How do investors learn as data becomes bigger? Evidence from a FinTech platform

Ahmed Guecioueur*

March 7, 2021

Abstract

Prior findings suggest that investors learn with experience. We study the complementary channel of learning from data, particularly the effects of making additional predictive signals available to investors. We analyse a panel of systematic traders' investment outcomes, sourced from a FinTech platform that organises trading contests under highly-controlled conditions that allow us to identify learning effects. Investor outcomes improve with experience, and this is also apparent when counterfactually assessing their trading decisions on historical data, suggesting that they make use of historical data to attain their objectives. Importantly, when additional predictive variables are added to the common part of investors' information sets, the individual-level dispersions of investors' performance outcomes narrow, while their relative performance outcomes improve at higher experience levels. To explain why this widening of their common dataset benefits experienced investors, we model an investor as choosing a portfolio by learning from historical data while also taking model uncertainty into account. The robust learner therefore ignores predictive signals with historical predictive contributions below a subjective model uncertainty threshold; we conjecture this threshold varies with experience.

Keywords: Learning, Information, Robustness, Systematic Trading, Big Data, FinTech
JEL Codes: G11, G14, G23, G41

*INSEAD, Boulevard de Constance, 77300 Fontainebleau, France, email: ahmed.guecioueur@insead.edu
I would like to thank Francisco Gomes, Joël Peress, Alexandra Roulet and Taisiya Sikorskaya for their helpful comments.

1 Introduction

The histories of signals available to investors have always *lengthened* with the passage of time: even in the earliest days of Wall Street, ticker tape machines rolled out updated prices for tradeable securities, and statistically the number of samples increased linearly with time. The information available to investors has often *widened* too, as additional predictive signals have become available (perhaps through a conscious information choice decision by the investor). The current interest in Big Data relates to the latter phenomenon: thanks to an explosion in data collection and advances in processing techniques, investors now have access to many more predictive signals. Figure 1 illustrates this distinction.

[Insert Figure 1 about here]

How do investors learn from these additional signals? Equivalently, how do investors learn as their datasets grow wider? In this paper we empirically quantify the effect of adding new predictive variables to the common part of investors' information sets, and disentangle this effect from previous findings on investor learning with experience (or "learning by doing").

We do so using a panel from a unique setting: a FinTech platform called Quantiacs that runs contests for what we term "systematic investors". These systematic investors design and submit trading strategies based on historical data from a "backtest" period. The resulting strategies take as inputs a common set of actual market data that is updated daily during a "live" trading period – based upon this data, each strategy chooses a portfolio of futures contracts without any input from the systematic investor, who has thus precommitted to following the systematic trading strategy during the live period.[1] Systematic trading styles are increasingly popular: for example, approximately 30% of all hedge fund assets under management were held by systematic (also called "quant") investors in late 2019.[2] In our institutional setting, individual investors ultimately seek an investment mandate from Quantiacs: after each contest ends, profit-sharing investments are offered to the top 3 contestants, as determined by a score that incentivises the systematic investors to maximize the Sharpe Ratios of their portfolio returns during the live period. After several years of running such contests, the platform suddenly introduced an additional set of macroeconomic signals to investors for them to use in making their trading decisions. This controlled change by the platform represents a sudden widening of the information available to a subset of the investors in our panel. We exploit this

[1] The platform uses the term "live" trading although no actual assets are managed so it would be more accurate to call this "paper" trading. Also, the "investors" are really "potential investors" as they do not receive any funding until the contests are over. We emphasise, however, that actual market data is always used.

[2] The Economist: *"March of the machines"*, Oct 5th 2019: `https://www.economist.com/briefing/2019/10/05/the-stockmarket-is-now-run-by-computers-algorithms-and-passive-managers`

institutional feature to detect whether and how investors' performance outcomes were affected, and therefore reason about their learning mechanisms.

In general, our institutional setting has several distinctive features that help us make inferences about investors' (unobservable) learning mechanisms based on their observed performance outcomes. Firstly, the common part of investors' information sets is known at all points in time and the platform's introduction of additional predictive signals may be taken to be exogenous, rather than relying on an endogenous information choice motive by investors. Agents have a common limitation on what information their trading strategies can make use of at all points in time, being restricted to the set of predictive signals provided by the trading platform, without the ability to base their systematic trading decisions on any information that is external to the platform. Investors also share the same preferences, as they are all incentivized to maximize their live-period Sharpe Ratios; this frees them from strategic considerations and puts them in a perfectly competitive setup because agents cannot affect each others' payoffs, whether through price impact or observability of actions. The contestants are also freed from consideration of some well-known drivers of portfolio choice that might confound our analysis, including wealth effects, background risk and horizon effects. Since the investors in our panel are individuals, not institutions, we therefore study how investors in general use historical data to choose their portfolios, under unique conditions that control access to a common information set. Our panel of performance outcomes consists of backtest & live-period Sharpe Ratios for systematic investors and the trading contests that they (repeatedly) participate in. An investor who improves her live-period Sharpe Ratio can be inferred as having learned how to do so, and therefore learning effects can be detected and related to data availability and experience.

We begin by confirming that investors learn with experience (i.e. participating in multiple contests over time) and also detect an overconfidence effect: relative outperformance in a previous contest is associated with poorer performance in the current contest. The prior empirical literature on investor learning has also found evidence of experience and overconfidence effects, in different settings.

The investors in our setting can make use of historical data provided to them by the trading platform. As the overarching goal of our study is to investigate how investors learn from data, we examine whether and how investors make use of this historical data: we find that, with experience, investors learn to improve their performance during contest backtest periods as well as the live trading periods. This suggests that investors do make use of the historical data. We provide an empirical framework to distinguish between this "in-sample" learning and "out-of-sample" learning – a distinction that has not yet been addressed in the literature.

Our main results are on how investor performance outcomes are affected when their common dataset widens. We find that making additional predictive variables available to investors is associated with a narrowing/decrease of the dispersions of her performance outcomes; this is consistent with learning models where an investor's posterior predictive variance narrows with additional information, rather than models where additional information leads an investor to explore among a wider set of candidate models. As for the levels of the performance outcomes (that investors seek to maximize), we find that experienced investors attain better performance outcomes, suggesting that they make use of the additional predictive signals that are available to them to better attain their objectives.

The performance outcomes of inexperienced investors do not change in a statistically significant manner, and we rationalise this finding by a fear of model uncertainty. We formulate a model where an investor aims to maximize her next-period performance by robustly learning the expected returns of the futures contracts in her investment universe from a dataset of historical signals and expected returns of similar historical futures contracts. A fear of worst-case model uncertainty leads the investor to "shrink" the parameter values that she learns based on the value of her subjective uncertainty parameter: the higher this threshold, the more predictive signals that may be discarded entirely. Therefore, an investor can make use of the historical data that is available to her while still choosing to ignore some subset of the available predictive signals, and the more she fears model uncertainty the more signals she will ignore. Excluding signals with predictive power may therefore hold her back from attaining her desired risk preferences: fear of model uncertainty when relying on historical data may work against the investor's ultimate goal. Furthermore, an investor who gains in experience may fear model uncertainty less, which can explain our findings on learning with experience. We provide supporting evidence for this mechanism by focussing on performance outcomes in the periods that immediately follow the release of macroeconomic signals to the trading platform.

The remainder of the paper is organised as follows. We first review the related literature and our study's contributions. In Section 2 we use a simple example to motivate how investors may learn using historical data, and formulate our testable empirical hypotheses. Section 3 gives more detail on our institutional setting and Section 4 describes our data. We present our empirical findings in Section 5. In Section 6 we explain our main empirical finding by developing a model of an investor who solves her portfolio choice problem using robust learning. Section 7 concludes.

Related literature

A recent line of research analyses the implications of data growth for financial markets, focussing on equilibrium outcomes (Farboodi and Veldkamp, 2020; Dugast and Foucault, 2020). We study this topic empirically, focussing on the behavior of individual investors.

We develop empirical hypotheses that build on two distinct approaches to modelling investor learning. The first approach in the literature is Bayesian updating, and this can be extended to incorporate parameter learning (Brennan, 1998; Barberis, 2000; Pástor and Veronesi, 2003, 2009), behavioral deviations like overconfidence (Gervais and Odean, 2001; Daniel et al., 2001) and robustness (Maenhout, 2004; Hansen and Sargent, 2008). An alternative approach to modelling learning is for the agent itself to select or alternate among various models (Hong et al., 2007; Barberis et al., 1998; Branch and Evans, 2010; Arthur et al., 1996), perhaps trading off between exploring new models against exploiting existing ones (Erev and Roth, 1998; Camerer and Ho, 1999). One of our empirical hypotheses seeks to discriminate between these approaches by testing whether systematic investors' performance outcomes widen or narrow at an individual level when additional signals are available to them.

Our study investigates how investors use a historical dataset to attain their objectives, so a particularly relevant approach is to model agents as econometricians – to borrow an analogy from Sargent (1993, pp. 21-23). The Bayesian models of Timmermann (1993) and Martin and Nagel (2019) may also be seen as belonging to this category. Martin and Nagel (2019) considered a risk-neutral agent that learns a model to predict asset payoffs based on historical data; we similarly model an agent that learns how to predict expected returns to attain its risk preferences. To do this, we lean on results at the intersection of the machine learning & optimisation literatures (Xu et al., 2010; Tibshirani, 1996) without assuming that investors make use of these techniques themselves. Gabaix (2014), Croce et al. (2015) and Molavi et al. (2021) modelled bounded rationality using notions of sparsity; we also model agents as ignoring a subset of their environment, but stemming from a different micro-foundation of robustness to model uncertainty. In a different setting, Hanna et al. (2014) found empirical evidence that agents may ignore some available data while learning.

Our study of investor learning and information is chiefly empirical in nature. A recurrent theme in empirical studies of information has been the challenge of quantifying just what "information" is. Fang and Peress (2009) were able to study information production and flows by measuring newsmedia coverage: such public information might be assumed to belong to the common part of investors' information sets. In general, though, measuring precisely what is in investors' information sets is difficult and so prior work has relied on indirect measures: for example, Hong et al. (2007) reasoned about changes in information sets based on changes in covariances of observable variables, Kacperczyk and Seru (2007) did so using regression R^2s, and Biais et al. (2010) proxied for investors' information possession by their realized profits. More recently, Chen et al. (2020) used exogenous changes in analyst coverage as "shocks to [the] information environment", while Cookson and Niessner (2020) condition on investors' trading styles to identify the component of their disagreement that is due to differing information sets.

By contrast, our institutional setting precisely controls the common information sets (including historical data) that investors make their trading decisions based upon, as well as their trading styles and other potential confounders.

It has long been accepted in economics that performance outcomes may improve with experience thanks to a "learning by doing" channel (Arrow, 1962). A number of finance studies have tackled the question of whether and how investors learn with experience by examining individual trading records: Nicolosi et al. (2009) used the same panel of US discount brokerage customers as Barber and Odean (2000) did, and found that investors' risk-adjusted returns and some measures of trading quality improve with a proxy for investor experience. Seru et al. (2010) used a Finnish panel of individual trading records to study how investors learn over time, seeking especially to disentangle the "learning by trading" channel from investor participation/attrition effects, and argued that the second channel was more important. Linnainmaa (2011) also sought to model selection, using a structural model of investors' beliefs and abilities with a panel of Finnish trading records, and arrived at a different conclusion: that estimates of investor performance may be *downward* biased, since investors are more likely to cease participation after a poor run of performance. Barber et al. (2020) argued that the notion of overconfidence is necessary to fully explain speculative trading. Like these studies, we are able to proxy for investor experience; furthermore, we can exactly measure (rather than proxy) investors' performance outcomes. We also more accurately identify learning effects because other potential confounders (preferences, information, trading horizons) are controlled for by our institutional setting. Importantly, we are able to go further than studying experience effects to also study (i) whether investors learn from historical data and (ii) how the addition of predictive signals affects their performance outcomes and hence their learning processes.

As well as information and learning, our study relates to the growing literature on the effects of individuals' beliefs on their financial decisions, as surveyed by Gomes et al. (2020). Giglio et al. (2021) provide evidence that investor beliefs affect their portfolio choices. In our setting, investors must form expectations on the first and second moments of asset returns in order to choose their portfolios, and we are able to observe *ex post* realizations of these.

Our study contributes to the literature on systematic trading, which has hitherto focussed on mutual funds (Abis, 2017) and hedge funds (Fung and Hsieh, 1997; Fabozzi et al., 2008; Khandani and Lo, 2011; Bhardwaj et al., 2014). Rather than studying funds' performance or characteristics, or their managers' incentives, we focus on the role of the individual humans who design systematic trading strategies, and how they learn to attain their objectives.

We join other studies that have sourced data from FinTech platforms in order to gain insights into economic questions: for example, Da et al. (2020) used crowd-sourced rank predictions of stock price dynamics collected by a FinTech platform that ran prediction contests to find evidence of extrapolative beliefs; Tang (2019) made use of data from peer-to-peer lending platforms. To our knowledge, the only paper to make use of trading contest data from a (now defunct) FinTech platform is by Wiecki et al. (2016). Outside the finance literature, such platforms have been the subject of business school case studies (Fleiss et al., 2017; Zheng, 2017).

2 Empirical framework

2.1 Motivation

In this study, we aim to use our unique setting to cleanly identify certain stylised facts about investor learning that should be particularly relevant to a Big Data world. In particular, our setting will allow us to distinguish between investor outcomes where the investors' datasets *widen* rather than simply *lengthen*.

It will be useful to have an empirical framework in mind when approaching our econometric analysis, yet in our literature review we saw that a wide array of different approaches have been taken to model investor learning. What can we say without imposing a particular learning model on investors? In our institutional setting, investors all aim to maximize the Sharpe Ratio of their portfolios during the "live" period that the trading contest is active, so their common objective can be analysed by appealing to the single-period mean-variance framework of Markowitz (1952). In fact, there is a direct analytical link between the solutions of the Sharpe Ratio maximization problem and of the mean-variance portfolio choice problem;[3] for our present purposes we simply note that the investor needs to know the first two moments of the returns of her tradable assets to attain her objective.

Simplifying to the case of a single asset, let us consider that an investor at time t has observed a realisation r_t of the asset's random return $\tilde{r}_t$. As we just discussed, she requires knowledge of the mean and variance of the next-period return $\tilde{r}_{t+1}$ in order to attain her objective. Simplifying a little further by assuming that the variance v is known and constant, the investor's problem reduces to learning the mean of the next-period return. We can model this unknown mean as a random variable itself,

$$\tilde{\mu}_{t+1} = \mu + \tilde{\varepsilon}_{t+1}, \tag{1}$$

where $\tilde{\varepsilon}_{t+1}$ is a zero-mean i.i.d. innovation and μ is an unknown parameter. The investor's *learning* problem is to estimate this unknown mean based on whatever information she has

[3] Under relatively mild assumptions on feasible portfolios and constraint sets, the solution to the Sharpe Ratio maximization problem is simply the tangency portfolio solution to the classical Markowitz (1952) mean-variance portfolio choice problem (Cornuéjols et al., 2018, pp. 102-103).

available, including historical data. In learning $\tilde{\mu}_{t+1} = \mu + \tilde{\varepsilon}_{t+1}$, the investor is interested in learning both components.

Based on a historical dataset of observed realisations (in this case, the single realisation r_t) and the known parameter v, the investor can learn the unknown parameter μ. This would result in good estimates of the mean of the unknown return $\tilde{\mu}_t$ (of which a single realisation was observed) and an apparently good historical performance of her portfolio choice based on her historical sample at time t. We will think of this as *in-sample learning*.

The investor's ultimate goal is for her portfolio to perform well tomorrow ($t + 1$), outside the historical sample of observed realisations. For this, the investor should not only attempt to learn the unknown parameter μ in this setting but also the random disturbance $\tilde{\varepsilon}_{t+1}$ that is set to occur tomorrow. Given some additional signal(s) to this disturbance, the investor can thus obtain a better estimate of $\tilde{\mu}_{t+1}$ overall. We will think of this as *out-of-sample learning*.

A unique feature of our contest dataset is that it reports the Sharpe Ratios (SRs) of contestants' entries on both the in-sample/backtest period (which contestants are also aware of) and the out-of-sample/live period (which contestants do *not* know at the time they develop and submit their systematic trading strategies, since these outcomes are calculated months later). We can thus empirically study both learning channels: studying in-sample (backtest period) SRs enables us to make inferences about investors' in-sample learning mechanisms, and studying out-of-sample (live period) SRs allows us to reason about their out-of-sample learning mechanisms.

2.2 Hypothesis development

We now develop a number of hypotheses to be tested empirically that are consistent with prior work and yet do not impose a particular learning model.

An individual investor's learning mechanisms are inherently unobservable, but we do observe the outcome that the investors in our panel are optimising for. Furthermore, we control for differences in horizons, preferences and information (thanks to our setting) and can also control for time effects and time-invariant individual effects (econometrically or by incorporating a relevant benchmark portfolio). With all these controls in force, therefore, if we observe an improvement in investor outcomes we may attribute it to learning. Based on the empirical evidence of Nicolosi et al. (2009), Seru et al. (2010), Linnainmaa (2011) and Barber et al. (2020), investor outcomes should improve with experience:

Hypothesis H1. *The learning outcome (out-of-sample performance) improves with experience.*

Barber et al. (2020) argued that investor learning and overconfidence mechanisms are jointly in play. In our setting, the measure of past relative performance is clear so we can also test for an overconfidence mechanism:

Hypothesis H2. *The improved out-of-sample performance with experience of **H1** is attenuated by overconfidence based on past relative performance.*

In section 2.1 we highlighted the complementary channels of in- and out-of-sample learning. We believe these mechanisms are also applicable to systematic investors in other settings and (more generally) to any class of investors who commit to engaging in trading activity based on *ex ante* forecasts or beliefs that are not revised over some period. This may include professional investors who are constrained by an institutional mandate to follow particular styles of trading or even trading rules. We therefore seek to disentangle which of these channels are at play, and present the following dual hypotheses:

Hypothesis H3 (dual). *The improvements in out-of-sample performance with experience of **H1** can be explained by a combination of the following:*

- ***a*** *improved in-sample performance.*
- ***b*** *improved generalisability of the in-sample performance to the out-of-sample environment.*

We now consider the important question of how investors learn in the presence of additional predictive signals. Our reasoning here is straightforward: utility-maximizing investors should make use of all available information in pursuit of their goal. As more information is added, therefore, we expect the levels of whatever objective they are maximizing to then increase. In our setting, this is the out-of-sample Sharpe Ratio:

Hypothesis H4. *When new variables are made available, out-of-sample performance increases.*

In order to maximize their out-of-sample performance outcomes, investors must form expectations on these future outcomes. In studying how investors learn, therefore, we are also interested in the nature of these expectations (or forecasts) that they form. Well-specified, unbiased statistical learning rules should converge to the truth as more information is processed. Bayesian learning rules, in particular, should produce forecasts that increase in their posterior precision. Investors who make use of such rules should therefore see the dispersions of their out-of-sample performance outcomes narrow at an individual investor level. On the other hand, investors may view the availability of a wider dataset as an opportunity to explore a more diverse set of strategies that produce very different results out-of-sample, thereby increasing the dispersion of their outcomes at an individual level. By taking the realized dispersion of an investor's out-of-sample performance outcomes to be a proxy for the dispersion of her forecasts, we may thus gain some insight into what (unobservable) learning or forecasting rules an investor actually uses. We therefore present the following alternating hypotheses:

Hypothesis H5 (alternating). *When new variables are made available, individual-level dispersions of investor outcomes either*

a *increase.*

b *decrease.*

We can distinguish between individual-level dispersions (above) and aggregate-level dispersions of investors' performance outcomes. Dugast and Foucault (2020) build a search model to characterise the equilibrium effects of "data abundance". One of that paper's empirical predictions is that, when investors have a common level of risk aversion, the dispersions of their skills are equivalent to the dispersions of the qualities of the predictors that they draw from their searches, and that this dispersion may increase with data abundance. Taking the aggregate-level dispersion of investors' performance outcomes as a proxy leads us to the following hypothesis:

Hypothesis H6. *When new variables are made available, aggregate-level dispersions of investor outcomes increase.*

3 Institutional setting

Systematic trading[4] involves designing and then implementing an algorithm that takes positions in various financial assets based upon a trading strategy that its (human) designers have specified at the outset. It has traditionally been associated with statistical arbitrage hedge funds (who buy and short portfolios of stocks) and so-called "Commodity Trading Advisors" (who trade futures contracts and other derivatives).

Quantiacs is a FinTech platform that runs trading contests for retail investors with the skills to implement a systematic trading strategy. Contestants upload code (in Matlab or Python) to implement a strategy that takes long or short positions in futures contracts, and each strategy's performance is assessed based on its in-sample ("backtest") Sharpe Ratio prior to the start of the contest, and its out-of-sample ("live") Sharpe Ratio during the contest period. The official scores assigned to entries incorporate the out-of-sample Sharpe Ratio, and so this incentivises traders to perform well out-of-sample. In-sample performance is determined from historical daily data and this is visible to traders as they backtest and fine-tune their strategy ahead of a contest launch. The Sharpe Ratios reported by the trading platform include the effect of simulated transaction costs, which the investors also perceive. Out-of-sample performance is based upon market data that arrives after the launch of a contest; i.e. during the Live period, and at this

[4]We use the terms "systematic trading" and "systematic investing" interchangeably. Other studies may also use the terms "algorithmic", "automated" or "quantitative" trading or investing.

point contestants are unable to modify their trading strategy in any way whatsoever. The trading strategy can continue to update portfolio positions at a daily frequency. The distinction between the Backtest and Live periods for a single contest is illustrated in Figure 2. The Live period of each contest lasts for approximately 3 months, and the 12 contests in our panel thus cover a number of years in out-of-sample/live calendar time.

[Insert Figure 2 about here]

To make the institutional setting more concrete, screenshots of the Quantiacs platform are used in Appendix A to illustrate the steps taken by an investor to code up, backtest and then submit a trading strategy entry to a contest.

Importantly, systematic trading strategies can only access a limited set of data inputs throughout their lifetime (both in- and out-of-sample) and this set of available predictive signals does not change over the lifetime of a contest. In fact, it has only widened once *in between* contests, and this event will be used to test two of our hypotheses.

The event in question is the addition of macroeconomic variables to investors' historical and live trading datasets after the end of the 7th contest and before the beginning of the 8th contest. These additional variables are listed in Appendix B. Out of a total of 12 contests, investors in the first 7 therefore had "narrower" datasets than investors in the final 5. It is possible for an investor to enter any contest that she pleases, and there is no entry fee. Contest live periods are non-overlapping.

We are able to observe individual investors' performance outcomes at an individual trading strategy level. For confidentiality reasons, we are unable to observe the positions taken by trading strategies, or the code used to implement them. The platform does not require contestants to provide demographic information or identifying characteristics.

4 Data

4.1 Contest leaderboard panel

Our sample consists of 12 trading contests[5] spanning a number of years. We can identify individual traders who may (and often do) take part in multiple contests over time in order to study learning dynamics. We can also exploit the panel structure of the leaderboard dataset to incorporate fixed effects in our analyses.

[5]Our sample does not include the first contest ever run by the platform because investors were asked to systematically trade US equities instead of futures contracts, making it inconsistent with all other contests. Our sample also excludes a very recent contest (named "Q13") because the available data does not report the negative out-of-sample Sharpe Ratios: using such censored data would lead to biased conclusions.

Table 1 displays some key characteristics of the 12 contests in our sample. Figure 3 visualises the distributions of Sharpe Ratios, conditioning only on the contest index.

[Insert Table 1 about here]

[Insert Figure 3 about here]

Figure 4 shows that a high proportion of contestants in each contest are first-time participants and that, furthermore, many participate only once. The latter fact is consistent with the prior literature and intuition discussed by Linnainmaa (2011). One important inference we can make from this chart is that investors do not appear to be noticeably more or less willing to participate in contests just before or after the new predictive variables were introduced in between contests 7 & 8.

[Insert Figure 4 about here]

While demographic data on the individual traders is not available, we can use their participation attributes to verify that the populations of investors are well-balanced across groups of investors with access to different sets of predictive variables. Table 2 displays the means of these contestant-level attributes, comparing the attributes of contestants who participated only before the additional predictive variables were made available, against the attributes of the contestants who participated only afterwards. The two contestant-level participation attributes represent a contestant's experience level, and a contestant's relative ranking in the previous contest he/she took part in (for contestants who participate in multiple contests). The differences are not statistically different to zero (even at a low significance level of, say, 90%). These groups will be used in a later analysis.

[Insert Table 2 about here]

Similarly, Table 3 compares contest-level attributes before and after the additional predictive variables were made available to investors. The three contest-level participation attributes represent the average level of experience of contestants per contest, the fraction of first-time participants, and the fraction of last-time participants. The relevant differences are also not statistically different to zero (including at low significance levels such as 90%).

[Insert Table 3 about here]

4.2 Macroeconomic announcement dates

Recall that the additional set of predictive signals added to the platform in between contests consists solely of macroeconomic variables (see Appendix B for the full listing). We are able to identify exactly which date each macroeconomic variable release was made available on the platform (for both backtesting & live trading periods); in fact, we can identify exactly which trading date return is the first to incorporate the information contained in each announcement, thanks to the controlled information set that the platform makes available to all trading strategies. We can use this to compare performance outcomes for the periods that immediately follow new releases of the additional predictive signals (i.e. macroeconomic releases) to the trading platform during the later trading contests.

4.3 Timeseries of daily returns for trading strategies

Our analysis will mostly rely on the comprehensive contest leaderboard panel that we have just described. We will also supplement it by exploiting a related dataset: timeseries of the daily returns of trading strategies during the live/out-of-sample trading periods of each contest. We can thus calculate some additional performance metrics for contestants' entries, in addition to the reported Sharpe Ratios on the leaderboard panel.

The limitation of this dataset is that a number of trading strategies from the contest leaderboard panel could not be matched to their corresponding timeseries of daily returns: 19% of the entries in our contest leaderboard panel could not be matched to return timeseries.

4.4 Futures contract prices for a benchmark portfolio

The trading platform provided contestants with historical daily market data for 88 futures (covering the backtest period) with which to formulate their trading strategies. These futures are listed in Appendix C. The platform used the same market data to compute in-sample/backtest Sharpe Ratios. Then, during the live period of each contest, the platform extended the timeseries of daily futures prices with the additional (actual) daily market data from the live period and fed them in as inputs to the systematic trading strategies running on the platform.

We have downloaded historical price data for the same universe of 88 futures, with the exception of Russell 2000 index futures, for which we simply used the index level.[6] We sourced

[6]We used the Russell 2000 index itself to proxy for the series of Russell 2000 index futures contracts because we were unable to source and combine price data for these futures contracts: the Russell 2000 index future has changed its listing multiple times between the ICE and CME exchanges and we do not have access to older historical data. We judged it more useful for our benchmark to possess a long timeseries of the underlying index over the full backtest and live periods. The downside is that the absence of the "basis" between the derivative price and underlying price means that contango/backwardation effects will be omitted, but we expect these to be

the data from Bloomberg, stitching together multiple actively-traded futures contracts with a (standard) difference adjustment.

We did so in order to construct a benchmark portfolio that we will use to compare contestants' performance outcomes against, in order to enable valid comparisons between contestants at different time periods in situations where panel fixed effects cannot be used.

The trading platform does not formally judge contestants against a benchmark, though it does provide each contestant with an example systematic trading strategy that takes equally-weighted long positions in each of the 88 futures contracts available on the platform. Benchmark indices exist for certain sectors – such as the 24-contract Goldman Sachs Commodity Index (GSCI) – but we are not aware of any benchmark for futures as an overall asset class. We have therefore elected to construct the benchmark as simply as possible, based on daily-rebalanced equally-weighted returns of long positions in the most-active contracts of the underlying 88 futures (or, in the special case of the Russell 2000 index, the future's underlying index itself). The use of long-only positions (similar to the GSCI commodity index methodology) is justified because the buyers of futures contracts receive delivery of the underlying physical/cash asset at maturity. The use of equally-weighted positions is justified by the absence of any alternative for weighting derivatives in such diverse underlying assets: unlike stocks or bonds, there is no fixed supply that limits the open interest that is possible in futures contracts.

5 Results

We now test our empirical hypotheses. The specifications in this section take investor outcomes to be their best out-of-sample Sharpe Ratios (SRs). (Appendix D.1 checks that our conclusions are robust to more complex functional forms and Appendix D.2 checks that they are robust to using the platform's official score rather than the out-of-sample Sharpe Ratios.)

5.1 Learning with experience

We first consider how investors learn as they gain in experience (by participating in more trading contests).

[Insert Table 4 about here]

Table 4 displays the results of regressing performance outcomes against the experience levels of the contestants. Performance outcomes are either the in-sample (backtest period) best Sharpe Ratios or the out-of-sample (live period) best Sharpe Ratios of the contestants i for

negligible on the most-active contract of a non-commodity future, as we have here.

each contest period t. A trader's experience is measured by the number[7] of contests she has participated in so far. The full panel specifications

$$\mathrm{SR}_{i,t}^{Best} = \beta \times \text{Contests experienced}_{i,t} + \lambda_i + \delta_t + \epsilon_{i,t} \tag{2}$$

include individual (contestant) fixed effects λ_i and time (contest) fixed effects δ_t: controlling for these two dimensions of unobserved heterogeneity should enable us to draw more accurate conclusions. In both OLS & panel specifications, the positive (and significantly non-zero) coefficient β on the number of contests so far is evidence for investor learning with experience.[8] The outcome variable in columns (3)-(4) is the live/out-of-sample performance outcome, so the positive estimates for β in those columns support our first hypothesis **H1** and confirm previous studies that find a learning with experience effect, such as Nicolosi et al. (2009) or Seru et al. (2010). The outcome variable in columns (1)-(2) is the backtest/in-sample performance outcome, so the positive estimates for β in those columns suggest that investors make use of the historical data that is available to them, since the counterfactual performances of their trading strategy entries on historical data improve concurrently.

[Insert Table 5 about here]

Given our ability to separately identify in-sample vs. out-of-sample performance outcomes, the next step of our analysis is to further study the relationship between these two outcomes. The regression results in Table 5 are variants of the following specification:

$$\begin{aligned} \text{Live SR}_{i,t}^{Best} &= \beta_1 \times \text{Contests experienced}_{i,t} + \beta_2 \times \text{Backtest SR}_{i,t}^{Best} + \beta_3 \times \text{Log Percentile}(\text{Score}_{i,t-1}^{Best}) \\ &+ \beta_4 \times \text{Contests experienced}_{i,t} \times \text{Backtest SR}_{i,t}^{Best} + \lambda_i + \delta_t + \epsilon_{i,t}. \end{aligned} \tag{3}$$

Throughout, the dependent variable is the live/out-of-sample performance outcome, and the backtest/in-sample performance variable now appears as a covariate on the right-hand-side of the specification. The small positive estimates for this term's coefficient β_2 indicate that improved out-of-sample performance is associated with improved in-sample performance. Furthermore, the magnitude of the coefficient β_2 is less than one, indicating that performance deteriorates out-of-sample, which agrees with intuition. When both individual and time-invariant heterogeneity are controlled for by fixed effects in column (8), the coefficient estimate is statistically significant. Column (8) also introduces a covariate for the prior relative performance

[7]Results are robust to the use of experience dummies rather than an integer variable: see Appendix D.1.

[8]Note that this channel of learning with experience is distinct from a statistical effect of using additional samples to form more accurate forecasts: while the number of contests a single investor participates in does increase with calendar time (i.e. length of historical samples), investors with equal levels of experience may have participated in different contests with different lengths of historical datasets available at the time the contests occurred.

measured by the (log) percentile of the official score that the investor's best previous entry had attained in the prior contest she participated in; the negative coefficient β_3 on this term is evidence for an overconfidence effect that acts counter to the main effect of learning with experience. This result supports **H2** and is consistent with the argument of Barber et al. (2020).

Returning to our dual hypotheses on the drivers of the learning with experience effect, the panel regression in column (9) in Table 5 introduces an interaction term between the highest in-sample SR and the experience variable: the lack of statistical significance on the coefficient of this interaction term β_4 indicates that learning with experience effects do not consist of an improvement in the ability of an in-sample performance figure to be generalised to the out-of-sample period (which would have manifested in a positive and significant coefficient). In other words, the improved out-of-sample performance outcomes that we noted previously are actually driven by improved in-sample performance. We conclude that an absence of improved generalisability suggests that it is in-sample learning (**H3a**) that underpins the positive effect of experience on investor outcomes and not a generalisation effect (**H3b**).

[Insert Table 6 about here]

Investors are able to submit multiple systematic trading strategies as entries to a contest, so we now study the determinants of this behavior using Poisson GLM regressions that take the number of entered strategies as the dependent variable, with the following form:

$$\begin{aligned}\text{Entries}_{i,t} \sim \textit{Poisson}\Big(\exp\big\{\, &\beta_1 \times \text{Contests experienced}_{i,t} + \beta_2 \times \text{Live SR}^{Best}_{i,t} \\ &+ \beta_3 \times \text{Log Percentile}(\text{Score}^{Best}_{i,t-1}) + \beta \times \text{Contestants}_t \,\big\}\Big). \end{aligned} \tag{4}$$

Regression results are displayed in Table 6. The positive estimates for β_1, β_2 indicate that submitting more entries in a contest is positively associated with both experience and an improved out-of-sample performance outcome. The positive estimates for β_2 may be due to this term capturing individual skill in the absence of fixed effects and it does not reduce the magnitude of the main effect. As for the overconfidence channel, it is unclear *a priori* whether the overconfidence effect would be associated with more entries (similar to the overtrading effects documented in the literature) or with fewer (since more entries in this situation are associated with better performance): the positive estimate for β_3 appears to favour the former explanation. The insignificant and near-zero estimate for β_4 may be evidence for the absence of any strategic effects during the current contests: investors neither enter more strategies nor fewer depending on how many other contestants are taking part.

The evidence so far supports **H1**, **H2** and **H3a**. We now examine the robustness of these results to a possible selection bias between the number of contests and the performance outcomes: since we observe attrition in our panel of investors, it is important to consider this issue. Unlike Seru et al. (2010) and Linnainmaa (2011), modelling participation and selection is not a primary object of our study, but it will be important for us to have confidence in our results to understand what direction any bias might introduce to our specifications. We therefore re-estimate the magnitude of learning with experience effects using a Heckman (1976)-type model, with standard errors computed according to Greene (1981). The first stage of the procedure is a probit model of the participation of our contestants in the next contest, and this is explicitly based upon a number of covariates that we specify. We argue that two of the covariates that we make use of in the first-stage selection equation satisfy an exclusion restriction: the mean of the Google Search Index for "Quantopian"[9] and the *relative* number of entries that the contestant has made in the *prior* contest compared to the current contest.

[Insert Table 7 about here]

The first-stage coefficients are shown in the top section of Table 7, and the estimates indicate that contestants are less likely to participate in the next contest as they gain in experience and as the rival trading platform (Quantopian) gains in attention, and are more likely to participate once more if they have participated more intensively in the prior contest or ranked relatively higher in the prior contest. The second-stage coefficients represent our outcome equation of interest, and these are shown in the middle section of Table 7. The coefficient on contests so far is positive in 3 of the 4 specifications we estimate. More importantly, the estimates of the ρ parameter in the second stage of the Heckman (1976) model are all negative, and therefore so are the coefficients on the Inverse Mills Ratio shown in the line below. (Note that the statistical significance actually improves in the robustness checks in Appendix D.2). The estimated sign on ρ therefore indicates the presence of *negative selection* in an OLS estimate of the magnitude of the effect of learning with experience. To put it another way, uncorrected OLS estimates of the magnitude of the learning effect would be biased *downwards*. This result is in agreement with the intuition of Linnainmaa (2011) and should reassure us that our earlier findings on learning with experience are not in fact driven by selection/attrition bias (or, as Seru et al. (2010) framed it, an investor learning about her type).

[9]Quantopian was one of Quantiacs' rival FinTech platforms for running trading contests (Fleiss et al., 2017; Zheng, 2017). The logic of our exclusion restriction is that the contestants on each platform may exhibit substitutability in which platform they participate in. It seems implausible to argue that search interest in and attention to a rival platform (Quantopian) could affect performance outcomes in our focal setting (Quantiacs) through any channel other than an effect on participation. Note that the rival Quantopian platform is now defunct (as announced on 29 October 2020) but was active throughout our sample period.

5.2 Learning with additional predictive variables

To sum up our analysis so far, we first revisited some effects that have been previously studied: learning with experience, the direction of selection bias, and the presence of overconfidence. We then provided some novel insights into the dual channels of in-sample and out-of-sample learning. We now turn to our primary concern: identifying what effects the availability of a wider dataset has on investor learning.

To cleanly identify the effects of investor learning, we divide contestants into Treatment & Control groups. Investors in the Control group have only participated in contests before the introduction of the additional macroeconomic variables to the Quantiacs platform, and investors in the Treatment group have only ever participated after their inclusion. Recall that we already verified in Section 4.1 that these two groups are well-balanced in terms of observable characteristics. To identify the effects of the additional predictive variables, we must assume that any changes to investors' learning dynamics is attributable to them having access to these additional predictive variables.

5.2.1 Effect on levels of performance outcomes

[Insert Table 8 about here]

The regressions in Table 8 analyze the relationship between investor experience and the availability of additional predictive signals (as covariates) and the out-of-sample performance outcomes (as the response), using variants of the following specification:

$$\begin{aligned}
&\text{Live SR}^{Best}_{i,t} - \text{Benchmark SR}_t \\
&= \beta_1 + \sum_{k=2}^{4}\big[\beta_k \times \mathbb{1}\{\text{Contests experienced}_{i,t} = k\}\big] + \gamma_1 \times \mathbb{1}\{\text{New variables available}_t\} \\
&\quad + \sum_{k=2}^{4}\big[\gamma_k \times \mathbb{1}\{\text{Contests experienced}_{i,t} = k\} \times \mathbb{1}\{\text{New variables available}_t\}\big] + \epsilon_{i,t}. \qquad (5)
\end{aligned}$$

The dummy variable named $\mathbb{1}\{\text{New variables available}_t\}$[10] indicates whether investors compete after the introduction of the additional predictive variables to the common parts of investors' information sets (i.e. whether investors belong to the Treatment group). In columns (3) & (6) this treatment dummy is interacted with dummies for investor experience to capture the effects of the availability of the new predictive variables on the performance outcomes of investors with comparable levels of experience between groups ($\gamma_k - \beta_k$).

[10]Note that since contestants are divided into two groups, we cannot identify individual-level fixed effects. Similarly, since the dummy for the availability of new predictors is time-invariant on a per-contest level, we can no longer identify contest-level fixed effects either. This leads us to use a benchmark portfolio instead.

The specifications in columns (1)-(3) simply take the out-of-sample SR to be the response variable. The remaining columns (4)-(6) adjust these performance outcomes by taking the excess of the out-of-sample SR over the SR from holding the benchmark index (equally-weighted buy-and-hold positions in contestants' tradeable universe of futures contracts) in order to permit valid comparisons across time periods. When the benchmark is taken into account in this way in column (6), we can see that the interactions between the dummy variable for treatment and dummies for investor experience have positive (and significantly non-zero) coefficients γ_k for higher levels of experience $k \geq 3$, in particular. The availability of wider data is thus associated with a steepening of investors' performance outcomes. In other words, we find evidence that more experienced investors make use of the additional predictive variables to better attain their objectives: our panel of investor outcomes supports **H4** for investors with higher levels of experience. This finding is a key empirical result.

[Insert Table 9 about here]

We now augment the previous specification (5) with two new covariates: the backtest/in-sample SR, and its interaction with the treatment dummy. Table 9 displays the regression results for variants of this new specification:

$$\begin{aligned} \text{Live SR}_{i,t}^{Best} - \text{Benchmark SR}_t \;=\; & \alpha + \beta_1 \times \text{Contests experienced}_{i,t} \\ & + \beta_2 \times \text{Backtest SR}_{i,t}^{Best} + \beta_3 \times \mathbb{1}\{\text{New variables available}_t\} \\ & + \beta_4 \times \text{Contests experienced}_{i,t} \times \mathbb{1}\{\text{New variables available}_t\} \\ & + \beta_5 \times \text{Backtest SR}_{i,t}^{Best} \times \mathbb{1}\{\text{New variables available}_t\} + \epsilon_{i,t}. \end{aligned} \tag{6}$$

As in the previous table, columns (4)-(6) take the benchmark SR into account as an implicit control. The estimated coefficients β_5 on the interaction term between the in-sample SR and treatment dummy are not significant, which indicates that the availability of the additional predictive variables does not seem to affect the generalisability of in-sample performance to the out-of-sample periods.

From studying the levels of the out-of-sample performance outcomes we can thus conclude that experienced investors appear to better attain their objectives when the additional predictive variables are available (**H4**). The availability of wider data has a beneficial effect that interacts with an investor's level of experience: we will rationalise this finding in Section 6.

5.2.2 Effect on the *ex post* realized moments of portfolio returns

An investor who seeks to maximize her out-of-sample portfolio Sharpe Ratio, as our contestants do, is faced with the joint problem of maximizing the out-of-sample returns of the portfolio while minimizing their variance; this is due to the construction of the Sharpe Ratio statistic. We observe timeseries of the daily returns of a subset of contestants' portfolios during the live period of each contest. Using these available timeseries, we thereby decompose the realized Sharpe Ratios by estimating the mean and standard deviation of these contestants' out-of-sample portfolio daily returns. We then relate these to experience levels and the availability of additional predictive variables, as before.

[Insert Table 10 about here]

We modify the main specification (5) by regressing the *ex post* estimated means of the out-of-sample daily returns of trading strategies against interacted dummies for the availability of additional predictive variables and investor experience levels:

$$\begin{aligned}
&\text{Mean}_{i,t}(\text{Best Live entry daily returns}) \\
&= \beta_1 + \sum_{k=2}^{3}\left[\beta_k \times \mathbb{1}\{\text{Contests experienced}_{i,t} = k\}\right] + \gamma_1 \times \mathbb{1}\{\text{New variables available}_t\} \\
&\quad + \sum_{k=2}^{3}\left[\gamma_k \times \mathbb{1}\{\text{Contests experienced}_{i,t} = k\} \times \mathbb{1}\{\text{New variables available}_t\}\right] + \epsilon_{i,t}. \qquad (7)
\end{aligned}$$

Regression results are displayed in Table 10, with the dependent variable in columns (1)-(3) constructed using raw returns, and the dependent variable in columns (4)-(6) adjusting for the benchmark portfolio's daily returns as an implicit control. Whether or not the benchmark adjustment is applied, the availability of additional predictive variables is associated with higher out-of-sample mean returns for more experienced investors, and the estimates for these coefficients γ_k are significantly different to zero.

[Insert Table 11 about here]

A similar modification of the main specification (5) and previous specification (7) is to regress the *ex post* estimated standard deviations of the out-of-sample daily returns of trading strategies against interacted dummies for the availability of additional predictive variables and investor experience levels:

$$\begin{aligned}
&\text{SD}_{i,t}(\text{Best Live entry daily returns}) \\
&= \beta_1 + \sum_{k=2}^{3}\left[\beta_k \times \mathbb{1}\{\text{Contests experienced}_{i,t} = k\}\right] + \gamma_1 \times \mathbb{1}\{\text{New variables available}_t\} \\
&\quad + \sum_{k=2}^{3}\left[\gamma_k \times \mathbb{1}\{\text{Contests experienced}_{i,t} = k\} \times \mathbb{1}\{\text{New variables available}_t\}\right] + \epsilon_{i,t}. \qquad (8)
\end{aligned}$$

Regression results are displayed in Table 11. As above, columns (1)-(3) use raw returns and columns (4)-(6) adjust for the benchmark portfolio's daily returns before estimating their standard deviation, in order to implicitly control for differing contest time periods. Once again, the conclusions do not depend on whether the benchmark adjustment is applied. While the estimates for coefficients γ_k in columns (3) & (6) again agree with intuition, this time they are mostly not significantly different to zero.

Our results indicate that the improvements in out-of-sample Sharpe Ratios (performance outcomes) that we previously associated with the availability of additional predictive variables (in Table 8, for example) can be attributed mainly to higher out-of-sample mean returns (Table 10). While we also detect lower out-of-sample standard deviations, these additional effects are mostly statistically insignificant (Table 11).

5.2.3 Effect on individual-level dispersions of performance outcomes

We now return to our analysis of individual investors' performance outcomes (out-of-sample Sharpe Ratios). We have so far studied the locations/levels of these performance outcomes, yet it is also important to study the dispersions of these outcomes, and how these dispersions relate to the availability of additional predictive variables: evidence on whether these dispersions widen or narrow with the availability of wider data is useful for understanding investors' learning mechanisms. We exploit the fact that an investor i may (potentially) enter multiple trading strategies j into a contest t, allowing us to measure individual investor-level dispersions.

[Insert Table 12 about here]

Table 12 displays regression results for specifications in which the response variables are these individual-level dispersions of investor outcomes:

$$\begin{aligned}\text{Range}_{i,t}(\text{SR}_{i,j,t}) \;&= \alpha + \beta_1 \times \mathbb{1}\{\text{New variables available}_t\} + \beta_2 \times \text{Contests experienced}_{i,t} \\ &+ \beta_3 \times \text{Entries}_{i,t} + \beta_4 \times \text{Backtest SR}^{Mean}_{i,t} + \beta_5 \times \text{Live SR}^{Mean}_{i,t} + \epsilon_{i,t} \qquad (9)\end{aligned}$$

Columns (1)-(4) use in-sample dispersions and columns (5)-(8) use out-of-sample dispersions. Crucially, the availability of additional predictive variables is associated with a decrease/narrowing in the individual-level dispersions of investors' out-of-sample performance outcomes, represented by the coefficient β_1 in columns (5)-(8). The direction of the effect is unchanged when controls are added in columns (7)-(8) for an investor's experience, the number of trading strategies entered into the contest or for the levels performance outcomes themselves. The main effect β_1 is statistically significantly different to zero when all controls are included, in column (8).

Therefore, the evidence supports **H5b** (narrower dispersions) rather than **H5a** (wider dispersions) for the out-of-sample performance outcomes. This is more consistent with models of investor learning in which investors exploit additional data to make more precise forecasts (such as Bayesian updating models) than with models in which investors explore a wider, more diverse set of candidate models when additional data is available.

5.2.4 Effect on aggregate dispersions of performance outcomes

[Insert Table 13 about here]

The previous *individual*-level results do not carry through to dispersions of *contest*-level aggregate outcomes, which we study using regressions of the following form:

$$\begin{aligned}\mathrm{SD}_t(\mathrm{SR}_{i,t}^{Best}) &= \alpha + \beta_1 \times \mathbb{1}\{\text{New variables available}_t\} \\ &\quad + \beta_2 \times \mathrm{Mean}_t(\text{Contests experienced}_{i,t}) + \beta_3 \times \mathrm{Mean}_t(\text{Entries}_{i,t}) + \epsilon_t. \qquad (10)\end{aligned}$$

Regression results are displayed in Table 13. The positive estimates for the coefficient β_1 in columns (1)-(2) indicate that the availability of additional predictive variables is associated with an increase/widening in the aggregate dispersions of investors' in-sample performance outcomes. There is no statistically discernible effect for the aggregate dispersions of the out-of-sample performance outcomes, in columns (3)-(4). Our conclusions are not affected by the inclusion of contest-level controls in columns (2) & (4). These results partially support **H6**. The model of Dugast and Foucault (2020) does not distinguish between in- and out-of-sample performance outcomes; our in-sample results are consistent with their model's predictions.

6 The value of additional predictive variables

Our empirical findings in Section 5.2.1 supported the hypothesis that investors would make use of additional predictive variables in order to attain their goals (of maximizing their out-of-sample Sharpe Ratios). Intriguingly, these statistically significant positive effects were detected for investors with higher levels of experience, in particular. In this section we seek to rationalise why these effects were not statistically significant for inexperienced investors.

In developing this hypothesis (**H4**) we made the assumption that investors would rationally use all available information and that doing so would aid them in maximizing their objectives. The performance outcomes for the more experienced investors did improve, suggesting that the additional predictive variables could indeed be beneficially used for this purpose. Evidence from a 2019 survey[11] also supports our intuition: 72% of investment firms stated they gained value from alternative data.

A potential explanation is that the inexperienced investors did not make full use of the historical data that was available to them. In the case of humans making decisions alone it would seem reasonable to attribute this to bounded rationality or some cognitive constraint on information processing capacity; however, the investors in our setting are required to systematically process all available signals and generate trades using computer code, so this may be less convincing.

We therefore propose the following explanation: it can be rational for decision-makers to ignore a subset of available predictive signals when they wish to make decisions that are robust to model uncertainty. This explanation is also compatible with a formal objective that investors should maximize their out-of-sample Sharpe Ratios over a single period, as our particular institutional setting requires. Furthermore, an investor who more greatly fears model uncertainty would use fewer predictive signals in attaining her objective, allowing us to link fear of model uncertainty to inexperience.

We now formalise this theory in a model of portfolio choice with learning, taking inspiration from Martin and Nagel (2019) and Hansen and Sargent (2008). We close this section by providing some supporting empirical evidence on how the systematic investors in our setting make use of additional predictive variables when they are made available.

6.1 Portfolio choice with learning

Since investors are incentivized by the contest setup to maximize their single-period out-of-sample Sharpe Ratios, we can define an investor's objective for a particular contest as

$$\max_{\boldsymbol{w}} \frac{\boldsymbol{\mu}^T \boldsymbol{w}}{\sqrt{\boldsymbol{w}^T \boldsymbol{\Sigma} \boldsymbol{w}}}, \tag{11}$$

that is, to maximize the Sharpe Ratio of a portfolio of weights $\boldsymbol{w}$ on all tradable futures contracts given a vector $\boldsymbol{\mu}$ of expected returns of those futures contracts and a variance-covariance matrix $\boldsymbol{\Sigma}$. There are no shorting constraints, and we omit any discussion of transaction costs. As we mentioned in Section 2.1, the solution to (11) may coincide with the solution to the classical Markowitz (1952) mean-variance portfolio choice problem, given some mild assumptions.

In reality, a learning problem arises because the moments of asset returns are not known. To make progress, let us assume that $\boldsymbol{\Sigma}$ is known and so the investor must therefore estimate

[11] Greenwich Associates: *"Demystifying Alternative Data"*, Q2 2019, `https://cdn.ihs.com/www/prot/pdf/0519/Demystifying-Alternative-Data-FINAL.pdf`

the unknown parameter vector $\boldsymbol{\mu}$ in order to achieve her objective.[12]

Considering just one of those futures contracts, the investor needs to estimate its unknown expected return μ in order to solve her portfolio choice problem (and to repeat the exercise for all other tradable futures contracts). She receives m signals $s_1, s_2, \ldots, s_m$ that can be used to form an estimate $\hat{\mu}$. We assume that she is aware of the true functional form of the relationship and that it is some linear combination of the signal values,

$$\mu = \sum_{i=1}^{m} b_i s_i = \boldsymbol{s}\,\boldsymbol{b}, \tag{12}$$

collecting the signal values in a row vector $\boldsymbol{s}$ and linear coefficients in a column vector $\boldsymbol{b}$.

Despite knowing the true functional form, the parameters $\boldsymbol{b}$ themselves are unknown to the investor, so she must proceed by *learning* them based on the historical data of similar futures contracts. Recall that all futures contracts mature at predefined dates; for example, there are four S&P500 E-mini futures contracts traded on the CME that mature each year (at predefined dates in March, June, September and December). Therefore an investor who wishes to estimate the expected return μ of a specific futures contract before it matures can make use of historical data for similar futures contracts that have matured in the past.

More formally, to learn these unknown parameters $\boldsymbol{b}$, the investor collects t previous realisations, each of which relates to the expected return v of some past futures contract and the corresponding prior signals available $s_1, s_2, \ldots, s_m$ at the time. Let us arrange these t sets of historical samples in t-dimensional column vectors $\boldsymbol{v}$ and $\boldsymbol{s}_1, \boldsymbol{s}_2, \ldots, \boldsymbol{s}_m$, respectively. For convenience, define the $t \times m$ data matrix $\boldsymbol{S} := [\boldsymbol{s}_1 \;\; \boldsymbol{s}_2 \;\; \ldots \;\; \boldsymbol{s}_m]$. Then the investor's learning problem is to determine the values of these unknown parameters based on the historical data that she observes. Defining $\mathrm{Err} : \mathbb{R} \times \mathbb{R} \to [0, +\infty)$ as some measure of error/deviance between its scalar arguments, her objective during the learning process will be to minimize the error/deviance between the observed historical expected returns $\boldsymbol{v}$ and her predictions based on the historical signal values $\boldsymbol{S}$:

$$\min_{\boldsymbol{b} \in \mathbb{R}^m} \mathrm{Err}(\boldsymbol{v}, \boldsymbol{S}\boldsymbol{b}) \tag{13}$$

Minimizing the errors in the estimated expected returns (13) directly improves the investor's ability to attain her objective: in fact, Best and Grauer (1991) showed analytically that either over- or under-estimating assets' expected returns can lead to large increases in either the mean or variance of the solution to the closely-related Markowitz (1952) portfolio choice problem.

[12]Empirically, it is well known that the first moment of asset returns is more difficult to estimate than the second moment. Theoretically, an asset's volatility can even be calculated exactly from continuously-observed returns (Back, 2017, pp. 594). Merton (1980) discusses these empirical and theoretical considerations in detail. Recall also that we made a similar assumption for the simple model in Section 2.1.

In our setting, then, both over- and under-estimates of the expected returns may lead to sub-optimal Sharpe Ratios. Knowing this, an investor who is learning to estimate expected returns should seek out a functional form for the error/deviance function in (13) that penalises both over- and under-estimates. This would aid her in pursuit of her ultimate objective: choosing a portfolio that maximizes the out-of-sample Sharpe Ratio.

6.2 Robust learning under model uncertainty

In characterising the investor so far, we have seen that she has well-defined preferences relating to Sharpe Ratio maximization, as required by our institutional setting. In the spirit of Sargent (1993) and Martin and Nagel (2019), she also acts as an econometrician who must learn about futures contracts' expected returns from historical data of similar contracts that have matured in the past. We now go further and consider what effects model uncertainty can have on her learning mechanisms.

Recall that the investor is aware of the functional form of any specific futures contract's expected return (with only the exact parameter values $\boldsymbol{b}$ unknown to her) and has access to a (potentially vast) set of historical signals and their realisations for similar futures contracts. However, she is painfully aware that these historical data only apply to other assets, albeit comparable ones, and is concerned at the validity of using these historical data to learn the expected return relationship that is of interest to her. Setting up her learning problem (13) requires a specification of how the investor deals with her problem of model uncertainty; we now formulate it in a manner that explicitly captures her fear of model uncertainty while also remaining tractable. In what follows, $||\cdot||_1$ and $||\cdot||_2$ denote the ℓ_1/taxicab and ℓ_2/Euclidean norms of a vector, respectively.

Definition 6.1. *A robust learner determines the unknown parameters $\boldsymbol{b}$ by considering the worst-case error that may result for any given choice of parameter values,*

$$\min_{\boldsymbol{b}\in\mathbb{R}^m} \max_{\boldsymbol{U}\in\mathcal{U}} ||\boldsymbol{v} - (\boldsymbol{S}+\boldsymbol{U})\boldsymbol{b}||_2, \tag{14}$$

where the model uncertainty can be interpreted as a matrix of signal-wise perturbations $\boldsymbol{U}$ that maximizes the ℓ_2 norm-based error for any choice of $\boldsymbol{b}$ and is constrained by an uncertainty set

$$\mathcal{U} := \left\{[\boldsymbol{u}_1\ \boldsymbol{u}_2\ \dots\ \boldsymbol{u}_m] : ||\boldsymbol{u}_i||_2 \le \delta_i\ \forall\ i = 1,\dots,m\right\} \tag{15}$$

that is characterised by a set of upper bounds $\delta_i \ge 0$ on the ℓ_2 norm of each possible signal-wise disturbance $\boldsymbol{u}_i$.

Remark 6.2. *Replacing the ℓ_2 norm in the objective* (14) *with a squared loss (to give a residual sum of squares) would not affect our subsequent results, thanks to the monotonicity of $x \mapsto x^2$.*

Definition 6.1 frames the investor's robust learning problem as attempting to minimize the worst-case prediction error while a malevolent opponent (Nature) conspires to maximize it up to the constraints permitted by $\mathcal{U}$. The higher the bounds δ_i, the worse the worst-case error and the more conservative the learner (as will be explained shortly). This is similar to the setup of a zero-sum game, or indeed the robust control setups of Hansen and Sargent (2008), whose framing has inspired our own.

Proposition 6.3. *A robust learner solves her original problem* (14) *by solving an equivalent formulation*

$$\min_{\boldsymbol{w} \in \mathbb{R}^m} ||\boldsymbol{v} - \boldsymbol{S}\boldsymbol{b}||_2^2 + \lambda||\boldsymbol{b}||_1, \tag{16}$$

where $\lambda \geq 0$ is a scaling of $\delta := \max_i \delta_i$ in (15).

Proof. Xu et al. (2010) Theorem 1 shows a direct equivalence between the objectives when the cost function is the ℓ_2 norm, and Appendix A of that paper shows a further equivalence to our squared-ℓ_2 norm formulation (16). □

Assumption 6.4. *Assume that the data matrix $\boldsymbol{S}$ is orthonormal: $\boldsymbol{S}^T\boldsymbol{S} = \boldsymbol{I}$.*

The purpose of this assumption[13] is to enable a closed-form solution to (16):

Proposition 6.5. *Under Assumption 6.4, the solution to the robust decision-maker's prediction problem is to predict*

$$\widehat{\mu} = \boldsymbol{s}\widehat{\boldsymbol{b}} \tag{17}$$

where $\boldsymbol{s}$ is an m-dimensional row vector consisting of the m predictive signals to the currently traded asset's expected return μ, and $\widehat{\boldsymbol{b}}$ is an m-dimensional column vector of learned parameters whose elements are defined by

$$\widehat{b}_k = \operatorname{sign}(\boldsymbol{s}_k^T\boldsymbol{v}) \max\left\{|\boldsymbol{s}_k^T\boldsymbol{v}| - \lambda, 0\right\} \tag{18}$$

Proof. Due to Tibshirani (1996), starting from the formulation (16). □

Remark 6.6. *To interpret* (18)*, note that $\boldsymbol{s}_k^T\boldsymbol{v}$ would simply be the parameter value that the decision-maker would have learned if she had used OLS regression to solve her problem instead, under Assumption 6.4.*

[13] To see that this modelling assumption is quite innocuous, recall that Gram-Schmidt orthogonalization can be applied to asset pricing factors to aid in their analysis: see Back (2017, pp. 142) for an example.

Following on from Remark 6.6, let us study the functional form (18) for a moment: the investor would learn parameter values $\widehat{\boldsymbol{b}}$ that are “shrunken” in comparison to what they would have been had she used OLS regression. This is an example of “soft-thresholding”, which we illustrate in Figure 5.

[Insert Figure 5 about here]

Importantly, the robust learner now takes as zero any parameters $\widehat{b}_k$ that empirically would have seemed to be very small in value ($\leq \lambda$, to be precise) if she had used OLS regression, thus effectively ignoring the corresponding signal s_k when she estimates $\widehat{\mu}$. In other words, the investor will ignore a signal s_k if its historical predictive contributions $|\boldsymbol{s}_k^T \boldsymbol{v}|$ fall short of a subjective threshold λ that captures the level of model uncertainty that she experiences. Therefore, the higher her model uncertainty λ, the more predictive signals she will ignore.

Remark 6.7. *By adding assumptions it is possible to view the model from a probabilistic perspective: Park and Casella (2008, Equation 3) write an equivalent Bayesian hierarchical model for the robust learning problem* (16). *Furthermore, Hans (2009, Equation 5) derives a posterior predictive distribution for $\widehat{\mu}$ that is a Gaussian with a variance that is a nonlinear function of the new signals $\boldsymbol{s}$, the historical data (signals $\boldsymbol{S}$ & expected returns $\boldsymbol{v}$) and the subjective parameter λ (indirectly).*

It is worth clarifying just what our present model of robust learning can explain: it can provide a clear rationale and intuition for why the investor may wish to ignore certain predictive signals $\boldsymbol{s}$ based on historical data $\boldsymbol{S}$ & $\boldsymbol{v}$ and a subjective model uncertainty parameter λ. Remark 6.7 points out that it does not provide testable predictions about the posterior predictive variance of $\widehat{\mu}$ due to its nonlinear dependence on historical data and the current signals, so other parts of the investor’s learning mechanisms lie outside the scope of this formal analysis.

6.3 Consequences of robust learning

Our model of robust learning illustrates a potential tradeoff between an investor’s desire to attain her risk preferences and her fear of model uncertainty due to a reliance on historical data. The investor learns from historical data but also fears the model uncertainty that results from her use of this historical data. As an important consequence, she may discard certain signals with predictive power even though they are part of her information set.

Fear of model uncertainty may therefore explain why inexperienced investors in our setting with access to additional predictive variables do not gain a statistically significant benefit in terms of their objective (out-of-sample Sharpe Ratio maximization) when compared to other inexperienced investors without access to those predictive signals. It may also explain why more experienced investors do gain such a benefit when compared to other experienced investors. We theorise that inexperienced investors who must learn from historical data may more greatly fear model uncertainty (i.e. have a high subjective threshold λ) and so may discard some predictive variables, leading them to fall short of attaining their risk preferences. Similarly, more experienced investors may fear model uncertainty less (i.e. have a lower subjective threshold λ) and therefore make use of more of the predictive signals that are available to them, helping them to better attain their objective.

Holding fixed the number of predictive signals, our theory of how investors learn from historical data may also explain our general finding that investors' performance outcomes improve with experience: as an investor gains in experience, she may fear model uncertainty less (i.e. reduce her subjective threshold λ) and therefore incorporate more predictive signals into her portfolio choice decision, which would help her attain a higher level of her objective.

Model uncertainty of this kind may also have implications for market efficiency: Proposition 6.5 and the subsequent discussion implies that any signal s_k to the asset's expected return may not find its way into the price if the historical signals $\boldsymbol{s}_k$ contributed to predicting (historical) expected returns by an amount that falls below some investor-specific subjective model uncertainty threshold. Of course, agents may be heterogenous in such a parameter and – outside our present setting – may have access to private predictive signals. Still, such a model seems worthy of more detailed study in the context of a wider analysis of market efficiency.[14]

Another consequence is that robust investors would make decisions based on biased expectations of returns, in a similar manner to the model of Martin and Nagel (2019), and this may lead to some form of return predictability. Such biased expectations are not necessarily irrational, as argued by Lim (2001) and normative theory[15] in statistics and machine learning.

Finally, since our robust learner may choose to ignore certain predictive signals when solving her portfolio choice problem, she is effectively maximizing her objective while adopting a simplified and sparse model of the world. This is consistent with the (non-financial) empirical evidence of Hanna et al. (2014). By employing the notion of sparsity, our model is also similar to the behavioral models of Gabaix (2014), Croce et al. (2015) and Molavi et al. (2021), except that we do not interpret our robust learners as being boundedly rational. Our model is also loosely related to that of Schwartzstein (2014) and other models of inattention (Gabaix, 2019).

[14] One relevant example of studying the equilibrium consequences of trades between multiple agents who learn from historical data is the paper by Balasubramanian and Yang (2020), in which individual agents face a similar forecasting problem to that studied by Martin and Nagel (2019).

[15] Hastie et al. (2009) repeatedly discuss a bias-variance tradeoff: that trading off bias against variance aids in making accurate predictions.

6.4 Evidence on the usage of additional predictive variables

We now provide empirical evidence in support of the proposed mechanism. Our analysis exploits the fact that all of the additional predictive variables that were suddenly made available by the trading platform are lower-frequency macroeconomic variables, in contrast to the daily market data signals. We match the daily returns of systematic investors' trading strategies to the exact trading dates at which new observations of the low-frequency macroeconomic variables were made available as inputs to investors' systematic trading strategies. We also match them to the simulated daily returns of trading strategies that were entered into earlier contests, when these variables were *not* actually available at the time, as a placebo.

Investors seek to maximize their trading strategy Sharpe Ratios using the data that is available to them at the time. Therefore, we can reason about whether they make use of macroeconomic predictive variables by examining their trading strategy performance outcomes immediately following macroeconomic variable releases (to the platform). Relatively better performance immediately following macroeconomic variable releases suggests that a trading strategy is more likely to be making use of these new variables, as compared to another trading strategy that performs relatively worse over the same short post-release period.

[Insert Table 14 about here]

We implement this analysis by conducting regressions of Sharpe Ratios[16] computed from trading strategy daily returns against interacted experience and macroeconomic data availability dummies, with the following form:

$$\begin{aligned}
&\text{Post-release SR}_{i,t} \\
&= \beta_1 + \sum_{k=2}^{3}\left[\beta_k \times \mathbb{1}\{\text{Contests experienced}_{i,t} = k\}\right] + \gamma_1 \times \mathbb{1}\{\text{New variables available}_t\} \\
&\quad + \sum_{k=2}^{3}\left[\gamma_k \times \mathbb{1}\{\text{Contests experienced}_{i,t} = k\} \times \mathbb{1}\{\text{New variables available}_t\}\right] + \epsilon_{i,t} \qquad (19)
\end{aligned}$$

Regression results are displayed in Table 14. Each column focusses on a (non-overlapping) period of 5 consecutive trading days following the addition of new observations of macroeconomic signals to the trading platform: the SR in the first column is computed from the 5 trading days immediately following macroeconomic releases, then the second column corresponds to the

[16] In this analysis, all trading strategy daily returns are in excess of matched benchmark index daily returns, which is equivalent to computing the Sharpe Ratio of daily returns with our equally-weighted index portfolio as the benchmark. Therefore, the riskless rate is ignored for simplicity, which is consistent with the Quantiacs trading platform's stated internal Sharpe Ratio calculation methodology. In unreported results, we repeated this analysis incorporating a riskless rate (proxied by 3mo Treasury bill rates from FRED) in addition to the benchmark index return, with no change to the qualitative pattern of regression coefficients.

next 5 trading days, and so on. The most important coefficients are those in column (1) of Table 14, as they correspond to the period immediately following macroeconomic releases to the platform. The estimated coefficients $\beta_1, \beta_2, \beta_3$ in the top three rows essentially convey the result of a placebo test, as they correspond to contests where the macroeconomic variables were *not* available on the platform: their negative values indicate poor performance in that sub-period, in agreement with the fact that investors do not make use of the macroeconomic variables (that they do not have access to yet). The estimated coefficients $\gamma_1, \gamma_2, \gamma_3$ in the bottom three rows are driven by contests where the additional predictive signals were indeed available: this time, they are positive and increasing as the investors participate in one additional trading contest (γ_2 in the fifth row), matching the general pattern that we have encountered in our main empirical results (Table 8).

Moving from columns (1) through to (4) of Table 14 increases the delay between the release of macroeconomic observations to the platform and the period upon which the Sharpe Ratios are estimated. In the first three rows (i.e. the placebo test), the estimates for the coefficients $\beta_1, \beta_2, \beta_3$ are either mixed in sign or (for β_3) increasing with the delay. The next three rows correspond to contests where the additional (macroeconomic) predictive signals were available to the investors: in contrast to the placebo rows above them, the estimates for the coefficients $\gamma_1, \gamma_2, \gamma_3$ in column (1) are higher than those in column (4), indicating that investors outperform in the period immediately following the additional data releases. This suggests that investors with access to additional predictive variables do indeed make use of them, to varying degrees.

Furthermore, the increase in performance from investors participating in a second contest compared to the first (i.e. $\gamma_2 - \gamma_1$, visible by comparing the fourth and fifth rows) is greater in column (1) than it is in column (4), indicating that investors learn to perform better with experience during the period immediately following the additional data releases, in particular. This observation supports the notion that, as investors learn with experience, they learn to make better use of the additional (low-frequency macroeconomic) variables rather being able to immediately make full use of them. Therefore, our empirical results are consistent with our hypothesized mechanism that, with experience, an investor's subjective model uncertainty decreases, leading her to make use of more of the predictive signals available to her in order to maximize her objective.

7 Conclusion

Using a novel panel of systematic investor outcomes from an institutional setting that controls investors' preferences, horizons and (crucially) the information they can use to make trading decisions, we studied how investors learn with experience, based upon historical data, and in response to additional predictive signals. As well as confirming existing results on investor learning (experience, attrition & overconfidence effects) we characterised the relative importance of in- and out-of-sample learning channels: the former better explains our observed learning effects.

Importantly, we considered the impact of additional predictive variables in the common parts of investors' information sets (i.e. a widening of this common dataset). We found that individual-level dispersions of investors' outcomes narrowed with the availability of these additional predictive variables, which has implications for models of investor learning. The levels of investors' performance outcomes improved with the availability of these additional predictive variables, for experienced investors in particular. This suggests that experienced investors were able to make use of the wider data to better attain their objectives.

To explain why inexperienced investors did not exhibit a statistically significant improvement in performance, we theorised that a fear of model uncertainty may lead them to choose to ignore some predictive variables. We formalised this explanation by developing a model of portfolio choice with robust learning, and suggested a link between an investor's fear of model uncertainty and her level of experience. Much like theories of "learning by doing" (Arrow, 1962; Seru et al., 2010), this model can explain why investors who make use of historical data improve their performance outcomes with experience; our model goes further by also explaining why performance outcomes steepen with experience when additional predictive signals are available to investors. We provided further empirical evidence in support of this mechanism.

Our findings contribute to the literatures on investor learning and the empirics of information in asset pricing – both of which are very relevant in a Big Data world – as well as towards a better understanding of the ever-growing category of systematic investors. We hope that our model of robust learning will be useful in other economic settings.

References

Abis, S., 2017. Man vs. machine: Quantitative and discretionary equity management.

Arrow, K. J., 1962. The economic implications of learning by doing. Review of Economic Studies 29, 155–173.

Arthur, W. B., Holland, J. H., LeBaron, B., Palmer, R., Tayler, P., 1996. Asset pricing under endogenous expectations in an artificial stock market. The economy as an evolving complex system II 27.

Back, K., 2017. Asset pricing and portfolio choice theory. Oxford University Press, second ed.

Balasubramanian, A., Yang, D. Y., 2020. Trading in high dimensions. Available at SSRN 3583217.

Barber, B. M., Lee, Y.-T., Liu, Y.-J., Odean, T., Zhang, K., 2020. Learning, fast or slow. Review of Asset Pricing Studies 10, 61–93.

Barber, B. M., Odean, T., 2000. Trading is hazardous to your wealth: The common stock investment performance of individual investors. Journal of Finance 55, 773–806.

Barberis, N., 2000. Investing for the long run when returns are predictable. Journal of Finance 55, 225–264.

Barberis, N., Shleifer, A., Vishny, R., 1998. A model of investor sentiment. Journal of Financial Economics 49, 307–343.

Best, M. J., Grauer, R. R., 1991. On the sensitivity of mean-variance-efficient portfolios to changes in asset means: Some analytical and computational results. Review of Financial Studies 4, 315–342.

Bhardwaj, G., Gorton, G. B., Rouwenhorst, K. G., 2014. Fooling some of the people all of the time: The inefficient performance and persistence of Commodity Trading Advisors. Review of Financial Studies 27, 3099–3132.

Biais, B., Bossaerts, P., Spatt, C., 2010. Equilibrium asset pricing and portfolio choice under asymmetric information. Review of Financial Studies 23, 1503–1543.

Branch, W. A., Evans, G. W., 2010. Asset return dynamics and learning. Review of Financial Studies 23, 1651–1680.

Brennan, M. J., 1998. The role of learning in dynamic portfolio decisions. Review of Finance 1, 295–306.

Camerer, C., Ho, T. H., 1999. Experience-weighted attraction learning in normal form games. Econometrica 67, 827–874.

Chen, Y., Kelly, B., Wu, W., 2020. Sophisticated investors and market efficiency: Evidence from a natural experiment. Journal of Financial Economics 138, 316 – 341.

Cookson, J. A., Niessner, M., 2020. Why don't we agree? Evidence from a social network of investors. Journal of Finance 75, 173–228.

Cornuéjols, G., Peña, J. F., Tütüncü, R., 2018. Optimization Methods in Finance. Cambridge University Press, second ed.

Croce, M. M., Lettau, M., Ludvigson, S. C., 2015. Investor information, long-run risk, and the term structure of equity. Review of Financial Studies 28, 706–742.

Da, Z., Huang, X., Jin, L. J., 2020. Extrapolative beliefs in the cross-section: What can we learn from the crowds? Journal of Financial Economics .

Daniel, K. D., Hirshleifer, D., Subrahmanyam, A., 2001. Overconfidence, arbitrage, and equilibrium asset pricing. Journal of Finance 56, 921–965.

Dugast, J., Foucault, T., 2020. Equilibrium data mining and data abundance.

Erev, I., Roth, A. E., 1998. Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria. American Economic Review pp. 848–881.

Fabozzi, F. J., Focardi, S. M., Jonas, C. L., 2008. On the challenges in quantitative equity management. Quantitative Finance 8, 649–665.

Fang, L., Peress, J., 2009. Media coverage and the cross-section of stock returns. Journal of Finance 64, 2023–2052.

Farboodi, M., Veldkamp, L., 2020. Long-run growth of financial data technology. American Economic Review 110, 2485–2523.

Fleiss, S., Kominers, S. D., Ughetta, C. B., 2017. Quantopian: A new model for active management, HBS Case Study 9-218-046.

Fung, W., Hsieh, D. A., 1997. The information content of performance track records: Investment style and survivorship bias in the historical returns of Commodity Trading Advisors. Journal of Portfolio Management 24, 30–41.

Gabaix, X., 2014. A sparsity-based model of bounded rationality. Quarterly Journal of Economics 129, 1661–1710.

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

Gervais, S., Odean, T., 2001. Learning to be overconfident. Review of Financial Studies 14, 1–27.

Giglio, S., Maggiori, M., Stroebel, J., Utkus, S., 2021. Five facts about beliefs and portfolios. American Economic Review (forthcoming).

Gomes, F., Haliassos, M., Ramadorai, T., 2020. Household finance. Journal of Economic Literature (forthcoming).

Greene, W. H., 1981. Sample selection bias as a specification error: A comment. Econometrica pp. 795–798.

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

Hans, C., 2009. Bayesian lasso regression. Biometrika 96, 835–845.

Hansen, L. P., Sargent, T. J., 2008. Robustness. Princeton University Press.

Hastie, T., Tibshirani, R., Friedman, J., 2009. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer Science & Business Media, second ed.

Heckman, J. J., 1976. The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models. In: *Annals of Economic and Social measurement*, NBER, vol. 5, pp. 475–492.

Hong, H., Stein, J. C., Yu, J., 2007. Simple forecasts and paradigm shifts. Journal of Finance 62, 1207–1242.

Kacperczyk, M., Seru, A., 2007. Fund manager use of public information: New evidence on managerial skills. Journal of Finance 62, 485–528.

Khandani, A. E., Lo, A. W., 2011. What happened to the quants in August 2007? Evidence from factors and transactions data. Journal of Financial Markets 14, 1–46.

Lim, T., 2001. Rationality and analysts' forecast bias. Journal of Finance 56, 369–385.

Linnainmaa, J. T., 2011. Why do (some) households trade so much? Review of Financial Studies 24, 1630–1666.

Maenhout, P. J., 2004. Robust portfolio rules and asset pricing. Review of Financial Studies 17, 951–983.

Markowitz, H., 1952. Portfolio selection. Journal of Finance 7, 77–91.

Martin, I., Nagel, S., 2019. Market efficiency in the age of Big Data.

Merton, R. C., 1980. On estimating the expected return on the market: An exploratory investigation. Journal of Financial Economics 8, 323–361.

Molavi, P., Tahbaz-Salehi, A., Vedolin, A., 2021. Model complexity, expectations, and asset prices.

Nicolosi, G., Peng, L., Zhu, N., 2009. Do individual investors learn from their trading experience? Journal of Financial Markets 12, 317–336.

Park, T., Casella, G., 2008. The Bayesian Lasso. Journal of the American Statistical Association 103, 681–686.

Pástor, L., Veronesi, P., 2003. Stock valuation and learning about profitability. Journal of Finance 58, 1749–1789.

Pástor, L., Veronesi, P., 2009. Learning in financial markets. Annual Review of Financial Economics 1, 361–381.

Sargent, T. J., 1993. Bounded Rationality in Macroeconomics. The Arne Ryde Memorial Lectures, Oxford University Press.

Schwartzstein, J., 2014. Selective attention and learning. Journal of the European Economic Association 12, 1423–1452.

Seru, A., Shumway, T., Stoffman, N., 2010. Learning by trading. Review of Financial Studies 23, 705–739.

Tang, H., 2019. Peer-to-peer lenders versus banks: substitutes or complements? Review of Financial Studies 32, 1900–1938.

Tibshirani, R., 1996. Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society: Series B (Methodological) 58, 267–288.

Timmermann, A. G., 1993. How learning in financial markets generates excess volatility and predictability in stock prices. Quarterly Journal of Economics 108, 1135–1145.

Wiecki, T., Campbell, A., Lent, J., Stauth, J., 2016. All that glitters is not gold: Comparing backtest and out-of-sample performance on a large cohort of trading algorithms. Journal of Investing 25, 69–80.

Xu, H., Caramanis, C., Mannor, S., 2010. Robust regression and Lasso. IEEE Transactions on Information Theory 56, 3561–3574.

Zheng, Y., 2017. Quants in Utopia? Quantopian and its crowd-wisdom hedge-fund model, HKU Case Study HK1109.

Table 1: High-level descriptors of the 12 contests in our data sample.

Contest index	Live trading period	Number of contestants	Mean entries per contestant
1	2014-12-01 – 2015-01-31	13	1.31
2	2015-07-01 – 2015-09-30	16	3.62
3	2015-10-01 – 2015-12-31	23	1.17
4	2016-01-01 – 2016-03-31	30	3.53
5	2016-04-01 – 2016-06-30	38	4.05
6	2016-08-01 – 2016-10-31	87	3.02
7	2017-01-01 – 2017-03-31	125	2.54
8	2017-04-15 – 2017-07-31	92	3.50
9	2017-10-01 – 2018-01-31	163	2.31
10	2018-02-01 – 2018-05-31	63	3.73
11	2018-07-01 – 2018-10-31	95	2.88
12	2019-01-01 – 2019-04-30	129	3.13
	Total:	874	2.92

Table 2: Participation attributes of individual contestants i at contests t, before and after the additional predictive variables were made available to contestants by the trading platform.

	Before (N=502)		**After** (N=289)			
	Mean	**Std. Dev.**	**Mean**	**Std. Dev.**	**Diff. in Means**	**p value**
Contests experienced$_{i,t}$	1.0796	0.3183	1.1135	0.4344	0.0340	0.2080
Percentile(Score$^{Best}_{i,t-1}$)	0.6445	0.2570	0.7175	0.2595	0.0730	0.3078

Note: the reported p-values are from two-sided t-tests.

Table 3: Participation attributes aggregated at the contest level (t), before and after the additional predictive variables were made available to contestants by the trading platform.

	Before (N=6)		**After** (N=4)			
	Mean	**Std. Dev.**	**Mean**	**Std. Dev.**	**Diff. in Means**	**p value**
Mean$_t$(Contests experienced$_{i,t}$)	1.2281	0.0394	1.3823	0.1531	0.1542	0.1357
Fraction of first-time contestants at t	0.8374	0.0482	0.8282	0.0491	-0.0092	0.7794
Fraction of last-time contestants at t	0.7774	0.0731	0.8366	0.0603	0.0593	0.2024

Note: the reported p-values are from two-sided t-tests. For mechanical reasons, the first contest is excluded (because its fraction of first-time contestants would be 1.0) and the last contest is excluded (because its fraction of last-time contestants would be 1.0).

Table 4: OLS & panel regressions of in-sample ("backtest") & out-of-sample ("live") performance outcomes against experience.

	Dependent variable:			
	Backtest $SR_{i,t}^{Best}$		Live $SR_{i,t}^{Best}$	
	OLS	*panel linear*	*OLS*	*panel linear*
	(1)	(2)	(3)	(4)
Contests experienced$_{i,t}$	1.161***	1.338***	0.445**	1.261***
	(0.055)	(0.505)	(0.178)	(0.456)
Intercept	✓		✓	
Contest FEs		✓		✓
Contestant FEs		✓		✓
Observations	874	874	874	874
R^2	0.156	0.024	0.035	0.040

Note: standard errors (in parentheses) are robust to heteroskedasticity and are double-clustered by contest and contestant. *p<0.1; **p<0.05; ***p<0.01

Table 5: OLS & panel regressions of out-of-sample performance against experience, backtest performance, and prior contest rank.

	Dependent variable:								
	Live $SR^{Best}_{i,t}$								
	OLS			*panel linear*					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Contests experienced$_{i,t}$	0.445**	−0.046	−0.044	0.504***	−0.258	−0.280	1.261***	0.912*	0.894
	(0.178)	(0.177)	(0.192)	(0.165)	(0.219)	(0.256)	(0.456)	(0.485)	(0.582)
Backtest $SR^{Best}_{i,t}$		0.048	0.050		0.034	0.008		0.136**	0.130
		(0.041)	(0.092)		(0.077)	(0.136)		(0.064)	(0.104)
Log Percentile($Score^{Best}_{i,t-1}$)		0.817**	0.818**		0.888*	0.876		−1.963**	−1.970**
		(0.381)	(0.384)		(0.526)	(0.554)		(0.838)	(0.890)
Contests experienced$_{i,t}$ × Backtest $SR^{Best}_{i,t}$			−0.001			0.008			0.003
			(0.022)			(0.032)			(0.037)
Intercept	✓	✓	✓						
Contest FEs				✓	✓	✓	✓	✓	✓
Contestant FEs							✓	✓	✓
Observations	874	124	124	874	124	124	874	124	124
R^2	0.035	0.045	0.045	0.044	0.049	0.049	0.040	0.088	0.088

Note: standard errors (in parentheses) are robust to heteroskedasticity and are double-clustered by contest and contestant. *p<0.1; **p<0.05; ***p<0.01

Table 6: Poisson GLM regressions of contestants' number of entries in a contest against experience and performance measures.

	Dependent variable:			
	$\text{Entries}_{i,t}$			
	(1)	(2)	(3)	(4)
$\text{Contests experienced}_{i,t}$	0.390^{***}	0.344^{***}	0.182^{***}	0.196^{***}
	(0.045)	(0.044)	(0.069)	(0.067)
$\text{Live SR}_{i,t}^{Best}$		0.166^{***}		0.086^{***}
		(0.029)		(0.032)
$\text{Log Percentile}(\text{Score}_{i,t-1}^{Best})$			0.777^{**}	0.677^{**}
			(0.344)	(0.335)
Contestants_t				-0.0005
				(0.002)
Intercept	✓	✓	✓	✓
Observations	874	874	124	124

Note: standard errors (in parentheses) are robust to heteroskedasticity.
*p<0.1; **p<0.05; ***p<0.01

Table 7: Heckit (i.e. Type II Tobit) two-stage selection models for implementing selection bias corrections to the regression of out-of-sample performance against experience. The first stage models the probability of participation, while the second stage models the outcome of interest.

Stage		*All contestants*		*Repeated contestants*	
		Live $SR_{i,t}^{Best}$	Live $SR_{i,t}^{Best}$	Live $SR_{i,t}^{Best}$	Live $SR_{i,t}^{Best}$
1. *Selection*	(Intercept)	1.95^{***}	-2.58^{***}	1.69^{***}	-0.35
		(0.18)	(0.53)	(0.44)	(0.79)
	Contests experienced$_{i,t}$	-0.88^{***}	0.33^{***}	-0.39^{***}	-0.25^{***}
		(0.05)	(0.06)	(0.07)	(0.08)
	Quantopian search index$_t$	-0.01^{**}	-0.01^{**}	0.00	0.00
		(0.00)	(0.00)	(0.01)	(0.01)
	Ratio of entries to contest mean$_{i,t-1}$	0.23^{***}	0.34^{***}	0.14^{***}	0.13^{**}
		(0.04)	(0.05)	(0.05)	(0.06)
	Log Percentile(Score$_{i,t-1}^{Best}$)		0.28^{**}		0.37^{**}
			(0.11)		(0.16)
2. *Outcome*	(Intercept)	-0.42^{***}	0.77	1.50^{***}	-0.05
		(0.13)	(2.94)	(0.39)	(3.07)
	Contests experienced$_{i,t}$	0.84^{***}	-0.20	0.20	0.17
		(0.17)	(0.21)	(0.23)	(0.27)
	Log Percentile(Score$_{i,t-1}^{Best}$)		0.50		0.41
			(0.54)		(0.69)
	Inverse Mills Ratio	-0.83^{***}	-0.81	-1.64	-1.91
		(0.32)	(0.56)	(1.32)	(1.86)
	σ	2.13	2.77	2.84	2.97
	ρ	-0.39	-0.29	-0.58	-0.64
	R^2	0.04	0.05	0.01	0.04
	Adj. R^2	0.04	0.03	-0.00	0.02
	Num. obs.	1482	745	233	174
	Censored	621	621	50	50
	Observed	861	124	183	124

Note: parentheses denote standard errors. $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Table 8: OLS regressions of out-of-sample performance against additional data availability & experience. The dependent variable in columns (1)-(3) does not use a benchmark, so there is no control for time periods; the dependent variable in columns (4)-(6) uses the excess out-of-sample Sharpe Ratio over the Sharpe Ratio of the benchmark index during that out-of-sample period, which effectively controls for time periods/contests. In all regressions, the only contestants under consideration are those that have competed entirely before or after the new predictive variables were added to the platform.

	Dependent variable:					
	Live $SR_{i,t}^{Best}$			Live $SR_{i,t}^{Best}$ − Benchmark SR_t		
	(1)	(2)	(3)	(4)	(5)	(6)
(Intercept)	−0.034	−0.363	−0.329	-1.459^{***}	-1.382^{***}	-1.346^{***}
	(0.334)	(0.250)	(0.248)	(0.351)	(0.422)	(0.428)
$\mathbb{1}\{\text{Contests experienced}_{i,t} = 2\}$	0.925^{**}	0.942^{***}	0.505	1.211^{***}	1.207^{***}	0.863^{**}
	(0.364)	(0.348)	(0.521)	(0.274)	(0.263)	(0.340)
$\mathbb{1}\{\text{Contests experienced}_{i,t} = 3\}$	1.500	1.356	−0.231	2.712^{**}	2.746^{**}	0.312
	(1.660)	(1.712)	(0.248)	(1.152)	(1.258)	(0.428)
$\mathbb{1}\{\text{Contests experienced}_{i,t} = 4\}$	3.074^{***}	3.056^{***}	2.749^{***}	3.880^{***}	3.884^{***}	2.360^{***}
	(0.609)	(0.564)	(0.248)	(0.562)	(0.584)	(0.428)
$\mathbb{1}\{\text{New variables available}_t\}$		0.520	0.466		−0.122	−0.178
		(0.524)	(0.536)		(0.645)	(0.659)
$\mathbb{1}\{\text{Contests experienced}_{i,t} = 2\} \times \mathbb{1}\{\text{New variables available}_t\}$			0.725			0.571
			(0.670)			(0.491)
$\mathbb{1}\{\text{Contests experienced}_{i,t} = 3\} \times \mathbb{1}\{\text{New variables available}_t\}$			1.763			2.695^{*}
			(1.908)			(1.446)
$\mathbb{1}\{\text{Contests experienced}_{i,t} = 4\} \times \mathbb{1}\{\text{New variables available}_t\}$			0.464			2.288^{***}
			(0.828)			(0.460)
Observations	790	790	790	790	790	790
R^2	0.029	0.045	0.048	0.052	0.053	0.057

Note: standard errors (in parentheses) are robust to heteroskedasticity and are double-clustered by contest and contestant.
*p<0.1; **p<0.05; ***p<0.01

Table 9: OLS regressions of out-of-sample performance against additional data availability, experience & in-sample performance. The dependent variable in columns (1)-(3) does not use a benchmark, so there is no control for time periods; the dependent variable in columns (4)-(6) uses the excess out-of-sample Sharpe Ratio over the Sharpe Ratio of the benchmark index during that out-of-sample period, which effectively controls for time periods/contests. In all regressions, the only contestants under consideration are those that have competed entirely before or after the new predictive variables were added to the platform.

	Dependent variable:					
	Live $SR_{i,t}^{Best}$			Live $SR_{i,t}^{Best}$ − Benchmark SR_t		
	(1)	(2)	(3)	(4)	(5)	(6)
Contests experienced$_{i,t}$	0.561	0.555	0.489*	0.949***	0.951***	0.686***
	(0.386)	(0.384)	(0.253)	(0.230)	(0.234)	(0.199)
Backtest $SR_{i,t}^{Best}$	0.099**	0.092**	0.215	0.083*	0.085*	0.226
	(0.043)	(0.043)	(0.146)	(0.048)	(0.048)	(0.229)
$\mathbb{1}$\{New variables available$_t$\}		0.488	0.435		−0.144	−0.492
		(0.520)	(0.900)		(0.639)	(0.895)
Contests experienced$_{i,t}$ × $\mathbb{1}$\{New variables available$_t$\}			0.103			0.386
			(0.590)			(0.351)
Backtest $SR_{i,t}^{Best}$ × $\mathbb{1}$\{New variables available$_t$\}			−0.130			−0.155
			(0.152)			(0.234)
Intercept	✓	✓	✓	✓	✓	✓
Observations	791	791	791	791	791	791
R^2	0.041	0.055	0.056	0.058	0.059	0.060

Note: standard errors (in parentheses) are robust to heteroskedasticity and are double-clustered by contest and contestant.
*p<0.1; **p<0.05; ***p<0.01

Table 10: OLS regressions of out-of-sample daily return means against additional data availability & experience. The dependent variable in columns (1)-(3) does not use a benchmark, so there is no control for time periods; the dependent variable in columns (4)-(6) adjusts the daily returns using the benchmark portfolio.

	Dependent variable:					
	Mean of Live daily returns (%) for contestant i's best entry at contest t					
	Raw returns			*Excess returns over the benchmark*		
	(1)	(2)	(3)	(4)	(5)	(6)
(Intercept)	-0.150^{**}	-0.008	0.002	-0.171^{**}	-0.019^{***}	-0.010^{*}
	(0.068)	(0.007)	(0.006)	(0.068)	(0.007)	(0.006)
$\mathbb{1}\{\text{Contests experienced}_{i,t} = 2\}$	0.158^{**}	0.153^{**}	-0.002	0.167^{**}	0.162^{**}	0.012
	(0.069)	(0.069)	(0.012)	(0.069)	(0.069)	(0.016)
$\mathbb{1}\{\text{Contests experienced}_{i,t} = 3\}$	0.163^{**}	0.220^{**}	-0.017^{***}	0.178^{***}	0.239^{**}	-0.020^{***}
	(0.069)	(0.098)	(0.006)	(0.069)	(0.099)	(0.006)
$\mathbb{1}\{\text{New variables available}_t\}$		-0.225^{**}	-0.240^{**}		-0.239^{**}	-0.254^{**}
		(0.101)	(0.108)		(0.101)	(0.108)
$\mathbb{1}\{\text{Contests experienced}_{i,t} = 2\} \times \mathbb{1}\{\text{New variables available}_t\}$			0.254^{**}			0.244^{**}
			(0.109)			(0.109)
$\mathbb{1}\{\text{Contests experienced}_{i,t} = 3\} \times \mathbb{1}\{\text{New variables available}_t\}$			0.271^{**}			0.296^{***}
			(0.109)			(0.108)
Observations	674	674	674	674	674	674
R^2	0.001	0.005	0.005	0.001	0.006	0.006

Note: standard errors (in parentheses) are robust to heteroskedasticity and are clustered by contest. *p<0.1; **p<0.05; ***p<0.01

Table 11: OLS regressions of out-of-sample daily return standard deviations against additional data availability & experience. The dependent variable in columns (1)-(3) does not use a benchmark, so there is no control for time periods; the dependent variable in columns (4)-(6) adjusts the daily returns using the benchmark portfolio.

	Dependent variable:					
	SD of Live daily returns (%) for contestant *i*'s best entry at contest *t*					
	Raw returns			*Excess returns over the benchmark*		
	(1)	(2)	(3)	(4)	(5)	(6)
(Intercept)	0.820***	0.696***	0.687***	0.904***	0.773***	0.760***
	(0.143)	(0.064)	(0.066)	(0.143)	(0.062)	(0.063)
$\mathbb{1}\{\text{Contests experienced}_{i,t} = 2\}$	−0.431***	−0.426***	−0.301***	−0.417***	−0.413***	−0.219**
	(0.149)	(0.146)	(0.094)	(0.148)	(0.145)	(0.092)
$\mathbb{1}\{\text{Contests experienced}_{i,t} = 3\}$	−0.489***	−0.539***	−0.230***	−0.457***	−0.510**	0.027
	(0.154)	(0.206)	(0.066)	(0.153)	(0.208)	(0.063)
$\mathbb{1}\{\text{New variables available}_t\}$		0.197	0.210		0.207	0.227
		(0.219)	(0.233)		(0.218)	(0.232)
$\mathbb{1}\{\text{Contests experienced}_{i,t} = 2\} \times \mathbb{1}\{\text{New variables available}_t\}$			−0.204			−0.316
			(0.248)			(0.245)
$\mathbb{1}\{\text{Contests experienced}_{i,t} = 3\} \times \mathbb{1}\{\text{New variables available}_t\}$			−0.352			−0.610***
			(0.241)			(0.236)
Observations	674	674	674	674	674	674
R^2	0.001	0.002	0.002	0.001	0.002	0.002

Note: standard errors (in parentheses) are robust to heteroskedasticity and are clustered by contest. *p<0.1; **p<0.05; ***p<0.01

Table 12: OLS regressions of individual (contestant-level) dispersions (calculated as the range, i.e. best/max minus worst/min) of contestants' Backtest and Live Sharpe Ratios, against additional data availability & experience.

	Dependent variable:							
	$\text{Range}_{i,t}(\text{Backtest SR}_{i,j,t})$				$\text{Range}_{i,t}(\text{Live SR}_{i,j,t})$			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\mathbb{1}\{\text{New variables available}_t\}$	0.193^{*}	0.114	0.134	0.114	-0.129	-0.216	-0.187	-0.257^{**}
	(0.107)	(0.087)	(0.097)	(0.091)	(0.167)	(0.163)	(0.118)	(0.114)
$\text{Contests experienced}_{i,t}$		0.666^{***}	0.357^{***}	-0.136^{*}		0.729^{***}	0.282^{**}	0.272^{**}
		(0.055)	(0.074)	(0.070)		(0.146)	(0.130)	(0.118)
$\text{Entries}_{i,t}$			0.123^{***}	0.055^{*}			0.178^{***}	0.179^{***}
			(0.042)	(0.029)			(0.023)	(0.023)
$\text{Backtest SR}_{i,t}^{Mean}$				0.863^{***}				
				(0.175)				
$\text{Live SR}_{i,t}^{Mean}$								0.157^{***}
								(0.037)
Intercept	✓	✓	✓	✓	✓	✓	✓	✓
Observations	874	874	874	874	874	874	874	874
R^2	0.002	0.088	0.209	0.550	0.001	0.115	0.396	0.414

Note: standard errors (in parentheses) are robust to heteroskedasticity and are double-clustered by contest and contestant.
$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Table 13: OLS regressions of the aggregate (i.e. contest-level) dispersions (standard deviations) of contestants' best Backtest and Live Sharpe Ratios, against additional data availability.

	Dependent variable:			
	$SD_t(\text{Backtest } SR^{Best}_{i,t})$		$SD_t(\text{Live } SR^{Best}_{i,t})$	
	(1)	(2)	(3)	(4)
$\mathbb{1}\{\text{New variables available}_t\}$	1.783***	1.075**	−0.259	0.263
	(0.448)	(0.502)	(0.325)	(0.374)
$\text{Mean}_t(\text{Contests experienced}_{i,t})$		4.511***		−3.416**
		(1.218)		(1.547)
$\text{Mean}_t(\text{Entries}_{i,t})$		−0.114		0.125
		(0.179)		(0.151)
Intercept	✓	✓	✓	✓
Observations	12	12	12	12
R^2	0.656	0.824	0.052	0.403

Note: standard errors (in parentheses) are robust to heteroskedasticity.
*p<0.1; **p<0.05; ***p<0.01

Table 14: OLS regressions of Sharpe Ratios (computed from daily returns in excess of benchmark returns) for trading week t following macroeconomic announcement dates (as recorded on the trading platform). The dependent variables are computed from trading strategy daily returns in excess of matched benchmark index daily returns.

	$SR_{i,t}$ computed on post-release periods			
	Trading Days 1-5	*Trading Days 6-10*	*Trading Days 11-15*	*Trading Days 16-20*
(Intercept)	-0.437^{***}	0.065	0.142^{***}	-0.403^{***}
	(0.024)	(0.059)	(0.030)	(0.016)
$\mathbb{1}\{\text{Contests experienced}_{i,t} = 2\}$	−0.036	0.054	0.044	0.001
	(0.054)	(0.104)	(0.100)	(0.044)
$\mathbb{1}\{\text{Contests experienced}_{i,t} = 3\}$	-0.131^{***}	−0.096	0.121^{***}	0.194^{***}
	(0.024)	(0.059)	(0.030)	(0.016)
$\mathbb{1}\{\text{New variables available}_t\}$	0.042	-0.136^{**}	−0.011	0.013
	(0.026)	(0.064)	(0.036)	(0.019)
$\mathbb{1}\{\text{Contests experienced}_{i,t} = 2\} \times \mathbb{1}\{\text{New variables available}_t\}$	0.778^{***}	0.670^{**}	0.691^{**}	0.484^{**}
	(0.183)	(0.278)	(0.281)	(0.194)
$\mathbb{1}\{\text{Contests experienced}_{i,t} = 3\} \times \mathbb{1}\{\text{New variables available}_t\}$	0.706^{***}	0.733^{***}	0.664^{***}	0.072
	(0.257)	(0.241)	(0.225)	(0.197)
Observations	674	674	674	674
R^2	0.166	0.159	0.165	0.068

Note: standard errors (in parentheses) are robust to heteroskedasticity and are double-clustered by contest and contestant.
*p<0.1; **p<0.05; ***p<0.01

Figure 1: Stylised illustration of how datasets may lengthen over time, or widen with the addition of one or more predictive signals.

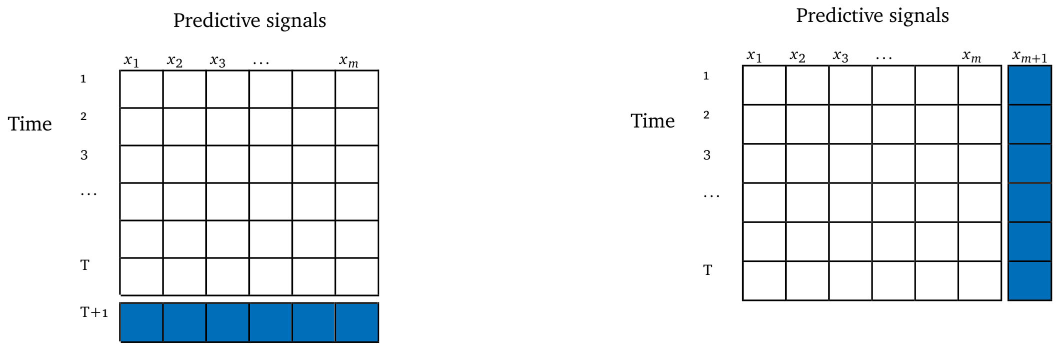


Figure 2: Stylised illustration of how the Backtest period for each contest expands over time, and the Live period is a predefined future period.

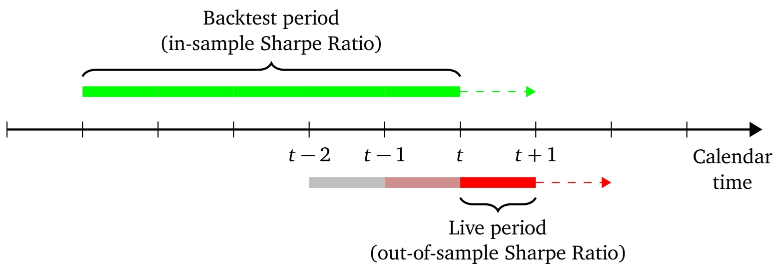


Figure 3: Box plots of the distributions of Live (out-of-sample) Sharpe Ratios, conditioning by contest index only. The second plot shows the subset of contestants' best entries per contest.

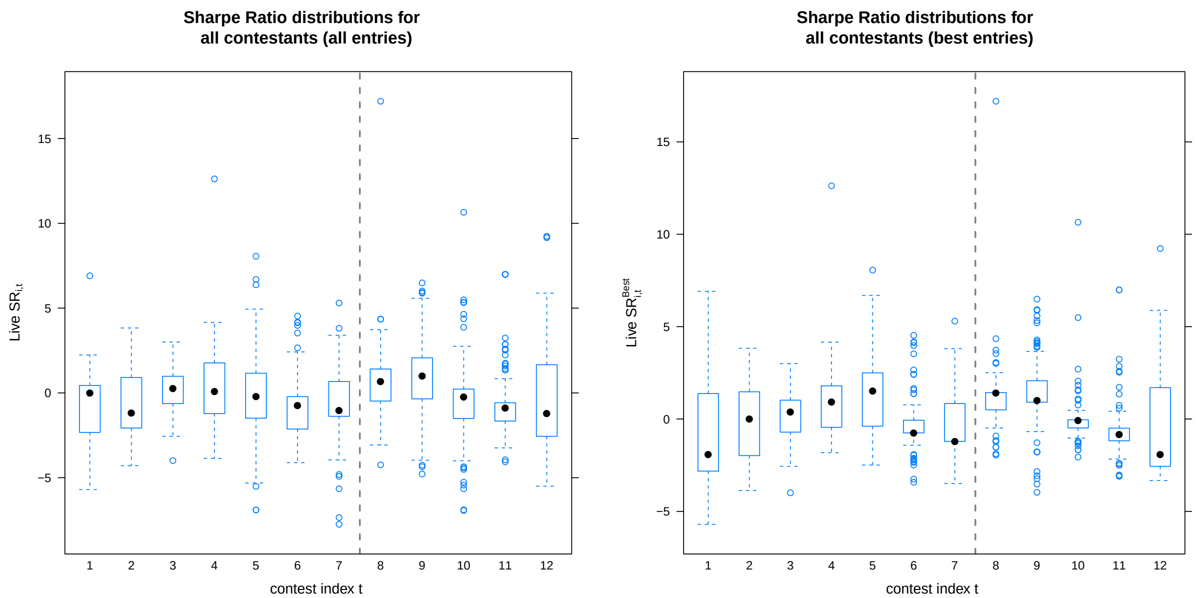


Figure 4: Proportions of contestants for whom this is the first contest, and those for whom this is the last contest (with breakdowns of the latter).

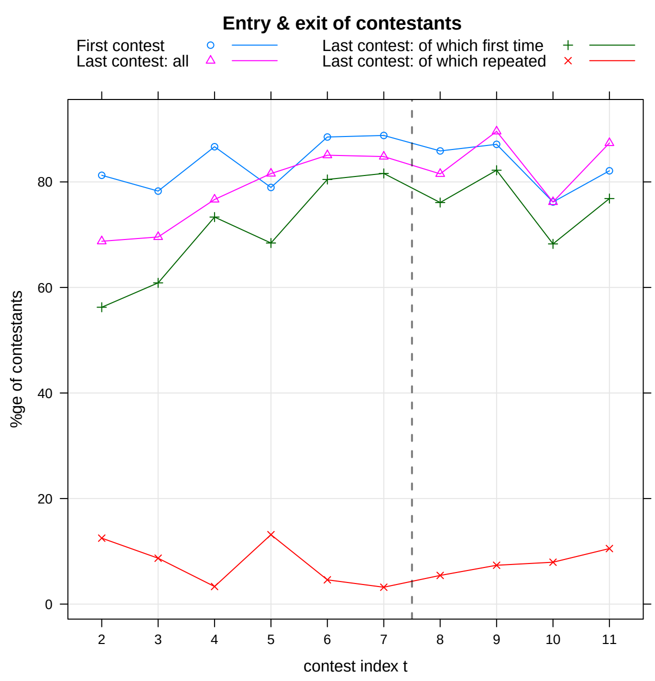
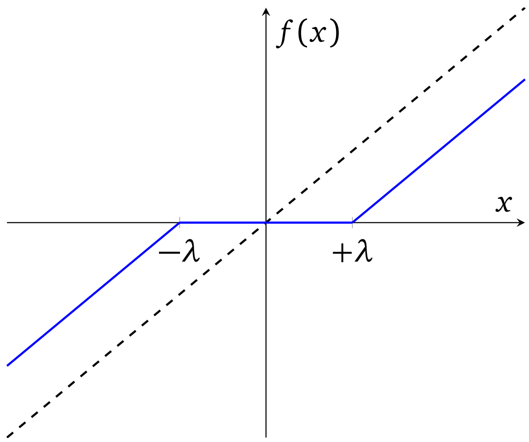


Figure 5: Illustration of soft-thresholding an input. The dashed line is an example input $f(x) = x$. The blue line $f(x) = \text{sign}(x) \max\{|x| - \lambda, 0\})$ is the result of soft-thresholding that input at threshold λ.



Appendix

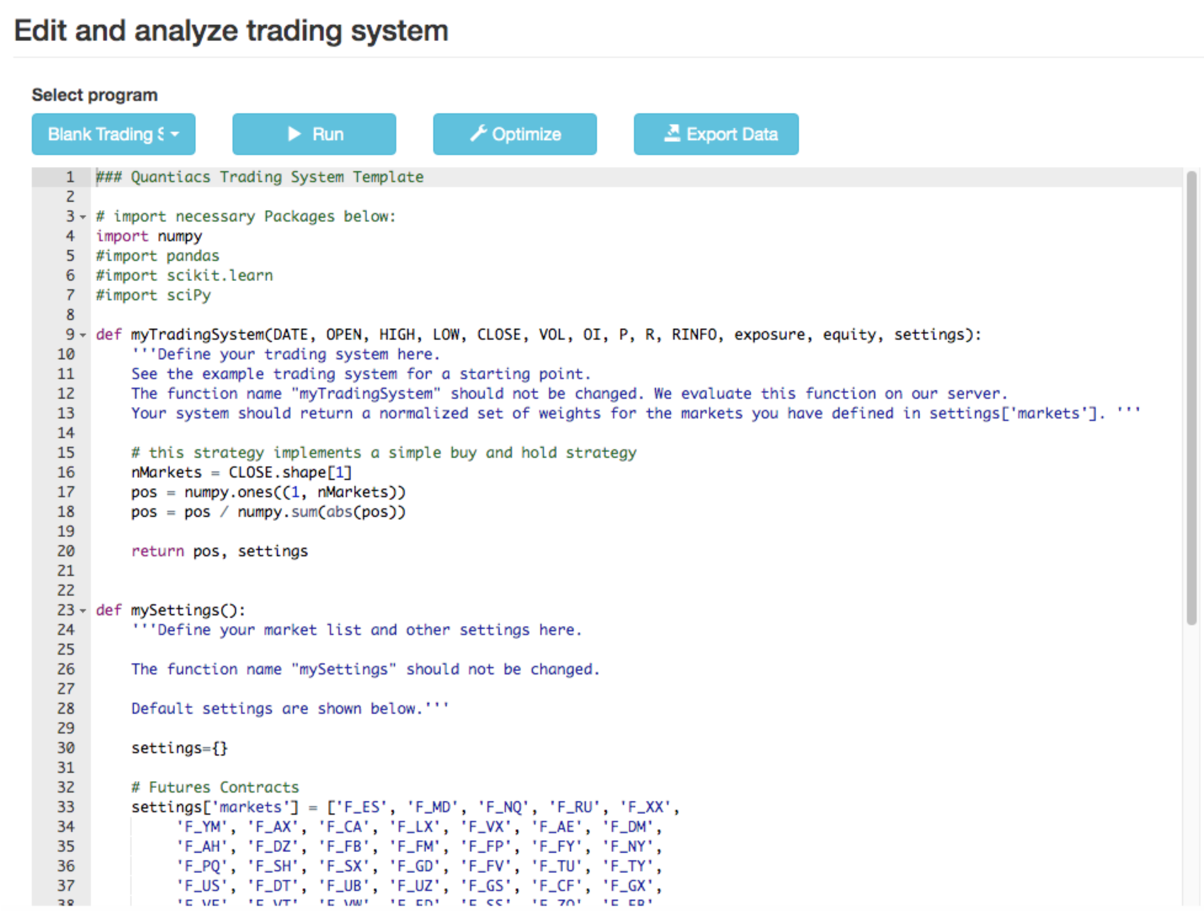
A Entering a trading contest

To make the institutional setting of Section 3 more concrete, we review the steps necessary to participate in a contest from the point of view of a systematic investor. Figure 6 shows an example of how to write (in Python) a systematic trading strategy that simply takes long positions in all available futures contracts. More sophisticated strategies are, of course, encouraged.

After coding up such a systematic trading strategy, the systematic investor can then use the Quantiacs platform to run a backtest of the strategy using historical market data. Figure 7 shows the output of such a backtest, for the long-only strategy of our above example. Various performance metrics are provided, but all are calculated on historical data, and we refer to this pre-contest period as the in-sample period.

The systematic investor may then decide to officially submit this trading strategy to a contest: if so, we call this the contestant's entry. Once a contest has begun (i.e. during the live/out-of-sample period), entries can no longer be modified, and the contestant has effectively committed to following that systematic trading strategy for the duration of the live/out-of-sample period. After the contest ends, the out-of-sample Sharpe Ratios are made available, the contestants are ranked, and results are displayed in a leaderboard, as in Figure 8.

Figure 6: Example of writing code (in Python) while logged into the trading platform in order to define a (simple long-only) systematic trading strategy, ahead of possible entry into a trading contest.




```
### Quantiacs Trading System Template

# import necessary Packages below:
import numpy
#import pandas
#import scikit.learn
#import sciPy

def myTradingSystem(DATE, OPEN, HIGH, LOW, CLOSE, VOL, OI, P, R, RINFO, exposure, equity, settings):
    '''Define your trading system here.
    See the example trading system for a starting point.
    The function name "myTradingSystem" should not be changed. We evaluate this function on our server.
    Your system should return a normalized set of weights for the markets you have defined in settings['markets']. '''

    # this strategy implements a simple buy and hold strategy
    nMarkets = CLOSE.shape[1]
    pos = numpy.ones((1, nMarkets))
    pos = pos / numpy.sum(abs(pos))

    return pos, settings


def mySettings():
    '''Define your market list and other settings here.

    The function name "mySettings" should not be changed.

    Default settings are shown below.'''

    settings={}

    # Futures Contracts
    settings['markets'] = ['F_ES', 'F_MD', 'F_NQ', 'F_RU', 'F_XX',
         'F_YM', 'F_AX', 'F_CA', 'F_LX', 'F_VX', 'F_AE', 'F_DM',
         'F_AH', 'F_DZ', 'F_FB', 'F_FM', 'F_FP', 'F_FY', 'F_NY',
         'F_PQ', 'F_SH', 'F_SX', 'F_GD', 'F_FV', 'F_TU', 'F_TY',
         'F_US', 'F_DT', 'F_UB', 'F_UZ', 'F_GS', 'F_CF', 'F_GX',
```

Figure 7: Example of backtest results for a systematic trading strategy, before an entry is submitted. At this point, a trading contest has not begun.

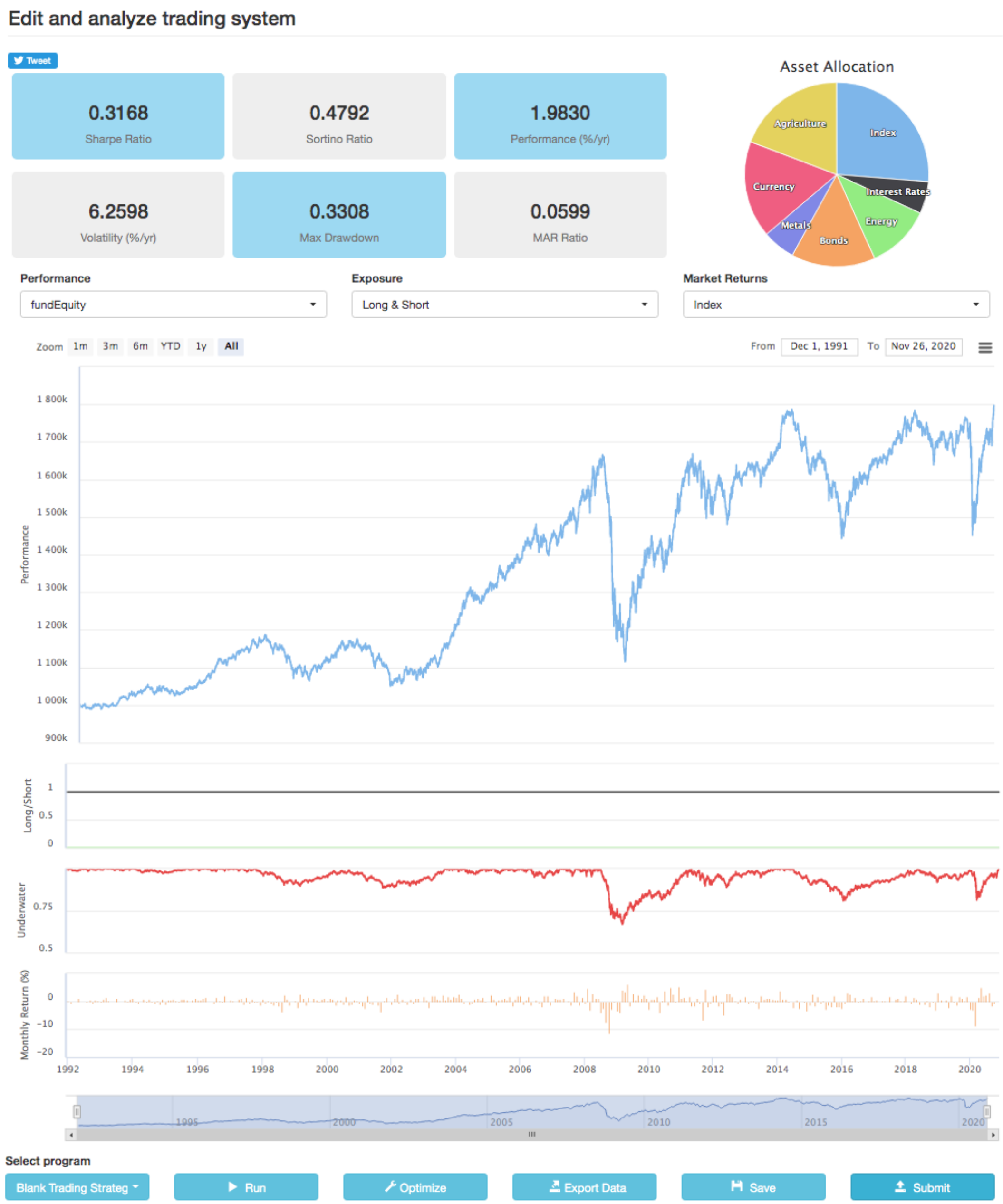
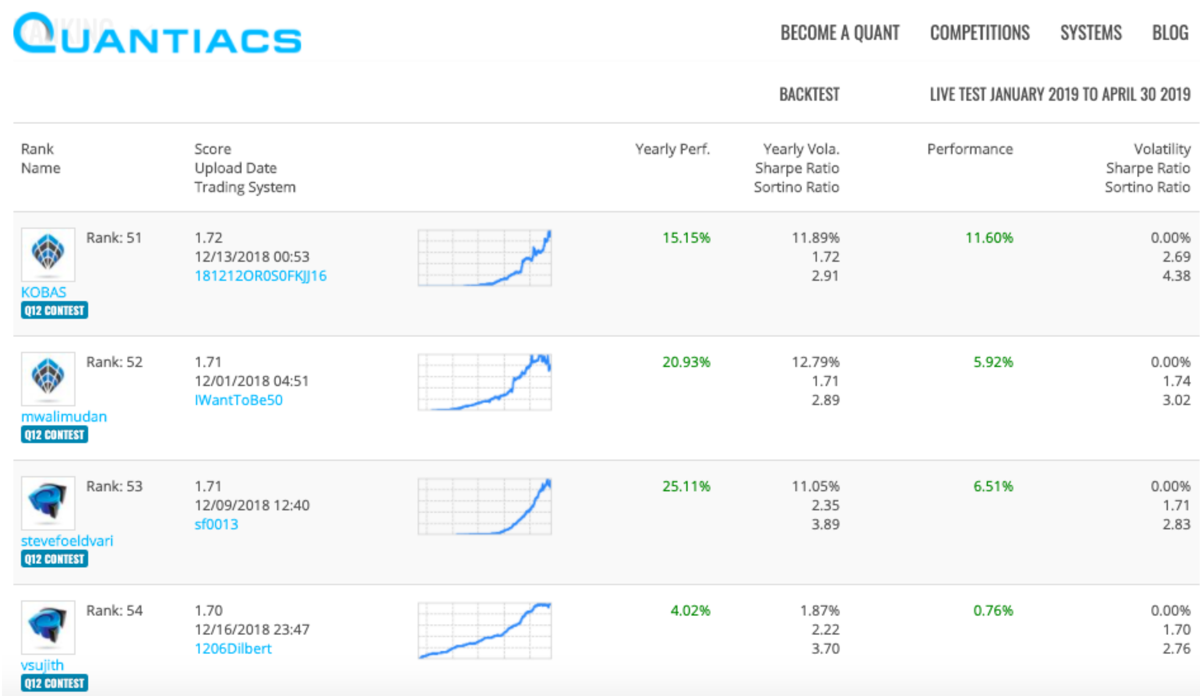


Figure 8: Excerpt from a contest leaderboard, showing both in-sample and out-of-sample Sharpe Ratios, after a trading contest has ended. The "live test" performance metrics were calculated once the contest period (1 January to 30 April 2019, in this case) had ended.


QUANTIACS

BECOME A QUANT | COMPETITIONS | SYSTEMS | BLOG

Rank Name	Score Upload Date Trading System	BACKTEST Yearly Perf.	Yearly Vola. Sharpe Ratio Sortino Ratio	LIVE TEST JANUARY 2019 TO APRIL 30 2019 Performance	Volatility Sharpe Ratio Sortino Ratio
Rank: 51 KOBAS Q12 CONTEST	1.72 12/13/2018 00:53 181212OR0S0FKJJ16	15.15%	11.89% 1.72 2.91	11.60%	0.00% 2.69 4.38
Rank: 52 mwalimudan Q12 CONTEST	1.71 12/01/2018 04:51 IWantToBe50	20.93%	12.79% 1.71 2.89	5.92%	0.00% 1.74 3.02
Rank: 53 stevefoeldvari Q12 CONTEST	1.71 12/09/2018 12:40 sf0013	25.11%	11.05% 2.35 3.89	6.51%	0.00% 1.71 2.83
Rank: 54 vsujith Q12 CONTEST	1.70 12/16/2018 23:47 1206Dilbert	4.02%	1.87% 2.22 3.70	0.76%	0.00% 1.70 2.76

B Macroeconomic variables added between contests

On 22 March 2017, Quantiacs announced in a blog post that 54 macroeconomic data series would be made available to all contests from the 8th contest onwards, for use in both backtesting and live trading. The observations available on the trading platform are almost all at a monthly frequency.[17] The variables are listed in Table 15.

Table 15: Macroeconomic variables added before the 8th contest.

	Macroeconomic variable	Quantiacs identifier
1	ADP Employment Change	USA_ADP
2	Average Hourly Earnings	USA_EARN
3	Average Weekly Hours	USA_HRS
4	Balance of Trade	USA_BOT
5	Business Confidence	USA_BC
6	Business Inventories	USA_BI
7	Capacity Utilization	USA_CU
8	Capital Flows	USA_CF
9	Challenger Job Cuts	USA_CHJC
10	Chicago Fed National Activity Index	USA_CFNAI
11	Chicago Pmi	USA_CP
12	Consumer Credit	USA_CCR
13	Consumer Price Index CPI	USA_CPI
14	Core Consumer Prices	USA_CCPI
15	Core Inflation Rate	USA_CINF
16	Dallas Fed Manufacturing Index	USA_DFMI
17	Durable Goods Orders	USA_DUR
18	Durable Goods Orders Ex Transportation	USA_DURET
19	Export Prices	USA_EXPX
20	Exports	USA_EXVOL
21	Factory Orders Ex Transportation	USA_FRET
22	Foreign Bond Investment	USA_FBI
23	Government Budget Value	USA_GBVL
24	Government Payrolls	USA_GPAY
25	Housing Index	USA_HI
26	Import Prices	USA_IMPX
27	Imports	USA_IMVOL
28	Industrial Production	USA_IP
29	Industrial Production Mom	USA_IPMOM
30	Inflation Rate	USA_CPIC

[17] Observations for three of the macroeconomic variables (USA_EXPX, USA_IMPX, USA_NFIB) were available only at a quarterly frequency for part of the backtest period, but were since made available at a monthly frequency (like the remaining variables) and the majority of their observations are at a monthly frequency.

	Macroeconomic variable	Quantiacs identifier
31	Inflation Rate Mom	USA_CPICM
32	Job Offers	USA_JBO
33	Labor Force Participation Rate	USA_LFPR
34	Leading Economic Index	USA_LEI
35	Manufacturing Payrolls	USA_MPAY
36	Manufacturing Production	USA_MP
37	Nahb Housing Market Index	USA_NAHB
38	Net Long Term Tic Flows	USA_NLTTF
39	NFIB Business Optimism Index	USA_NFIB
40	Non Farm Payrolls	USA_NFP
41	Non Manufacturing PMI	USA_NMPMI
42	Nonfarm Payrolls Private	USA_NPP
43	NY Empire State Manufacturing Index	USA_EMPST
44	Pending Home Sales	USA_PHS
45	Philadelphia Fed Manufacturing Index	USA_PFED
46	Producer Prices	USA_PP
47	Producer Prices Change	USA_PPIC
48	Retail Sales MoM	USA_RSM
49	Retail Sales YoY	USA_RSY
50	Retail Sales Ex Autos	USA_RSEA
51	Richmond Fed Manufacturing Index	USA_RFMI
52	Total Vehicle Sales	USA_TVS
53	Unemployment Rate	USA_UNR
54	Wholesale Inventories	USA_WINV

C Futures contracts

This appendix lists the futures contracts that participants' strategies are able to trade. For each future, the trading platform selected a single contract maturity that was available for trading at each point in time. Our procedure for downloading this historical data and constructing a benchmark futures portfolio was detailed in Section 4.4.

Table 16: Tradable futures on Quantiacs that we used to construct a benchmark index. (Note that we used the underlying Russell index for row 33).

	Quantiacs ticker	Name	Type	Matching Bloomberg base code
1	F_AD	Australian Dollar	Currency	AD Curncy
2	F_BO	Soybean Oil	Agriculture	BO Comdty
3	F_BP	British Pound	Currency	BP Curncy
4	F_C	Corn	Agriculture	C Comdty
5	F_CC	Cocoa	Agriculture	CC Comdty
6	F_CD	Canadian Dollar	Currency	CD Curncy
7	F_CL	WTI Crude Oil	Energy	CL Comdty
8	F_CT	Cotton	Agriculture	CT Comdty
9	F_DX	US Dollar Index	Currency	DX Curncy
10	F_EC	Euro FX	Currency	EC Curncy
11	F_ED	Eurodollars	Interest Rate	ED Comdty
12	F_ES	E-mini S&P 500 Index	Index	ES Index
13	F_FC	Feeder Cattle	Agriculture	FC Comdty
14	F_FV	5-year Treasury Note	Bond	FV Comdty
15	F_GC	Gold	Metal	GC Comdty
16	F_HG	Copper	Metal	HG Comdty
17	F_HO	Heating Oil	Energy	HO Comdty
18	F_JY	Japanese Yen	Currency	JY Curncy
19	F_KC	Coffee	Agriculture	KC Comdty
20	F_LB	Lumber	Agriculture	LB Comdty
21	F_LC	Live Cattle	Agriculture	LC Comdty
22	F_LN	Lean Hogs	Agriculture	LH Comdty
23	F_MD	E-mini S&P 400	Index	FA Index
24	F_MP	Mexican Peso	Currency	PE Curncy
25	F_NG	Natural Gas	Energy	NG Comdty
26	F_NQ	E-mini Nasdaq 100 Index	Index	NQ Index
27	F_NR	Rough Rice	Agriculture	RR Comdty
28	F_O	Oats	Agriculture	O Comdty
29	F_OJ	Orange Juice	Agriculture	JO Comdty
30	F_PA	Palladium	Metal	PA Comdty
31	F_PL	Platinum	Metal	PL Comdty
32	F_RB	Gasoline	Energy	XB Comdty

	Quantiacs ticker	Name	Type	Matching Bloomberg base code
33	F_RU	Russell 2000	Index	RTY Index
34	F_S	Soybeans	Agriculture	S Comdty
35	F_SB	Sugar	Agriculture	SB Comdty
36	F_SF	Swiss Franc	Currency	SF Curncy
37	F_SI	Silver	Metal	SI Comdty
38	F_SM	Soybean Meal	Agriculture	SM Comdty
39	F_TU	2-year Treasury Note	Bond	TU Comdty
40	F_TY	10-year Treasury Note	Bond	TY Comdty
41	F_US	30-year Treasury Bond	Bond	US Comdty
42	F_W	Wheat	Agriculture	W Comdty
43	F_XX	Dow Jones STOXX 50	Index	VG Index
44	F_YM	E-mini Dow Jones Industrial Average	Index	DM Index
45	F_AX	DAX	Index	GX Index
46	F_CA	CAC40	Index	CF Index
47	F_DT	EURO Bond	Bond	RX Comdty
48	F_UB	EURO Bobl	Bond	OE Comdty
49	F_UZ	EURO Schatz	Bond	DU Comdty
50	F_GS	10-Year Long Gilt	Bond	G Comdty
51	F_LX	FTSE 100 Index	Index	Z Index
52	F_SS	3-Month Short Sterling	Interest Rate	L Comdty
53	F_DL	Milk Class III	Agriculture	DA Comdty
54	F_ZQ	30-Day Fed Funds	Interest Rate	FF Comdty
55	F_VX	Volatilty Index	Index	UX Index
56	F_AE	AEX Index	Index	EO Index
57	F_BG	Gas Oil	Energy	QS Comdty
58	F_BC	Brent Crude Oil	Energy	CO Comdty
59	F_LU	Rotterdam Coal	Energy	XA Comdty
60	F_DM	MDAX Index	Index	MF Index
61	F_AH	Bloomberg Commodity Index	Index	DN Index
62	F_CF	10y Swiss Note	Bond	SWC Comdty
63	F_DZ	TechDAX	Index	DP Index
64	F_FB	DJ Stoxx Bank 600	Index	BJ Index
65	F_FL	Chicago Ethanol	Energy	CUA Comdty
66	F_FM	Stoxx Europe Mid 200	Index	SXR Index
67	F_FP	OMX Helsinki 25	Index	OT Index
68	F_FY	Stoxx Europe 600	Index	SXO Index
69	F_GX	Euro BUXL	Bond	UB Comdty
70	F_HP	Natural Gas Penultimate	Energy	ZA Comdty
71	F_LR	Brazilian Real	Currency	BR Curncy
72	F_LQ	Newcastle Coal	Energy	XW Comdty
73	F_ND	New Zealand Dollar	Currency	NV Curncy

	Quantiacs ticker	**Name**	**Type**	**Matching Bloomberg base code**
74	F_NY	Nikkei 225	Index	NI Index
75	F_PQ	PSI20	Index	PP Index
76	F_RR	Russian Ruble	Currency	RU Curncy
77	F_RF	EURO FX/Swiss Franc	Currency	RF Curncy
78	F_RP	EURO FX/British Pound	Currency	RP Curncy
79	F_RY	EURO FX/Japanese Yen	Currency	RY Curncy
80	F_SH	Swiss Mid Cap	Index	S1 Index
81	F_SX	Swiss Market	Index	SM Index
82	F_TR	South African Rand	Currency	RA Curncy
83	F_EB	3-Month EuriBor	Interest Rate	ER Comdty
84	F_VF	5-Year Euro Swapnote	Bond	T Comdty
85	F_VT	10-Year Euro Swapnote	Bond	P Comdty
86	F_VW	2-Year Euro Swapnote	Bond	RW Comdty
87	F_GD	Goldman Sachs Commodity Index	Index	GI Index
88	F_F	3-Month EuroSwiss	Interest Rate	ES Comdty

D Robustness checks

D.1 Using experience dummies

In this section of the appendix we replace the number of contests integer variable with a set of dummies in the appropriate regression specifications to verify that our conclusions are robust to a relaxation of the linear functional form. The direction of the relationship is confirmed.

The robustness results are as follows:

- Table 17 modifies the specification of Table 4.
- Table 18 modifies the specification of Table 5.

Table 17: OLS & panel regressions of in-sample ("backtest") & out-of-sample ("live") performance outcomes against experience.

	Dependent variable:			
	Backtest $SR_{i,t}^{Best}$		Live $SR_{i,t}^{Best}$	
	OLS	*panel linear*	*OLS*	*panel linear*
	(1)	(2)	(3)	(4)
$\mathbb{1}\{\text{Contests experienced}_{i,t} = 2\}$	2.434***	2.639***	1.391***	1.322**
	(0.810)	(0.929)	(0.320)	(0.555)
$\mathbb{1}\{\text{Contests experienced}_{i,t} = 3\}$	3.846***	3.574***	1.129	1.644**
	(0.660)	(1.355)	(0.785)	(0.792)
$\mathbb{1}\{\text{Contests experienced}_{i,t} = 4\}$	4.940***	4.085**	1.866***	4.998***
	(1.795)	(1.905)	(0.399)	(1.415)
$\mathbb{1}\{\text{Contests experienced}_{i,t} = 5\}$	5.223***	4.448**	1.892**	6.279***
	(1.142)	(2.093)	(0.883)	(1.929)
$\mathbb{1}\{\text{Contests experienced}_{i,t} = 6\}$	4.044***	6.643**	1.759	7.740***
	(0.921)	(2.728)	(1.404)	(2.569)
$\mathbb{1}\{\text{Contests experienced}_{i,t} = 7\}$	1.919**	5.530*	0.753	9.646***
	(0.759)	(2.940)	(0.505)	(3.097)
$\mathbb{1}\{\text{Contests experienced}_{i,t} = 8\}$	−0.154***	6.035*	−1.194***	9.397***
	(0.037)	(3.388)	(0.328)	(3.345)
Intercept	✓		✓	
Contest FEs		✓		✓
Contestant FEs		✓		✓
Observations	874	874	874	874
R^2	0.214	0.094	0.058	0.101

Note: standard errors (in parentheses) are robust to heteroskedasticity and are double-clustered by contest and contestant. *p<0.1; **p<0.05; ***p<0.01

Table 18: Panel regressions of out-of-sample performance against experience, backtest performance, and prior contest rank, with interactions.

	Dependent variable:			
	Live $SR^{Best}_{i,t}$			
	(1)	(2)	(3)	(4)
$\mathbb{1}\{\text{Contests experienced}_{i,t} = 2\}$	1.322**	1.137*		
	(0.563)	(0.585)		
$\mathbb{1}\{\text{Contests experienced}_{i,t} = 3\}$	1.644**	1.394*	−0.160	−0.123
	(0.804)	(0.822)	(0.758)	(0.871)
$\mathbb{1}\{\text{Contests experienced}_{i,t} = 4\}$	4.998***	4.713***	2.759***	3.329**
	(1.436)	(1.470)	(1.014)	(1.341)
$\mathbb{1}\{\text{Contests experienced}_{i,t} = 5\}$	6.279***	5.968***	3.834***	3.738**
	(1.958)	(1.973)	(1.478)	(1.633)
$\mathbb{1}\{\text{Contests experienced}_{i,t} = 6\}$	7.740***	7.276***	4.814**	6.999***
	(2.608)	(2.677)	(1.925)	(2.022)
$\mathbb{1}\{\text{Contests experienced}_{i,t} = 7\}$	9.646***	9.259***	6.282***	6.278**
	(3.144)	(3.183)	(2.343)	(2.625)
$\mathbb{1}\{\text{Contests experienced}_{i,t} = 8\}$	9.397***	8.975***	5.793**	6.840**
	(3.395)	(3.459)	(2.664)	(2.674)
Backtest $SR^{Best}_{i,t}$		0.070	0.164**	0.180**
		(0.048)	(0.074)	(0.088)
Log Percentile($\text{Score}^{Best}_{i,t-1}$)			−2.011**	−2.077**
			(0.934)	(0.954)
Backtest $SR^{Best}_{i,t}$ × $\mathbb{1}\{\text{Contests experienced}_{i,t} = 3\}$				0.035
				(0.110)
Backtest $SR^{Best}_{i,t}$ × $\mathbb{1}\{\text{Contests experienced}_{i,t} = 4\}$				−0.014
				(0.072)
Backtest $SR^{Best}_{i,t}$ × $\mathbb{1}\{\text{Contests experienced}_{i,t} = 5\}$				0.111
				(0.198)
Backtest $SR^{Best}_{i,t}$ × $\mathbb{1}\{\text{Contests experienced}_{i,t} = 6\}$				−0.322**
				(0.139)
Backtest $SR^{Best}_{i,t}$ × $\mathbb{1}\{\text{Contests experienced}_{i,t} = 7\}$				0.443
				(0.451)
Contest FEs	✓	✓	✓	✓
Contestant FEs	✓	✓	✓	✓
Observations	188	188	124	124
R^2	0.101	0.109	0.215	0.261

Note: standard errors (in parentheses) are robust to heteroskedasticity and are double-clustered by contest and contestant. *p<0.1; **p<0.05; ***p<0.01

D.2 Using the official score

In this section of the appendix we repeat our main results using the official scores that participants receive instead of out-of-sample Sharpe Ratios. An entry's score is the minimum of its in- and out-of-sample Sharpe Ratios, so the official score values will sometimes be higher than the out-of-sample Sharpe Ratios. Our conclusions all carry through when using the official scores instead of out-of-sample Sharpe Ratios.

The robustness results are as follows:

- Table 19 modifies the specification of Table 5.
- Table 20 modifies the specification of Table 6.
- Table 21 modifies the specification of Table 7.
- Table 22 modifies the specification of Table 18

Table 19: OLS & panel regressions of official score against experience, backtest performance, and prior contest rank.

	Dependent variable:								
	$\text{Score}_{i,t}^{Best}$								
	OLS			*panel linear*					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Contests experienced$_{i,t}$	0.392***	0.071	0.016	0.457***	0.040	0.035	0.955***	0.827***	0.842***
	(0.120)	(0.086)	(0.070)	(0.094)	(0.090)	(0.087)	(0.230)	(0.217)	(0.251)
Backtest $\text{SR}_{i,t}^{Best}$		0.107***	0.063		0.151***	0.145***		0.105***	0.110***
		(0.033)	(0.075)		(0.036)	(0.044)		(0.032)	(0.042)
Log Percentile($\text{Score}_{i,t-1}^{Best}$)		0.548*	0.543*		0.131	0.128		−1.367***	−1.361***
		(0.288)	(0.290)		(0.340)	(0.350)		(0.383)	(0.377)
Contests experienced$_{i,t}$ × Backtest $\text{SR}_{i,t}^{Best}$			0.016			0.002			−0.002
			(0.020)			(0.010)			(0.020)
Intercept	✓	✓	✓						
Contest FEs				✓	✓	✓	✓	✓	✓
Contestant FEs							✓	✓	✓
Observations	874	124	124	874	124	124	874	124	124
R^2	0.072	0.228	0.232	0.095	0.263	0.263	0.082	0.172	0.172

Note: standard errors (in parentheses) are robust to heteroskedasticity and are double-clustered by contest and contestant. *p<0.1; **p<0.05; ***p<0.01

Table 20: Poisson GLM regressions of contestants' number of entries in a contest against experience and performance measures.

	Dependent variable:			
	Entries$_{i,t}$			
	(1)	(2)	(3)	(4)
Contests experienced$_{i,t}$	0.390***	0.260***	0.182***	0.161**
	(0.045)	(0.049)	(0.069)	(0.067)
Score$_{i,t}^{Best}$		0.382***		0.261***
		(0.043)		(0.048)
Log Percentile(Score$_{i,t-1}^{Best}$)			0.777**	0.380
			(0.344)	(0.255)
Contestants$_t$				−0.0001
				(0.002)
Intercept	✓	✓	✓	✓
Observations	874	874	124	124

Note: standard errors (in parentheses) are robust to heteroskedasticity.
*p<0.1; **p<0.05; ***p<0.01

Table 21: Heckit (i.e. Type II Tobit) two-stage selection models for implementing selection bias corrections to the regression of official score against experience. The first stage models the probability of participation, while the second stage models the outcome of interest.

		All contestants		*Repeated contestants*	
Stage		$\text{Score}_{i,t}^{Best}$	$\text{Score}_{i,t}^{Best}$	$\text{Score}_{i,t}^{Best}$	$\text{Score}_{i,t}^{Best}$
1. *Selection*	(Intercept)	1.95^{***}	-2.58^{***}	1.69^{***}	-0.35
		(0.18)	(0.53)	(0.44)	(0.79)
	Contests experienced$_{i,t}$	-0.88^{***}	0.33^{***}	-0.39^{***}	-0.25^{***}
		(0.05)	(0.06)	(0.07)	(0.08)
	Quantopian search index$_t$	-0.01^{**}	-0.01^{**}	0.00	0.00
		(0.00)	(0.00)	(0.01)	(0.01)
	Ratio of entries to contest mean$_{i,t-1}$	0.23^{***}	0.34^{***}	0.14^{***}	0.13^{**}
		(0.04)	(0.05)	(0.05)	(0.06)
	Log Percentile($\text{Score}_{i,t-1}^{Best}$)		0.28^{**}		0.37^{**}
			(0.11)		(0.16)
2. *Outcome*	(Intercept)	-0.99^{***}	-0.20	0.20	-0.57
		(0.08)	(1.71)	(0.25)	(2.00)
	Contests experienced$_{i,t}$	0.75^{***}	-0.06	0.37^{**}	0.33^{*}
		(0.11)	(0.12)	(0.15)	(0.18)
	Log Percentile($\text{Score}_{i,t-1}^{Best}$)		0.41		0.23
			(0.31)		(0.45)
	Inverse Mills Ratio	-0.76^{***}	-0.78^{**}	-1.95^{**}	-2.16^{*}
		(0.19)	(0.33)	(0.84)	(1.27)
	σ	1.33	1.66	1.88	2.09
	ρ	-0.57	-0.47	-1.04	-1.03
	R^2	0.09	0.13	0.05	0.12
	Adj. R^2	0.09	0.11	0.04	0.10
	Num. obs.	1482	745	233	174
	Censored	621	621	50	50
	Observed	861	124	183	124

Note: parentheses denote standard errors. $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Table 22: Panel regressions of official scores against experience, backtest performance, and prior contest rank, with interactions.

	Dependent variable:			
	$\text{Score}^{Best}_{i,t}$			
	(1)	(2)	(3)	(4)
$\mathbb{1}\{\text{Contests experienced}_{i,t} = 2\}$	0.992***	0.735**		
	(0.323)	(0.331)		
$\mathbb{1}\{\text{Contests experienced}_{i,t} = 3\}$	1.716***	1.367***	0.522	0.727*
	(0.439)	(0.470)	(0.395)	(0.422)
$\mathbb{1}\{\text{Contests experienced}_{i,t} = 4\}$	3.649***	3.251***	2.034***	1.802**
	(0.793)	(0.812)	(0.621)	(0.892)
$\mathbb{1}\{\text{Contests experienced}_{i,t} = 5\}$	4.716***	4.283***	3.274***	2.978***
	(1.068)	(1.040)	(0.825)	(0.882)
$\mathbb{1}\{\text{Contests experienced}_{i,t} = 6\}$	4.840***	4.192***	2.890**	2.933**
	(1.388)	(1.398)	(1.220)	(1.294)
$\mathbb{1}\{\text{Contests experienced}_{i,t} = 7\}$	7.110***	6.570***	5.124***	4.387***
	(1.571)	(1.570)	(1.341)	(1.662)
$\mathbb{1}\{\text{Contests experienced}_{i,t} = 8\}$	6.967***	6.379***	5.062***	4.643***
	(1.705)	(1.719)	(1.500)	(1.595)
Backtest $\text{SR}^{Best}_{i,t}$		0.098***	0.138***	0.148**
		(0.034)	(0.042)	(0.067)
Log Percentile($\text{Score}^{Best}_{i,t-1}$)			−1.213***	−1.206***
			(0.431)	(0.421)
Backtest $\text{SR}^{Best}_{i,t} \times \mathbb{1}\{\text{Contests experienced}_{i,t} = 3\}$				−0.061
				(0.073)
Backtest $\text{SR}^{Best}_{i,t} \times \mathbb{1}\{\text{Contests experienced}_{i,t} = 4\}$				0.004
				(0.040)
Backtest $\text{SR}^{Best}_{i,t} \times \mathbb{1}\{\text{Contests experienced}_{i,t} = 5\}$				0.006
				(0.113)
Backtest $\text{SR}^{Best}_{i,t} \times \mathbb{1}\{\text{Contests experienced}_{i,t} = 6\}$				−0.077
				(0.065)
Backtest $\text{SR}^{Best}_{i,t} \times \mathbb{1}\{\text{Contests experienced}_{i,t} = 7\}$				0.158
				(0.265)
Contest FEs	✓	✓	✓	✓
Contestant FEs	✓	✓	✓	✓
Observations	188	188	124	124
R^2	0.139	0.197	0.288	0.322

Note: standard errors (in parentheses) are robust to heteroskedasticity and are double-clustered by contest and contestant. *p<0.1; **p<0.05; ***p<0.01

Table 23: OLS regressions of individual (contestant-level) dispersions (calculated as the range, i.e. best/max minus worst/min) of contestants' official scores, against additional data availability.

	Dependent variable:			
	$\text{Range}_{i,t}(\text{Score}_{i,t})$			
	(1)	(2)	(3)	(4)
$\mathbb{1}\{\text{New variables available}_t\}$	-0.159	-0.233^{**}	-0.212^{***}	-0.212^{***}
	(0.107)	(0.097)	(0.063)	(0.060)
$\text{Contests experienced}_{i,t}$		0.624^{***}	0.291^{***}	0.291^{***}
		(0.111)	(0.107)	(0.106)
$\text{Entries}_{i,t}$			0.133^{***}	0.133^{***}
			(0.020)	(0.020)
$\text{Score}_{i,t}^{Mean}$				0.001
				(0.036)
Intercept	✓	✓	✓	✓
Observations	874	874	874	874
R^2	0.003	0.183	0.516	0.516

Note: standard errors (in parentheses) are robust to heteroskedasticity and are double-clustered by contest and contestant. *p<0.1; **p<0.05; ***p<0.01