the reliability of assertions drawn from the data.

The content is solely the opinion of SyndicateRoom and/or other contributors and research from third parties. It should not therefore be relied upon in making any investment decisions.

You should not invest in any investment product unless you understand the nature of it, along with the extent of your exposure to risk. You should be satisfied that any product or service is suitable for you given your financial position and investment objectives. Where appropriate, you should seek advice from a financial advisor in advance of making investment decisions

Risk warning: [Please click here to read the full risk warning.](#)

Investing in early-stage businesses involves risks, including illiquidity, lack of dividends, loss of investment and dilution, and it should be done only as part of a diversified portfolio. SyndicateRoom is targeted exclusively at sophisticated investors who understand

these risks and make their own investment decisions. Tax relief depends on an individual's circumstances and may change in the future. In addition, the availability of tax relief depends on the company invested in maintaining its qualifying status. Past performance is not a reliable indicator of future performance. You should not rely on any past performance as a guarantee of future investment performance.

This page has been approved as a financial promotion by Syndicate Room Ltd, which is authorised and regulated by the Financial Conduct Authority (No. [613021](#)).

We use cookies to improve our service. By continuing to use this site you are agreeing to their use. [Find out more.](#)