Technology portfolio modeling, generative AI, and optimal fund design

Marcelo Rinesi

2023-11-27

Technology portfolios: absolute and relative valuations

First, a disclaimer: our use of mathematical notation is in the pursuit of clarity and to facilitate "importing" ideas from related fields - no interesting mathematical machinery will be used or quantitative results implied[1].

We look at companies, including companies at the design stage, as *portfolios of technologies*, so their overall technological capability is modeled as

$$\sum_{i \in 1..N} w_i a_i$$

With w_i relative technological investment weights ($\sum w_i = 1$ although later we can consider differences in total budget) and a_i the individual capabilities of each technology, which we model as additive random variables with a_i expected utility and δ_i implementation difficulty. Conceptually, δ_i imposes a form of non-linearity: $w_i < \delta_i$ makes it very unlikely for the utility of a_i to be realized at all.

When total budgets are low enough that we can ignore the $w_i >> \delta_i$ case, it's tempting to think of the valuation of this technology portfolio as simply

$$V = \sum_{i \in 1..N} w_i a_i$$

But this doesn't apply well to early investments in the tech industry. While a bank considering a loan might only look at the overall competence of the business, i.e. its absolute technological capabilities, tech-oriented investors are unlikely to be interested in a company that's at or below the average tech level in an industry, no matter how high this average level might be.

Writing $\bar{w_i} = \frac{1}{Z} \sum_{k \in 1..Z} w_i^k$ the average weight of a_i in the universe of competitors, we model this as

$$V = \alpha \sum_{i \in 1..N} (w_i - \bar{w_i}) a_i + (1 - \alpha) \sum_{i \in 1..N} w_i a_i$$

Here α represents how competitive is the valuation model, that is, how much valuation depends on the difference between a company's overall technological capabilities and those of the average industry competitor. The choice of α if of course deliberate: it's the excess performance of a portfolio of technologies measured against the "market average."

[1] At least not at this stage.

Optimizing a portfolio of technologies

So much for the overall setting. What does standard portfolio theory tells us about how companies should choose their technology portfolios?

A first observation is that the role of the risk-free asset in standard finance is taken here by the analogous zero-difficulty technology. We pick a_0 to refer by an abuse of notation both to the set of zero-difficulty technologies ("electricity," "GMail," etc) and to their overall utility. By definition, $\delta_0 = 0$ or close enough.

We can characterize non-tech companies as those with $w_0 \approx 1$; they aren't companies that don't *use* technology, but they are the ones that don't use any *difficult* technology: a neighborhood laundromat might use a 4G cellphone to call me to let me know there's a problem with my laundry, but that doesn't make it a tech company. This is an optimal strategy for tech-risk/difficulty-averse companies with $\alpha = 0$.

This is worth emphasizing: *Large technological capabilities alone don't make something a tech company, specially in large a_0 industries and societies.*

Another important fact to understand tech portfolio dynamics is that a_0 isn't static: technological change manifests not just as new, high-utility/high-difficulty technologies, but also as the normalization of existing technologies. The zero-difficulty technological capabilities of a contemporary company are much higher in absolute terms than they were at any time in the past, and we expect this to continue to be so in the future. *The basic assumption of the modern economy, and in fact our civilization, is that a_0 grows over time.*

This has strong implications for technology portfolios. Consider what happens when a_0 shifts to $a'_0 > a_0$ — again, this isn't an hypothetical or infrequent event but rather the underlying "pulse" of the economy. For $\alpha = 0$ companies this is, straightforwardly, a direct win. For $\alpha \approx 1$ companies, which is a better model for startups and the "tech" industry in general

$$V_{a'_0} - V_{a_0} = (a'_0 - a_0)(w_0 - \bar{w_0})$$

Perhaps unexpectedly, we see that tech companies benefit the most from steady zero-free technology improvements whenever they are *overweight* zero-difficulty a_0 technologies compared to their valuation peers. This isn't the whole story —appetite for risk and difficulty is profitable for a reason— but it's still important to note that, *ceteris paribus*, the faster you think zero-difficulty technology capabilities will rise, the more weight you should consider giving to a_0. The analogous case in finance is a commonplace: the higher the risk-free interest rate, the more weight you want to give to bonds in your portfolio.

The analogy between a_0 and the risk-free interest rate is flawed in other ways, though. The secular rise in a_0 is *good* for the economy as a whole: it's a big part of how we define technological progress. But we see a similar form of investment crowd-out in the fact that, due to standard efficiency arguments, for any technology with $\delta_i > 0$ and $a_i < a_0$ the optimal w_i is zero: it doesn't make sense to invest in difficult and risky technologies that underperform the zero-difficulty technology bundle, and as the utility of the latter rises over time, many specific technology bets become irrational. We could improve very much vacuum tube-based computer technology, but it doesn't make sense to.

To clarify some of the implications of this we make three formal assumptions. The first one is that we have ordered the technologies a_i according to δ_i, so $i < k \rightarrow \delta_i < \delta_k$. Second, by the argument above, we ignore technologies outside the efficiency frontier — except for a_0. That is, we will assume that no technology is dominated, in the sense that

$$\forall i > 0 : \nexists 1 \leq k < i : a_k > a_i$$

A third assumption looks innocuous but isn't: we take the $a_{i>0}$ difficult technologies to be *uncorrelated*: successful application of one is independent from any other; w_i, in other words, only counts towards a_i. Despite this simple statement the correlation structure of technologies in a portfolio can be extremely complex, difficult to ascertain, and depend on very specialized domain knowledge. And yet it's hard to imagine how one can optimize a portfolio well without knowing the correlations between its components!

From here on we will assume that we have already done the difficult an extremely important work of decorrelating our set of technologies, not because this is trivial but because it's a prerequisite to optimization but one that can only be done case by case.

Going back to the importance of a_0, consider again the general portfolio

$$\sum_{i \in 1..N} w_i a_i$$

We define R as the minimal k such that $a_k > a_0$. For any efficient portfolio, then, $w_i = 0$ for $0 < i < R$. Between the zero-difficulty technological baseline and the fields of possible technological bets there's a barren land of technologies less useful than zero-difficulty tech[2]. As a_0 rises over time, then, so does R. All things being equal, the higher the a_0 zero-difficulty baseline, the more you have to move into high-δ, high-a technologies before you find one that is rational to invest on.

A first look at generative AI

We can look at the impact of generative AI (a_g) in this framework. Not *only*, and perhaps not *mostly* at how it works in practice, but as to how it's *perceived*; our model shows when and why perception can be more valuable than performance. We take it to be or to be perceived to be at this point in time as having large a_g and initially large but quickly diminishing δ_g. This has the following stylized implications:

- Early in the process, very large δ_g meant that few $\alpha = 1$ companies, and no $\alpha = 0$ ones, invested on it. High realized a_g led to even larger changes in *relative* valuations, as $\bar{w_g} = 0$ and w_g was very high (either due to concentrated bets or of the sheer size of their budget) in the original crop of generative AI companies.
- As δ_g falls, generative AI becomes a dominant substitute for more of the technology portfolio budget; this is the exponential adoption phase, as more and more $\alpha = 1$ organizations (but also lower-α ones) start using it. This leads to changes in relative valuations, too, but as $\delta_g \to 0$, $\bar{w_g} \to w_g$; intuitively, "if you waited until it's easy for you to do it, you've waited until it's easy for your competitors to do it too." This offers an stylized explanation for the fact that the later crowd of generative AI companies, and their customers, hasn't gotten as much of a valuation bump as would have been expected given the usefulness of the technology.
- As δ_g approaches zero, generative AI is also becoming part of the a_0 suit of technologies: nowadays it's hard to find mass products from large, zero-difficulty, infrastructural tech companies that *don't* leverage some form of generative AI. This implies a rise in a_0, which, as we have seen, benefits the most $\alpha = 0$ companies, and among $\alpha = 1$ ones those with $w_0 > \bar{w_0}$.

In this stylized model an optimal strategy for a company informed by generative AI requires two responses:

- Make sure that the a_0 suit of zero-difficulty technologies it uses leverages generative AI adequately (organizations and individuals have strong inertia in their tool selection, so staying at the efficiency frontier isn't automatic ever in the zero-difficulty case).
- Adjust their technology portfolios for a higher R.

The second point is probably not stressed enough in most commentary about generative AI: Its impact, like that of any new low-δ, high-a technology goes beyond "make sure you're using it." It *shifts upward* the minimum level of risk you are willing to take, or rather the minimum level of utility you are aiming for, in any efficient technology portfolio. For $\alpha << 1$ companies this might put R beyond what they are comfortable with - the new a_0' is so high that the optimal choice is to not invest on any risky technologies and to effectively become $\alpha = 0$. For $\alpha = 1$ companies the optimal response is the opposite: more difficult and more ambitious bets (or more ambitious and therefore, at the efficiency frontier, more difficult bets) with a high enough to increase the overall utility of the portfolio relative to others who have gotten "for free" the $a_0 \to a_0'$ upgrade and benefited more from it due to a higher w_0.

[2] All of this rests on crude one-dimensional additive utilities, but this does no more violence to reality than most simple economic models, and it's hopefully at least suggestive of real trade-offs.

Optimizing funds, not companies

We can say more about technology portfolio design, this time at the level of the *portfolio of companies*. The relevant financial intuition is that for a given level of risk appetite a basket of uncorrelated options is preferable to an option on the basket as a whole. To make this more precise, recall that a startup investment only succeeds when valuation is a high multiple of investment; it's by definition a high risk/high reward trade. In terms of our simplified model, this translates to requiring high *relative* valuations. This means that the part of our portfolio that contributes to relative valuation will have the form

$$\sum_{j \in J} (w_j - \bar{w_j})\alpha_j$$

Where for every $j \in J$, $j > R$, $\bar{w_j} \approx 0$ and α_j is high compared to competitors' portfolio valuations.

At this point the non-linearity of $w_i \alpha_i$ with respect to δ_i becomes critical. Consider the technology index M^k, the largest i such that $w_i^k > 0$. This is the technology with the highest value a_M^k upon realization an startup k has in their portfolio (for our portfolio we'll just use a_M and w_M), but also the highest δ_M. Consider also a fixed budget $W = \sum w_i$ that's relatively similar for all competitors [3].

How large can w_M be compared to W to maximize relative valuation? The answer depends on multiple factors, particularly risk appetite, but there's an slightly different question that's easier to answer: How large can a_M be compared to W to maximize relative valuation in an environment of multiple competitors with very large risk appetites?

This case is clearer. Call a_W the technology with the lowest δ such that $\delta_W \approx W$. Any portfolio that significantly distributes W across multiple technologies will have $w << W \approx \delta_W$ and therefore, if efficient, $w_W = 0$ This means in turn that the *best technology* it can hope for is $a_M << a_W$. Assuming indifference to risk, the dominant strategy is instead $a_M = a_W$; in other words, to put all your budget W into the highest-a with, plausibly, $\delta < W$ but not too far below it - and with, as far as you know, $w^k = 0$ or at least (if inefficient) $w^k < \delta$ for every competitor k.

The informal argument is also clear enough: if you don't care about risk and are competing against others who don't care about risk either, you take the largest bet you can reasonably have a chance of winning because with anything less than that you lose the competition even if you win the bet.

To add a bit of pragmatism, we acknowledge that startups need *some* technology other than the one they are betting on: even high-end labs have email. By the same concentration-of-budget/risk argument as above, we relax the optimal technology portfolio to

$$w_0 a_0 + w_M a_M$$

With w_0 as low as pragmatically possible, and M such that $W - w_0 \approx \delta_M$.

Funds, of course, do care about risk. They have a budget W of their own, and want to deploy it across $J > 1$ different uncorrelated bets (if $J = 1$, that's not a fund, that's a founder). But we do know how the optimal way to deploy $\frac{W}{J}$ looks like, so the optimal fund, as a set of technology portfolios, will be

$$\{w_{j_0} a_0 + w_{j_1} a_{j_1}\}_{j \in J}$$

With $j_1 \neq j_1'$ for $j \neq j'$ and therefore a_{j_1} and $a_{j'_1}$ uncorrelated between different companies, w_{j_0} as low as possible in every company, and $w_{j_1} \approx \delta_{j_1} \approx \frac{W}{J} - w_{j_0}$ [4].

[3] Many strategic considerations become much easier, or at least very different, if you can outinvest all your relevant competitors, so that case isn't in our scope of analysis.

[4] This describes equal weighting between companies; it's conceptually straightforward to extend the idea when it's otherwise, and it doesn't impact our main conclusions.

Informally, under plausible conditions the optimal early stage fund then is a set of startups each going all-in on a single, uncorrelated technology that's the most difficult one they can plausibly hope to realize with their budget (minus some budget for run-the-business zero-difficulty technologies).

This is not how most early stage startup companies look like! The cultural and psychological expectation is for startup companies, and founders even more so, to be wide-ranging neophilic experimentalists constantly looking for new cutting-edge difficult technologies to prototype and incorporate into their companies; financial pragmatism, in this milieu, means "without blowing up too much the total budget."

Cultural expectations aside —and it's important to note that cultural expectations are enough, and perhaps the most relevant factor in any case— this isn't obviously wrong as an strategy for individual companies. Relative valuation is the most critical factor at certain points, but the short-term concern of a startup is continued survival, which is driven to a large degree by *absolute* valuation, i.e., absolute technological capabilities. And just as a standard game theoretic argument suggests that narrow bets are optimal for the usual relative valuation case, for absolute valuations the standard financial argument favors a balanced portfolio

$$\sum w_i a_i$$

with $w_i > 0$ for multiple technologies with middling a_i and δ_i ; there are cultural and competitive constraints that make low-δ technologies unsuitable, but a too-large δ_i risks either wastefulness or more risk than desired by the company (as $w_i \approx \delta_i$ is required for a chance of a_i to be realized).

This is optimal for founders and startups once they are inside the fund, and it's to a degree even baked into what funds look for in founders and their companies (multiple advanced technologies add to the perception of an advanced startup) but it's in most cases more conservative than implied by $\alpha = 1$ fund strategies.

Familiarity breeds (some) valuation

We can build a basic model of "cultural and psychological expectations" by adding a third term to the general valuation model:

$$V = \alpha V_\alpha + \beta V_\beta + \gamma V_\gamma$$

$$V_\alpha = \sum_{i \in 1..N} (w_i - \bar{w_i}) a_i$$

$$V_\beta = \sum_{i \in 1..N} w_i a_i$$

$$V_\gamma = \sum_{i \in 1..N} (w_i - \bar{w_i}) {w_i}^\star$$

Where $\{{w_i}^\star\}$ describe the technology portfolio of a "flagship" or reference company: the most famous, largest, most successful, etc. company in a given competitive environment, and V_γ is therefore higher the closer the portfolio is to this reference company versus the average portfolio in the reference environment.

Real-world observation suggests that γ is often far from zero; in many ways it operates as a cognitive shortcut, so it's naturally most relevant for quick intuitive evaluations, and it's often a very large factor in access to relevant media, conferences, etc. Reasonably, founders (and funds operating under the same logic) will seek to maximize V_γ at the stage where obtaining attention (e.g. to seek first or early-stage funding) is most critical, and keep it high until V_α is high and realized enough to be visible on its own to the then-relevant audiences (late-stage funds, public markets, acquirers, etc).

There's a logic here that makes this a good strategy in the short term. Flagship companies often have high V_α and/or V_β (which is what makes them a flagship company) and therefore maximizing V_γ by shifting $\{w_i\} \to \{w_i{}^\star\}$ increases the other two valuation terms [5].

However, competitors old and new are in the same situation and with access to the same information channels. So $\{\bar{w}_i\} \to \{w_i{}^\star\}$ at the same time, and $\sum(w_i - \bar{w}_i) \to 0$. This impacts portfolio valuation in two different ways:

- $V_\gamma \to 0$: the closer everybody is to a reference company, the less this matters to your valuation - it's just what the competitive universe looks like.
- $V_\alpha \to 0$: the closer everybody is to a reference company, the closer they are to each other, and the lower the possible relative valuations.

Reviewing this mechanism, we can summarize it as γ-optimal portfolios having an intrinsic tendency to lose performance over time in both large-γ and large-α valuation environments. As online (and, in general, high-frequency) attention mechanisms tend to approach the large-γ model, this contributes to a boom-and-bust cycle where a high-V_α company begets high-V_γ followers with, initially and for the earliest adopters, somewhat high V_γ and high V_α, but with both valuation terms falling quickly as technological and profile advantages disappear.

Locating and pursuing a newly successful flagship company (going from "the Uber of X" to "the OpenAI of Y") just repeats the same dynamic.

A second look at generative AI

To finish these notes we can use this extended model to take a closer look at generative AI. A key point is that "generative AI" isn't a single technology [6]. Consider a set of generative AI technologies $\{a_{g_j}\}$ with g_j ordered according to j, and a $a_{g_0} > a_0$ technology representing whatever is built-in of generative AI in existing zero-difficulty technology (e.g. as filters on image edition software, or as manually-reviewed autocomplete tools in text editors).

We assume a constant γ factor $w_g{}^\star$ for all a_{g_j} : in other words, the reputational impact of a generative AI technology at this point is driven by the fact that it's *a* generative AI technology, not *which sort of* generative AI technology it is.

As every efficient technology portfolio is already at its budget limit, we will look at $\Delta^{i,j}$ defined as the the valuation impact of setting $w_i = 0$ and $w_{g_j} = w_i$ — i.e., moving the budget for a_i to a new generative AI technology a_{g_j}.

$$\Delta^{i,j} = \alpha(V_\alpha{}^{i,j} - V_\alpha) + \beta(V_\beta{}^{i,j} - V_\beta) + \gamma(V_\gamma{}^{i,j} - V_\gamma)$$

The first term is

$$\alpha((w_i - \bar{w}_{g_j})a_{g_j} - (w_i - \bar{w}_i)a_i)$$

The second term

$$\beta w_i(a_{g_j} - a_i)$$

And the third term

$$\gamma((w_i - \bar{w_{g_j}})w^\star_{g_j} - (w_i - \bar{w}_i)w_i{}^\star)$$

[5] In practice, there's sometimes a misunderstanding of the true reasons behind the technological performance of a company, specially at different scales.

[6] This is even more so for "AI," which is frustratingly often used as a synonym.

The β term is straightforward: it tells us that the absolute improvement in performance is proportional to the difference in capabilities between a technology and its "generative AI replacement." In particular, as $a_{g_0} > a_0$ by assumption, $\Delta^{0,0} > 0$: at a bare minimum, companies can improve their absolute valuation by replacing any zero-effort technology with the generative AI-enabled version. The impact of other changes depend on $\delta_{g_j} \leq \delta_i$: moving budget from an old to a new technology only makes sense if the new technology is better *but* it's also no more difficult to implement. This is always the case for $\delta = 0$, but no necessarily true otherwise.

The γ term is slightly trickier. In general — *at the moment* — $w^\star_{g_j} > w_i{}^\star$, so the change in this valuation term will be positive whenever $(w_i - \bar{w_{g_j}}) \geq (w_i - \bar{w_i})$: that is, you outinvest the average competitor on the new technology as much as you outinvested them (if at all) in the old technology. Like every γ-driven strategy, it can be very powerful in the short term and allow to pass through certain filters, but it's time- and context-sensitive, and has decreasing returns over time.

The α term, the change in relative valuation, is the most relevant to the $\alpha = 1$ case. It depends, of course, on the realized technological utility of both a_i and a_{g_j}, but it's also influenced by the pre-existing w weights. Optimization here in a risk-averse portfolio requires moving budget from technologies a_i to equally difficult a_{g_j} technologies that are under-invested in the average competitor compared to a_i, ideally with $a_{g_j} > a_i$, but with this of course not being *sufficient*.

This is for the generalized portfolio, which we have shown is usually suboptimal from a fund's point of view. So we recalculate $\Delta^{i,j}$ for a fund-optimal company portfolio $w_0 a_0 + w_M a_M$ (with $\delta_0 = 0$, $\bar{w_M} = 0$, and $\alpha \approx 1$ by the usual arguments) given the existence of new generative AI technologies a_{g_0} and a_{g_M} . We know that $w_0 a_{g_0} > w_0 a_0$ in a difficulty-/risk-free way, so we can focus on the second part of the α term:

$$\Delta = (w_M - \bar{w_{g_M}})a_{g_M} - w_M a_M$$

It's clear that if $a_{g_M} \leq a_M$ then $\Delta < 0$. We can't also have $\delta_{g_M} > \delta_M$ (or at least not significantly) as, by our model of an optimal fund company portfolio, $W \approx w_M \approx \delta_M$, and investing on a_{g_M} with $\delta_{g_M} > \delta_m$ will make it unlikely to realize a_{g_M} at all.

But the same argument applies to everybody else: either nobody else is investing on w_{g_M} (and therefore $\Delta > 0$ with the assumptions above) or there's a competitive evaluation sub-universe with $\bar{w_{g_M}} \approx a_{g_M} \approx w_{g_M}$ which means $\Delta \approx -w_M a_M << 0$. The same applies in general: an optimal fund portfolio company, in this abstract model, shouldn't bet in more than one technology, and it shouldn't bet on a technology competitors with approximately the same budget are betting as well [7].

We rephrase our answer to the optimal fund response to generative AI. An optimal fund approaching the ideal of a basket of uncorrelated high-return technology options, should

- Shift w_0 budget to w_{g_0} for all its companies; in other words, make sure their zero-difficulty infrastructure is up-to-date with generative AI.
- Shift w_M to w_{g_M} where $a_{g_M} > a_M$ and $\delta_{g_M} \approx \delta_M \approx W$ *and no other competitor is doing the same*, which is unlikely in the current environment. More importantly, δ for a given technology is heavily influenced by the expertise and experience of a team [8], so this budget transfer would in practice by very inefficient; recall that we are in the high-δ region of the technology space, and lack of expertise can be fatal in a way that doesn't apply in the low-δ region. (And once you're in the low-δ region, you're no longer optimizing for $\alpha = 1$).
- Shift technology portfolio design criteria to ensure $a_M \geq a_g$ for the largest g with $\delta_g \approx W$. In other words, generative AI has shifted in some areas the efficient frontier of technological bets, and, at least in the pure $\alpha = 1$ high relative valuation environment, made some portfolios no longer optimal [9].

[7] Much hinges on the definition of what makes two technologies the same or not; this is the technology decorrelation problem that turns out to be critical for this sort of technology portfolio engineering.

[8] Exploring this is a matter for a different note.

[9] As a simplified model conclusion; there are of course sunk investments, search and reputational costs, etc.

Informally, the optimal post-generative AI company in a fund uses generative AI-enabled zero-difficulty technology it did not develop in-house (but can be supported by the fund in a zero marginal cost basis) and concentrates its budget on a single technology bet that's plausibly within its possibilities, but also significantly more capable in a suitable sense than an equally expensive generative AI project. Generative AI therefore both narrows and expands the set of viable companies for a fund at a given investment level: even technology bets that were optimal before are no longer viable, but at the same time high-δ technologies that were, without general awareness, correlated with generative AI have seen their difficulty lowered and are now potential optimal bets. Identifying this newly opened $\delta_M \approx W$ frontier is a way to describe the core challenge of startup design in the new technological environment.