Predictably Bad Investments: Evidence from venture capitalists*

Diag Davenport

Chicago Booth

June 14, 2022

Abstract

Do institutional investors invest efficiently? To study this question I combine a novel dataset of over 16,000 startups (representing over $9 billion in investments) with machine learning methods to evaluate the decisions of early-stage investors. By comparing investor choices to an algorithm's predictions, I show that approximately half of the investments were predictably bad—based on information known at the time of investment, the predicted return of the investment was less than readily available outside options. The cost of these poor investments is 1000 basis points, totalling over $900 million in my data. I provide suggestive evidence that over-reliance on the founders' background is one mechanism underlying these choices. Together the results suggest that high stakes and firm sophistication are not sufficient for efficient use of information in capital allocation decisions.

*I owe many thanks to Rafael Batista, Berkeley Dietvorst, Nicholas Herzog, Sam Hirshman, Jessica Jeffers, Steve Kaplan, Yueran Ma, Cade Massey, Sendhil Mullainathan, Devin Pope, James Ross, Donovan Rowsey, Josh Schwartzstein, and attendees at annual meeting of the Society for Judgment and Decision Making for helpful comments and conversations. This research was generously funded by the John and Serena Liew Fellowship Fund at the Fama-Miller Center for Research in Finance, University of Chicago Booth School of Business.

1 Introduction

The most influential companies of today all began as startups in need of capital. Investors looking to capitalize on the next generation of superstar firms want to do two things: fund startups that will eventually be superstars and avoid lemons. In other words they face a prediction problem (Kleinberg et al. 2015). Previous work has explored what influences investor choices[1] as well as descriptive and causal evidence on the financial performance of investor portfolios[2]. However, little work has been able to tie the two together to identify systematic choices while quantifying their impact on performance. For example, in recent work, Zhang (2020) studies investor discrimination through field experiments, but it isn't clear whether this behavior is costly to the investors. This paper provides evidence that approximately half of VC investments are predictably bad relative to several outside options thus costing firms over $900 million.

Consistent with the predictive nature of investor decisions, I take a machine learning approach to evaluating the quality of their decisions. I first train an algortihm based on early information about each firm to predict the late-stage exit of the firm. The central eeconometric challenge with this exercise is the selective label problem (Kleinberg et al. 2018a)—we see the performance of firms that received investment, but we don't see the performance that would have been realized by firms that didn't receive investment, if they had in fact received investments. I leverage the fact that the problem is one-sided and focus my evaluation on the firms that received investments. Were those investments good predictions or false positives? Importantly, I take the ex ante view in answering these questions.

Using data from Pitchbook on the universe of startups accepted into the top accelerator programs, I first document new descriptive evidence on the venture capital (VC) industry, documenting major accelerators, investors, and markets that startups participate in. Next, I find that a standard off-the-shelf algorithm can strongly discern between startups that are more or less likely to have exits. That an algorithm can detect a statistically reliable pattern opens the scope for investors to have returns not by luck, but by selecting on quality. Indeed, I observe a nominal return of 79% among VCs, which outperforms the S&P500 by 15

[1] Many related things are known in the venture capital space. Important work includes Gompers et al. (2015) who surveyed 79 private equity investors with combined AUM of over $750B about their practices in firm valuation, capital structure, governance, and value creation. An incomplete list of other factors that have been explored includes product-market fit (Hellmann and Puri 2000), syndication strength (Lerner 1994), CEO personality characteristics (Kaplan and Sorensen 2017, Kaplan et al. 2012), pitch delivery (Hu and Ma 2021), and perceived passion (Gompers et al. 2020).

[2] For example, the effect of VC monitoring on performance (Hellmann and Puri 2002; Lerner 1995; Kaplan and Stromberg 2001, Kaplan and Strömberg 2004).

percentage points over the same period. To evaluate room for further outperformance, I use the algorithm to build counterfactual portfolios forgone by investors by pruning predictably bad investments and opting for a more standard outside option (a public equity or bond instrument). I find that despite the fact VCs outperform the market on average, the returns are driven by the top half of the predicted quality distribution. By dropping the bottom half of investments and instead investing in the market, returns would have increased by 7 to 41 percentage points. This qualitative finding is robust to a set of outside options (stock market or bond market). Together the results suggest that there is significant room to improve how venture capitalists select into investments.

Finally, I investigate why investors consistently invest in companies that are predictably doomed to fail. I provide evidence for one key bias: overweighting of human capital information. I find that investors overweight the importance of founder characteristics, especially for lower-quality firms. I explore several mechanisms underlying these poor investments. I assume throughout the main analyses that investors and their principals attempt to maximize returns, making the deviations I document persistent mistakes. I discuss relaxations of this assumption in Section 6 and explore alternative interpretations.

Notably, my main empirical analyses are closely related to independent but contemporaneous work from Lyonnet and Stern (2022) (LS) as both closely follow Mullainathan and Obermeyer (2022). While my qualitative conclusions are similar to those of LS, I note several key differences. First, I emphasize that investors as well as the researchers studying them have focused too much on the jockey and not enough on the horse; instead of documenting specific biases about how founder information is prioritized, I find that founder information as a category is of second order importance, relative to information on the business itself (i.e., the horse). Second, I observe the actual investments made by investors, which allows me to calculate actual returns and estimate the economic magnitude of returns left on the table. Third, I restrict attention to the consideration set of venture capitalists by focusing on startups that participate in top accelerators, whereas LS considers a highly inclusive definition of startups, studying (a representative sample of) *all* new companies in France. Finally, I study startups competing for funding in the US VC landscape which is significantly more developed and competitive than that of France. Showing that the results hold in the US underscores the persistence and importance of my documented biases in highly competitive markets.

This work contributes to several literatures. First, a growing literature on **behavioral industrial organization** has assessed whether biases can persist in equilibrium in markets with large stakes, sophistication, and expertise. (List 2003, List 2006, Benz and Meier 2008, Strulov-Shlain 2021). In particular, I emphasize deviations that behavioral firms are making from a benchmark of profit maximization as in DellaVigna and Gentzkow (2019), Bloom and Van Reenen (2007), Di Maggio et al. (2022) and Hortaçsu and Puller (2008). I show that highly costly behavior can persist even among firms with strong financial incentives. Startup investors appear to be susceptible to the same attentional and information processing biases observed in consumers by both psychologists and economists. (Gabaix 2019, Howell 2020, Kahneman 2011, Enke 2020)

Methodologically this paper further demonstrates the value of **machine learning to answer economic questions**. My results suggest that venture capital investing is yet another setting that can be improved by the human+machine paradigm. (e.g., Dawes et al. 1989, Kleinberg et al. 2018a) In addition to being practically useful to investors, this approach has generated useful research insights across several fields in economics, such as health, law and economics, labor, and development, but research on venture investing has yet to fully leverage these tools.[3] (Mullainathan and Obermeyer 2019, Kleinberg et al. 2018b, Jean et al. 2016, Bansak et al. 2018). In addition, since the most influential data to the model is text, my approach adds to a newer, methodological literature on "text as data". (Gentzkow et al. 2019, Gentzkow and Shapiro 2010, Erel et al. 2021, Ke et al. 2019, Kogan et al. 2009).

Finally, the source, persistence, and opportunities for future returns has long been a fundamental question in **finance**. (Fama 1965) Of particular interest is whether returns can persist, which can be framed as distinguishing between investors getting the prediction problem right or just getting lucky. (e.g., Kaplan and Schoar 2005, Nanda et al. 2020) My results suggest that, at least in private markets, there is significant scope for persistent performance, which would be realized by investors if they relied less on the founders, informing the "horse versus jockey" debate. (Kaplan et al. 2009, Zingales 2000, Gompers et al. 2020, Rajan 2012)

The rest of this paper proceeds as follows. Section 2 provides a framework for assessing the startup landscape. Section 3 describes the data and how I use it in an ML framework. Section 4 documents new empirical facts about the startup lifecycle. Section 5 evaluates realized and forgone returns. Section 6 explores the underlying causes and how these mistakes persist. Section 7 concludes.

[3] Previous work has used machine learning to predict a variety of outcomes (Ghassemi et al. 2020; Ying et al. 2021, Bento 2018, Żbikowski and Antosiuk 2021, Krishna et al. 2016, Hu and Ma 2021) but none of these have taken into consideration what information was plausibly available to investors and then used that to evaluate investor actual choices.

2 Framework

2.1 Startup Lifecycle

To guide the empirical analysis, I summarize the startup lifecycle with a simple two-period game (depicted in figure 1):

1. **Prediction policy problem**: Early-stage investor evaluates the startup and decides whether to invest

2. **Outcome realization**: Startup realizes its fate. A set of late stage investor decides to invest giving the early stage investor an opportunity to liquidate the investment and realize gains, if any.

Formally, startups, indexed by i, are evaluated by investors, indexed by j, in period 1 and have a late stage exit in period 2. Further, let there be characteristics of the startup X (which are observed by both the investor and the researcher) and Z (which are observed only by the investor); both sets of characteristics may change over time. The key choice for the investor is to choose a level of funding $f \geq 0$. Then, in period 2, the startup realizes an exit with probability p, which obeys

$$p_i = \theta_{i,t=2} + T_{i,t=2} \tag{1}$$

where $\theta_{it} = \theta(X_{it}, Z_{it}) \geq 0$ is a measure of the intrinsic quality of the firm and $T_{it} = T(X_{it}, Z_{it}|f_{t=1}) \geq 0$ captures the treatment effect of the funding f the startup received in the first period.[4]

As is standard, the investor is unable to directly observe θ and T. However, in this setting the difficulty for the investor is magnified because θ is a function of $(X, Z)_{i,t=2}$ which is potentially different than the information available to the investor at the time of investment $(X, Z)_{i,t=1}$.[5]

2.2 Investor Choice and Payoff

For a given level of funding f, the firm's expected return is

$$r(\theta, T) = pM$$

[4] I assume T is nonnegative, but make no assumption about its shape.

[5] The calendar time that typically separates period 1 and 2 is five years

where M is the multiple on invested capital the investor would receive in the event of a successful exit. Return-maximizing investors then face a straightforward investment decision: invest in firm i if $r(\theta, T) \geq \rho$ where ρ is the expected return from some outside option. Investors then need to form some prediction of r,

$$\hat{r}^h(X, Z, b)$$

which is a function of the information (X,Z) available to the investor as well as a vector of potential behavioral distortions b, including non standard preferences, belief distortions, or attentional constraints.

2.3 Evaluating Investor Choice

An alternative to forming human predictions of r is to build an algorithm that can generate predictions $\hat{r}^a(X)$. This approach has several benefits to the researcher. First, it employs an analysis that the firm could have and may have done in order to improve investments. Second, such an algorithm has the potential to proxy complex prediction rules made by humans. Third, the algorithm is not subject to b. That is, there is no opportunity for human distortions to influence predictions.

However, there are also two key drawbacks to this approach as discussed in (Kleinberg et al. 2015):

1. Selective labels: don't know counterfactuals since neither T nor θ are known
2. Selection on unobservables: Z is unobserved and potentially influencing exits and investments.

To address both problems I form predictions on those firms that received investment and assess which of those should have been avoided because $\hat{r}^a(X) < \rho$. In other words, my approach allows me to identify false positives; no further statements can be made about 'missed opportunities' without further assumptions, which I leave to future research.

To see how this approach addresses the central econometric challenge, note the relationship between p from equation (1) and the data we observe. Regardless of whether the startup received investment, we can define a binary realization Y of exiting. Then for any cell X, $\bar{Y}$ is an unbiased estimator of p iff all startups in the set received funding. For any cell X, $\bar{Y}$ is an unbiased estimator of θ iff no startups in the set received funding. In other words, without directly estimating the treatment effect, by construction we know that treated firms have realized their treatment effect T, whereas unfunded firms are relying strictly on their inherent type θ. This is an instantiation of the selective label problem with a mixed prediction problem (Kleinberg et al. (2015)). As a result, the algorithm generating $\hat{r}^a$ can only credibly be estimated and evaluated on firms that did in fact receive funding.

With this framework in mind I will take systematic decisions to invest when $\hat{r}^a < \rho$ as evidence of a predictably bad investment.

3 Data and Methods

The sole data source for this paper is Pitchbook Data. Pitchbook is a subscription data provider widely used by investors for information on deals, companies, and other investors in private capital markets. This data is generally useful to academics because it provides a thorough view of the private market, but it is especially useful for this paper because I will rely on data readily available to the investors I study, emphasizing the extent to which investors misuse available (albeit costly) information. There are three categories of information that I synthesize from the Pitchbook data: *finances (e.g., Revenue, EBITDA, total capital raised), founder information (e.g., educational background and previous experience)*, and *company/product description*. I describe the construction of the sample and relevant variables below.

3.1 Defining the sample

I begin by identifying every startup that participated in any of the top 100 accelerator and incubator programs between 2009 and 2016.[6] Many of the companies will be familiar to the reader and include Airbnb, Doordash, Stripe, Dropbox, Coinbase, Instacart, and Uber. For each of those 16,054 firms, I then construct a dataset of all equity deals known to Pitchbook within the first five years of completing their accelerator program. One key motivation for pre-defining a set of firms of interest and then following their valuations over time is to avoid any survivorship bias which would severely limit the interpretation of any results.

Because my primary empirical strategy relies on prediction techniques, it is important to distinguish between the subsample used for training the algorithm and the subsample used to evaluate it. All models are trained on companies that completed accelerators before 2014. To avoid overfitting concerns, all results and figures are based exclusively on the hold out

[6] Using the Pitchbook data, investors who are tagged to the Accelerator/Incubator investor type produces about 7,000 investors which I sort by the number of Accelerator/Incubator DEALS they have participated on, then pulled the top 100

data that covers 2014-2016. There are several advantages to taking a time-based sample splitting approach. First, it allows me to mechanically avoid any hindsight bias. Second, the empirical exercise simulates what forward-looking investors could have done in real time. And third, the analysis implicitly tests for the stability of predictiveness over time.

Notably there are two reasons why predictive quality may suffer when moving from the training set to the test set. The first is the traditional overfitting concern. The second is changes in the covariance structure between X and Y. These two concerns make the current exercise a conservative test of the predictability of exits and returns. But ultimately, the goal is not to reach a certain threshold of predictive accuracy per se. Rather, my central goal is to assess and improve investor decision quality and by building an algorithm based on information that investors had access to, I can simulate different human+algorithm policies that were implementable.

3.2 Raw predictors

For each funding deal, Pitchbook provides numerical data on the financials of the firm and qualitative information about the CEO, *at the time of the deal*. This time-specific nature of the data is central because my primary empirical strategy will make predictions of late-stage outcomes exclusively using information available to post-incubator investors. Accordingly, nearly all predictors will come from the financials and deal information as of participation in the incubator. This includes numerical information such as Revenue and EBITDA, but also includes the terms of participation in the incubator such as the level of investment and the equity stake that the incubator will take in the startup. Collectively I refer to these variables as *financial information*. The data also contain text fields that include biographical and educational information about the CEO, which I will refer to collectively as founder information.[7]

The one data point that is not specific to the incubator deal is the company description. Pitchbook does not store or provide a time-varying description of the company. Instead, for each startup I have a current description of the firm, its product, and activities, independent of its health or status. I use this information in my analyses as a way to extract information about the company's product and thus refer to this field as product information. A key assumption in this approach is that the descriptions found in Pitchbook do not evolve as

[7] To make use of the rich information embedded in the text fields, I employ a step of borrowing techniques from natural language processing in the machine learning literature. Specifically I vectorize the text fields using 1 and 2 grams and reweight using a TF-IDF algorithm, all within the training data. The resulting transformations are then applied to the test set. See Gentzkow et al. 2019 for a recent review on empirical research relying on transformations of text data.

a function of early-stage funding or late-stage success which would mechanically make the descriptions predictive. To the best of my knowledge the company descriptions available to me today are not systematically different from the descriptions available to early-stage investors when they would have been engaged in due diligence. Further, all results are robust to the exclusion of the product information.

Tables 1 and 2 provide summaries of the sample and the key predictors.

3.3 Outcomes

There are two categories of outcomes that will be central to the main results: post-incubator early-stage investment and late-stage exit.

I define early-stage investment as any equity deal in the Pitchbook data within two years of incubator completion that is categorized as "Series A", "Series B", "Seed Round", or "Angel (Individual)" in the Deal Type. For much of the analysis I will collectively refer to investors at this stage as early-stage investors though some analyses will separately identify investors in Series A or B rounds as venture capitalists and other investors as Angel/Seed. 34% of the startups receive early-stage investments.

I then define late-stage exit by assessing whether a startup has any of three eligible late-stage deals/transactions: initial public offering (IPO), merger or acquisition (MA), or any funding round that is categorized by Pitchbook as Series C or later (C+) within five years of accelerator completion. This is a fairly liberal definition of exit because it neither requires a liquidation of a stake nor a positive return. Justifying an IPO as an exit is straightforward as the startup is publicly sold and the early investors have a clear opportunity to liquidate and profit. MA transactions are murkier. While the investor may have the opportunity to sell their shares, it is not necessarily expected that there will be a positive return from the transaction. With an eye towards this I will incorporate in my main analyses measures of financial return, usually multiple on invested capital (MOIC), along with binary indicators of success in order to avoid misinterpreting situations where the early-stage investor successfully sells a stake in a MA transaction albeit at a steep economic loss. C+ transactions present the opposite problem. I limit the eligible C+ transactions to those late-stage rounds that are classified by Pitchbook as "Up" rounds to avoid including deals where the the startup has raised more money, though at a lower valuation, but I cannot observe whether the investor has the opportunity to liquidate their ownership and therefore realize any gains from the transaction. Nonetheless, recent evidence suggests that for the Series C round and later, there is ample opportunity to sell shares on a secondary market. Using this three-pronged definition, 3% of the overall sample has a successful exit and 9% of the startups that receive early-stage investment have a successful exit.

In the sample there are 5,440 early-stage investments totalling over \$9.3B dollars and 497 late stage transactions totalling over \$105B in company value.

3.4 Missingness

Many of the predictors are missing not at random. If my objective was to provide estimates for the influence of any particular variable on an outcome, this missingness would be a major vulnerability for interpreting the results. But since my only goal is to find a predictive signal, the absence of information is potentially a valuable signal in itself. Accordingly, any observation with a missing value for any variable is coded with a filler value so it can be flexibly handled in the algorithm learning stage.

There is also missingness in some of the outcomes which needs to be handled differently. If a startup does not have a Pitchbook-documented late-stage exit by the end of the five-year horizon, then I assume the value of the firm is zero to reflect the realized valuation by the investors and their principals. If a startup has a late-stage exit but the valuation is not known, I also assume it is equal to zero. If the initial stake of the early-stage investor is unknown then I assume they purchased 30% of the company, which is standard for Series A investors.

3.5 Constructing an algorithm

A traditional econometric approach might consist of estimating an OLS model and then using that model to form $\hat{y}$ predictions. This traditional approach will not suffice in this setting for two reasons. First, most of the data is text and OLS would require taking a strong stance on how to code the data. Second, and perhaps more importantly, instead of imposing linearity in the relationships, more flexible algorithms allow for arbitrary relationships and interactions between all the variables and take a data-driven approach to assessing which interactions and relationships are reliable. To that end, I train a model using XGBoost on numerical transformations of the text data combined with the native numerical data.

3.6 Calculating performance and returns

The primary metric that I use to evaluate returns is “multiple on invested capital” (MOIC), which is common in both academic research and industry. It’s calculation is simple:

$$M = \frac{\text{Value of investment}}{\text{Initial investment level}}$$

I describe the specific calculations and considerations for each term in turn.

3.6.1 Initial investment level

For analyses that include MOIC, I restrict attention to those companies where Pitchbook provides data on the total invested equity in the early funding round following accelerator participation. No adjustments or imputations are imposed on these initial investments.

3.6.2 Value of investment

The basic equation for calculating the value of the investment is

$$\text{Value of investment} = \text{Value of firm}_{t=2} * \text{Initial ownership }\% * \text{Dilution factor}$$

Several of these quantities are not provided in the data and therefore need to be imputed or assumed. First, the valuation of the firm at t=2 in model time is defined as the latest valuation associated with a "late-stage" round by the end of 2020. If no such rounds exist then the company valuation is set to zero. If such a round exists but the valuation is not provided by Pitchbook, I set it to zero. Next, Pitchbook typically provides the initial ownership stake of the early stage investors. If it does not then I assume the early stake is 30%. Finally, I assume a constant dilution factor of .75.

4 Descriptive Statistics on Startups, Investors, and Performance

4.1 Accelerators by the numbers

Accelerators are seen as a launch pad for many startups. This section documents new data on how accelerator cohorts have performed from 2009-2016. Table 3 provides several basic statistics for the top 5 accelerators in the Pitchbook data.[8]. The key statistic is the sum of the post-money valuations for firms launched from a given accelerator. The top two (Rocketspace and Y Combinator) support several common, though conflicting intuitions

[8]Appendix X shows the same table for the top 100

about startup investing. First, Rocketspace, while only investing in 3 companies in my data, is ranked at the top. It's sole non-zero investment is Uber. This supports the notion that one right-tail outlier can define the performance of the entire portfolio. The second investor in the table is Y Combinator which has a much broader scope, investing in over 400 companies over the course of my data. These two investors together account for over 70% of the market value of companies in my data, suggesting either a strong ability for these accelerators to discern on type θ or a strong treatment effect T of the community, mentoring and advising that they aim to provide.

4.2 Investors and performance

Table 4 presents data similar to that in Table 3, instead focusing on the investors who seek out alum from the accelerators. The names of the investors in the top firms will be familiar to practitioners: Benchmark, Seqouia, and Accel. These firms again appear to have some ability to discern since they make up relatively small shares of total investment, but comprise outsized shares of market value of portfolio companies. For example, Sequoia Capital only makes up 1% of invested early-stage equity in my data, but its portfolio companies make up 6% of market value in my data.

4.3 Markets

Finally, Table 5 presents the markets that are most represented among accelerator participants. Column "% of All", for example, shows the distribution of markets among all accelerator participants. I interpret differences in the columns as measures of how much certain markets are preferred at different stages of the startup lifecycle. For example, 8.5% of the full sample operates in the "Application Software" market. Whereas only 5.9% of the firms receiving early stage investments are in the Application Software market. This crudely suggests that Application Software is underrepresented in early stage investments and that investors appear to slightly prefer the opportunities in other markets to those in Application Software.

Haven given a broad, descriptive overview of an important collection of startups and the investors who fund them, I now turn to evaluating the quality of their decisions.

5 Evaluating Exits and Returns

5.1 Startup success is predictable

5.1.1 Algorithm can discern success rates

I begin by showing a monotonic relationship between the predicted success of startups in the hold-out set and their observed success, thus demonstrating that success is predictable early in the life cycle of a company.[9] The algorithm has found signal. That startup success is predictable at all creates scope for savvy investors to have persistent returns which remains an outstanding question in the literature.

Figure 2 compares the predicted success measure to the outcome for which it is trained: a binary indicator for IPO, MA, or C+ funding round. The x-axis rank orders startups by their predicted success and the y-axis shows the fraction of firms that achieve an exit. Those firms deemed poor investments by the algorithm successfully attain exits at a rate of 5%, whereas the top predicted investments exit at a rate X times higher, despite the fact that I test the model on companies it has never seen, thus avoiding overfitting concerns. What's especially striking is the implication that what was predictive from 2009-2014 (what the model was trained on) remains predictive on later data (what the model is tested on), allaying concerns about distribution shifts over time.

That the model can predict the outcome it was trained to predict is informative for evaluating the model itself, but does not speak directly to the investors' decision problem. Investors want to maximize the dollar value of the return, not simply the probability of an exit. I improve the analysis by showing the distribution of valuations in each bin of predicted success. Figure 3 shows two striking empirical facts. First, companies with higher predicted success are more likely to have positive valuations. Second, the split between positive, low-value and positive high-valuations shifts dramatically to the right as one gets to higher levels of predicted success.[10]

5.1.2 What predicts success?

What is the algorithm using to predict exit so successfully? Discerning how the inputs to an algorithm relate to its predictions is notoriously difficult. For ease of interpretation, I project the algorithm's predictions onto common variables discussed in the literature, including level

[9] Given the selective label problem I restrict attention to the 5,440 startups that received investments.

[10] For the reader interested in more traditional measures of predictive quality from the machine learning literature, the model's ROC curve is shown in Figure 4.

of funding, proxies for founder education quality, and gender. Table 6 reports the projections onto the model's raw predictions for all startups; Table 7 reports the projections onto the model prediction percentiles among the firms that did receive early investments.

The projections are consistent with the existing work that suggests founder education and prior investments increase the probability of successfully exiting later in the startup lifecycle. However, I emphasize that nothing in these projections can or attempts to speak to a causal interpretation of X on Y. Instead they give some guidance on what is or could be used to form $\hat{Y}$.

5.1.3 On failure rates

Given how well the predicted success positively correlates with various well-observed measures of actual success, one might expect predicted success to negatively correlate with a notion of failure. In addition to the funding rounds data described in the Data section, Pitchbook also provides data on two events that indicate business failure: bankruptcy declaration and documented shut down. I show the relationship between failure and predicted success in Figure 5. Counterintuitively, the figure seems to show no clear relationship between the firm's predicted success and its probability of failing. If true, such a finding might be fatal to the validity of the model. However, given the high observability of the success outcomes and the high potential for mismeasurement and selection issues in the failure data, I interpret this contradiction as evidence that there are severe reporting biases in failure. It appears that those firms that had good chances of success are more likely to formally shutdown their business *in an observable way*. I therefore caution future research against interpreting the failure data as a source of truth without accounting for endogenous reporting.

5.2 Assessing returns

In the following subsections I will evaluate returns through several strategies

5.2.1 Observed returns

The exit rate is higher in the firms predicted to do well by the algorithm, but the returns aren't as straightforward. Returns do not vary consistently with predictions as seen in Figure 6.

Implies contracts and investment terms matter and possibly contribute to returns

One way to explore the quality of the investments is to see if they differ by investor type.

To do this I decompose figure 6 to compare individual investors (Angels) and institutional investors (venture capitalists). These two groups differ in several important ways. One prominent difference is that Angels invest their own money whereas VCs primarily invest other people's money and thus have broader scope for agency problems.

Figure 7 shows that the individual investors overall have lower returns than institutional investors suggesting that the sophistication, access to investments, and incentive scheme of the VC is enough to overcome the potential agency problems.

These results show that VCs outperform two common and intuitive benchmarks: the return of other investors and the average return of the outside option. However, as the framework in section 2 suggests, neither of these benchmarks is grounded in economic theory and therefore do not speak to the true optimal benchmark.

Knowing the true optimal benchmark, that is asserting the exact valuation and investment each investor should have for a given startup is not feasible in my data. The core challenge to such an exercise would be assessing counterfactual contract specifications, which requires knowing the function relating success to funding ($\frac{dT}{df}$). Without estimating inherent types and treatment effect functions, evaluating counterfactual investment contracts on the intensive margin is not possible. However, evaluating counterfactuals on the extensive margin is feasible—what would've happened to f (the observed investment) if it wasn't invested in a given startup and instead was placed in an outside option.

The following subsection explores how well firms are performing relative to something closer to a second-best benchmark by comparing *each investment* against an outside option to construct a return-maximizing portfolio.

5.2.2 Forgone returns: Counterfactual bond portfolio

In the first set of counterfactuals I consider, the VC chooses between an observed investment contract or a hypothetical outside option. In particular, I consider a hypothetical bond that pays 8% simple interest.

One thought experiment to demonstrate how investors might improve their performance with an algorithm is to have the algorithm filter out the bottom k% of human startup selections and replace them with the outside option. Figure 8 graphs the results of such an analysis. Each point on the x-axis indicates a different scenario in which the aggregate investor invests only in the top 1-k% of companies according to predicted success. For example, the left most point is the return realized if the investor invests in the top 100 percent (i.e., all investments), thus yielding the exit rate that was actually observed. Then moving to the right, we observe how the VC portfolio multiple would change as a result of having an increasingly selective k, all the way up to only investing in the top 1 percent of startups.

5.2.3 Forgone returns: Counterfactual SP portolfio

Another approach is to consider a real outside option: the stock market. In Figure 9 I graph the resulting analysis. Instead of witholding investment as above, the investor invests the money in the SP500 index. This exercise is very similar to the Public Market Equivalent (Kaplan and Schoar 2005), but differs in that I do not directly observe cash flows.

In this analysis I find qualitatively similar results, though the returns to selectivity are higher since the return to the stock market option is higher than that of the hypothetical bond.

The two counterfactual portfolio analyses yield three takeaways. First, current VC returns are significantly higher than simple returns from less risky investments. Second, VCs could reliably drop up to 50% of their investments and improve returns by avoiding low-return startups. Third, VCs cannot trim investments all the way up to the top 1% because eventually they start swapping high quality investments for lower returns. There are at least two interpretations of these empirical facts. First, is a literal and extreme interpretation: investors might do well to adopt a selection model close to that of CalTech, selecting only a few, highly promising startups. An alternative interpretation is that these analyses provide suggestive evidence that investors should be seeking out better deals terms. Under either interpretation I conclude that VC returns could be up to 1,000 basis points higher if investors avoided more predictably bad investments, on either the intensive or extensive margin.

5.2.4 Robustness

One prominent limitation in the approach above is the dependence of the analysis on a single ex post realization of the world. To address the concern that my results may not be robust to another draw from the superpopulation of portfolio outcomes under a given investment strategy, I conduct a bootstrapping procedure where I draw a sample of candidate startups (with replacement), implement a constant selectivity decision rule, repeat 2,000 times and report a distribution of statistics stemming from the procedure. Table 8 provides suggestive evidence that the hill-shaped relationship in Figure 9 is not due to chance. Though the estimates are noisy, the vast majority of the range, under any calculation, has the correct sign. Figure 11 provides further support by showing the distribution of MOIC that would have been realized under the current investment strategy versus a more (ex ante) selective one. This figure shows that the returns to switching to a more selective strategy were positive in over 95% of the bootstrapped samples.

6 Unpacking VC decisions

In this section I explore the underlying decision model of VCs in order to understand why they invest in predictably bad investments.

6.1 Cream of the crop vs bottom of the barrel

In Section 5.1 I construct an algorithm to predict whether a startup will have a successful exit. In this section I will train algorithms for a different outcome: will the startup receive an early-stage investment. In particular, I will build an early-decision algorithm on the bottom 30% of firms according to the late-exit algorithm. Then I will build a second algorithm to predict early investments on the top 30% of firms according to the late-exit algortihm. Figure 12 plots the coefficients of each of the models and compares them to each other. Despite the fact that none of the coefficients carry causal interpretations, there are several suggestive takeaways. First, nearly all points lie directly on the axes, suggesting that firms use totally different criteria when selecting good and bad investments. Second, the word tokens along the vertical axis appear to differ systematically from those along the horizontal axis—the model trained on the worst firms appears to prioritize founder details, whereas the model trained on the best firms appears to prioritize product details. When making good investments, investors appear to bet on the horse, but when making bad investments they appear to be betting on the jockey.

6.2 A formal test of overweighting

Are investors relying too much on the jockey? Following Mullainathan and Obermeyer (2022), I test for overweighting of a founder cue by building an algorithm to predict late outcome strictly using the founder information. Then we can ask whether two companies that are equally likely to succeed according to X are differentially likely to receive investment based on $X_{education}$. Since X is a proper superset of $X_{education}$, the null corresponds to no overweighting of the founder information. Table 9 strongly rejects that null and suggests that investors overweight education in general, but especially for lower quality firms.

6.3 Are Investors Making Mistakes?

I have shown evidence that indicates investors are systematically underperforming and systematically over-relying on certain signals of quality. The findings are consistent with misprediction and attentional biases and that is my preferred interpretation. However, other mechanisms could be at play and that would make it inappropriate to label these choices mistakes. For example, the incentive structure may not be well-aligned between principal and investor, and/or investors could have a rich set of preferences that includes reasons to tradeoff financial returns in favor of say interpersonal relationships, administrative ease, or simple effort reduction. These 'omitted payoffs' are known to complicate the interpretation of humans deviating from what an algorithm might recommend. For example, investors may get direct, non-monetary utility from investing in some kinds of companies or founders, or they may simply dislike the psychic costs of discerning between the best and worst companies they are likely to invest in. Further, orthogonal to their own preferences, investors may face external constraints from the capital providers. This could create a situation in which investors invest in the companies that are easiest to defend or justify.

6.4 The source and persistence of biases

I further develop my view on the underlying mechanisms by briefly exploring several plausible economic and institutional features in this setting that allow biases to persist. I argue here that the VC space represents a perfect storm of reasons why investors are likely unaware of their forgone returns and why they may not do much about them even if they did know.

6.4.1 Conventional wisdom

Business icons from Ray Kroc to Jeff Bezos have captured our attention and lend credence to the mythology surrounding founders and their startups. Hit TV shows such as "Shark Tank" and industry norms around "Pitch Day" events suggest an ethos built around founder charisma and personal persuasion. In his best-selling book "Good to Great", Collins (2009) asserts "First Who, Then What" and the idea has been absorbed into the zeitgeist. The ubiquity of this perspective is a true success story of "model persuasion" (Schwartzstein and Sunderam 2021). Investors seem convinced that the founder-first model of the world is the correct one. This likely facilitates investors neglecting to notice features that are predictive and a feedback loop of never noticing or learning persists, consistent with the model and evidence presented in Hanna et al. (2014).

6.4.2 Performance is relative

Investors select an investment strategy, observe returns, and ask the question (in some form)—did we do well? The answer, of course, depends on the endogenous definition of success. What do investors compare their performance to? There are several intuitive candidates: an internal hurdle rate, last year's performance, investor's expectations, alternative investments, or one's competitors. This paper takes a different approach from all of these candidates. Instead of comparing the ex post returns of the chosen strategy to other ex post returns, I compare a set of available strategies using ex ante measures of quality. This choice leads to a difference in sign as to whether investors are "outperforming". How investors choose the relevant counterfactual or reference group for performance is understudied in the literature and is ripe for future research. I argue that who and what investors compare themselves to is likely a significant influence on portfolio choice, especially if investors are "satisficers" and not maximizers (Simon 1955).

6.4.3 Feedback

Startup investing is a "wicked environment" (Hogarth et al. 2015). There is enormous selection in the distribution of outcomes we observe in the world, most success happens over long time horizons filled with factors and events that may seem relevant, and the underlying distribution of founders and companies changes over time. That an algorithm can find predictive signal suggests that a disciplined statistical process can find useful information in the large space of X, but the confounds of the environment may make those insights opaque to human investors.

6.4.4 Why don't investors build algorithms?

Despite their potential, we don't see investors universally adopting algorithms for investment decisions. There are at least two factors that could plausibly contribute to this lack of adoption: institutional constraints and behavioral preferences. For example, company inertia and preferences of limited partners could limit profit-improving technology adoption. Alternatively, investors may be overconfident in their abilities (Moore and Healy 2008), they may intrinsically value agency, or have an inherent algorithm aversion (Dietvorst et al. 2015).

7 Conclusion

This paper offers evidence that sophisticated investors in high-stakes markets can sustainably make systematically poor choices. I offer one strategy to limit such bad decisions: have a human make a set of recommendations to an algorithm which then culls the options that are not likely to reach a benchmark threshold. Future research should explore interventions to reduce the misprediction of high-return opportunities and whether investors, their principals, or startup founders are the source of the suggested overemphasis on human capital.

References

Bansak, K., J. Ferwerda, J. Hainmueller, A. Dillon, D. Hangartner, D. Lawrence, and J. Weinstein (2018). Improving refugee integration through data-driven algorithmic assignment. *Science 359*(6373), 325–329.

Bento, F. R. d. S. R. (2018). *Predicting start-up success with machine learning*. Ph. D. thesis.

Benz, M. and S. Meier (2008). Do people behave in experiments as in the field?—evidence from donations. *Experimental economics 11*(3), 268–281.

Bloom, N. and J. Van Reenen (2007). Measuring and explaining management practices across firms and countries. *The quarterly journal of Economics 122*(4), 1351–1408.

Collins, J. (2009). Good to great-(why some companies make the leap and others don't).

Dawes, R. M., D. Faust, and P. E. Meehl (1989). Clinical versus actuarial judgment. *Science 243*(4899), 1668–1674.

DellaVigna, S. and M. Gentzkow (2019). Uniform pricing in us retail chains. *The Quarterly Journal of Economics 134*(4), 2011–2084.

Di Maggio, M., D. Ratnadiwakara, and D. Carmichael (2022). Invisible primes: Fintech lending with alternative data. Technical report, National Bureau of Economic Research.

Dietvorst, B. J., J. P. Simmons, and C. Massey (2015). Algorithm aversion: people erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology: General 144*(1), 114.

Enke, B. (2020). Wysiati. *Quarterly Journal of Economics*.

Erel, I., L. H. Stern, C. Tan, and M. S. Weisbach (2021). Selecting directors using machine learning. *The Review of Financial Studies 34*(7), 3226–3264.

Fama, E. F. (1965). The behavior of stock-market prices. *The journal of Business 38*(1), 34–105.

Gabaix, X. (2019). Behavioral inattention. In *Handbook of Behavioral Economics: Applications and Foundations 1*, Volume 2, pp. 261–343. Elsevier.

Gentzkow, M., B. Kelly, and M. Taddy (2019). Text as data. *Journal of Economic Literature 57*(3), 535–74.

Gentzkow, M. and J. M. Shapiro (2010). What drives media slant? evidence from us daily newspapers. *Econometrica 78*(1), 35–71.

Ghassemi, M., C. Song, and T. Alhanai (2020). The automated venture capitalist: Data and methods to predict the fate of startup ventures. In *AAAI KDF Workshop*.

Gompers, P., S. N. Kaplan, and V. Mukharlyamov (2015). What do private equity firms say they do? Technical report, National Bureau of Economic Research.

Gompers, P. A., W. Gornall, S. N. Kaplan, and I. A. Strebulaev (2020). How do venture capitalists make decisions? *Journal of Financial Economics 135*(1), 169–190.

Hanna, R., S. Mullainathan, and J. Schwartzstein (2014). Learning through noticing: Theory and evidence from a field experiment. *The Quarterly Journal of Economics 129*(3), 1311–1353.

Hellmann, T. and M. Puri (2000). The interaction between product market and financing strategy: The role of venture capital. *The review of financial studies 13*(4), 959–984.

Hellmann, T. and M. Puri (2002). Venture capital and the professionalization of start-up firms: Empirical evidence. *The journal of finance 57*(1), 169–197.

Hogarth, R. M., T. Lejarraga, and E. Soyer (2015). The two settings of kind and wicked learning environments. *Current Directions in Psychological Science 24*(5), 379–385.

Hortaçsu, A. and S. L. Puller (2008). Understanding strategic bidding in multi-unit auctions: a case study of the texas electricity spot market. *The RAND Journal of Economics 39*(1), 86–114.

Howell, S. T. (2020). Reducing information frictions in venture capital: The role of new venture competitions. *Journal of Financial Economics 136*(3), 676–694.

Hu, A. and S. Ma (2021). Persuading investors: A video-based study. Technical report, National Bureau of Economic Research.

Jean, N., M. Burke, M. Xie, W. M. Davis, D. B. Lobell, and S. Ermon (2016). Combining satellite imagery and machine learning to predict poverty. *Science 353*(6301), 790–794.

Kahneman, D. (2011). *Thinking, fast and slow.* Macmillan.

Kaplan, S. N., M. M. Klebanov, and M. Sorensen (2012). Which ceo characteristics and abilities matter? *The Journal of Finance 67*(3), 973–1007.

Kaplan, S. N. and A. Schoar (2005). Private equity performance: Returns, persistence, and capital flows. *The journal of finance 60*(4), 1791–1823.

Kaplan, S. N., B. A. Sensoy, and P. Strömberg (2009). Should investors bet on the jockey or the horse? evidence from the evolution of firms from early business plans to public companies. *The Journal of Finance 64*(1), 75–115.

Kaplan, S. N. and M. Sorensen (2017). Are ceos different? characteristics of top managers. Technical report, National Bureau of Economic Research.

Kaplan, S. N. and P. Stromberg (2001). Venture capitals as principals: Contracting, screening, and monitoring. *American Economic Review 91*(2), 426–430.

Kaplan, S. N. and P. E. Strömberg (2004). Characteristics, contracts, and actions: Evidence from venture capitalist analyses. *The Journal of Finance 59*(5), 2177–2210.

Ke, Z. T., B. T. Kelly, and D. Xiu (2019). Predicting returns with text data. Technical report, National Bureau of Economic Research.

Kleinberg, J., H. Lakkaraju, J. Leskovec, J. Ludwig, and S. Mullainathan (2018a). Human decisions and machine predictions. *The Quarterly Journal of Economics 133*(1), 237–293.

Kleinberg, J., H. Lakkaraju, J. Leskovec, J. Ludwig, and S. Mullainathan (2018b). Human decisions and machine predictions. *The quarterly journal of economics 133*(1), 237–293.

Kleinberg, J., J. Ludwig, S. Mullainathan, and Z. Obermeyer (2015). Prediction policy problems. *American Economic Review 105*(5), 491–95.

Kogan, S., D. Levin, B. R. Routledge, J. S. Sagi, and N. A. Smith (2009). Predicting risk from financial reports with regression. In *Proceedings of Human Language Technologies: The 2009 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, pp. 272–280.

Krishna, A., A. Agrawal, and A. Choudhary (2016). Predicting the outcome of startups: less failure, more success. In *2016 IEEE 16th International Conference on Data Mining Workshops (ICDMW)*, pp. 798–805. IEEE.

Lerner, J. (1994). The syndication of venture capital investments. *Financial management*, 16–27.

Lerner, J. (1995). Venture capitalists and the oversight of private firms. *the Journal of Finance 50*(1), 301–318.

List, J. A. (2003). Does market experience eliminate market anomalies? *The Quarterly Journal of Economics 118*(1), 41–71.

List, J. A. (2006). The behavioralist meets the market: Measuring social preferences and reputation effects in actual transactions. *Journal of political Economy 114*(1), 1–37.

Lyonnet, V. and L. H. Stern (2022). Venture capital (mis) allocation in the age of ai. *Fisher College of Business Working Paper* (2022-03), 002.

Moore, D. A. and P. J. Healy (2008). The trouble with overconfidence. *Psychological review 115*(2), 502.

Mullainathan, S. and Z. Obermeyer (2019). A machine learning approach to low-value health care: wasted tests, missed heart attacks and mis-predictions. *NBER Working Paper* (w26168).

Mullainathan, S. and Z. Obermeyer (2022). Diagnosing physician error: A machine learning approach to low-value health care. *The Quarterly Journal of Economics 137*(2), 679–727.

Nanda, R., S. Samila, and O. Sorenson (2020). The persistent effect of initial success: Evidence from venture capital. *Journal of Financial Economics 137*(1), 231–248.

Rajan, R. G. (2012). Presidential address: The corporation in finance. *The Journal of Finance 67*(4), 1173–1217.

Schwartzstein, J. and A. Sunderam (2021). Using models to persuade. *American Economic Review 111*(1), 276–323.

Simon, H. A. (1955). A behavioral model of rational choice. *The quarterly journal of economics 69*(1), 99–118.

Strulov-Shlain, A. (2021). More than a penny's worth: Left-digit bias and firm pricing. *Chicago Booth Research Paper* (19-22).

Ying, D., G. Wua, and J. Lia (2021). Comparison of different machine learning methods for predicting the success of the startups. Technical report, Sinovation Ventures.

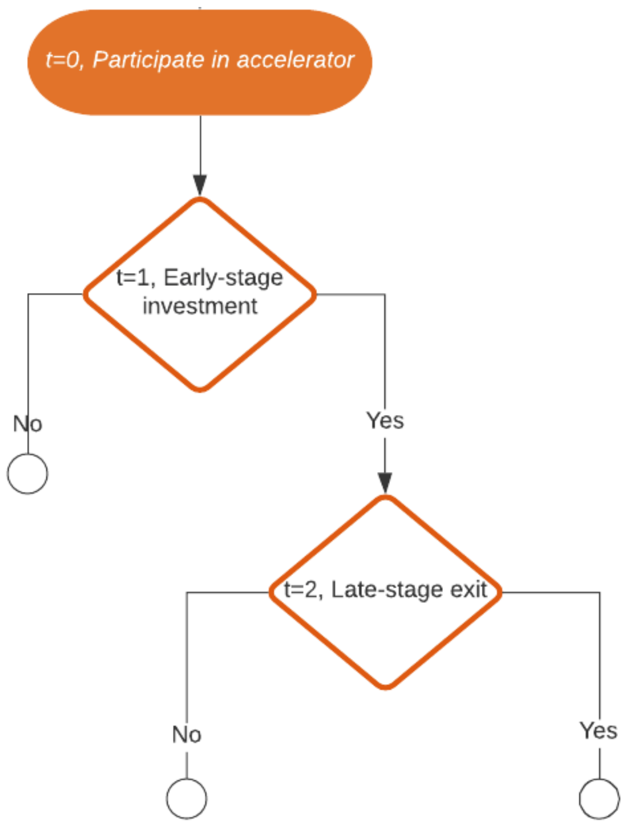
Żbikowski, K. and P. Antosiuk (2021). A machine learning, bias-free approach for predicting business success using crunchbase data. *Information Processing & Management 58*(4), 102555.

Zhang, Y. (2020). Discrimination in the venture capital industry: Evidence from two randomized controlled trials. *arXiv preprint arXiv:2010.16084*.

Zingales, L. (2000). In search of new foundations. *The journal of Finance 55*(4), 1623–1653.

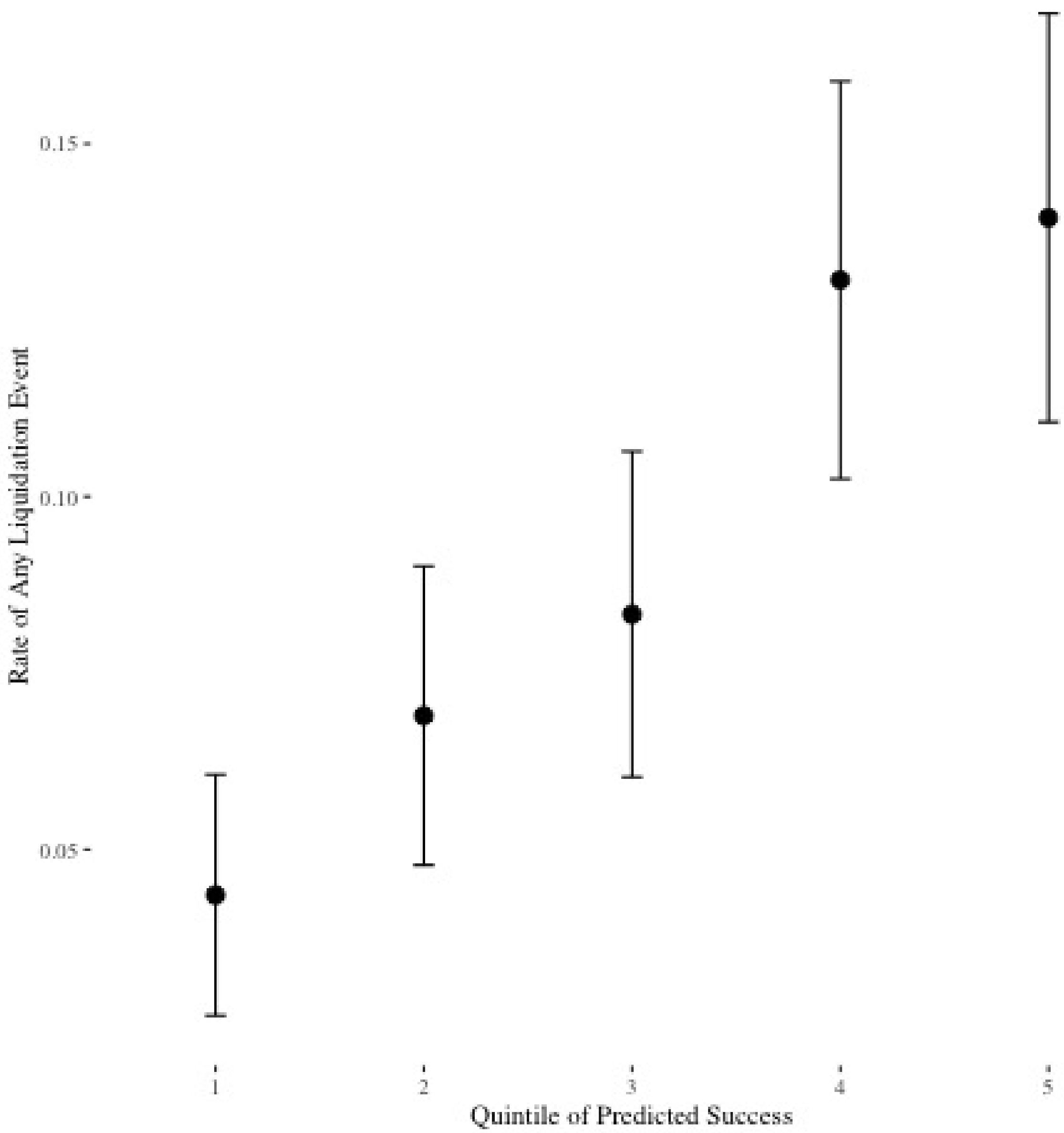
A Figures and Tables

Figure 1: Timeline of events



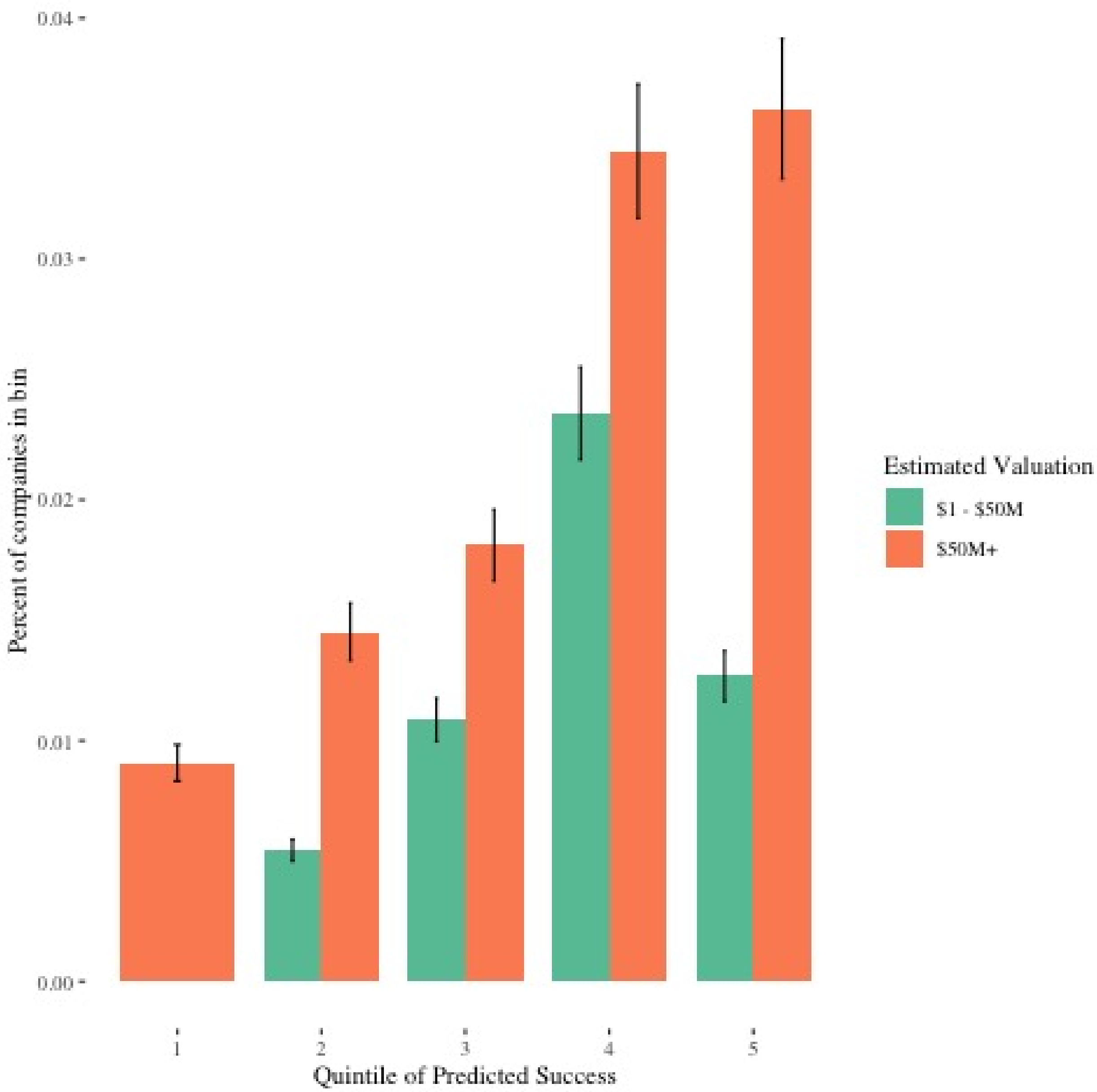
Note: t=0 defines the sample; All companies that participated in any of the top 100 US incubators or accelerators before 2019 are followed through the end of 2020. Two key outcomes are measured: early-stage investment (t=1) and late-stage exit (t=2). Only information available as of the incubator stage (t=0) is used to predict the two outcomes.

Figure 2: Late-Stage Outcomes



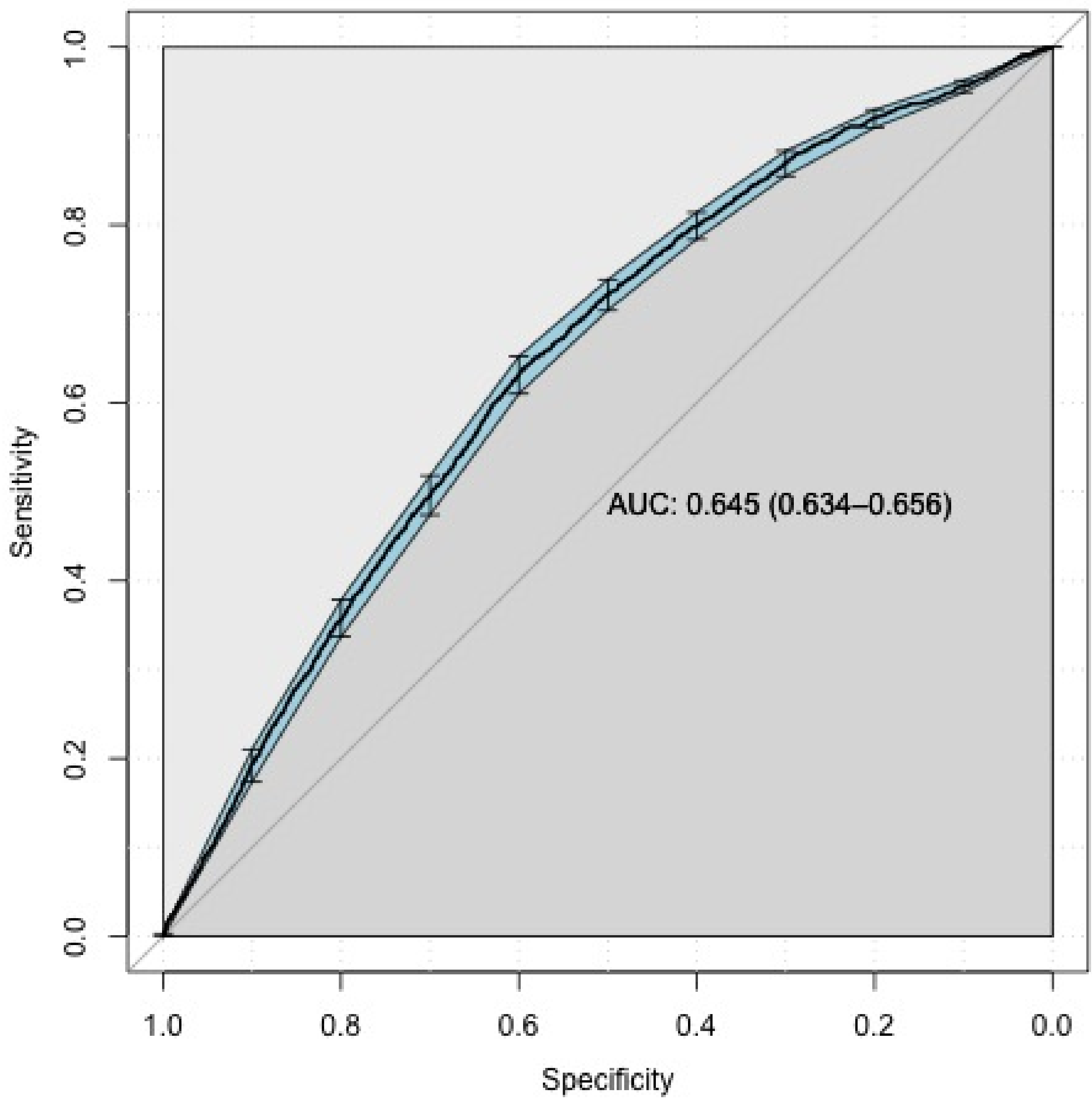
Note: Relative frequency of exit shown by quintile of predicted success produced by the ML predictions. Exit is defined as an IPO, acquisition, or Series D or later transaction. Companies with no late-stage transactions are valued at \$0. Sample is subsetted to those companies that received early-stage VC investments. Predicted success deciles are defined within the subsample.

Figure 3: Distribution of valuation outcomes



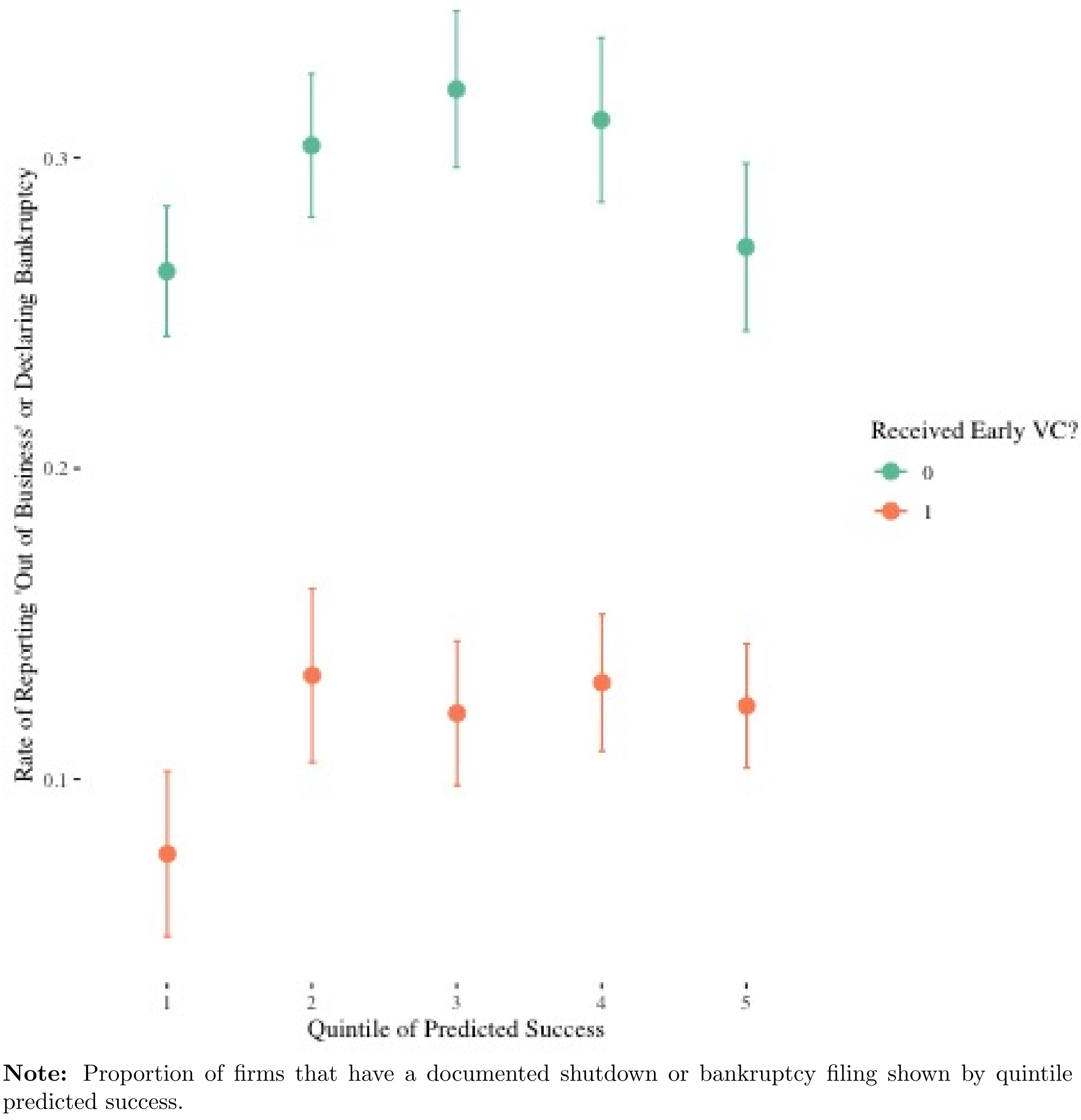
Note: Relative frequency of each valuation bin shown by decile of predicted success produced by the ML predictions. Companies with no late-stage transactions are valued at $0 and are thus not shown. Sample is subsetted to those companies that received early-stage VC investments. Predicted success deciles are defined within the subsample.

Figure 4: ROC-Curve



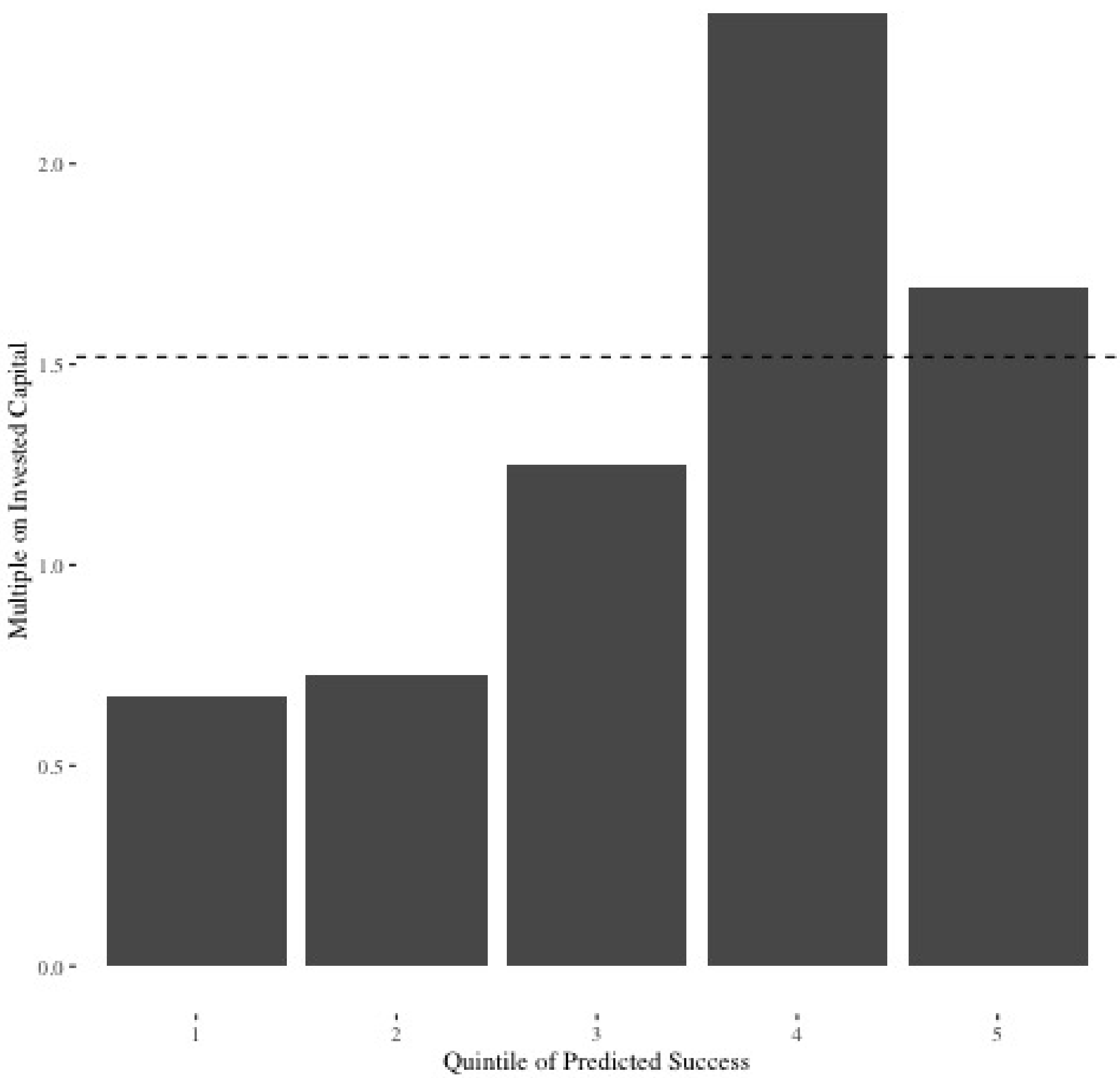
Note: This graph shows the Receiver Operating Characteristic Curve for the algorithm trained to predict late-stage exit. The area under the curve (AUC) is a common measure of the predictive quality of the algorithm. An AUC of 0.5 is equivalent to random guessing and improvement over 0.5 indicates an ability to predict with some level of accuracy.

Figure 5: Documented failures



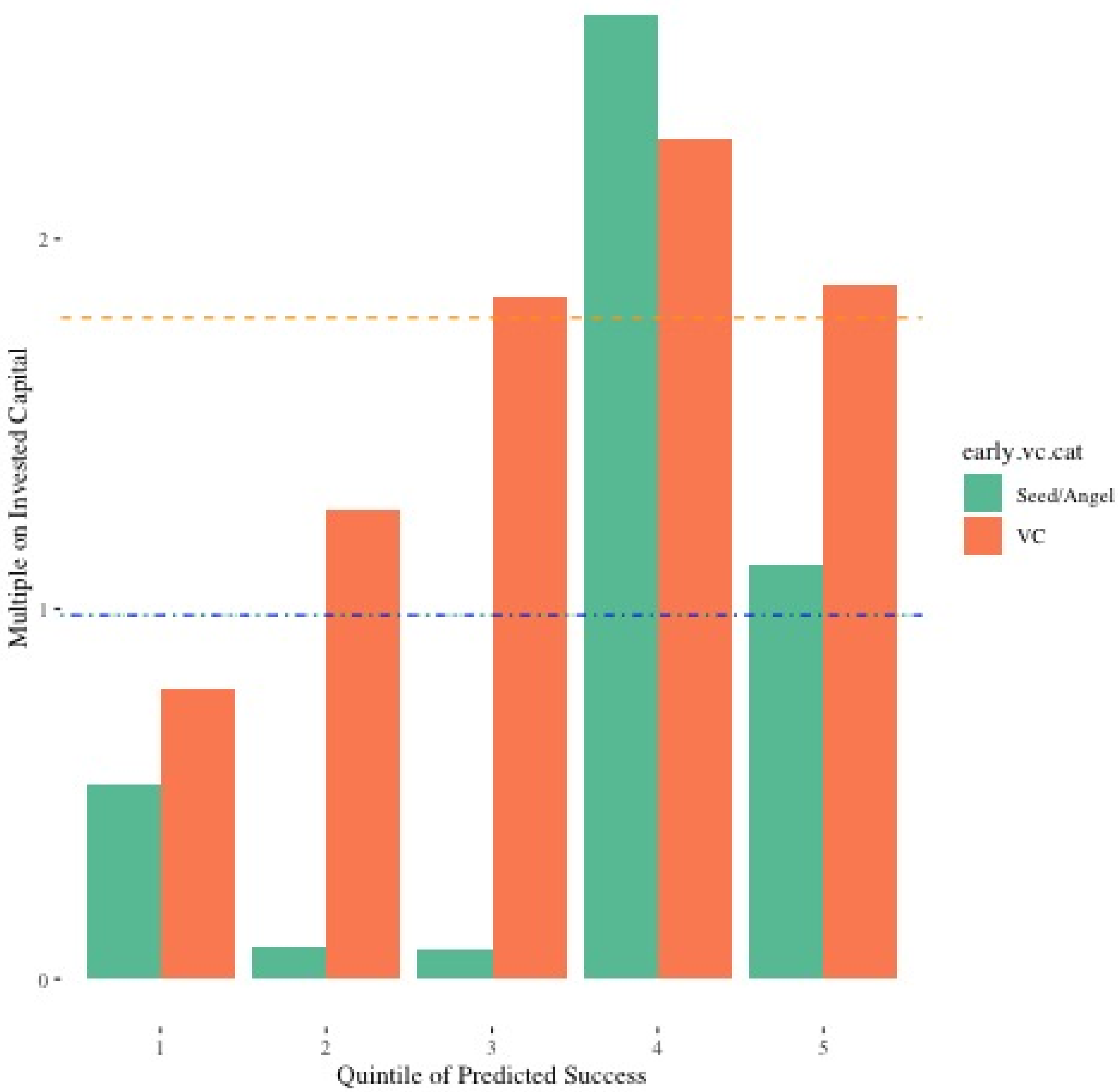
Note: Proportion of firms that have a documented shutdown or bankruptcy filing shown by quintile of predicted success.

Figure 6: Returns by predicted quality



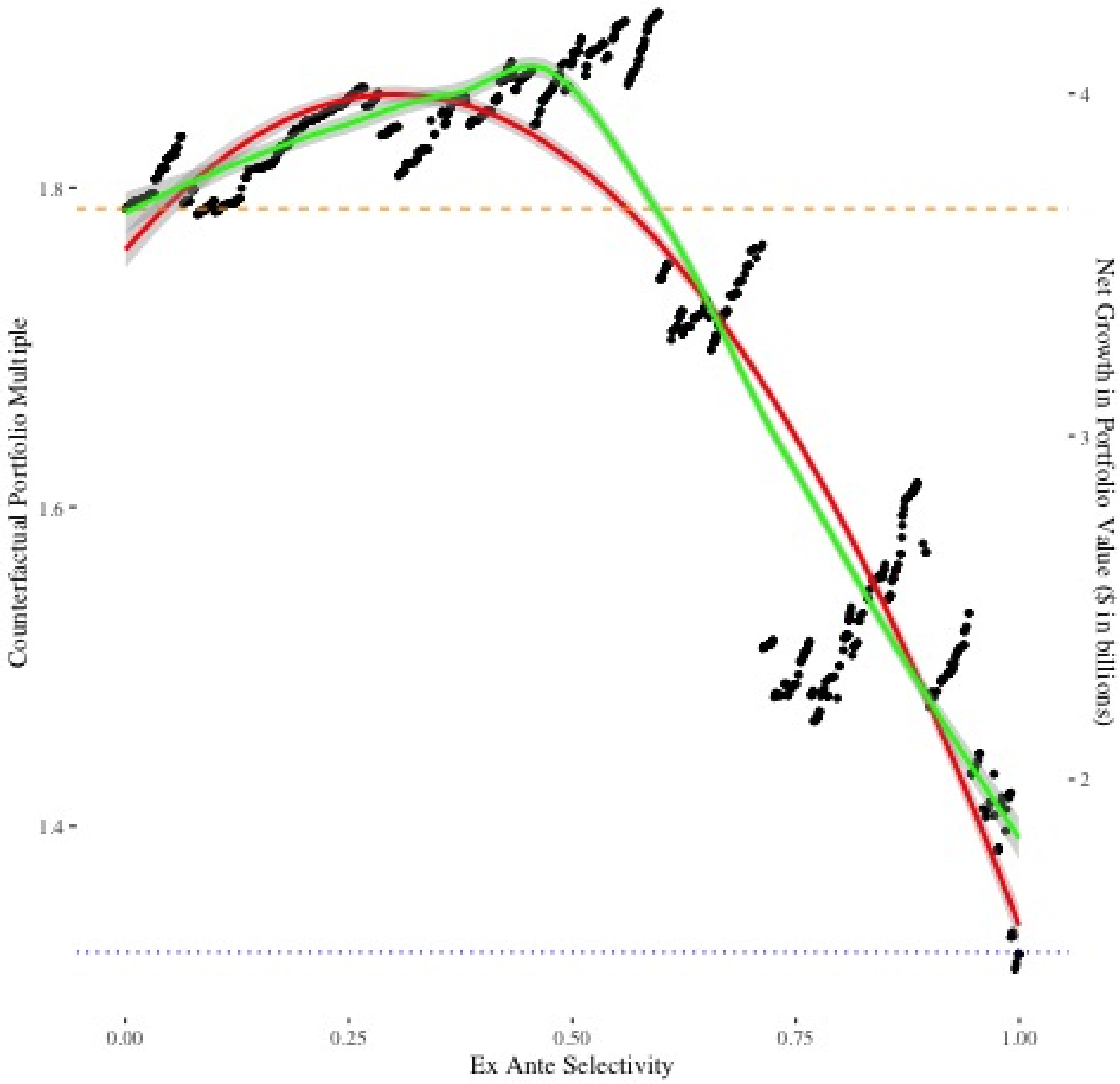
Note: This graph shows the multiple on invested capital (MOIC) for firms in each quintile of predicted success. MOIC is defined by the ratio between the five-year valuation of the stake in the company and the level of the initial investment for that stake.

Figure 7: Returns by predicted quality and investor type



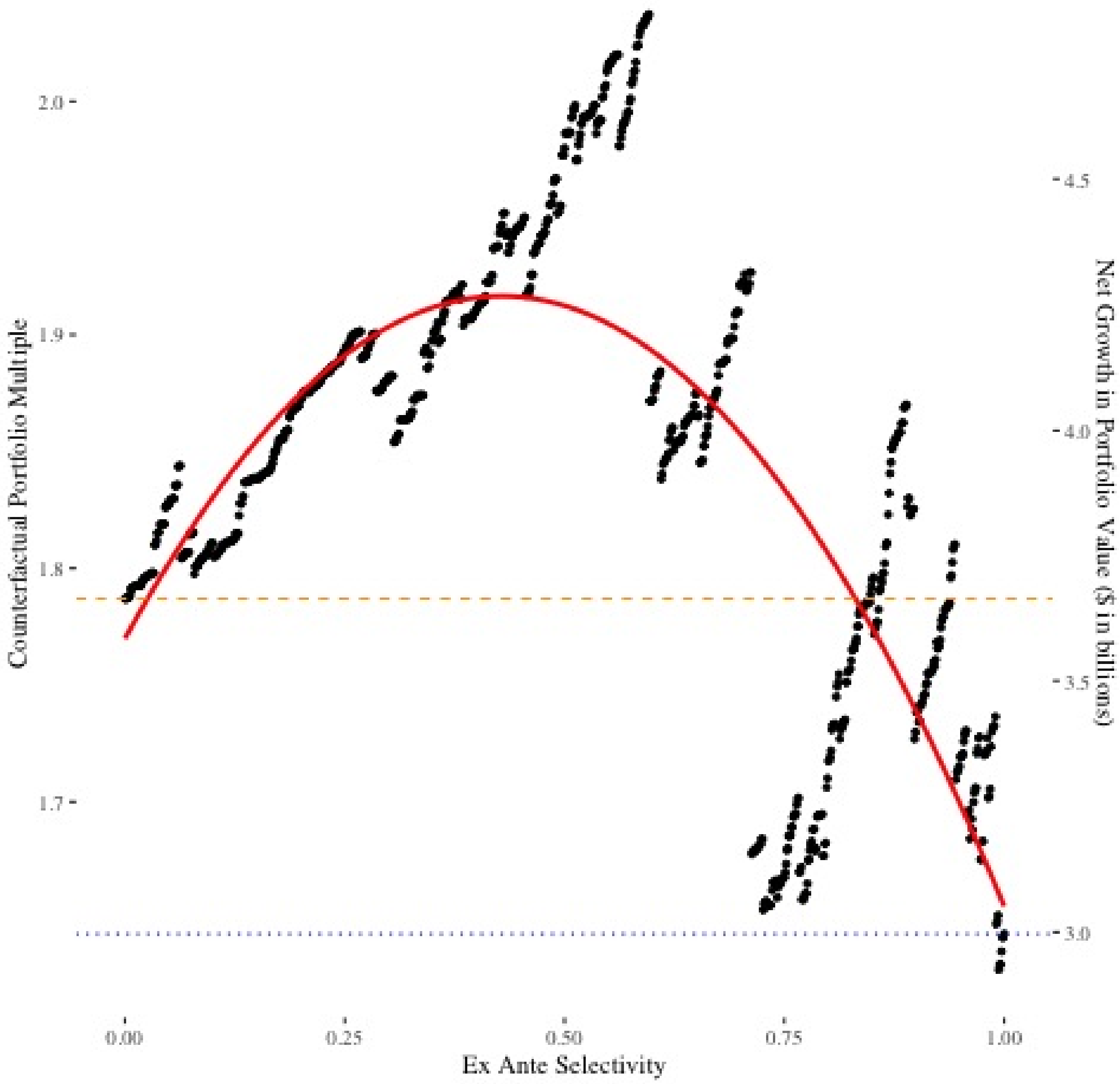
Note: This graph shows the multiple on invested capital (MOIC) for firms in each quintile of predicted success separately for individual investors ("Seed/Angels") and institutional investors ("VC"). MOIC is defined by the ratio between the five-year valuation of the stake in the company and the level of the initial investment for that stake. The orange, dashed line is the overall MOIC for instituional investors and the blue dot-dash line is the overall MOIC for individual investors.

Figure 8: Counterfactual Payoff with Fewer Mistakes



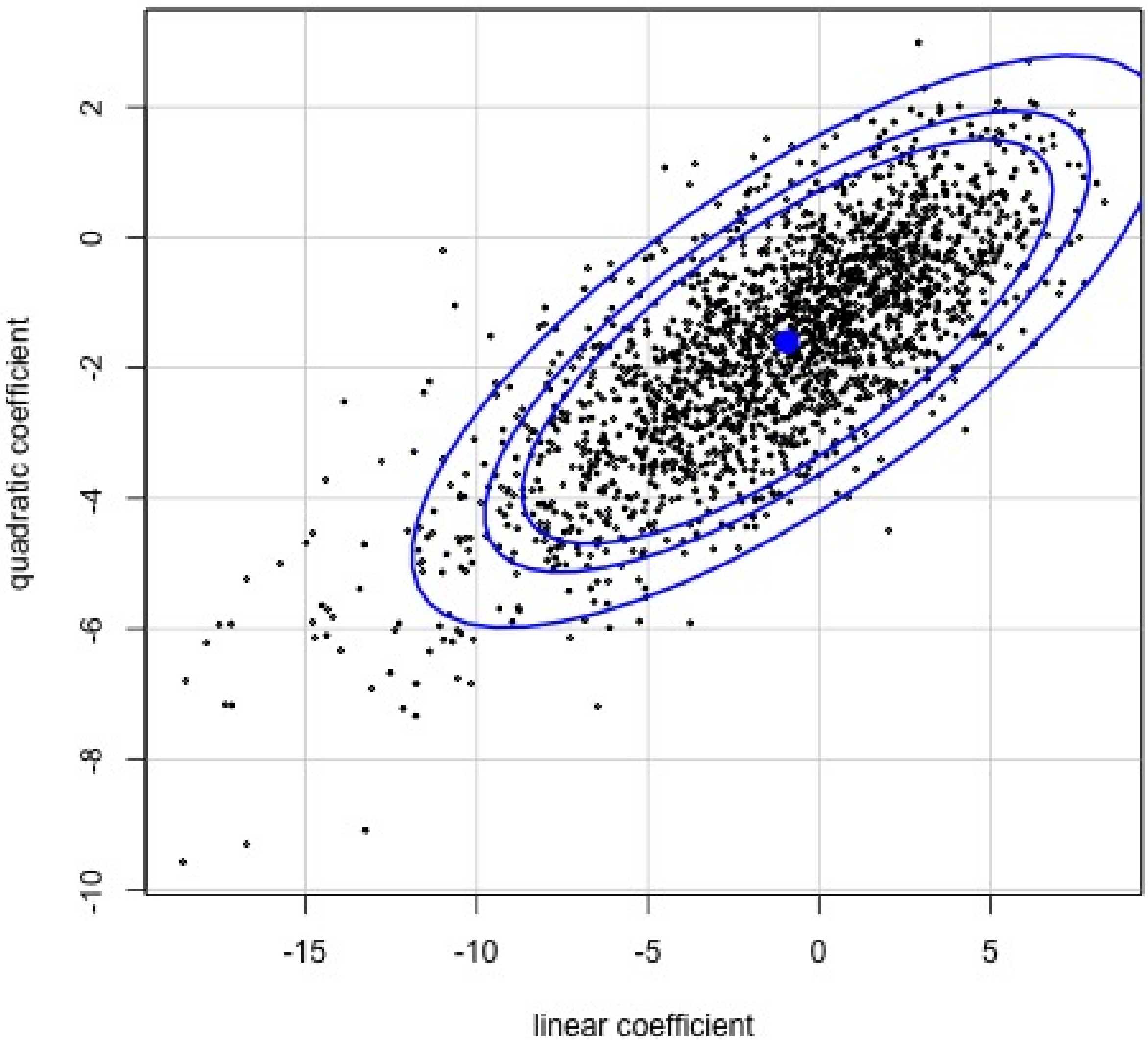
Note: Nominal return on investment that would have been realized by VCs had they had selected the outside option instead of investing in the bottom k percent of startups. Values above $1B are replaced with a $1B valuation. The outside option is assumed to be invested in a bond that pays 8% interest per year. The orange dashed line is the actual multiple on invested capital (MOIC). The blue dotted line is the MOIC the VC's would have realized if they instead invested all money in a 7% bond. Improvements above the orange dashed line indicate forgone returns that were predictable. A total of $6.25 billion in initial investments is captured in the graph, which provides the mapping between the two y axes.

Figure 9: Counterfactual Payoff with SP Outside Option



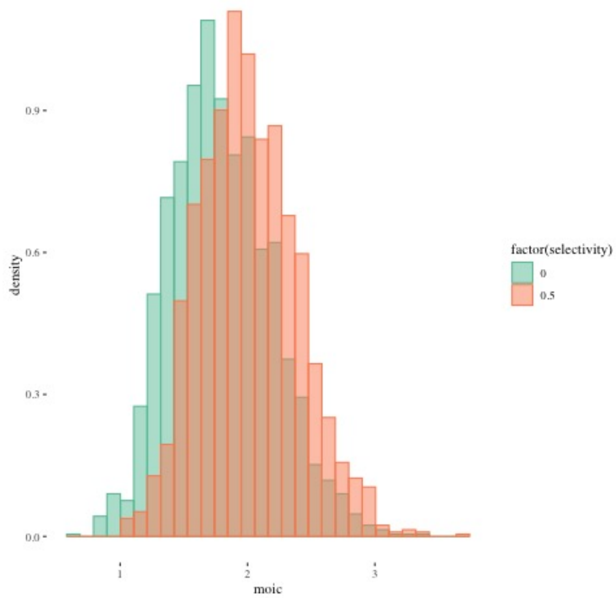
Note: Nominal return on investment that would have been realized by VCs had they only invested in the top k percent of startups. Any money that is not invested in a startup is assumed to be invested in the SP 500 index. The red dashed line is the actual multiple on invested capital (MOIC). The blue dotted line is the MOIC the VC's would have realized if they instead invested all money in the market. Improvements above the red dashed line indicate loss of alpha due to mistakes. A total of $6.25 billion in initial investments is captured in the graph, which provides the mapping between the two y axes.

Figure 10: Coefficient uncertainty

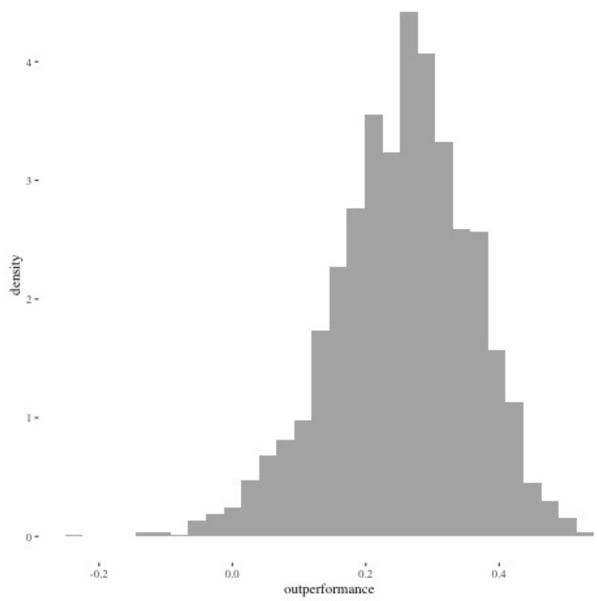


Note: This graph shows the uncertainty in the estimation of the curvilinear relationship between ex ante selectivity and returns (MOIC). Negative values for the y axis (the quadratic coefficient) indicate a hill-shaped relationship. The concentric circles correspond to the 90th, 95th, and 99th confidence regions, respectively. Estimates are produced by bootstrapping.

Figure 11: Bootstrapped estimates of MOIC and returns to selectivity



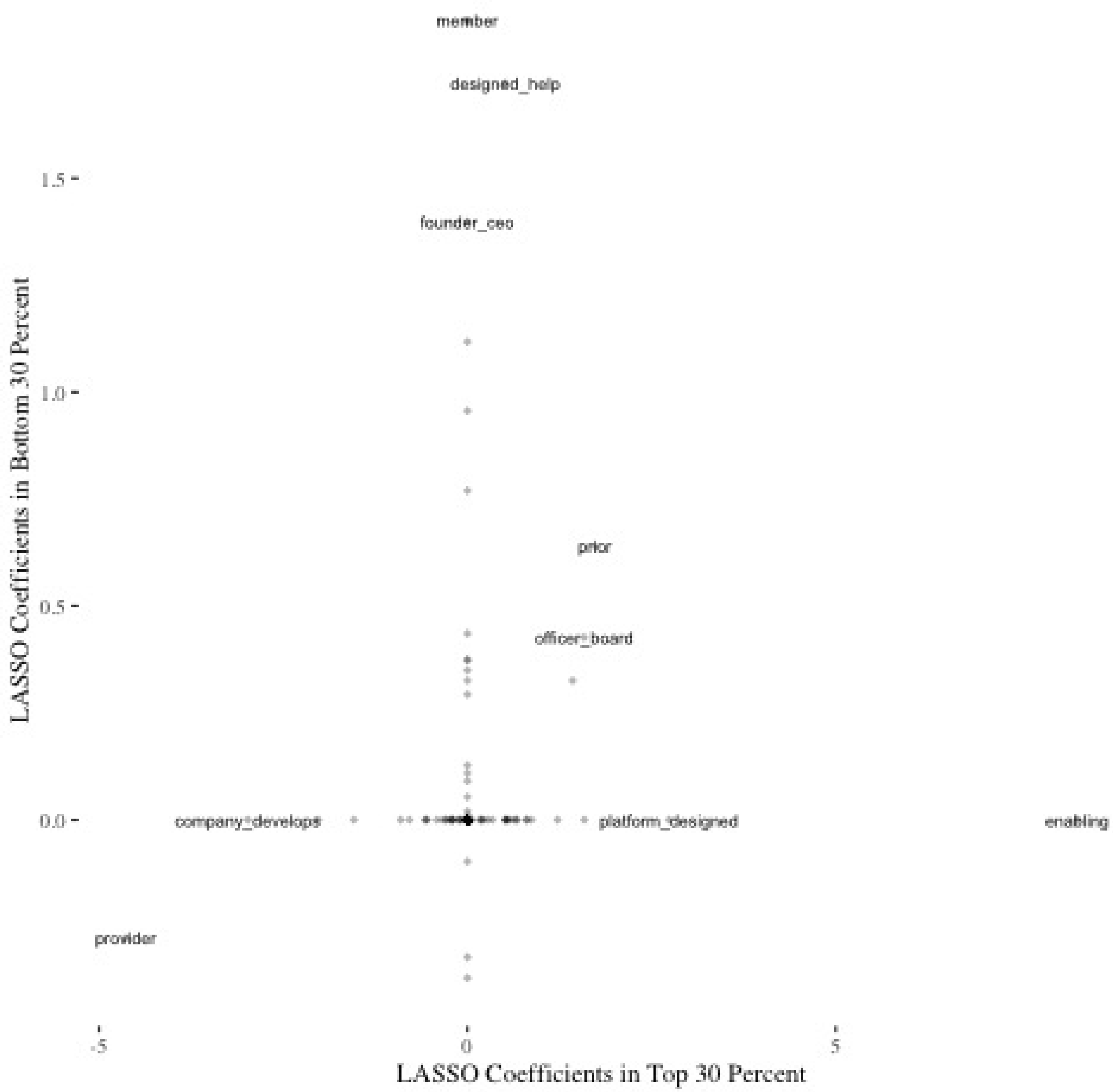
(a) Returns



(b) Relative outperformance

Note: This graph quantifies the distribution of performance gains derived from a bootstrapping procedure. In panel (a) the green bars indicate the distribution of MOIC that is realized under the current investment strategy, while the orange bars indicate the distribution of returns under a strategy of dropping the lower half of startups (according to the ex ante quality measure) in favor of the SP 500. The right panel (b) takes the paired difference between the two distributions in (a) and therefore graphs the distribution of the benefit of being selective (up to 50%) relative to the chosen portfolio of VC investors (selectivity = 0%).

Figure 12: What Predicts Investments in Good vs Bad Investments?



Note: I build two LASSO models with the same feature set **X** to predict which companies will receive an early-stage investment: once on companies in top 30 percent of algorithmic success predictions (the wheat) distribution and once on companies in the bottom 30 percent of the algorithmic success predictions (the chaff). I plot the results of the two regressions where each point represents a variable, the x axis is the coefficient from the first regression and the y axis is the coefficient from the second. Nearly all points cluster around one of the axes though not at the origin, indicating the importance of features in one set is orthogonal to features important in the other. Investors have totally different criteria for the different quality firms. For example, "founder_serves" is a large, positive predictor for both sets. On the other hand "founded" is a strong negative predictor among the cream but it has no predictive power among the chaff.

Table 1: Summary statistics: Firm Characteristics by Early Stage

		No early VC		Early VC	
		Mean	Std. Dev.	Mean	Std. Dev.
Has late stage exit?		0.0	0.0	0.1	0.3
Number of funding deals		1.4	0.9	1.6	1.0
Accelerator funding amount ($M)		0.1	0.2	0.1	0.3
Accelerator stake (%)		16.4	16.6	13.6	13.8
Accelerator equity investment ($M)		0.1	0.2	0.1	0.3
Revenue ($M)		8.3	57.3	166.8	2310.5
EBITDA ($M)		-0.5	8.7	-0.5	8.2
		N	Pct.	N	Pct.
Has CEO bio?	FALSE	745	7.0	70	1.3
	TRUE	9869	93.0	5370	98.7
Female founder?	FALSE	9807	92.4	4880	89.7
	TRUE	807	7.6	560	10.3
Has CEO education?	FALSE	3817	36.0	868	16.0
	TRUE	6797	64.0	4572	84.0
CEO has MBA?	FALSE	9543	89.9	4591	84.4
	TRUE	1071	10.1	849	15.6
CEO has M7 MBA?	FALSE	10046	94.6	4866	89.4
	TRUE	568	5.4	574	10.6
CEO has PhD?	FALSE	9742	91.8	4928	90.6
	TRUE	872	8.2	512	9.4
Known industry code	FALSE	4	0.0	0	0.0
	TRUE	10610	100.0	5440	100.0
Has company description	FALSE	2	0.0	0	0.0
	TRUE	10612	100.0	5440	100.0

Note: Sample statistics. Accelerator recipients define the sample. "Early VC" indicates the subsample of firms that received an early-stage investment after the accelerator; "No early VC" firms did not receive an investment. All fields correspond to data available when a given startup is accepted into an accelerator.

Table 2: Summary statistics: Firm Characteristics by Late Stage

	No late-stage exit		Late-stage exit	
	Mean	Std. Dev.	Mean	Std. Dev.
Raised A or B round after accelerator?	0.1	0.3	0.4	0.5
Early-stage invested equity ($M)	0.5	2.3	3.9	6.7
Early-stage invested capital ($M)	2.1	4.5	3.9	6.7
Has IPO exit?	0.0	0.0	0.0	0.1
Has merger/acquisition exit?	0.0	0.0	0.8	0.4
Post-money valuation at exit			679.5	4182.5
Shut-down or bankruptcy?	0.3	0.5	0.0	0.2

Note: Sample statistics. Recipients of early-stage funding define the subsample for this table. “Late-stage exit” indicates the subsample of firms that achieved a late-stage exit; “No late-stage exit” firms did not achieve an exit.

Table 3: Top Accelerators

	InvestorName	Alum Mkt. Value ($M)	% of Total	# of investments
1	RocketSpace	51,000	0.53	3
2	Y Combinator	22,407	0.24	437
3	Plug & Play Tech Center	3,959	0.04	134
4	JLABS	3,686	0.04	38
5	MasterCard Start Path	1,700	0.02	2
6	Impact USA	1,400	0.01	4
7	Wharton Venture Initiation	1,200	0.01	13
8	StartX (US)	1,151	0.01	64
9	Agoranov	1,062	0.01	33
10	Techstars	771	0.01	280

Note: This table documents the 10 most successful accelerators in the data. Success is defined as the total five-year market value of all startups that have completed the accelerator’s programming (”Alum Mkt. Value”). ”% of Total” captures the percentage of all market valuations in my data that are attributable to a given accelerator. ” of investments” captures the number of startups that complete a given accelerator and then raise an early-stage round within two years.

Table 4: Top Investors

	Investor Name	Invested ($M)	% of Invst.	Mkt.Value ($M)	% of Total Mkt. Value	Count
1	Benchmark	14.8	0.00	51,000	0.52	2
2	Sequoia Capital	67.9	0.01	5,564	0.06	6
3	Accel	200.6	0.03	5,348.48	0.05	11
4	Index Ventures	71.0	0.012	4,030.06	0.04	5
5	Andreessen Horowitz	139.2	0.02	3,720	0.04	18

Note: This table documents the early-stage investors associated with the most valuable firms. "Mkt.Value" captures the total five-year market value of startups for which a given investor served as the lead investor in the startups first funding round after the accelerator. "% of Total Mkt. Value" captures the fraction of all valuations in my data that are associated with an early stage investment by a given investor. "Invested" and "% of Invst." capture the total funding by an investor that I observe and the fraction that a given investor's funding constitutes of the total funding in my data, respectively.

Table 5: Top Markets

	Market	#	% of All	% of Early	% of Late
1	Business/Productivity Software	2,038	0.127	0.186	0.167
2	Application Software	1,362	0.085	0.059	0.145
3	Media and Information Services (B2B)	834	0.052	0.061	0.056
4	Social/Platform Software	823	0.051	0.033	0.044
5	Information Services (B2C)	637	0.040	0.040	0.034
6	Financial Software	609	0.038	0.059	0.062
7	Electronics (B2C)	446	0.028	0.034	0.020
8	Educational Software	405	0.025	0.027	0.014
9	Other Services (B2C Non-Financial)	360	0.022	0.017	0.014
10	Entertainment Software	303	0.019	0.016	0.018

Note: This table documents the representation of markets among different sets of companies. The first column of percentages shows the distribution of markets among all accelerator participants. The second column shows the distribution of markets among those companies receiving early stage investments. The third column shows the distribution of markets among the companies that achieved an exit.

Table 6: ML-OLS Projection

	Dependent variable:				
	Algorithm predicted success				
	(1)	(2)	(3)	(4)	(5)
(Pre-accel. funding)^2	-0.00002^{***}				-0.00002^{***}
	(0.00000)				(0.00000)
MBA CEO		0.030^{***}			0.027^{***}
		(0.001)			(0.001)
M7 MBA CEO			0.020^{***}		0.009^{***}
			(0.002)		(0.002)
Female founder				0.007^{***}	0.005^{***}
				(0.001)	(0.001)
Pre-accel. funding	0.003^{***}	0.001^{***}	0.001^{***}	0.001^{***}	0.002^{***}
	(0.0002)	(0.0001)	(0.0001)	(0.0001)	(0.0002)
Constant	0.047^{***}	0.044^{***}	0.046^{***}	0.047^{***}	0.042^{***}
	(0.0004)	(0.0004)	(0.0005)	(0.0005)	(0.0005)
Observations	4,578	4,578	4,578	4,578	4,578
R^2	0.041	0.141	0.057	0.027	0.166
Adjusted R^2	0.041	0.141	0.057	0.027	0.166

*p<0.1; **p<0.05; ***p<0.01

Note: This table reports regression results to approximate the relationship between the algorithm's prediction and simple variables previously explored in the literature. The algorithm's success prediction is the predicted probability of success $p \in [0, 1]$.

Table 7: ML-OLS Projection

	Dependent variable:									
	ML Predicted Success Percentile					ML Predicted Success Probability				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
(Prior funding)^2	-0.001^{***}				-0.001^{***}	-0.0001^{***}				-0.0001^{***}
	(0.0001)				(0.0001)	(0.00002)				(0.00002)
MBA CEO		0.211^{***}			0.189^{***}		0.028^{***}			0.028^{***}
		(0.017)			(0.018)		(0.002)			(0.002)
M7 MBA CEO			0.137^{***}		0.058^{***}			0.012^{***}		0.001
			(0.021)		(0.021)			(0.003)		(0.003)
Female CEO				0.028	0.021				0.001	-0.0001
				(0.022)	(0.021)				(0.003)	(0.003)
Prior funding	0.048^{***}	0.018^{***}	0.018^{***}	0.018^{***}	0.047^{***}	0.005^{***}	0.002^{***}	0.002^{***}	0.002^{***}	0.005^{***}
	(0.005)	(0.002)	(0.002)	(0.002)	(0.004)	(0.001)	(0.0003)	(0.0003)	(0.0003)	(0.001)
Constant	0.555^{***}	0.541^{***}	0.558^{***}	0.570^{***}	0.519^{***}	0.051^{***}	0.049^{***}	0.052^{***}	0.053^{***}	0.047^{***}
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Observations	1,535	1,535	1,535	1,535	1,535	1,535	1,535	1,535	1,535	1,535
R^2	0.069	0.121	0.061	0.035	0.157	0.048	0.120	0.035	0.021	0.144
Adjusted R^2	0.067	0.120	0.060	0.034	0.154	0.047	0.119	0.033	0.019	0.141

$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Note: This table reports regression results to approximate the relationship between the algorithm's prediction and simple variables previously explored in the literature. The subsample used for this analysis excludes startups that did not receive early-stage investments. The dependent variable for Columns (6)-(10) is the algorithm's predicted probability of success p. The dependent variable for (1)-(5) is the percentile of p (predicted probability of success) within the set of firms that received early-stage investments.

Table 8: Bootstrapped Confidence Intervals for Quadratic Relationship

Confidence Level	Normal	Percentile	BCa
90%	(-5.405, 0.670)	(-5.708, 0.335)	(-6.939, -0.170)*
95%	(-5.987, 1.252)	(-6.459, 0.726)	(-8.019, 0.171)

Note: This table reports the bootstrapped confidence intervals for the coefficient on the quadratic term in the relationship between ex ante selectivity and returns (MOIC). Negative signs indicate a bill shaped relationship. For each confidence level, I calculate the bootstrapped interval by three different procedures. * indicates the interval does not include 0.

Table 9: Human capital overweighting

	Dependent variable:				
	Received early stage investment?				
	(1)	(2)	(3)	(4)	(5)
Full model prediction	3.430***	2.783***	7.107***	0.631*	2.371***
	(0.175)	(0.219)	(0.476)	(0.334)	(0.785)
Edu model prediction		1.357***	4.990***	0.582**	1.820***
		(0.275)	(0.449)	(0.281)	(0.578)
Full*Edu model prediction			−62.335***		−19.079**
			(6.097)		(7.794)
Constant	0.213***	0.165***	−0.050**	0.146***	0.075**
	(0.008)	(0.013)	(0.025)	(0.016)	(0.033)
Non-linear Risk Control?	No	No	No	Yes	Yes
Observations	10,452	10,452	10,452	10,452	10,452
Adjusted R^2	0.035	0.037	0.047	0.056	0.057

*p<0.1; **p<0.05; ***p<0.01

Note: This table reports regression results from correlating several algorithmic success predictions to a binary dependent variable indicating if the startup received an early-stage investment (=1) or not (=0). The "Full model prediction" is the algorithm's predicted success when trained on all available variables. The "Edu model prediction" is the algorithm's predicted success when trained only on founder's education information. The non-linear risk controls add fixed effects for the quintile of ML-predicted success.