

Babe Ruth, Power Laws, and What It All Means

Josh Nussbaum

13-17 minutes



One of the advantages of a fellow like [Warren] Buffett, whom I've worked with all these years, is that he automatically thinks in terms of decision trees and the elementary math of permutations and combinations....

- Charlie Munger during his 1994 talk at USC entitled "[A Lesson on Elementary, Worldly Wisdom As It Relates To Investment Management & Business](#)"

It's been well documented that venture capital is an asset class predicated on the success of a small portion of investments as it relates to the overall portfolio. Peter Thiel frequently talks about this [power law](#), Chris Dixon has written about the "[Babe Ruth Effect](#)" in venture capital, and Marc Andreessen recently tweeted about it when he said "... [remember startups and VC are a game of outliers, not averages.](#)"

As Peter Thiel [notes](#), "The more a VC understands this skew pattern, the better the VC." I don't think there are all that many VC's out there that don't understand the skew pattern. What Peter (and Chris and Marc) understand that many others don't is how to build a portfolio with proper risk/reward that lends itself to this power law and achieves returns amongst the best in the entire industry.

At Metamorphic we often say that it has never been easier to start a company and never been harder to scale one. As a result, being an early stage firm means we see a lot of deals every year. It is our job to make investments that we believe have a high chance of success and will provide strong returns for our investors.

Contrary to popular belief (I'm kidding) VC's aren't soothsayers and they don't see the future. So why is it that some funds consistently outperform others?

Some might say there is luck involved (and I won't argue against that) as deal flow and referrals can often be the results of happenstance or serendipity. Most would say that the best VC platforms and brands that help entrepreneurs in a differentiated way over long periods of time results in seeing the best companies, which in turn results in making the best investments and as a result top quartile returns will follow. This is certainly the case, however if you take a look at the best performing funds with the best brands, it's likely that they look at upwards of 3,000 companies a year.

Maybe a few investments per fund (depending on the number the fund makes) are no-brainer investments and pattern recognition also helps, but decisions aren't typically easy. This is true more often during the earliest stages, where companies always have holes and flaws and there is little to zero data that can be analyzed. Paul Graham has said that great companies often look like bad ideas in the beginning. This is why investors look closely at specific sectors, form theses, and dig deeply into emerging technologies and trends. It's also one reason why operating experience in specific sectors and types of businesses has become a common requisite in newer GPs (this is one reason but it's secondary to having the experience to both sympathize with entrepreneurs and help companies which is crucial in winning deals). Understanding the intricacies of these businesses (pattern recognition) helps investors see the "forest through the trees" in taking early risk.

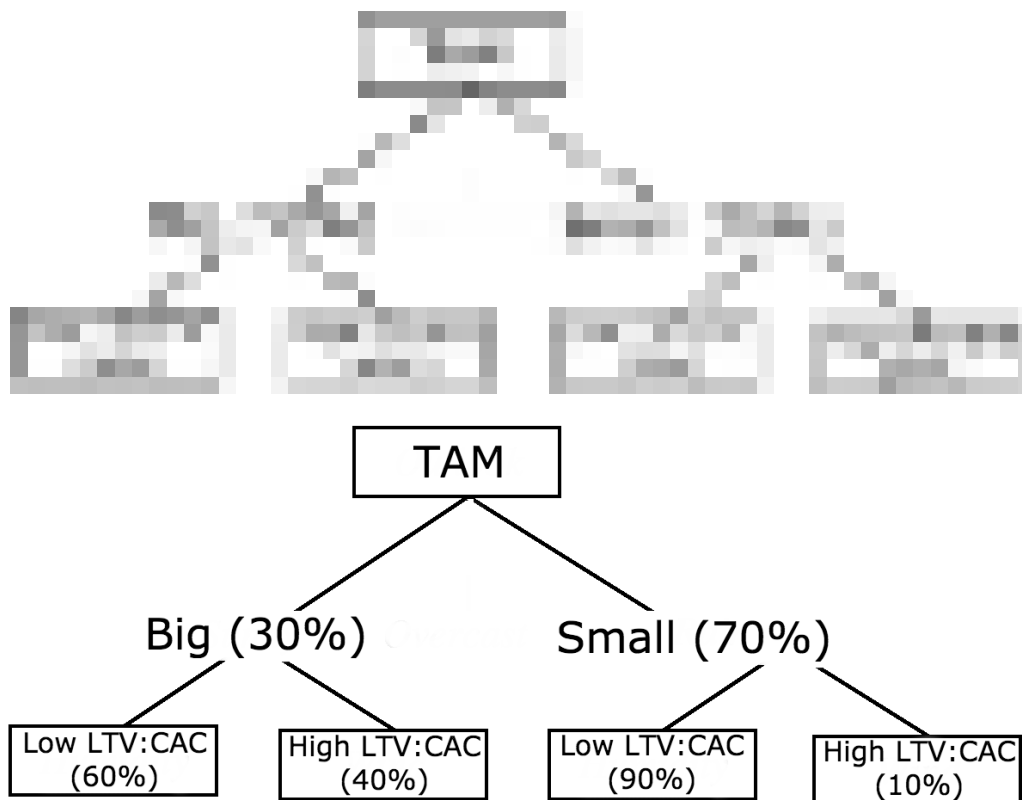
The above has been written about at length in identifying good VCs and while it's very much the case (in my limited experience at least), what is so scarcely discussed is why great investors pick x investment over y investment. Peter Thiel's power law is mentioned often but what is rarely discussed is how the truly great investors optimize to best capitalize on the skew pattern and the overall construction of a portfolio. Warren Buffett best described the venture strategy when he [said](#):

... If significant risk exists in a single transaction, overall risk should be reduced by making that purchase one of many mutually-independent commitments. Thus, you may consciously purchase a risky investment — one that indeed has a significant possibility of causing loss or injury — if you believe that your gain, weighted for probabilities, considerably exceeds your loss, comparably weighted, and if you can commit to a number of similar, but unrelated opportunities. Most venture capitalists employ this strategy. Should you choose to pursue this course, you should adopt the outlook of the casino that owns a roulette wheel, which will want to see lots of action because it is favored by probabilities, but will refuse to accept a single, huge bet.

As Michael Mauboussin pointed out in his paper on the “[Babe Ruth Effect](#)” (and Chris Dixon later cited), our brains are wired to avoid losses and therefore investors understand looking for grand slams, but have a difficult time doing so when it means the strikeout has a high likelihood instead. This is where decision trees come into the picture and why Charlie Munger attributes much of Warren Buffet’s success to his ability to quickly think in terms of them.

As an example, lets look at Uber. Here is a company that on day one is attempting to disrupt a highly regulated industry, with a capital intensive model, and probably looked like a logistical nightmare in the early days. Both founders had past success, but mostly in more traditional Silicon Valley businesses that didn’t require the number of stakeholders and moving parts that are core to Uber’s business model. Now a large percentage of investors would turn down the investment citing the reasons above. A small percentage would invest as a call option because of the team and the vision. But an even smaller percentage would quickly approach (most likely unconsciously through pattern recognition) the investment in decision tree form. It’s very high risk for a number of reasons but because you have a great team that slightly reduces the risk. However, the magnitude of the outcome is potentially so massive that if you invested in 25 companies with the same value of the investment opportunity, chances are you’d have one extremely impactful outcome and as a result, a well-performing fund.

To provide an example, take a look at the decision tree I created below evaluating an imaginary company. Keep in mind that I’m keeping this example very simple to demonstrate how a decision tree works.



Peer-to-Peer Guitar Marketplace Decision Tree

The company I’m evaluating is a P2P marketplace for renting guitars. What I did below was start with the probability a large and small TAM. Lets say that the retail market for buying new guitars is \$4 Billion. As a result the likelihood of a TAM that is big enough is relatively low (I used 70% for this example) but maybe this platform expands the market (30%).

From each of those branches, I evaluated the likelihood of a high LTV:CAC versus a low one. On the side of a large TAM I assigned the likelihood of a low LTV:CAC a 60% probability assuming that the supply will be fragmented from demand (if you own a guitar you won't need to rent one) and users are unlikely to use the platform frequently (at some point they would just buy a guitar). I gave the probability of a high LTV:CAC a 40% likelihood because while not it's not likely, maybe a large TAM and highly targeted user base will result in lower cost user acquisition.

On the other side of the tree, I gave a low LTV:CAC a 90% likelihood due to the issues stated above combined with a small TAM. To calculate the likelihood of these scenarios I multiply the probabilities of each scenario to find the probability-weighted outcomes or what is known in probability theory as the *expected values*. Given this tree, the most likely scenario by an order of magnitude (63%) is small TAM, low LTV:CAC. This is an easy pass for the investor who is making these assumptions, however TAM size and LTV:CAC aren't anywhere near conclusive ways to evaluate a marketplace. If you want to see how complicated this can get, take a look at [Bill Gurley's 10 Factors To Consider When Evaluating Digital Marketplaces](#) each of which are criteria to evaluate these marketplace businesses.

Now where this gets tricky is assigning monetary values to each outcome. This is where venture capital as an asset class is quite different and Peter Thiel's power law comes into the picture. Not only do you want to calculate the likelihood of success evaluating each company, but success needs to be impactful enough to significantly impact returns relative to fund size and ownership percentage. Bill Gurley did an excellent job articulating Uber's potential market size [here](#), but this is where many investors get tripped up because it's so difficult to make accurate predictions. In order to evaluate a digital marketplace using Bill Gurley's criteria and decision tree theory, you need branches for each of the 10 factors, while also assigning potential company value to each scenario and on top of it all, you need to be accurate enough with your criteria to make the right decisions. To once again quote [Charlie Munger](#),

"It's not supposed to be easy. Anyone who finds it easy is stupid."

Decision trees represent the *expected monetary value of the probability-weighted average of the outcomes*. The best of the best, the top 1% of best performing venture investors instinctually make decisions based on probabilities of success while properly understanding the magnitude of the outcome if success is achieved.

Lower risk investments with higher probability for success, but where success isn't massive, don't typically make good venture investments because the time horizon to reach liquidity for investors in early stage companies is lengthy (except for the rare large early exit which is difficult to plan for) and therefore the asset class is only worthwhile for LPs if they can earn significant multiples on their capital. It also goes without saying that there is no such thing as an early stage venture investment with a high likelihood of success. Given the interpersonal dynamics, market risks, and unforeseen hurdles and roadblocks core to all early stage startups, the probability of success in any case can only be so high.

While these decision trees are useful for evaluating each investment on an isolated basis, truly great investors understand that each investment decision is a branch of the fund's decision tree and impacts likelihood of strong returns. The idea to write this post came after reading [USV's updated investment thesis](#). Here is a thesis-driven firm that forms and publishes very clear and concise ideations on the types of businesses and markets they're investing in. USV invests in a number of companies that fit in their thesis of where they believe the best opportunities to find the best companies that will be built over the course of the next several years. Knowing full well that VC is a [hits business](#), USV makes a number of investments within their thesis, and by doing so, the risk/reward each investment improves the probability-weighted average of the outcomes as a whole. By accurately predicting where the next massive businesses will be built and leveraging the thought process behind decision trees and expected value on multiple levels, USV has consistently performed amongst the very best in the entire industry.

I recently had a discussion with a friend about markets that have five or six funded competitors with similar strategies and the thought process behind investors funding later market entrants. My belief is that bad investors do this because they really like the idea and didn't have the opportunity to invest in the first or even second or third mover, but decide they want a horse in the race regardless. Average to good VC's might try and mitigate risk by investing in multiple companies that are one of many in different markets and therefore probability says they'll be in at least one market leader. Great VC's making these investments are able to weigh probabilities and risks, tipping the scales in their favor through an unfair advantage they recognize in the team or strategy and their overall portfolio.

The hardest concept to internalize here is that a VC portfolio in itself is a decision tree with branches weighing each potential investment and follow-on investment. This is what is so complicated about the minds of great investors. The best VC's aren't deliberately creating decision trees. It's completely natural

and unconscious. I'd also argue that misunderstanding this concept is a key reason why non-traditional private investors missed on the valuations of the "unicorns".

This is why I believe such a small cohort of VC investors have traditionally made up such a large percentage of returns in the asset class. This is referred to as "Matthew Effects" whereas the more good investments VC's make, the better their network becomes, the more helpful they become and the more entrepreneurs want to work with them. The other key attribute of this is pattern recognition and the reason VC is often referred to as an apprenticeship business. Pattern recognition isn't just about learning to read people, companies and business models but also being able to approximate probabilities as it relates to key risks, potential black swan scenarios, and magnitude of potential outcome. The decision tree model isn't quite as useful without accuracy or properly attributing risk/reward (as evidenced by Bill Gurley's Uber market analysis).

Good venture investors have differentiated networks, roll up their sleeves, provide real, useful value to companies, and have built great brands which allows them to see a high number of quality deals. But the truly great investors understand how to take all of that and build a portfolio with proper risk and reward, providing a higher likelihood of hitting the grand slam (or multiple home runs) and therefore continue to perform fund after fund, new technology after new technology, and cycle after cycle.

Notes:

- All of my comments are based on inferences. I have no inside knowledge of any investors or companies mentioned above.
- I didn't go into detail on the long list of other attributes of great investors because it's been written about and discussed at length. It's no coincidence that the same investors are involved with so many great companies. Investment decisions are a small part of the equation.