Algorithm Aversion: Theory and Evidence from Robo-Advice*

Cynthia Pagliaro[1] Tarun Ramadorai[2] Alberto G. Rossi[3]

Stephen Utkus[3] Ansgar Walther[2]

[1]Vanguard, [2]Imperial College, [3]Georgetown University

Preliminary. Please Do Not Cite or Distribute without the Authors' Permission

December 5, 2022

Abstract

Algorithms hold great potential to lower costs and democratize access to a wide variety of consumer services. How do humans interact with algorithms, and what are the major barriers to algorithmic adoption? We answer these questions using a structural model applied to unique data that captures interactions between human clients and "hybrid" robo-advisors that offer different levels and standards of human counseling to complement algorithmic investment. The model features three dimensions of investors' algorithm aversion, all of which can be influenced by human advice, namely: a per-period disutility of dealing with the algorithm, pessimism about the algorithm's ability to manage assets, and uncertainty about the algorithm's performance. We estimate the model's parameters using quasi-random variation in the matching of clients with human advisors generated by mechanical allocation rules. We find evidence that algorithm aversion is mainly driven by ongoing disutility and uncertainty, and that human advice is important in retaining investors in robo-advice during market downturns.

*We thank Brian Martin at Vanguard for useful conversations on institutional details, and Paige Weisman for excellent research assistance.

Electronic copy available at: https://ssrn.com/abstract=4301514

Introduction

Algorithms are becoming ubiquitous. For a wide variety of common tasks, we routinely rely on computers to take automated decisions by applying rule-sets and codified processes to data. In finance, such automation holds enormous potential to lower costs and democratize access to high-quality asset and liability management and financial advice—services routinely delivered by humans and often accessible only to the wealthy (Gomes, Haliassos, and Ramadorai, 2021; D'Acunto and Rossi, 2022).

To achieve such democratization, algorithms must be deployed at scale. However, there is evidence that there may be limits to the widespread adoption of algorithmic solutions arising from human discomfort. Psychology experiments have shown that people prefer inferior human forecasters to superior algorithmic forecasters even after observing performance differences (Dietvorst, Simmons, and Massey, 2015). This type of "algorithm aversion" shows up in various settings, but there is little consensus about the psychological underpinnings of this phenomenon (Burton, Stein, and Jensen, 2020).

In this paper, we use a structural model combined with novel evidence from the economically important field setting of robo-advising to identify the factors underpinning algorithm aversion.[1] Our structural model features a per-period disutility of dealing with an algorithm, as well as a "learning" channel through which the human client refines their understanding of the value that the algorithm provides over time and across states of the economy. We adapt the model to describe the decisions made by an investor that delegates their financial portfolio to a robo-advisor that implements an algorithmic solution. We identify the parameters of this model using a unique dataset from a large global provider of "hybrid" robo-advice, which complements algorithmic portfolio optimization solutions with human support and advice.

The model adapts a standard dynamic portfolio choice framework (Campbell and Viceira (1999)) to analyze algorithm aversion in robo-advice. The investor in the model chooses the fraction of their wealth that they allocate to the robo-advisor, which generates a return stream modeled as a generic risky asset paying a positive risk premium. We introduce two changes to the standard setup to capture

[1] Robo-advising is a uniquely useful venue to understand algorithm aversion because of its current scale and expected growth——estimated at over $10 trillion under management over the next decade (see, e.g., Deloitte and Touche, "The expansion of robo-advisory in wealth management," 2016).

Electronic copy available at: https://ssrn.com/abstract=4301514

both a preference and a learning channel of algorithm aversion. The preference channel is modelled as the investor facing a per-period disutility cost for investing in the robo-advisor, which captures the ongoing psychological costs of dealing with an algorithm. The learning channel is modelled by positing that the investor is uncertain about the expected return generated by the algorithm. The investor learns about this return as time evolves, updating her prior belief using Bayes rule applied to the sequence of returns generated by the algorithm. This focus on learning creates similarities between our work and Timmermann (1993), Brennan and Xia (2001), and Pástor and Veronesi (2003).

This model setup allows for three conceptually distinct sources of algorithm aversion: the ongoing *disutility* of dealing with the robo-advisor, investors' prior *pessimism* about expected returns generated by the algorithm, and investors' prior *uncertainty* about expected returns. We further assume that the investor in the model has access to human advice to complement their use of the algorithmic solution; all three facets of algorithm aversion can be affected by human advice. In particular, depending on their type, human advisors can mitigate the investor's ongoing disutility, shift the investor's prior expectations, and affect the investor's prior uncertainty. Through this effect on uncertainty, human advisors affect the investor's rate of learning about the algorithm's ability to perform the complex task of portfolio optimization. However, the human advisor cannot affect the "true" return distribution, which is in keeping with the standard institutional features of robo-advising and the empirical setting to which we map our model.

These model features have two possible interpretations. First, different prior beliefs across different types of human advisors could themselves be the result of rational Bayesian updating. More specifically, imagine that the investor starts with an initial (meta-)prior about the algorithm's ability to deliver high expected returns, and assume that a human advisor can convey credible information to explain the true ability of the algorithm. In this case, the starting points of investors advised by advisors with different characteristics can be thought of as investors' updated beliefs conditional on such information. Second, different prior beliefs can also arise in a behavioral model, in which human advisors can use salesmanship and persuasion to move investors' beliefs, even if they convey no hard information (Gennaioli, Shleifer, and Vishny (2015)).

In addition, it is worth highlighting that the value of (algorithmic) advice can be interpreted

Electronic copy available at: https://ssrn.com/abstract=4301514

more broadly. In our formal model, we define it chiefly in terms of the expected returns it generates. In reality, there are many possible dimensions of value-added, for example, helping investors set goals for retirement, or providing behavioral coaching and "financial peace of mind" (Pagliaro and Utkus, 2019; Rossi and Utkus, 2019). Our model can be viewed as incorporating these benefits, while converting them into (equivalent) units of expected return. That said, we do provide some evidence that uncertainty about returns is a dimension of algorithm aversion in the setting that we study.

Key outcomes in the model, such as investor retention in the robo-advising service and the fraction of the portfolio delegated to the algorithm, vary with the type of human advice. In theory, variation in these outcomes can therefore identify the effects of the "deep" model parameters that capture the different components of algorithm aversion. In our empirical setting, to accurately estimate the model's parameters, we require that the advisor type is randomly assigned to the investor in the sense that advisor type assignment is orthogonal to the investor's degree of algorithm aversion.

Robo-advising is one of the largest real-world settings in which algorithms are used to provide customized solutions at scale to large numbers of households The data on robo-advising that we employ comes from Personal Advisor Services (PAS), a division of Vanguard. The service it provides is "hybrid," complementing algorithmic robo-advice with human advisory input.[2] In the service we analyze, the algorithm manages the investment portfolio, while the human advisor interacts with investors to help them understand what the algorithm does and advises them on additional auxiliary services such as cash flow, tax, and estate planning. The data provide information on trades, positions, demographic characteristics, and investor-advisor interactions for a large set of previously self-directed investors who participated in the robo-advisor between January 2015 and March 2018.

An ideal experiment for estimating algorithm aversion would randomly assign clients to identical portfolios, managed either by a robo-advisor ("treatment") or a traditional human advisor ("control"). Our empirical setting contains exogenous variation that approximates this ideal experiment, and we use it to back out the model parameters. More specifically, after eliciting investors' characteristics (age, income, investment horizon, and preferences) the robo-component of the robo-advisor (the algorithm) formulates a financial plan and an associated investment portfolio. The sign-up process then requires

[2]Hybrid robo-advisors allow investors differing levels of interaction with human financial advisors in addition to the algorithmic solution, as opposed to "pure" robo-advisors which simply provide access to the algorithmic solution.

Electronic copy available at: https://ssrn.com/abstract=4301514

the investor to schedule an appointment with a human financial advisor who explains the financial plan, completes onboarding, and provides ongoing support to investors. Critically, the assignment of investors to advisors follows mechanical rules. This assignment is driven by workload balancing imperatives rather than any assessment of advisor type. Once the current client "load" of a given advisor is accounted for, we confirm that there is no empirical relationship between the historical retention rate of clients assigned to a given advisor and the assignment of new clients to the advisor.

This quasi-random assignment allows us to use historical retention (conditional on current client load) to measure advisor type. We can think of clients randomly assigned human advisors with past low retention as closer to the "treatment" group in the ideal experiment, and clients randomly assigned a past-high-retention advisor as closer to the "control" group. Leveraging this reasoning, we use the treatment effect of random assignment to advisors of different types to draw inferences about the parameters of our structural model of algorithm aversion.

In reduced-form, the data reveal that assignment to advisors of different types predicts future client retention. Put differently, assignment to low-retention advisors predicts a greater likelihood of exiting robo-advising than assignment to a high-retention advisor. In addition, we find that human advice is particularly effective during periods of market turbulence when extreme signals of returns hit investors. Indeed, high-retention advisors' clients are less likely to quit robo-advising during periods of market distress than low-retention advisors' clients.

These observed differences allow us to identify the different dimensions of algorithm aversion. We derive an estimation equation that maps these reduced-form estimates to the structural parameters of the model. Interpreting our empirical findings through this lens, we derive several further economic insights. First, the presence of a high-retention human advisor removes a large proportion of the uncertainty component of algorithm aversion. Indeed, our estimates are consistent with a model in which a high-retention human advisor removes over 90% of the effect of investors' prior about the expected returns generated by the algorithm.

We also find that high-retention human advisors reduce clients' propensity to quit in benign market conditions by around 25%. In the model, this estimate can be the result of either the pessimism dimension of algorithm aversion (high-retention human advisors correct pessimistic priors about ex-

Electronic copy available at: https://ssrn.com/abstract=4301514

pected performance), or the disutility dimension (a completely automated service imposes a greater psychological burden on investors). To get a sense of the relative importance of these two effects, we employ an auxiliary prediction of the model, namely, that the effect of advice conditional on market performance should be weaker on experienced robo-advising clients, who have substantially converged on their posterior beliefs about the ability of the algorithm. We confirm that this holds in the data, where experienced clients, regardless of their advisor type, react less to market turbulence. This finding is consistent with experienced clients having less to learn about the algorithm's ability to deliver returns.

This finding leads us to a second economic insight. Having confirmed that experienced clients' beliefs are close to convergence to the algorithm's true ability, it is natural to argue that prior pessimism should not be important for these clients' decisions going forward. In models of learning, the weight that investors place on their prior beliefs converges to zero as they become more experienced. We can therefore extract a cleaner estimate of the disutility component of algorithm aversion by studying the behavior of experienced clients. Repeating our baseline empirical estimation in a sample of experienced clients, we find that high-retention human advisors still reduce the baseline propensity to quit by about 23%. While more comprehensive structural estimation of model parameters would shed additional light on the decomposition between the disutility and pessimism components of algorithm aversion, our estimates using this strategy suggest that the disutility component is relatively more important.

Overall, our empirical work provides evidence that ongoing disutility and uncertainty are the most salient dimensions of algorithm aversion in this important empirical setting. While we do map the model's parameters to our empirical work, in this draft of the paper we do not engage in full-blown structural estimation of the parameters. We intend to pursue this as the next natural step in our agenda and use such structurally-estimated parameters to evaluate counterfactual mechanisms to make human-algorithm matching most efficient in the presence of algorithm aversion.

Our work contributes to several strands of the finance and economics literatures. First, we contribute to the growing literature about the costs and benefits of dealing with algorithms and the digital revolution (see, for example, Allcott, Gentzkow, and Song (2021); Niszczota and Kaszás (2020); Castelo, Bos, and Lehmann (2019); Fuster et al. (2022)). Second, we contribute to the growing litera-

Electronic copy available at: https://ssrn.com/abstract=4301514

ture on robo-advising (see, for example, D'Acunto, Prabhala, and Rossi (2019); Reher and Sun (2019); Rossi and Utkus (2019, 2020)). This literature is now beginning to explore the role of trust in robo-advice, linked to the broader literature on the role of trust in financial institutions and stock market participation (see, e.g., Guiso, Sapienza, and Zingales (2008)). Third, we contribute to the household finance literature, which seeks ways to both incentivize households to participate in financial markets and do so in an efficient, well-diversified manner (see, e.g., Campbell (2006); Guiso and Sodini (2013); Gomes, Haliassos, and Ramadorai (2021)). Fourth, our work is related to the more general literature on the quality of financial advice (see, e.g., Linnainmaa et al. (2018); Egan, Matvos, and Seru (2019). In this setting, our focus is primarily on the ability of advisers to assuage investors' concerns about the algorithmic solution.

The remainder of this paper is organized as follows. Section 1 presents the model and its testable implications. Section 2 describes the institutional details of our empirical setting and the details of our identification strategy. Section 3 contains our main empirical results, and Section 4 concludes.

1 Model

1.1 Environment

Our model considers a population of investors $i \in [0, 1]$. Each investor is endowed with initial wealth W_0^i, makes investment decisions at dates $t = 0, 1, \ldots T-1$, and consumes her final wealth W_T^i at date $t = T$. At each investment date, investor i decides whether to participate in an algorithmic portfolio management service (the "algorithm" or "robo-advisor"). We write $\gamma_t^i = 1$ if investor i participates at date t, and $\gamma_t^i = 0$ if she opts out.

An investor who opts out invests all her wealth in an outside portfolio, which produces a constant return $\bar{R}$.[3] A participating investor invests a fraction $\alpha_t^i \geq \underline{\alpha}$ of her wealth with the algorithm, where $\underline{\alpha}$ is the minimum investment required for participation,[4] and all remaining wealth in the outside

[3]The assumption that the outside portfolio is risk-free has two possible interpretations. First, it can represent a world in which each investor's outside option is cash investment. Second, it can represent the certainty equivalent return on an investor's risky outside portfolio. The latter interpretation maps exactly to our model if we assume that outside returns and advised returns are uncorrelated. Below, we allow for unobserved heterogeneity in the outside return $\bar{R}$ across consumers. Moreover, in a behavioral model, the outside option could also represent investors' possibly misguided beliefs about what they can achieve, not necessarily feasible return.

[4]If $\underline{\alpha} = 0$, then this requirement simply imposes the realistic constraint that investors cannot short-sell the portfolio

Electronic copy available at: https://ssrn.com/abstract=4301514

portfolio. The algorithm yields stochastic returns R^i_{t+1} between dates t and $t+1$ and charges a fee of $C > 0$ per unit of assets under management. Therefore, the net return after fees on investments in the algorithm is $R^i_{t+1} - C$.

Each investor also has access to a human advisor who guides her through the process of algorithmic investment. Investors are randomly matched to human advisors, who are of different types. We let $h^i = 0$ denote the event that investor i is matched with a type-0 (relatively lower effectiveness) human advisor. This state of the world corresponds more closely to pure algorithmic investment and evaluation by the investor with no human guidance. Conversely, we write $h^i = 1$ if investor i is matched with a type-1 (relatively higher effectiveness) human advisor. Our model allows contact with such type-1 advisors to reduce investors' inherent aversion to algorithmic advice, or to move their beliefs about the quality of the algorithmic service. We describe these channels in detail below.

1.1.1 Preferences

Investor i's lifetime utility is given by

$$\log(W^i_T) - f\left(h^i\right) \cdot \sum_{t=0}^{T-1} \gamma^i_t. \tag{1}$$

In addition to the standard logarithmic utility from consuming her final wealth, the investor suffers a fixed utility cost of $f\left(h^i\right) > 0$ for any period where she uses the robo-advice. We interpret this parameter as a psychological aversion to dealing with an algorithm, which can vary with the type of human advice. For example, if $f(0) > f(1)$, then a type-1 human advisor alleviates an investor's disutility to algorithmic advice.

1.1.2 Beliefs

Defining the log return on the robo portfolio as $r^i_{t+1} \equiv \log R^i_{t+1}$, we assume that

$$r^i_{t+1} = \theta + \sigma u^i_{t+1},\ u^i_{t+1} \sim N(0,1). \tag{2}$$

generated by the algorithm. In practice, providers of algorithmic advice tend to impose a minimum investment $\underline{\alpha} > 0$.

Electronic copy available at: https://ssrn.com/abstract=4301514

In this equation, θ is the average log return, which serves as a measure of the algorithm's quality, σ is a parameter measuring its volatility, and u_{t+1}^i denotes a standard Gaussian shock that is independent across time periods. As discussed above, one can alternatively interpret θ as a composite of expected log returns, on one hand, and the per-period utility benefit of other dimensions of financial advice, on the other hand. Examples include the utility benefits of goal-setting, planning for retirement and bequests, behavioral coaching and generating trust/confidence (Pagliaro and Utkus, 2019; Rossi and Utkus, 2019).

The investor is uncertain about the algorithm's quality θ.[5] At date 0, she views quality as a random variable with a Gaussian prior distribution:

$$\theta \sim N\left(m_0\left(h^i\right), v_0\left(h^i\right)\right) \tag{3}$$

Subsequent beliefs, at dates $t > 0$, are updated using Bayes rule after observing the history of returns $I_t^i = \left\{r_1^i, \ldots, r_t^i\right\}$. These updates give rise to standard Kalman filtering equations, which we derive below.

In this specification, the type of human advice can affect both the mean and variance of investors' prior beliefs. For example, if $m_0(1) > m_0(0)$, then a type-1 human advisor convinces the investor that the robo advisor has high average returns. If $v_0(1) < v_0(0)$, then a type-1 advisor generates tighter/more confident beliefs.

Importantly, we do not allow the *true* quality θ of the algorithm to depend on h^i. Instead, we model it as an inherent property of the algorithm, which is fixed across investors and unaffected by human intervention.

This environment has two possible interpretations. First, different prior beliefs across type-1 and type-0 advisors at date 0 could themselves be the result of rational Bayesian updating. If the investor starts with an initial (meta-)prior about θ and if a human advisor can convey credible information to the investor regarding the (constant) true θ, then one can think of $m_0\left(h^i\right)$ and $v_0\left(h^i\right)$ as investors' updated beliefs conditional on such information. Second, different prior beliefs can also arise in a

[5] Our setup, in which the mean return is uncertain for a known level of volatility σ, yields a particularly tractable characterization of investors' choices. It also captures the general idea that there is uncertainty about the risk-return trade-off, or Sharpe ratio, that a robo advisor is able to achieve.

Electronic copy available at: https://ssrn.com/abstract=4301514

behavioral model, in which human advisors can use salesmanship and persuasion to move investors' beliefs, even if they convey no hard information. However, even if the difference in priors is given a behavioral interpretation, we assume that investors' updated beliefs from date 1 onwards satisfy Bayes' rule under the assumption that the prior specified in (3) is true.

1.1.3 Dimensions of algorithm aversion

We have modeled a setting where investors can be matched either with a type-0 human advisor (a situation akin to "pure" robo advice) or a type-1 advisor. The object of interest in our theoretical and empirical analysis will be the difference between investors' preferences and beliefs in these two scenarios. Concretely, our primary goal is to estimate three quantities from the model.

The first is the effect of type-1 human advice on the per-period *disutility* induced by having to deal with the algorithm:

$$f(1) - f(0).$$

If this difference is negative, so that fixed costs are lowered by human input, then it measures a pure form of psychological algorithm aversion.

The second measure is the importance of the type of human-advisor on investors' prior expectations of returns from the algorithm, which is given by

$$m_0(1) - m_0(0).$$

This difference describes potential investors' *pessimism* about the algorithm's ability to manage assets. If it is positive, then type-1 human advice makes investors more optimistic on average about the algorithm's performance.

The final quantity of interest is the effect of human input on investors' perceived variance of expected returns, given by

$$v_0(1) - v_0(0).$$

This difference measures investors' *uncertainty* about the algorithm's performance. It is not clear *a priori* whether one would expect algorithm aversion to generate higher or lower prior variance. For

Electronic copy available at: https://ssrn.com/abstract=4301514

example, if an investor is extremely uncertain about the algorithm's performance and a type-1 advisor reduces this uncertainty, then the difference would be negative. On the other hand, if an investor is initially convinced that the algorithm's performance is terrible (i.e., if her prior is tight around a low mean), then a type-1 advisor might be able to *decrease* her confidence in this negative view and the difference above would be positive. Ultimately, the distinction between these possible cases is an empirical question, which we explore below.

1.2 Characterization of optimal choices

1.2.1 Optimization problem

An investor's problem is to choose participation decision γ_t^i and investment α_t^i to maximize her lifetime utility:

$$\max_{\{\gamma_t^i, \alpha_t^i \geq \underline{\alpha}\}_{t=1}^{T-1}} E_0 \left[\log(W_T^i) - f\left(h^i\right) \cdot \sum_{t=1}^{T-1} \gamma_t^i \right]$$

This problem is subject to two constraints. First, the evolution of her wealth obeys the inter-temporal budget constraint

$$W_{t+1} = W_t \left[\bar{R} + \gamma_t^i \alpha_t^i \left(R_{t+1}^i - C - \bar{R} \right) \right].$$

The investor earns the outside return $\bar{R}$ on her portfolio. If she participates in robo-advice service ($\gamma_t^i = 1$) and invests fraction α_t^i of her wealth, then she earns an additional excess return $R_{t+1}^i - \bar{R}$ on this investment.

The second constraint embeds the assumption that an investor who opts out of the algorithm cannot re-join at a later date:

$$\gamma_t^i = 0 \Rightarrow \gamma_s^i = 0 \text{ for all } s > t.$$

This constraint, which is realistic in our empirical setting, is economically meaningful because it generates an option value associated with using the algorithm. Even relatively pessimistic investors may have an incentive to continue to participate while they learn about the algorithm's quality because opting out would preclude future participation.

Electronic copy available at: https://ssrn.com/abstract=4301514

1.2.2 Potential outcomes notation

In order to clearly discuss how we empirically identify the parameters of this model, we adopt a potential outcomes framework. Concretely, we solve the investors' maximization problem for each potential value of human advisor type $h^i = h \in \{0, 1\}$. For every investor $i \in I$, this gives rise to two sets of optimal decisions: $\gamma_t^i(1)$, for investor i's potential participation decision at date t when matched to a type-1 human advisor, and $\gamma_t^i(0)$, for investor i's potential participation decision at date t when matched to a type-0 human advisor.

Similarly, we adopt potential outcomes notation to define investors' beliefs, which are affected by the type of human advisors via our specification of priors in Equation (3). For example, we write $E\left[\theta|I_t^i, h^i = 1\right]$ and $E\left[\theta|I_t^i, h^i = 0\right]$ for investor i's expected value of θ, conditional on information I_t^i at date t and the type h^i of their human advisor.

1.2.3 Characterization of beliefs

Each investor i faces uncertainty about the quality θ of the algorithmic advisor. Investors extract signals about θ from their observation of the history $I_t^i = \left\{r_1^i, \ldots, r_t^i\right\}$ of realized returns. Since log returns are normally distributed and investors start with the normal prior in Equation (3), their posterior beliefs about θ given I_t are also described by a normal distribution. Hence, we can characterize it in terms of its mean and variance, which we define as follows:

$$m_t^i(h) \equiv E\left[\theta|I_t^i, h^i = h\right],$$
$$v_t^i(h) \equiv Var\left[\theta|I_t^i, h^i = h\right].$$

Notice, once again, that both mean and variance are defined as potential outcomes. Each investor has a potentially different mean and variance at date t (conditional on the same realized return series I_t^i) depending on the type h of their human advisor. Using these definitions, we can characterize the evolution of beliefs.

Electronic copy available at: https://ssrn.com/abstract=4301514

Lemma 1. *Investors' posterior means satisfy the stochastic filtering equation:*

$$m_{t+1}^{i}(h) = m_{t}^{i}(h) + \underbrace{\frac{Cov_{t}^{h}[\theta, r_{t+1}^{i}]}{Var_{t}^{h}[r_{t+1}^{i}]}}_{\equiv\ Gain\ g_t(h)} \cdot \underbrace{(r_{t+1}^{i} - m_{t}^{i}(h))}_{\equiv\ News\ \eta_{t+1}^{i}(h)}. \tag{4}$$

Moreover, their posterior variances $v_t^i(h) = v_t(h)$ are independent of i and satisfy the deterministic equation

$$\frac{1}{v_t(h)} = \frac{1}{v_0(h)} + t\frac{1}{\sigma^2}. \tag{5}$$

Finally, the gain parameter is related to posterior variances as follows:

$$g_t(h) = \frac{v_t(h)}{v_t(h) + \sigma^2}. \tag{6}$$

Equation (4) characterizes the evolution of investors' expectations about algo quality θ. Conditional on the prior beliefs induced by human advice in Equation (3), this expression follows directly from a standard Kalman filtering equation. At date $t+1$, the investor starts with her previous expectation $m_t^i(h)$ and updates it in proportion to new information $\eta_{t+1}^i(h)$, which is measured by the gap between realized and expected returns. The responsiveness of beliefs to new information is measured by the Kalman gain parameter $g_t(h)$.[6]

An important property of the Kalman gain $g_t(h)$ is that it directly encodes differences in prior variance/confidence that are generated by different human advisors. Indeed, it is easy to see that $g_t(1) > g_t(0)$ if and only if $v_0(1) > v_0(0)$. The key economic implication is that human advice determines investors' sensitivity to information. At all dates $t \geq 1$, an investor with a type-1 human advisor is more (less) responsive to new information if and only if this advisor endowed them with a more (less) dispersed prior.

[6] Alternatively, one can (loosely) interpret Equation (4) as a statistical regression, in which the investor projects θ onto realized returns. The Kalman gain is the regression coefficient—as usual, given by the ratio of covariance to variance—in this projection.

Electronic copy available at: https://ssrn.com/abstract=4301514

1.2.4 Approximate recursive solution

We derive a tractable approximation to investors' optimal choices by casting their maximization problem in recursive form and by employing a log-linear approximation to their intertemporal budget constraints.

The state variables in any investor's problem are the current date t, their current log wealth $w = w_t^i \equiv \log W_t^i$, their current belief $m = m_t^i(h)$ about algo quality, and an indicator $z \in \{0, 1\}$ for whether the investor is still enrolled in algorithmic advice (i.e., $z = 0$ if the investor has opted out at any preceding date). The investor's choices given these state variables are her participation $\gamma \in \{0, 1\}$ and their investment share α in the algorithm.

The laws of motion for state variables m and z are as follows:

$$m' = m + g_t(h)\,\eta', \tag{7}$$

$$z' = \gamma z, \tag{8}$$

where $r' = r_{t+1}^i$ denotes next period's log return on algorithmic investment. As is standard, we let primes denote the state variables for the next period. Equation (7) is the equivalent of the Kalman filter (in Lemma 1) in recursive form. Equation 8 tracks the future enrollment status z', which is equal to one if and only if the investor was previously enrolled ($z = 1$) and continues to participate in the current period ($\gamma = 1$).

The approximate law of motion for the evolution of wealth is as follows:

$$w' \simeq w + \bar{r} + \underbrace{\alpha\left(m + \eta' - c - \bar{r}\right)}_{\text{Net excess return}} + \underbrace{\frac{1}{2}\alpha(1-\alpha)\sigma^2}_{\text{nonlinear adjustment}}, \tag{9}$$

where we have further defined $c = \log C$ and $\bar{r} = \log \bar{R}$. We show in the appendix that this approximation, which applies the approach of Campbell et al. (2002) to a setting with incomplete information, holds exactly in a model with continuous portfolio rebalancing. The first term, $w + \bar{r}$, denotes next period's log wealth in a scenario where all wealth is invested in the outside portfolio. The second term adds the excess return, net of fees c, due to the share α of the portfolio that is invested in the

Electronic copy available at: https://ssrn.com/abstract=4301514

algorithm. Since $m + \eta' \equiv r'$, this return could equivalently be written in terms of realized returns as $r' - c - \bar{r}$. However, our formulation in terms of m and η' is useful because it highlights that, from the perspective of an investor with imperfect information and conditional on current beliefs m, the news shock η' drives the evolution of both beliefs and wealth.[7] The final term in (9) represents an adjustment for the Jensen's inequality. It reflects the fact that the log of an average (the portfolio return) is greater than the average of the logs (the algorithm and outside returns). This term serves to represent the inherent risk aversion of an investor with log utility.

We let $\mathcal{V}_h(t, m, w, z)$ denote the investor's maximized continuation value as a function of the state variables. Consistent with our potential outcomes notation, this is again defined separately for each potential value h of human advisor type. Continuation values must satisfy the Bellman equation:

$$\mathcal{V}(t, m, w, z; h) = \begin{cases} \max_{\alpha \geq \underline{\alpha}, \gamma} E_h \left[-\gamma f(h) + \mathcal{V}_h(t+1, m', w', z'; h) \middle| t, m, w, z, \alpha, \gamma \right], & t = 0, \ldots, T-1 \\ w, & t = T, \end{cases} \tag{10}$$

where the expectation is taken under the laws of motion defined above. The first case in Equation (10) is the standard principle of optimality. Since there is no intermediate consumption, the only flow utility experienced by an investor at date $t < T$ is the fixed cost of using the algo. Today's continuation value is equal to this flow utility plus the maximized expectation of next period's continuation value. The second case denotes the terminal condition that, when reaching date T, investors simply consume their wealth W and therefore enjoy utility $w = \log W$.

We conjecture that the value function can be written as the separable function

$$\mathcal{V}_h(t, m, w, z, h) = w + (T - t)\bar{r} + zJ(t, m; h) \tag{11}$$

The motivation for this conjecture is that a log investor's lifetime utility is given by the log of current wealth plus the sum of her maximized log portfolio returns. If the investor is not enrolled at date t ($z = 0$) in algorithmic advice, then Equation (11) states that she simply earns her outside return $\bar{r}$ for the remaining $T - t$ periods. If she is enrolled, then the function $J(t, m; h)$ represents the option

[7] This property is a well known feature of continuous time models with learning, e.g., Brennan (1998) and Veronesi (2000). Our derivations in Online Appendix A.1 demonstrate that it carries over to discrete-time models, modulo our log-linear approximation.

Electronic copy available at: https://ssrn.com/abstract=4301514

value of using the algorithm from date t onward. We now derive the following result:

Proposition 1. *Investors' value function satisfies Equation 11. The option value $J(t,m;h)$ is determined by the recursive equation:*

$$J(t,m;h) = \begin{cases} \max_{\alpha \geq \underline{\alpha}, \gamma} \gamma \left[\alpha \left(m + \frac{1}{2}\sigma^2 - \bar{r} - c \right) - \frac{1}{2}\alpha^2\sigma^2 - f(h) + E\left[J(t+1, m'; h) \,|t, m \right] \right], & t = 0, \ldots, T-1 \\ 0, & t = T, \end{cases} \tag{12}$$

where the expectation $E\left[J_h(t+1, m')\,|t, m\right]$ is computed using the normal distribution

$$m'|t, m \sim N\left(m, v_t(h)\, g_t(h)\right). \tag{13}$$

Proposition 1 verifies our conjecture and characterizes the option value $J(t,m;h)$ of using the algorithm from date t onwards. Equation (12) is particularly clear in the case where the minimum investment share $\underline{\alpha} \to 0$. In this case, our characterization for $t < T$ simplifies to

$$J(t,m;h) = \max\left\{ \underbrace{\frac{1}{2}S(m)^2 - f(h)}_{\text{squared Sharpe ratio less fixed cost}} + \underbrace{E\left[J(t+1, m', h)|t, m, h\right]}_{\text{future value of enrolment}}, 0 \right\}, \tag{14}$$

where the option value of using the algorithm from date t onwards consists of two components. The first term is a measure of the algorithm's expected performance between t and $t+1$, net of the fixed costs of using the algorithm. Performance is measured in terms of the algorithm's expected Sharpe ratio, which is defined as:

$$S(m) = \max\left\{ \frac{m + \frac{1}{2}\sigma^2 - r_f - c}{\sigma}, 0 \right\}. \tag{15}$$

The second term is the option value of remaining enrolled at date $t+1$, as measured by the expected value of using the algorithm from date $t+1$ onwards. Equation (14) expresses the option value as the sum of these two terms or as zero if that sum is negative, which reflects the fact that the investor has the option to quit at date t.

With this characterization, we can directly derive the investor's optimal (potential) participation and investment decisions.

Electronic copy available at: https://ssrn.com/abstract=4301514

Corollary 1. *If investor i has a human advisor of type h, then her participation decision at date t satisfies*

$$\gamma_t^i(h) = \begin{cases} 1, & \text{if } J_h\left(t, m_t^i(h)\right) \geq 0, \\ 0, & \text{otherwise} \end{cases} \tag{16}$$

If $\gamma_t^i(h) = 1$, then optimal portfolio choice is

$$\alpha_t^i(h) = \max\left\{\frac{m_t^i(h) + \frac{1}{2}\sigma^2 - \bar{r} - c}{\sigma^2}, \bar{\alpha}\right\}. \tag{17}$$

The first part of the corollary states that it is optimal to continue using the algorithm if and only if the option value of doing so, evaluated at current beliefs $m = m_t^i(h)$, is strictly positive. The second part characterizes the optimal portfolio allocation. If the minimum investment constraint $\alpha \geq \underline{\alpha}$ is not binding, then the optimal investment is given by a standard myopic rule as in Merton (1969) and Campbell et al. (2002), by which the optimal share of wealth invested in the algorithm is the ratio of expected excess returns to variance.[8]

This result also clarifies the role of log utility in our analysis. The optimal portfolio shares in (17) do not contain any hedging demand. In models with more general preferences, investors who learn about expected returns can invest either more or less than the myopic share due to a desire to hedge the risk of bad news (i.e., a drop in beliefs m, which also correlates with a drop in wealth). Another assumption that is key for this formula is the fact that the outside return $\bar{r}$ is deterministic. With a stochastic outside portfolio, the investor faces a choice between two risky assets. While both of these extensions have been characterized analytically in the existing literature, our ultimate goal is structural estimation, and we choose to abstract from these effects for tractability.

1.3 Estimation equations

In this section, we make three additional assumptions and show that we are able obtain closed-form estimation equations in our model, which we then implement in the data. First, we assume that the

[8] In the discrete-time approximation, the expected excess return includes a correction $\frac{1}{2}\sigma^2$ for the properties of the log-normal distribution. Campbell and Viceira show that $m_t^i(h) + \frac{1}{2}\sigma^2 - \bar{r} - c = \log E_t\left[\frac{1+R_{t+1}^i - C}{1+\bar{R}}\right]$, thus reflecting the expected difference in log returns that matters in the investor's optimization problem.

Electronic copy available at: https://ssrn.com/abstract=4301514

model has only three periods with $T = 2$. Second, in order to facilitate estimation via a discrete choice model, we assume that investors have random utility at date 1. Investor i perceives her outside return at date 1 to be $\bar{r} + \xi_1^i$, where ξ_1^i is a random variable drawn from the cumulative distribution F_ξ, independently across investors i. Finally, for simplicity, we assume that the minimum investment requirement is small and let $\underline{\alpha} \to 0$.

As in our empirical exercise, we focus on investors' decision to quit at date 1, conditional on having signed up at date 0. The following result derives empirical predictions for this decision:

Proposition 2. *In the model with $T = 2$ and random utility, if investor i participates in algo advice at date 0, then she quits at date 1 (i.e., $\gamma_1^i(h) = 0$) if and only if*

$$\alpha + \beta\left(-r_1^i\right) + \gamma h + \delta h \times \left(-r_1^i\right) + \xi_1^i \geq 0, \tag{18}$$

where the coefficients in Equation (18) are defined as follows:

$$\alpha = \underbrace{\bar{r} + c - \frac{1}{2}\sigma^2 + 2\sigma\sqrt{f(0)} - m_0(0)\left(1 - g_0(0)\right)}_{\textit{intercept / baseline propensity to quit}} \tag{19}$$

$$\beta = g_0(0) \tag{20}$$

$$\gamma = \underbrace{\sigma\left(\sqrt{2f(1)} - \sqrt{2f(0)}\right)}_{\textit{disutility difference}} - \underbrace{\left[m_0(1)\left(1 - g_0(1)\right) - m_0(0)\left(1 - g_0(0)\right)\right]}_{\textit{prior mean difference}} \tag{21}$$

$$\delta = \underbrace{g_0(1) - g_0(0)}_{\textit{differential Kalman gain}} \tag{22}$$

Proposition 2 shows that the decision to participate (or, conversely, to quit) at date 1 is driven by a standard linear discrete choice equation. The intercept coefficient α defines investors' baseline propensity to participate, which corresponds to investor i's potential (random) outcome with a type-0 human advisor ($h = 0$).

In addition, the decision to participate is driven by three variables and associated coefficients. First, the coefficient β captures the effect of experiencing poor performance (i.e., negative returns).

Electronic copy available at: https://ssrn.com/abstract=4301514

From this coefficient, one can extract a direct estimate of the Kalman gain for investors with a type-0 human advisor.

Second, the coefficient γ measures the treatment effect of having a type-1 human advisor, i.e., of switching from $h = 0$ to $h = 1$. This coefficient is the combination of two structural effects of human type on investment choices. The first term in Equation (21) is the difference in the monotone transformation $\sqrt{2f(h)}$ of the disutility of using robo-advice. The second term is the difference in prior means between investors with type-1 and type-0 human advisors, where each potential mean value $m_0(h)$ is scaled by the weight $(1 - g_0(h))$ that the investor places on her prior when forming beliefs at date 1.[9] At first glance, we cannot separately estimate the effect of human advice on fixed costs and prior means in a three-period model. However, the weight $(1 - g_0(h))$ that investors place on their prior means declines deterministically if period 1 in the model is interpreted as "longer", i.e., if investors observe a more informative realized return series before making the decision to quit at date 1. As a result, we will obtain an approximate empirical estimate of the size of the fixed cost difference by estimating γ separately for investors who have had a relatively long experience of realized returns from their algorithm portfolio.

The final coefficient δ measures the interaction of a type-1 advisor with negative returns, thus capturing the heterogeneous treatment effect of human advice for investors that have experienced poor performance. This coefficient corresponds to a structural estimate of the differential Kalman gain between investors with type-1 and type-0 advisors. From the discussion following Lemma 1, we can deduce that one would see $\delta < 0$ if and only if type-1 human advice decreases investors' sensitivity to new information or, equivalently, decreases the variance of their prior beliefs about the algorithm's expected returns θ.

Before moving to the data, it is useful to discuss a final issue related to the identification of the discrete choice model we have outlined. In principle, one can estimate the coefficients $\{\alpha, \beta, \gamma, \delta\}$ using a standard econometric discrete choice model (e.g., a linear probability model if the distribution of ξ_1^i

[9] To see this connection, notice that the Kalman filter in Lemma 1, for date $t = 1$, can equivalently be written as a weighted average of the prior mean and the first realized return:

$$m_1^i(h) = m_0(h)(1 - g_0(h)) + r_1^i g_0(h),$$

so that the weight on prior beliefs is indeed equal to $(1 - g_0(h))$.

Electronic copy available at: https://ssrn.com/abstract=4301514

is uniform, or a Logit model if it is a logistic/extreme value distribution). Before taking this approach to the data, we discuss two issues related to reduced-form and structural identification. Equation (18) is only a hypothetical model that one would be able to estimate by observing *both* potential outcomes for each investor i. Clearly, each investor in practice is assigned only one human advisor of type h^i, and we observe only the potential outcome associated with this assignment. In order to obtain unbiased estimates in Equation (18), we must make the standard assumption of *selection on observables,* which states that the assignment of h^i in the data is random, conditional on any investor characteristics that we can control for in estimation:

$$\gamma_t^i(0), \gamma_t^i(1) \perp h^i \big| \text{ controls } x_t^i$$

Below, we discuss in detail why this condition is likely to be satisfied in our empirical setting.

2 Institutional Setting and Data

We use data from a large US-based hybrid robo-advisor called Personal Advisor Services (PAS), a division of Vanguard, to estimate algorithmic aversion. The service is "hybrid" as it complements algorithmic robo-advice with human advisory input. More broadly, robo-advisors are commonly classified as either pure or hybrid robo-advisors. Pure robo-advisors do not feature any substantive interactions between investors and human financial advisors, whereas hybrid robo-advisors allow investors differing levels of interaction with human financial advisors. In our setting, the algorithm manages the investment portfolio, while the human advisor interacts with investors to help them understand what the algorithm does and advises them on additional auxiliary services, such as opening IRA accounts and estate planning.

2.1 Characteristics of Robo-advised Investors

Our data contains information on trades, positions, demographic characteristics, and investor-advisor mapping for previously self-directed investors who signed up for the robo-advisor service. The trade data include all trades placed by the robo-advisor over the period January 2015 through March 2018.

Electronic copy available at: https://ssrn.com/abstract=4301514

The position data contain associated monthly holdings observations, i.e., the data track both investments and trades. The data also track the dates on which investors initiated, enrolled, implemented, and quit the advice service. The demographic characteristics include investor marital status, age, and gender. The investor-advisor mapping data track the dates, times, length, and the initiator (advisor or investor) of all advisor-investor interactions, including meetings and phone calls.

Table 1 reports summary statistics (mean, standard deviation, and percentiles of the distribution) of different variables characterizing robo-advised investors in the sample recorded 12 months after investors signed up for the service. Panel A focuses on investors' demographic characteristics. The average investor is 64 years old, 59% of the users are male, and 72% are married. The robo-advised population comprises older, wealthier investors who are more gender-balanced than datasets commonly used to study trading behavior. In contrast with the demographics in these data, average investor age is 51 in the brokerage trading data employed by both Gargano and Rossi (2017) and Barber and Odean (2001), and women constitute lower fractions in those datasets, with 27% in Gargano and Rossi (2017) and 21% in Barber and Odean (2001).

Table 1 Panel B shows details of investors' portfolio allocations. The mean allocation to the robo-advisor of $758,000 is substantial, and more than 50% larger than the median ($478,890), reflecting significant right-skewness. There are 8 assets on average in each account, comprising individual mutual funds, stocks, bonds, and ETFs, with the bulk of the investors being invested in only six assets and only 25% of the investors having more than 9 assets in their portfolio. This reflects the different "glide paths" to which investors are assigned by the algorithm; we discuss this in greater detail below.

Table 1 Panel B also shows that the average investor has 59% of his/her portfolio invested in equities, 39% in bonds, and 2% in cash (i.e., money market mutual funds). There is cross-sectional dispersion in these fractions, once again reflecting the limited number of portfolio glide paths into which the algorithm allocates investors. At the 25th percentile, investors have 49% of their wealth in equities, while the 75th percentile ranked equity holdings fraction is 69%.

Stocks and bonds are primarily held through mutual funds rather than single stocks, as can be seen from the mean of 95% of wealth invested in mutual funds. Indeed, only 2% of investors have any wealth in individual stocks or ETFs (1% in each). Finally, while this is not reported in the table, only

Electronic copy available at: https://ssrn.com/abstract=4301514

a negligible number of investors have direct exposure to corporate bonds and options.

Mutual fund holdings can be decomposed by fund strategy. The bottom of Panel B shows that 83% of mutual fund holdings are in indexed mutual funds, and 32% of mutual fund holdings are in funds that invest internationally. Only a negligible number of investors invest in mutual funds with standalone emerging markets exposure.

Table 1 Panel C focuses on fees and transactions. The average management fee across all mutual fund holdings is 7 bp. The expense ratios are also low, on average 9 bp with a median of 8 bp. The third row of Panel C shows that the average turnover ratio of the mutual funds in the portfolio is 27%. In terms of active transactions, there are on average 2.6 transactions placed per month, an average of $15,567 dollars, a reflection of periodic portfolio rebalancing by the robo-advisor.

2.2 Robo-advisor Characteristics and Returns of Advised Clients

Over the sample period, the robo-advisor classifies investors into five risk glide paths based on their financial objectives, investment horizons, and demographic characteristics. Although confidentiality prohibits disclosing the details of the algorithm that generates the investors' portfolio allocations, we report facts that help to explain broad features of the allocations and the performance of the robo-advised portfolios.

We compute average annualized monthly portfolio returns for all PAS investors. PAS investors can be categorized into two tiers based on their investment commitment. Higher-tier investors have a dedicated investment advisor, while lower-tier investors receive support from a pool of human advisors. (Section 2.3 describes the (quasi-random) process of assignment of the dedicated human advisors to higher-tier clients, lower-tier investors receive support based on scheduling/availability from the pool). In this analysis, we co-mingle the two groups because their investment performance is virtually identical. For comparison, we also compute the average returns for non-robo-advised clients in Vanguard who interacted with PAS at some point. This includes the returns of clients who later signed up for PAS and those who did not subsequently sign up after considering the service.

We show the cross-sectional distributions of both sets of average returns in Figure 1(a), with the average returns of robo-advised investors in blue and those of self-directed investors in red. The plot

Electronic copy available at: https://ssrn.com/abstract=4301514

shows that robo-advised investors have higher mean returns than non-robo-advised investors, i.e., the blue distribution is shifted to the right relative to the red distribution.[10]

There is considerable cross-sectional dispersion in both distributions and no visible reduction in the variance of the average returns of robo-advised investors, which range from 0% to 20% per year. However, the substantially greater mean creates a clear increase in Sharpe ratios for robo-advised investors. In terms of our theoretical model in the previous section, this increase corresponds to the case where the true expected return θ achieved by the algorithm is larger than the return $\bar{r}$ that investors are able to achieve without advice, which is an important assumption in our structural analysis.[11]

To better understand the sources of variation in the returns of robo-advised investors, we conduct the following three-step procedure. In the first step, following Balasubramaniam et al. (2022), we compute principal components (PCs) of the equity share of 1,700 investors for which we have 36 continuous months of robo-advised returns. The first PC explains 46% of the variation in equity shares, and the first five PCs explain 81% of this variation. These results are the first indication of a pronounced factor structure in the robo-advisor portfolio allocations, arising from the small number of glide-paths to which investors are assigned. In the second step, we regress the time series of the equity share for each of the 55,000 investors in our data on the first five PCs estimated in the first step. Finally, in the third step, we cluster investors' loadings in five groups using a k-nearest neighbors estimation procedure. 82.37% of the investors are classified in the first group, followed by the second and third groups with 13.51% and 3.37% of the investors, respectively. Finally, the last two groups contain less than 1% of the advised investors. All in all these results suggest that, while 5 risk glide paths exist, the majority of investors end up on two glide paths over our sample period.

We complement this simple analysis of equity shares by regressing monthly investor total portfolio returns on the market portfolio, investors' equity share, and the interaction of these two variables. This regression has an R^2 of over 75%, highlighting that the equity-bond allocation decision and the

[10] The red distribution exhibits prominent bunching at zero as some self-directed investors have the entirety of their portfolio in cash.

[11] Our model is technically well-defined even in the case where $\theta < \bar{r}$, and investors may choose to participate in robo-advice if their prior expectations about θ are (wrongly) optimistic. However, this case also implies that, as investors obtain more information about the true θ, 100% of investors will eventually quit robo-advice. For this reason, the case with $\theta > \bar{r}$ is substantially more interesting for the empirical setting, in which a large proportion of clients do not quit.

Electronic copy available at: https://ssrn.com/abstract=4301514

variation in aggregate equity returns together explain the major share of variation in returns seen in the data. We report realized monthly returns against the predicted returns from this regression in a binned scatterplot, Figure 1(b), which shows that the two quantities line up very closely along the 45° line.

Given the robo-advising glide paths, the cross-sectional variation seen in the equity share and returns is likely an outcome of investors' preferences and demographic characteristics which the algorithm translates into different portfolio allocations. While we do not have information regarding investors' risk preferences, we do have investors' age. We, therefore, regress monthly client returns on the market portfolio, investors' age, and their interaction. This regression has an R^2 of over 71%, which is very close to the R^2 obtained using equity shares. The corresponding binned scatterplot of realized vs. fitted returns reported in Figure 1(c) suggests that any risk preference variation over and above that which is correlated with age likely plays a minor role in determining the portfolio allocation implemented by the algorithm.

Figure 1(d) provides a different perspective. The black line shows the distribution of cross-sectionally demeaned average returns on all robo-advised portfolios. The red and blue solid lines show the residuals from the regressions reported above. The substantial reduction in the variance of the distribution of these residuals shows that equity share and age explain a majority of the variation in average portfolio returns, which means that performance is in large part homogenous for investors with similar demographic characteristics.

Figure 1(d) also shows dashed red and dashed blue distributions, which include advisor fixed effects into the regressions of performance on equity share and age, respectively. If advisors were instrumental in affecting the portfolio allocations recommended by the algorithm, we would expect these two distributions to have smaller variance than the corresponding solid distributions of residuals from models that do not include advisor fixed effects. Instead, we find that the dashed blue and red distributions are virtually identical to the respective solid distributions, suggesting that human advisors play virtually no role in affecting investors' portfolio allocation. In the rest of the paper, we use these hybrid interactions to assess the degree to which human advisors can help alleviate algorithm aversion. This identification relies in part on human advisors helping the investor to understand what

Electronic copy available at: https://ssrn.com/abstract=4301514

the machine is doing, rather than directly affecting portfolio strategy. Figure 1(d) highlights that this is an accurate characterization of our setting. This supports our model setup and subsequent parameter estimation.

2.3 Quasi-Random Assignment of New Investors to Advisors

Over the sample period the robo-advisor matched investors with human financial advisors. The assignment and onboarding process began by eliciting investors' characteristics—i.e., age, income, investment horizon, and preferences. Using this information, the robo-component of the robo-advisor (i.e., the algorithm) formulated a financial plan and an associated investment portfolio. The sign-up process then required the investor to schedule a time to meet with a human financial advisor who explained the financial plan and completed onboarding.

The assignment of clients to advisors was quasi-random, based on the match between a prospective client's time schedule and the availability of the advisors in the system. The availability of advisors was driven by workload balancing imperatives rather than any assessment of advisor type (i.e. advisor skills, characteristics, or ability to retain clients).

Once the current client "load" of a given advisor is accounted for, we confirm below that there is no empirical relationship between the historical retention rate of clients assigned to a given advisor and the assignment of new clients to the advisor. We also show that the characteristics of clients are "balanced" across the dimension of the historical retention rates of the advisors to which they are quasi-randomly assigned.

2.4 Measuring the Effect of Human Advisors

We seek to understand how human input affects algorithm aversion in this setting. Measuring human advisor type is challenging for a number of reasons. First, financial advisor type is likely the combination of many different but complementary traits, including subjective assessments of advisor communication and relationship-management skills. Second, the skills needed to be a successful "plain vanilla" financial advisor may be different from those needed as a hybrid robo-financial advisor, so it is difficult to rely on "standard" personality assessments or financial competence scores in this novel

Electronic copy available at: https://ssrn.com/abstract=4301514

setting. Third, an important determinant of customer satisfaction in the financial domain is portfolio performance, which is controlled by the robo-advising algorithm rather than human advisors, so using performance-related measures to measure advisor type—a common strategy in financial economics—is unavailable in this setting.

To overcome these challenges, we rely on revealed preferences, measuring advisor type using advisor-specific average client retention rates. More specifically, in each period and for each advisor, we compute the percentage of all investors that remain with robo-advice. While there are clearly many different factors determining these retention rates, we assume that at least some component of the cross-advisor differences in these retention rates is correlated with advisor type.

Figure 3(a) reports the percentage of clients retained by each advisor, scaled in such a way that the advisors with the highest retention are assigned a value of 100. The figure shows that the cross-advisor dispersion in client retention is substantial. A possible concern with this figure is that the large dispersion we document may be due to outliers caused by small base effects—i.e., advisors that left the firm soon after they were hired, so they never got to manage a large portfolio of investors. In Figure 3(b), we remove from the analysis advisors in the lowest 15% of the distribution in terms of clients under management throughout the sample period and still find relatively large dispersion in retention rates across advisors. Finally, in Figures 3 (c) and (d), we show in the full sample of advisors, as well as in the sample that excludes the advisors in the bottom 15% of the distribution in terms of clients under management, that the positive relation between retention rates and total clients advised is restricted to the domain of advisors with a relatively small number of clients.

We use advisor-specific client retention to measure advisors' type in what follows, without taking a stance on the specific advisor skills that increase retention in our empirical work. In the model, we map type-1 advisors to the three dimensions of algorithm aversion that we defined in the previous section: reductions in continuing client *disutility* arising from algorithm aversion, or belief-based reductions in either *pessimism* or *uncertainty* about the algorithm's ability to manage assets.

The institutional details outlined above indicate that the assignment of new clients to advisors during the sample period was random with respect to advisor type. Assignment appeared to depend on advisor capacity and availability but not on advisors' skills or characteristics. In Online Appendix A.2,

Electronic copy available at: https://ssrn.com/abstract=4301514

we validate the quasi-random assignment of clients to advisors. We first show that the characteristics of clients are "balanced" across the dimension of the retention rates of the advisors to which they are assigned. We then show, using a variety of alternative specifications, that the number of clients under management predicts client assignment to advisors while past advisor retention does not.

3 Empirical Estimates

Having established that client assignment to advisors is random conditional on the number of clients under management, we now proceed to estimate the causal effect of the type of human advice (as measured by historical client retention) on retaining investors in robo-advice.

3.1 Non-Parametric Survival Estimates

We begin our analysis by constructing a real-time measure of advisor retention, computed recursively only using past data in each time period. The retention of advisor i at time t is calculated as the ratio between the clients who remain with robo-advice and the total number of clients assigned to advisor i up to time t. We then lag this quantity to ensure that there is no mechanical correlation between measured historical retention at time t and the probability of clients remaining with robo-advice in period t.

Next, we categorize advisors as type-1 (high retention) or type-0 (low retention) in each period using the period-specific cross-sectional median retention rate as the breakpoint. Finally, we obtain nonparametric estimates of the survival function for clients assigned to type-1 and type-0 advisors using the Kaplan-Meier estimator. We estimate:

$$\widehat{S}(t) = \Pi_{i:t_i \leq t} \left(1 - \frac{d_i}{n_i}\right), \tag{23}$$

where t_i is a time when at least one investor quits robo-advice, d_i is the number of clients quitting robo-advice at time t, and n_i is the number of individuals that stayed with robo-advice (i.e., neither quit nor censored) up to time t_i.

We report the results of these estimates graphically in Figure 4. Figure 4(a) reports these results

Electronic copy available at: https://ssrn.com/abstract=4301514

computed using all advisors in the data. As a crude way to control for the non-random assignment associated with different numbers of clients under management, we repeat our estimates excluding from the sample advisors in the bottom 5%, 20%, and 40% of the distribution of clients under management in Figures 4 (b) through (d). In all figures, the blue lines (and associated 95% confidence intervals) show survival rates for clients that are assigned to type-1 advisors, and the red lines show survival rates for investors assigned to type-0 advisors.

The figures show that investors assigned to type-1 advisors exhibit significantly higher future retention than those assigned to type-0 advisors. In Figure 4 (a), 94% of investors in the type-1 advisor group and 90% of investors in the type-0 advisor group stay with robo-advising at the one-year mark. Indeed, these survival estimates diverge for the two groups over time. At the three-year mark, 91% of the investors assigned to type-1 advisors are still with robo-advice. The corresponding value for type-0 advisors is 86%.

Reassuringly, Figures 4 (b) through (d) show very similar patterns. In all cases where we restrict the number of clients managed by each advisor, we find a significant difference in retention between type-1 and type-0 advisor groups, confirming that our measure of human advisor type plays an important role in explaining future retention into robo-advising.

Investors with high wealth levels, specifically those above $1M, are assigned to advisors with greater seniority and experience, which might act as an additional confound. Figure 4 (e) re-estimates our results using only investors with wealth above $1M, and the estimates are reassuringly similar estimates to the baseline.

A possibility is that the recursive construction of historical advisor retention introduces a mechanical relation between this proxy for advisor type and future measured retention at the advisor level. To check robustness, we conduct a placebo simulation exercise. In the original dataset, we replace the true attrition dates of each investor with a random date drawn from a uniform distribution, effectively breaking any potential relation between advisor characteristics and investor retention. Using these placebo dates, we recompute the recursive measure of retention at the advisor level and recalculate Kaplan-Meier survival estimates for the investors assigned to low and high retention advisors. We report these placebo estimates in Figure 4 (f), which shows that simulated advisor retention does not

Electronic copy available at: https://ssrn.com/abstract=4301514

predict future client retention, reassuring us that there is no mechanical relation at work. While we do not report this here, this finding is also supported when we estimate advisor retention in the first half of the sample and predict retention in the second half of the sample.

3.2 Cox Proportional Hazard Estimates

To estimate the economic magnitude of the effect of human advisor type on investor retention, we estimate a Cox proportional hazard model. To satisfy the assumptions of the Cox proportional hazard model, we first estimate the smoothed hazards to see if they are parallel. We report the results for all advisors as well as excluding advisors in the bottom 5%, 20% and 40% of the distribution in terms of clients under management in Figures 5 (a) through (d).

In all cases, we find very similar findings. Starting from Figure 5 (a), the proportional hazard is higher for investors assigned to type-0 advisors. For this group of investors, the hazard estimates (red line and confidence interval) increase and peak at 6.5% at the one-year mark only to decline and reach their bottom of 3.2% at the three-year mark. The estimates then become more imprecise as they are computed on fewer investors. The blue line (and confidence interval) shows hazard estimates for those assigned to type-1 advisors. These follow similar dynamics at lower levels: the hazard estimates peak at 3.5% at the one-year mark, then they decline and bottom out at 2.5% in year three. The estimates subsequently become more imprecise, as evidenced by the larger confidence intervals. The patterns in Figures 5 (b) through (d) are similar and broadly consistent for advisors with different numbers of clients.

We then estimate a Cox proportional-hazard model of the following form on the full set of advisors:

$$h(t|x_j) = h_0(t) \cdot exp(x_j\beta),$$

where we estimate the model using a "type-1" dummy as the only regressor. The coefficient estimate $\beta = -0.455$, implying that the ratio of the hazard between those clients assigned to type-0 and type-1 advisors is $exp(-0.455) = 0.63$. In other words, those assigned to type-1 human advisors have a 37% lower hazard than those assigned to type-0 human advisors.

We also use advisor retention as a continuous variable in a robustness exercise. We take "100 minus

Electronic copy available at: https://ssrn.com/abstract=4301514

the percentage of past clients quitting" as the regressor of interest and re-estimate the Cox model. Using this specification, we obtain a highly statistically significant coefficient $\beta = -0.02$. Economically, the estimate implies that the hazard ratio between those with a higher-retention advisor (a 1% higher retention rate) and a lower-retention advisor is $exp(-0.02) = 0.98$. The Cox model estimates are very similar when we exclude advisors in the bottom 5%, 20%, and 40% of the distribution in terms of clients under management. These estimates confirm the causal impact of human advisor type on investor participation in robo-advising, and we can link these estimates back to our model of algorithm aversion.

3.3 Estimating the Effect of Human Advice Across Market Conditions

Our results thus far are not conditional on market conditions. As we illustrated using the theoretical model, the effect of assignment to a type-1 or type-0 advisor is particularly significant in periods of market distress, especially in the case when investors are learning about the ability of the algorithm to deliver value. To test for this mechanism, we estimate regressions of the following form at the monthly frequency:

$$\begin{aligned} Dummy_quit_{i,t} = \alpha + \beta\ I_{\{MKT_RET_{t-1}<0\}} + \gamma\ I_{\{Type1_Advisor_{i,t-1}=1\}} \\ + \delta\ I_{\{MKT_RET_{t-1}<0\}} \times I_{\{Type1_Advisor_{i,t-1}=1\}} + \epsilon_{i,t}, \end{aligned} \tag{24}$$

where $Dummy_quit_{i,t}$ is equal to 1 if investor i quits robo-advising in month t and 0 otherwise, $I_{\{MKT_RET_{t-1}<0\}}$ is an indicator variable equal to 1 if the market return is negative in month $t-1$ and 0 otherwise, and $I_{\{Type1_Advisor_{i,t-1}=1\}}$ is equal to 1 if investor i is assigned to an advisor with past retention above the median and 0 otherwise.

The α coefficient measures investors' unconditional monthly quit probability. The β coefficient estimates the conditional increase in attrition from robo-advice in periods in which the market has performed poorly, the γ coefficient measures the differential quit rate associated with clients assigned to type-1 advisors, and finally, the δ parameter measures the extent to which the differential quit rate across clients assigned to type-1 and type-0 advisors varies with market conditions.

The results of estimating this equation are reported in Table 2, where we multiply all the coefficient

Electronic copy available at: https://ssrn.com/abstract=4301514

estimates by 100 to express magnitudes in percentage points. Starting from the first column, which uses the returns on the CRSP value-weighted index as a proxy for market returns, the α coefficient reveals an unconditional 0.388% quit probability per month. This roughly translates to an annual attrition rate of $0.388\% \times 12$=4.656%, which is in line with the attrition rate that we report in Figure 4. The β coefficient shows that attrition from robo-advising increases by 0.125 percentage points in poor market conditions. This increase is large, translating to a 0.125/0.388=32% increase in attrition in such times. The γ coefficient shows that being assigned to type-1 advisors reduces attrition by 0.98% per month, confirming the results from our survival analysis. Finally, the coefficient of interest δ is negative and significant, showing that type-1 advisors reduce client attrition the most when markets are down.

In Figure 6 we use the regression estimates from Equation (24) to compute attrition across different market conditions and advisor characteristics. Investors assigned to type-1 advisors have an attrition probability of 0.3% per month, irrespective of market conditions. Investors assigned to type-0 advisors, on the other hand, have a probability of attrition of 0.39% in good market conditions and a probability of attrition of 0.51% in bad market conditions. These results suggest that type-1 advisors reduce attrition both conditionally and unconditionally. Once again, type-1 advisors are particularly important in bad market conditions.

The results in the first column of Table 2 use the returns on the CRSP value-weighted index as a proxy for market conditions. The rationale is that investors are likely to use returns on the overall market to assess current economic conditions. To make sure our results are not driven by this choice of market proxy, in the second column we use the returns on the portfolio of each individual investor in the previous month as a proxy for current economic conditions. Overall, the results are very similar and, if anything, the statistical significance of the δ coefficient is higher in the second column compared to the first.

The results reported in Table 2 impose the threshold for "bad" market returns at zero. However, the worse market conditions are, the more valuable human financial advisors are likely to be. To test whether this is the case, we re-estimate Equation 24 by allowing the threshold for the market return to vary from -5% to 4%. We then report graphically the estimates of the parameter of interest δ. Figure

Electronic copy available at: https://ssrn.com/abstract=4301514

6(c) shows that when market returns are extremely low, the δ coefficient is even more significant and negative, consistent with type-1 advisors being even more important in reducing attrition when market distress is intense.

3.3.1 Using Past Volatility as Conditioning Information

Another natural measure of market uncertainty and distress is volatility. To test whether advisor type is differentially important in periods of high volatility, we re-estimate Equation 24 after replacing periods of negative market returns with periods of high volatility, where the latter are defined as periods where the volatility exceeds median volatility computed over the full sample. The results from this exercise are reported in the last two columns of Table 2. In column (3) of this table, monthly volatility is computed using daily returns on the CRSP value-weighted index. In column (4) investor volatility is computed using the realized daily returns of each individual client portfolio.

Overall, the results are both economically and statistically in line with the ones in the first two columns of Table 2. In all cases, we find that type-1 advisors reduce investors' robo-advising attrition both unconditionally as well as in periods of market distress.

3.4 Structural Estimates of Algorithm Aversion

Our reduced-form estimates in the previous section are an empirical analog of the discrete choice model we characterized in Proposition 2 of our theoretical analysis. The proposition provides a useful mapping from the reduced-form parameters $\{\alpha, \beta, \gamma, \delta\}$ in our attrition regression (24) to the deep parameters of the model. We use this mapping to extract structural estimates of the different dimensions of algorithm aversion in our theory. Before we begin this exercise, we validate an auxiliary prediction of the model, namely, that investors learn about the quality of the robo-advisor over time.

3.4.1 Experience and Learning

One of the key predictions of the model is that, as users become more familiar with the robo-advisor, they are less sensitive to spells of poor performance. As a test of the model, we take the universe of investors and split them into low and high experience groups, where experience is computed as the

Electronic copy available at: https://ssrn.com/abstract=4301514

number of months investors have been enrolled in robo-advice. We then estimate Equation (24) across the two groups and report the results in Table 3.

The first column of the table shows that short tenure/low experience investors are very sensitive to investment performance. Spells of bad market performance increase these investors' attrition probability by 0.196 percentage points. For low-experience investors, advisor type has an important effect both unconditionally and conditionally: the γ and δ coefficients are both negative and significant.

The second column of Table 3 shows that long-tenure/high-experience investors behave very differently. They are not more likely to quit robo-advising during periods of poor market performance, and advisor type does not reduce attrition in periods of poor market performance, as the δ coefficient is negative but statistically insignificant. That said, even for these investors, unconditionally, advisor type does have an important effect in reducing attrition, as the γ coefficient is negative and significant.

3.4.2 Uncertainty Dimension of Algorithm Aversion

Proposition 2 provides a very clear mapping from our reduced-form estimates to investors' perceived uncertainty about the returns generated by the algorithm. On one hand, the coefficient β on poor performance is equal to the Kalman gain, or sensitivity to new information, among investors with a type-0 human advisor. On the other hand, the coefficient δ, or the interaction between bad returns and the type-1 advisor indicator, measures the difference in Kalman gains between clients with type-1 and type-0 human advice.

Moreover, we can recall from Lemma 1 that investors' Kalman gains are an increasing function of the prior variance $v_0(h)$ that is induced by a human advisor with type $h \in \{0, 1\}$. In the appendix, we use this insight to derive the following formula:

$$\frac{v_0(1) - v_0(0)}{v_0(0)} = \frac{\frac{\beta+\delta}{1-\beta-\delta}}{\frac{\beta}{1-\beta}} - 1 \tag{25}$$

The left-hand side of Equation (25) is the percentage change in variance brought about by switching from a type-0 to a type-1 human advisor in the model. The right-hand side expresses this quantity in terms of our empirical estimates. To interpret this expression, notice that the Kalman gain of an

Electronic copy available at: https://ssrn.com/abstract=4301514

investor with a type-0 advisor is β, and we can show that the variance perceived by this investor is proportional to $\beta/(1-\beta)$. Similarly, the Kalman gain of an investor with a type-1 advisor is $\beta+\delta$, and the perceived variance is proportional to $(\beta+\delta)/(1-\beta-\delta)$. Evaluating the percentage difference between these two scenarios yields Equation (25).

Evaluating this expression using the coefficient estimates from our baseline specification, in Table 2 column (1), we obtain the following estimate:

$$\frac{v_0(1)-v_0(0)}{v_0(0)} \simeq -0.9577 \tag{26}$$

This estimate suggests that a type-1 human advisor reduces investors' perceived uncertainty about the algorithm's performance by about 96%. Therefore, our empirical estimates are consistent with a model where effective human advice removes the vast majority of the uncertainty associated with automated portfolio choice. This result gives a strong indication that the *uncertainty* component of algorithm aversion is an important one. Moreover, our structural estimates suggest that the logic of the "money doctors" view of investment advice in Gennaioli, Shleifer, and Vishny (2015), in which human advice serves mainly to induce reductions in perceived risk, may extend to the modern business model of hybrid robo-advice.

3.4.3 Pessimism and Disutility Dimensions of Algorithm Aversion

Proposition 2 further provides an interpretation for the γ coefficient in our regression Equation (24). The coefficient corresponds to the unconditional reduction, i.e. the reduction independent of performance, in clients' propensity to quit robo-advice and measures the combination of two structural forces. First, γ contains the change in the fixed disutility of using robo-advice achieved by a type-1 human advisor (the second term in Equation (21)). Second, γ contains the corresponding change in the prior mean beliefs about algorithm performance (the first term in Equation (21). To get a sense of the economic magnitude of these two forces, one can compare our estimates of γ to the baseline propensity to quit, which is captured by the intercept coefficient α. According to our estimates in Table 2, column (1), we find that a type-1 human advisor changes the propensity to quit, in percentage

Electronic copy available at: https://ssrn.com/abstract=4301514

terms, by about $\gamma/\alpha \simeq -25\%$. This result implies that human advice removes about a quarter of the joint effect of the pessimism and disutility dimensions of algorithm aversion.

However, in order to disentangle these pessimism and disutility dimensions, we require further empirical moments. In our auxiliary test of the model's implications above, we argued that the expectations of experienced investors appear to have converged, in the sense that they are no longer sensitive to the new information contained in realized returns. In terms of the Kalman filtering equations in our model, convergence corresponds to a situation where experienced investors, who have had access to a long series of evidence, no longer place any weight on their prior beliefs when forming expectations about performance. This is useful because, in Equation (21), which maps the γ coefficient to structural quantities, the effect of prior pessimism is multiplied by the factor $1 - g_0(h)$, for $h \in \{0, 1\}$, that investors place on their prior beliefs. A situation in which beliefs have converged corresponds to $g_0(h) \simeq 1$, so that these weights are close to zero. Combining these insights, we can argue that an estimate of γ among experienced investors will be a good approximation of the *first* term in Equation (21), which measures the structural difference in disutility between investors with type-1 and type-0 human advisors. Taking our estimates for experienced investors from Table 3, we find that $\gamma/\alpha \simeq -23\%$, which is quite close to the quantity we found in the sample of all investors. This finding suggests that most of the effect captured by γ is driven by the disutility dimension of algorithm aversion.

4 Conclusion

We study algorithm aversion using a model of robo-advising adoption. In the model, investors experience a per-period disutility of dealing with an algorithm, as well as a "learning" channel in which they refine their understanding of the performance of the robo-advisor over time and across states of the economy. We pair this structural model with empirical work on unique data from a large US hybrid robo-advisor.

In the robo-advisor we study, as with other such hybrid services, the algorithm manages the investment portfolio, while the human advisor interacts with investors to help them understand what the algorithm does and provides auxiliary advice on issues such as estate planning. A key feature of

Electronic copy available at: https://ssrn.com/abstract=4301514

this setting is that the assignment of investors to advisors follows mechanical rules driven by workload balancing imperatives rather than any assessment of advisor type. This means that once the current "load" of a given advisor is accounted for, the assignment of new clients to this advisor is orthogonal to the historical client retention of the advisor (a useful proxy for advisor type).

This quasi-random assignment of clients to different types of advisors allows us to cleanly map our empirical estimates to the parameters of the model. We find that this measure of advisor type predicts the future retention rate of clients that are assigned to them, which can be mapped to both the learning channel and the ongoing disutility channel in the model. We also find that high-retention advisors' clients are less likely to quit robo-advising during periods of market turbulence than lower-retention advisors' clients, which more cleanly maps to the learning channel. Finally, we find that experienced clients, regardless of their advisor type, react less to market turbulence—consistent with them having less to learn about the algorithm's ability to deliver returns. The attrition rates of such experienced clients also deliver a cleaner measure of the effects of ongoing disutility from dealing with the algorithm.

Our results make the notion of algorithmic aversion more transparent, separating this puzzling phenomenon into both learning and disutility components. In future drafts, we intend to rigorously structurally estimate the parameters of the model and use these parameters to evaluate counterfactuals to make human-algorithm matching more effective in the presence of algorithm aversion.

Electronic copy available at: https://ssrn.com/abstract=4301514

References

Allcott, H., M. Gentzkow, and L. Song. 2021. Digital addiction. Working Paper, National Bureau of Economic Research.

Balasubramaniam, V., J. Y. Campbell, T. Ramadorai, and B. Ranish. 2022. Who owns what? a factor model for direct stock holding. *Journal of Finance* .

Barber, B. M., and T. Odean. 2001. Boys will be boys: Gender, overconfidence, and common stock investment. *Quarterly Journal of Economics* 116:261–92.

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

Brennan, M. J., and Y. Xia. 2001. Stock price volatility and equity premium. *Journal of monetary Economics* 47:249–83.

Burton, J. W., M.-K. Stein, and T. B. Jensen. 2020. A systematic review of algorithm aversion in augmented decision making. *Journal of Behavioral Decision Making* 33:220–39.

Campbell, J. Y. 2006. Household finance. *The Journal of Finance* 61:1553–604.

Campbell, J. Y., and L. M. Viceira. 1999. Consumption and portfolio decisions when expected returns are time varying. *The Quarterly Journal of Economics* 114:433–95.

Campbell, J. Y., L. M. Viceira, L. M. Viceira, et al. 2002. *Strategic asset allocation: portfolio choice for long-term investors*. Clarendon Lectures in Economic.

Castelo, N., M. W. Bos, and D. R. Lehmann. 2019. Task-dependent algorithm aversion. *Journal of Marketing Research* 56:809–25.

D'Acunto, F., N. Prabhala, and A. Rossi. 2019. The promises and pitfalls of robo-advising. *Review of Financial Studies* .

D'Acunto, F., and A. G. Rossi. 2022. Robo-advising. *Palgrave Macmillan Handbook of Technological Finance* .

Electronic copy available at: https://ssrn.com/abstract=4301514

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

Egan, M., G. Matvos, and A. Seru. 2019. The market for financial adviser misconduct. *Journal of Political Economy* 127:233–95.

Fuster, A., P. Goldsmith-Pinkham, T. Ramadorai, and A. Walther. 2022. Predictably unequal? the effects of machine learning on credit markets. *The Journal of Finance* 77:5–47.

Gargano, A., and A. G. Rossi. 2017. Does it pay to pay attention? *Working Paper* .

Gennaioli, N., A. Shleifer, and R. Vishny. 2015. Money doctors. *The Journal of Finance* 70:91–114.

Gomes, F., M. Haliassos, and T. Ramadorai. 2021. Household finance. *Journal of Economic Literature* 59:919–1000.

Guiso, L., P. Sapienza, and L. Zingales. 2008. Trusting the stock market. *the Journal of Finance* 63:2557–600.

Guiso, L., and P. Sodini. 2013. Household finance: An emerging field. In *Handbook of the Economics of Finance*, vol. 2, 1397–532. Elsevier.

Linnainmaa, J. T., B. Melzer, A. Previtero, and S. Foerster. 2018. Financial advisors and risk-taking. *Working Paper* .

Merton, R. C. 1969. Lifetime portfolio selection under uncertainty: The continuous-time case. *The review of Economics and Statistics* 247–57.

Niszczota, P., and D. Kaszás. 2020. Robo-investment aversion. *Plos one* 15:e0239277–.

Pagliaro, C., and S. Utkus. 2019. Assessing the value of advice. *Vanguard Research* .

Pástor, L., and P. Veronesi. 2003. Stock valuation and learning about profitability. *The Journal of Finance* 58:1749–89.

Reher, M., and C. Sun. 2019. Automated financial management: Diversification and account size flexibility. *Journal of Investment Management* .

Electronic copy available at: https://ssrn.com/abstract=4301514

Rossi, A. G., and S. Utkus. 2019. The needs and wants in financial advice: Human versus robo-advising. *Working Paper* .

Rossi, A. G., and S. P. Utkus. 2020. Who benefits from robo-advising? evidence from machine learning. *Evidence from Machine Learning (March 10, 2020)* .

Timmermann, A. G. 1993. How learning in financial markets generates excess volatility and predictability in stock prices. *The Quarterly Journal of Economics* 108:1135–45.

Veronesi, P. 2000. How does information quality affect stock returns? *The Journal of Finance* 55:807–37.

Electronic copy available at: https://ssrn.com/abstract=4301514

Electronic copy available at: https://ssrn.com/abstract=4301514

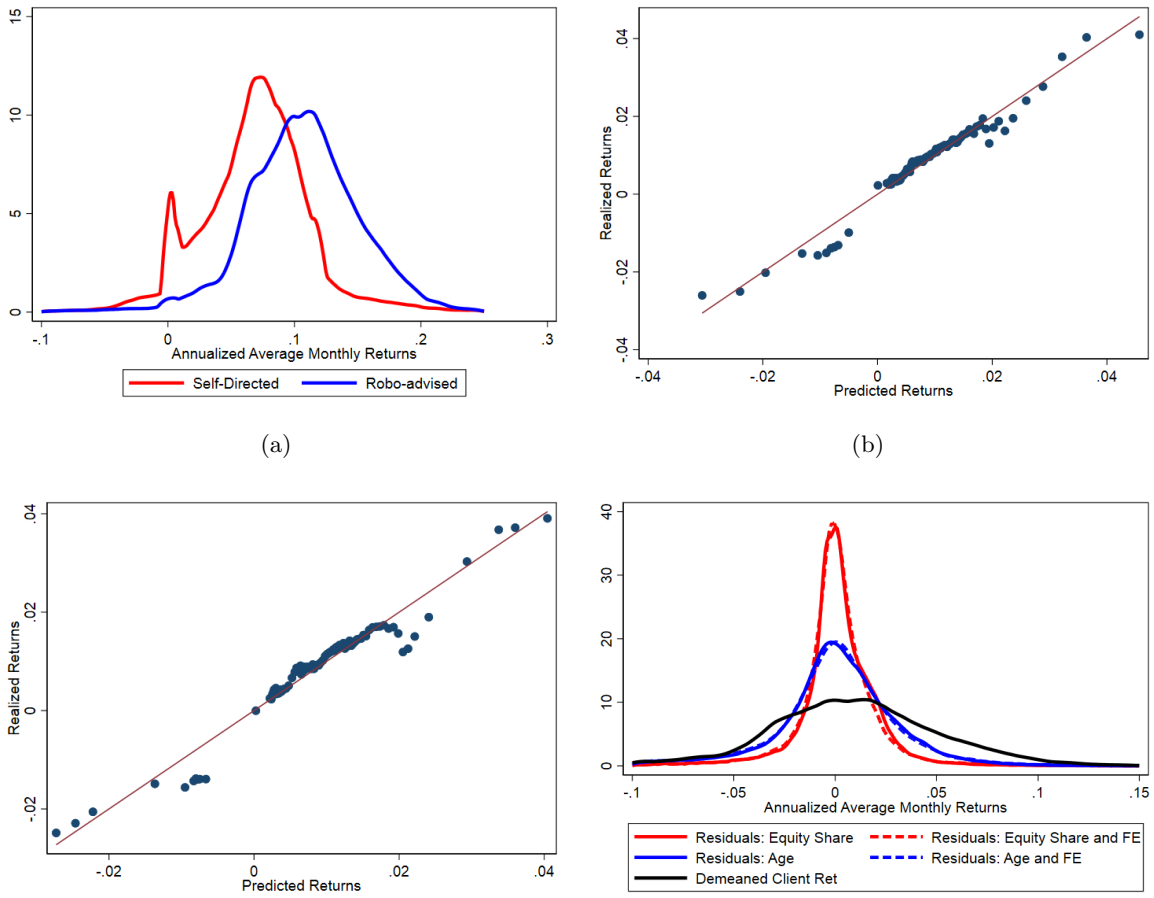


Figure 1: This figure reports results on investors' portfolio performance. We compute the cross-section of investment performance across robo-advised and non-robo-advised investors in Subfigure (a). Subfigures (b) and (c) relate monthly client returns to predicted returns on the basis of investors' equity share and age, respectively. Subfigure (d) shows the extent to which investor returns are determined by their equity share, demographic characteristics, and their assigned advisor.

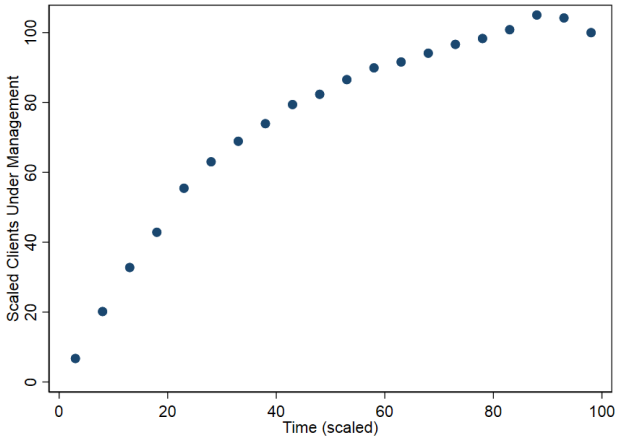


Figure 2: This figure reports reports the scaled median number of clients per advisor throughout the advisors' lifecycle.

Electronic copy available at: https://ssrn.com/abstract=4301514

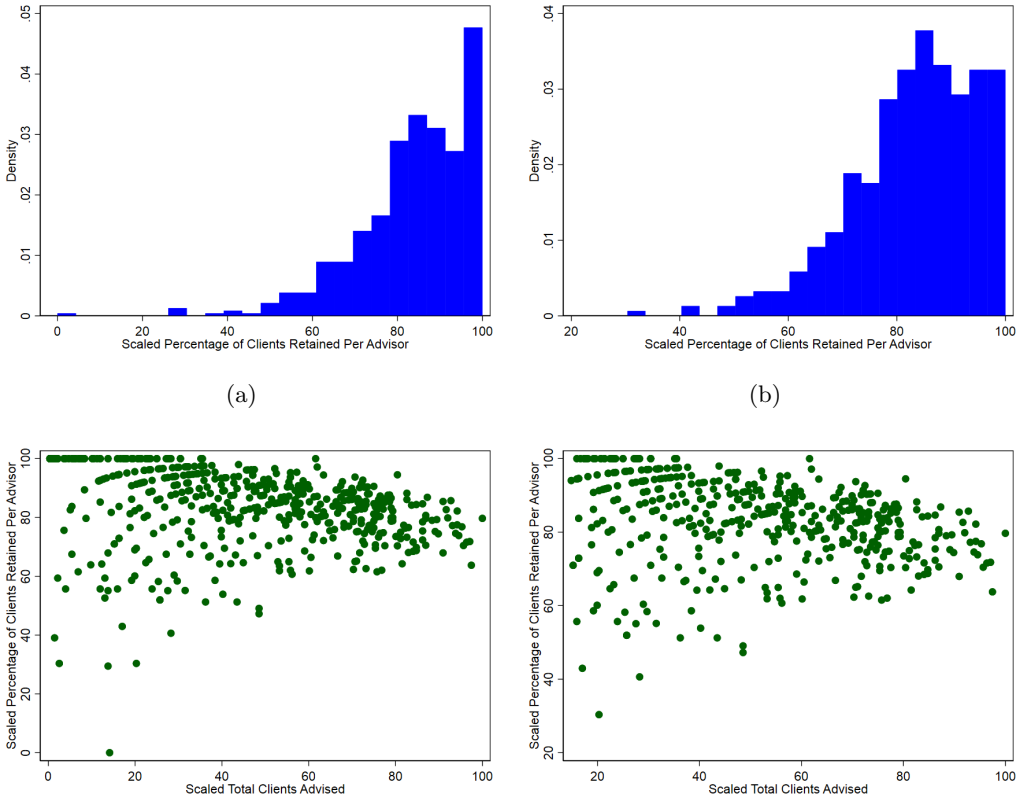
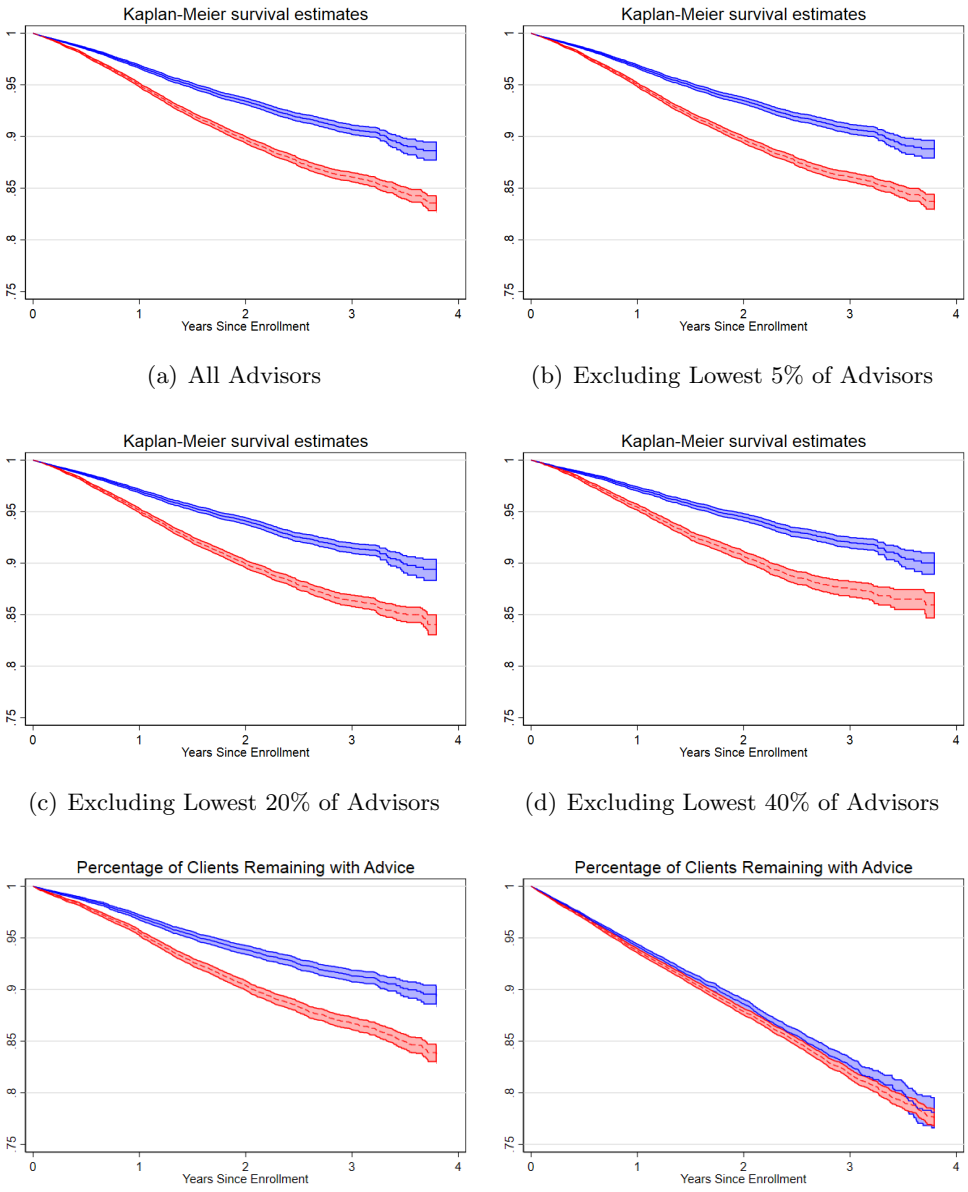


Figure 3: This figure reports summary statistics of investor retention in robo-advice. Subfigure (a) and (b) report a histogram of the scaled cross-sectional variation in client retention across advisors, with the first focusing on all advisors and the second removing from the analysis those advisors in the bottom 15% of the distribution in terms of clients under management throughout the sample. In subfigures (c) and (d) we show the relationship between the scaled number of investors advised by each advisor and the percentage of clients retained per advisor, with the first plot focusing on all advisors and the second removing from the analysis those advisors in the bottom 15% of the distribution in terms of clients under management throughout the sample.

Electronic copy available at: https://ssrn.com/abstract=4301514



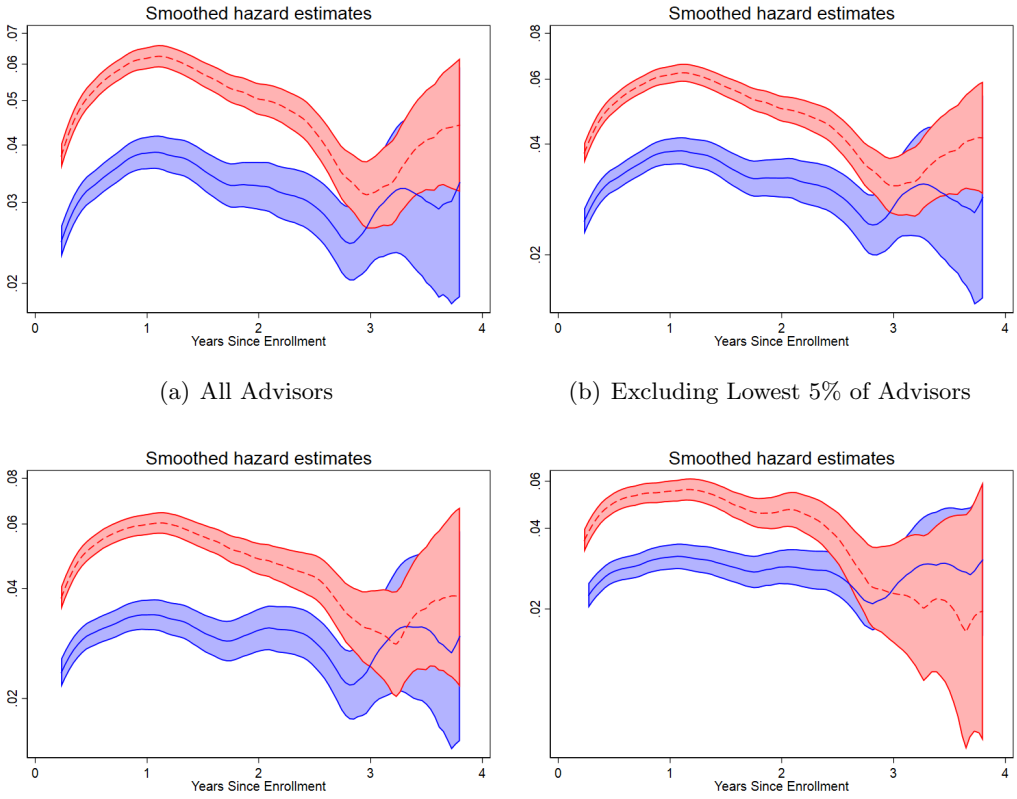
(a) All Advisors (b) Excluding Lowest 5% of Advisors

(c) Excluding Lowest 20% of Advisors (d) Excluding Lowest 40% of Advisors

(e) Clients with more than $1M in assets (f) Simulation Under Null of No Advisor Skills

Figure 4: This figure shows survival plots for investors assigned to advisors of different types. Investors associated with type-1 advisors are in blue. Those with type-0 advisors are in red. The results are computed across all advisors in subfigure (a); excluding advisors in the lowest 5%, 20%, and 40% of the distribution in terms of clients under management in subfigures (b), (c), and (d), respectively; and including only advisors who cater to clients with more than $1M in assets in subfigure (e). In subfigure (f) we report results for a simulation exercise that imposes the null of no advisor skills.

Electronic copy available at: https://ssrn.com/abstract=4301514



(a) All Advisors

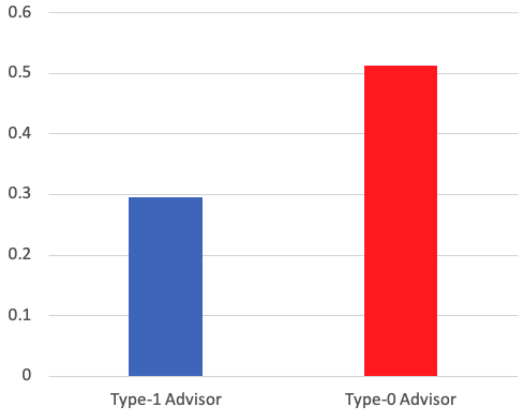
(b) Excluding Lowest 5% of Advisors

(c) Excluding Lowest 20% of Advisors

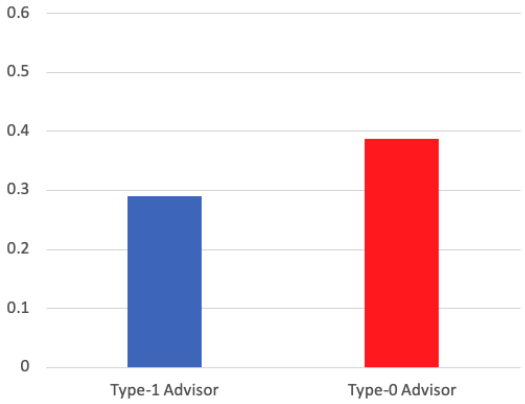
(d) Excluding Lowest 40% of Advisors

Figure 5: This figure shows smooth hazards for investors assigned to advisors of different types. Investors associated with type-1 advisors are in blue. Those with type-0 advisors are in red. The results are computed across all advisors in subfigure (a), and excluding advisors in the lowest 5%, 20%, and 40% of the distribution in terms of clients under management in subfigures (b), (c), and (d), respectively.

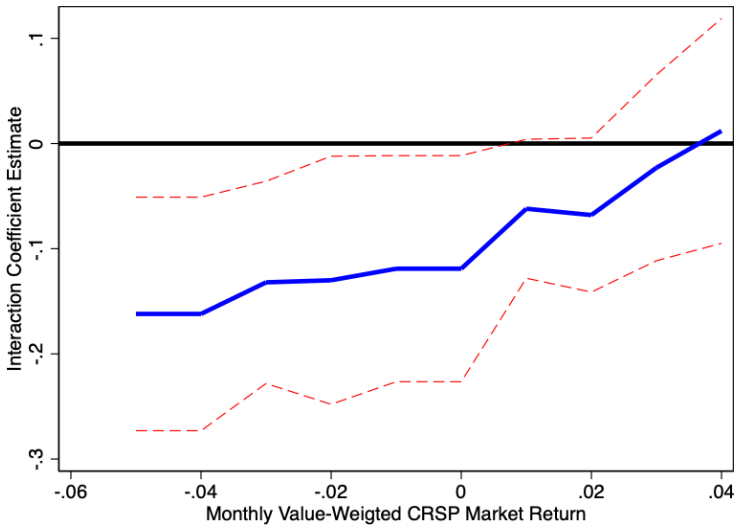
Electronic copy available at: https://ssrn.com/abstract=4301514



(a) Poor Market Conditions



(b) Good Market Conditions



(c) Results Varying the Definition of Market Distress

Figure 6: This figure reports the monthly attrition (in percentages) of investors assigned to advisors of different types, conditioning on different market conditions in subfigures (a) and (b). In subfigure (c), we re-estimate Equation 24 by allowing the threshold for the market return to vary from -5% to 4%. We then report the estimates of the parameter of interest, δ, and the associated 95% confidence intervals.

Electronic copy available at: https://ssrn.com/abstract=4301514

Table 1. Demographic and Portfolio Characteristics of Advised Investors

Panel A. Demographic Characteristics

	N	mean	sd	p25	p50	p75
Age	54,328	63.81	12.39	57 .00	65.00	72.00
Male	54,747	0.59	0.49	0.00	1.00	1.00
Married	54,747	0.72	0.45	0.00	1.00	1.00

Panel B. Portfolio Allocation

	N	mean	sd	p25	p50	p75
Wealth	54,747	758,340	821,023	210,784	478,890	981,327
NumAssets	54,747	7.95	4.91	5.00	6.00	9.00
PctEquityShare	54,742	0.59	0.18	0.49	0.57	0.69
PctBondShare	54,742	0.39	0.18	0.28	0.41	0.50
PctCashShare	54,742	0.02	0.05	0.00	0.00	0.01
PctMutualFunds	54,744	0.95	0.10	0.96	1.00	1.00
PctCash	54,744	0.02	0.05	0.00	0.00	0.01
PctStocks	54,744	0.01	0.04	0.00	0.00	0.00
PctETF	54,744	0.01	0.03	0.00	0.00	0.00
PctIndex	54,747	0.83	0.18	0.75	0.86	1.00
PctInternational	54,723	0.32	0.09	0.28	0.34	0.36
PctEmerging	54,723	0.00	0.00	0.00	0.00	0.00

Panel C. Transactions and Fees

	N	mean	sd	p25	p50	p75
MgtFee	54,717	0.07	0.02	0.06	0.06	0.07
ExpRatio(×100)	54,707	0.09	0.03	0.08	0.08	0.10
TurnRatio	54,685	0.27	0.12	0.19	0.28	0.34
Transaction	54,747	2.63	4.09	0.00	1.00	3.00
Volume ($)	54,747	15,567	50,131	0.00	244	4,000

This table reports demographic characteristics and portfolio allocation behavior of investors 12 months after signing up for advice. The results are computed at the investor level and include all account types, that is, taxable and non-taxable (IRA) accounts. Panel A reports demographic characteristics, panel B focuses on portfolio characteristics, and panel C focuses on transactions and fees paid. For each variable, we report the number of accounts used in the computations, the mean, the standard deviation, and the 25^{th}, 50^{th}, and 75^{th} percentiles of the distribution.

Electronic copy available at: https://ssrn.com/abstract=4301514

Table 2. The Effect of Human Advice Across Market Conditions

	CRSP Return	Investor Return	CRSP Volatility	Investor Volatility
Bad Market	0.125**	0.132***	0.115***	0.103***
	(2.59)	(3.35)	(2.91)	(3.26)
Type1_Advisor	-0.098***	-0.092***	-0.069***	-0.080***
	(-5.65)	(-4.94)	(-3.24)	(-3.96)
Interaction	-0.119**	-0.109**	-0.086**	-0.061**
	(-2.17)	(-2.35)	(-2.68)	(-2.11)
Constant	0.388***	0.380***	0.349***	0.353***
	(18.99)	(17.43)	(17.54)	(16.36)
Clustering	Date&User	Date&User	Date&User	Date&User
R-square	0.00011	0.00013	0.00014	0.00013
N	830,268	830,268	830,268	830,268

This table reports coefficient estimates of the following baseline regression estimated at the monthly frequency:

$$Dummy_quit_{i,t} = \alpha + \beta\ I_{\{BadMarket_{t-1}=1\}} + \gamma\ I_{\{Type1_Advisor_{i,t-1}=1\}} + \delta\ I_{\{BadMarket_{t-1}=1\}} \times I_{\{Type1_Advisor_{i,t-1}=1\}} + \epsilon_{i,t},$$

where $Dummy_quit_{i,t}$ is equal to 1 if investor i quits robo-advising in month t and 0 otherwise, $I_{\{BadMarket_{t-1}<0\}}$ is an indicator variable equal to 1 if market conditions are bad in month $t-1$ and 0 otherwise, and $I_{\{Type1_Advisor_{i,t-1}=1\}}$ is equal to 1 if investor i is assigned to an advisor with a past retention above the median and 0 otherwise. We multiply all the coefficient estimates by 100 so they are expressed in percentages. The standard errors are clustered by user and date. We use CRSP return, Investor Return, CRSP Volatility and Investor Volatility as proxies for market conditions in columns (1) through (4).

Electronic copy available at: https://ssrn.com/abstract=4301514

Table 3. The Effect of Human Advice Across Market Conditions
Long and Short Tenure Investors

	Short Tenure	Long Tenure
Bad Market	0.196***	0.018
	(3.29)	(0.33)
Type1_Advisor	-0.098***	-0.092***
	(-5.08)	(-3.71)
Interaction	-0.120**	-0.103
	(-2.03)	(-1.59)
Constant	0.337***	0.437***
	(17.28)	(15.93)
Clustering	Date&User	Date&User
R-square	0.00023	0.00006
N	434,947	395,321

This table reports coefficient estimates of the following baseline regression estimated at the monthly frequency:

$$Dummy_quit_{i,t} = \alpha + \beta\ I_{\{MKT_RET_{t-1}<0\}} + \gamma\ I_{\{Type1_Advisor_{i,t-1}=1\}} + \delta\ I_{\{MKT_RET_{t-1}<0\}} \times I_{\{Type1_Advisor_{i,t-1}=1\}} + \epsilon_{i,t},$$

where $Dummy_quit_{i,t}$ is equal to 1 if investor i quits robo-advising in month t and 0 otherwise, $I_{\{MKT_RET_{t-1}<0\}}$ is an indicator variable equal to 1 if the market return is negative in month $t-1$ and 0 otherwise, and $I_{\{Type1_Advisor_{i,t-1}=1\}}$ is equal to 1 if investor i is assigned to an advisor with a past retention above the median and 0 otherwise. We multiply all the coefficient estimates by 100, so they are expressed in percentages. We split the sample of robo-advising users into two groups on the basis of their tenure and report the results for short tenure in column (1) and long tenure in column (2). In both cases, the standard errors are double-clustered by users and dates.

Electronic copy available at: https://ssrn.com/abstract=4301514

Online Appendix

(Not for publication)

Online Appendix A.1 Model derivations

Online Appendix A.1.1 Proofs

Proof of Proposition 1

The result follows from guessing and verifying that a value function with the functional form in (11) satisfies the Bellman equation. In addition, we need to establish that the conditional distribution of $m' = m_{t+1}^i(h)$ conditional on $m = m_t^i(h)$ is as defined in the Proposition. To see this, note that from the Kalman filtering equations, we have

$$\begin{aligned} m_t^i(h) &= m_{t-1}^i(h) + g_{t-1}(h) \cdot \eta_t^i(h) \\ &= m_{t-1}^i(h) + g_{t-1}(h)\left(r_t^i - m_{t-1}^i(h)\right) \\ &= m_{t-1}^i(h) + g_{t-1}(h)\left(\theta + \sigma u_{t+1}^i - m_{t-1}^i(h)\right). \end{aligned}$$

The expectation of $m_t^i(h)$ one period ahead is

$$\begin{aligned} E_t^i\left[m_{t+1}^i(h)\right] &= m_t^i(h) + g_t(h)\left(E_t^i\left[\theta + \sigma u_{t+1}^i\right] - m_t^i(h)\right) \\ &= m_t^i(h), \end{aligned}$$

using the fact that $E_t^i[\theta] = m_t^i(h)$ and $E_t^i\left[u_{t+1}^i\right] = 0$. The variance of $m_t^i(h)$ one period ahead is

$$\begin{aligned} Var_t^i\left[m_{t+1}(h)\right] &= Var_t^i\left[g_t(h)\left(\theta + \sigma u_{t+1}^i\right)\right] \\ &= g_t(h)^2\left(Var_t^i[\theta] + \sigma^2\right) \\ &= g_t(h)^2\left(v_t(h) + \sigma^2\right). \end{aligned}$$

Electronic copy available at: https://ssrn.com/abstract=4301514

Substituting the definition of $g_t(h)$, we get

$$\begin{aligned} Var_t^i\left[m_{t+1}^i(h)\right] &= \left(\frac{v_t(h)}{v_t(h)+\sigma^2}\right)^2\left(v_t(h)+\sigma^2\right) \\ &= \frac{v_t(h)}{v_t(h)+\sigma^2}v_t(h) \\ &= g_t(h)\,v_t(h). \end{aligned}$$

It follows that the distribution of $m' = m_{t+1}^i(h)$ conditional on $m = m_t^i(h)$ is

$$m'|t, m \sim N\left(m, v_t(h)\,g_t(h)\right),$$

as required.

Proof of Proposition 2

When $\underline{\alpha} = 0$, Corollary 1 along with Equations (14) and (15) imply that investor i quits at date 1 if and only if

$$\begin{aligned} &\frac{1}{2}S(m)^2 - f(h) \leq 0 \\ \Leftrightarrow \max&\left\{\frac{m+\frac{1}{2}\sigma^2-\left(\bar{r}+\xi_1^i\right)-c}{\sigma}, 0\right\} \leq \sqrt{2f(h)}, \end{aligned}$$

where $m = m_1^i(h)$, and where we have substituted the random outside return $\bar{r}+\xi_1^i$. This inequality always holds if $S(m) = 0$ and the investor chooses $\alpha = 0$. If $S(m) > 0$, then this inequality reduces to

$$m_1^i(h) + \frac{1}{2}\sigma^2 - \bar{r} - \xi_1^i - c \leq \sigma\sqrt{2f(h)}$$

Substituting from Equation (4) for $m_1^i(h)$, we obtain

$$m_0(h) + g_0(h)\left(r_1^i - m_0(h)\right) + \frac{1}{2}\sigma^2 - \bar{r} - \xi_1^i - c \leq \sigma\sqrt{2f(h)}$$

Electronic copy available at: https://ssrn.com/abstract=4301514

or, rearranging,

$$\bar{r}+c-\frac{1}{2}\sigma^2++g_0\left(h\right)\left(-r_1^i\right)+2\sigma\sqrt{f\left(h\right)}-m_0\left(h\right)\left(1-g_0\left(h\right)\right)+\xi_1^i\geq 0$$

Since $h\in\{0,1\}$, we can write this equivalently as

$$\begin{aligned}&\underbrace{\bar{r}+c-\frac{1}{2}\sigma^2+2\sigma\sqrt{f\left(0\right)}-m_0\left(0\right)\left(1-g_0\left(0\right)\right)}_{\equiv\alpha}\\&+\underbrace{g_0\left(0\right)}_{\equiv\beta}\left(-r_1^i\right)\\&+\underbrace{\left\{\sigma\left(\sqrt{2f\left(1\right)}-\sqrt{2f\left(0\right)}\right)-\left[m_0\left(1\right)\left(1-g_0\left(1\right)\right)-m_0\left(0\right)\left(1-g_0\left(0\right)\right)\right]\right\}}_{\equiv\gamma}h\\&+\underbrace{\left[g_0\left(1\right)-g_0\left(0\right)\right]}_{\equiv\delta}h\left(-r_1^i\right)+\xi_1^i\geq 0,\end{aligned}$$

which yields the required result.

Online Appendix A.1.2 Analysis of mapping from reduced-form to structural parameters

This appendix describes how we derive estimates of the reduction in uncertainty brought about by a type-1 human advisor from our reduced-form estimates. From the Kalman filtering equations in Lemma 1, we obtain that

$$\begin{aligned}g_0\left(h\right)&=\frac{v_0\left(h\right)}{v_0\left(h\right)+\sigma^2}\\&=\frac{1}{1+\frac{\sigma^2}{v_0(h)}},\end{aligned}$$

which further implies

$$\begin{aligned}\frac{\sigma^2}{v_0\left(h\right)}&=\frac{1}{g_0\left(h\right)}-1\\&=\frac{1-g_0\left(h\right)}{g_0\left(h\right)}\end{aligned}$$

Electronic copy available at: https://ssrn.com/abstract=4301514

or

$$v_0(h) = \sigma^2 \frac{g_0(h)}{1 - g_0(h)}$$

and

$$\frac{v_0(1)}{v_0(0)} = \frac{\frac{g_0(1)}{1-g_0(1)}}{\frac{g_0(0)}{1-g_0(0)}}.$$

In terms of our reduced form estimates, Proposition 2 shows that we have

$$g_0(0) = \beta$$
$$g_0(1) = \beta + \delta,$$

so that we can write the percentage change in prior variance with a type-1 human advisor in terms of reduced-form estimates as

$$\begin{aligned}
\frac{v_0(1) - v_0(0)}{v_0(0)} &= \frac{\frac{\beta+\delta}{1-\beta-\delta}}{\frac{\beta}{1-\beta}} - 1 \\
&= \frac{(\beta+\delta)(1-\beta)}{\beta(1-\beta-\delta)} \\
&= \frac{\beta + \delta - \beta^2 - \beta\delta}{\beta - \beta^2 - \beta\delta} \\
&= \frac{\delta}{\beta(1-\beta-\delta)}.
\end{aligned}$$

Online Appendix A.1.3 Derivation of approximate intertemporal budget constraints

Consider a continuous-time environment in which the per-unit value P_t^i of the robo advisor's portfolio for investor i follows the geometric Brownian motion

$$\frac{dP_t^i}{P_t^i} = \left(\theta + \frac{1}{2}\sigma^2\right) dt + \sigma dZ_t^i. \tag{27}$$

Solving this stochastic differential equation over the time interval from t to $t+1$, we can characterize the discrete-time returns on the robo advisor's portfolio as

$$r_{t+1} \equiv \log \frac{P_{t+1}}{P_t} = \theta + \sigma u_{t+1}^i, \tag{28}$$

Electronic copy available at: https://ssrn.com/abstract=4301514

where we have defined $u_{t+1}^i \equiv Z_{t+1}^i - Z_t^i$. This characterization demonstrates that this model is consistent with our assumption that discrete-time returns are log-normal with mean θ and variance σ^2.

Investor wealth follows the process (dropping i superscripts for simplicity)

$$\begin{aligned}\frac{dW_t}{W_t} &= \alpha_t \left(\frac{dP_t}{P_t} - cdt \right) + (1 - \alpha_t)\, \bar{r} dt \\ &= \left[\bar{r} + \alpha_t \left(\theta + \frac{1}{2}\sigma^2 - c - \bar{r} \right) \right] dt + \alpha_t \sigma dZ_t. \end{aligned} \tag{29}$$

Assuming that the portfolio share α_t is held constant between dates t and $t+1$, we once again obtain a standard geometric Brownian motion over this interval, which has solution

$$\log \frac{W_{t+1}}{W_t} = w_{t+1} - w_t = \left[\bar{r} + \alpha_t \left(\theta + \frac{1}{2}\sigma^2 - c - \bar{r} \right) - \frac{1}{2}\alpha_t^2 \sigma^2 \right] + \alpha_t \sigma u_{t+1}. \tag{30}$$

Rearranging (30) and substituting (28), we get

$$\begin{aligned} w_{t+1} &= w_t + \bar{r} + \alpha_t \left(\theta + \sigma u_{t+1} - c - \bar{r} \right) + \frac{1}{2}\alpha_t \left(1 - \alpha_t \right) \sigma^2 \\ &= w_t + \bar{r} + \alpha_t \left(r_{t+1} - c - \bar{r} \right) + \frac{1}{2}\alpha_t \left(1 - \alpha_t \right) \sigma^2 \qquad (31) \\ &= w_t + \bar{r} + \alpha_t \left(m_t + \eta_{t+1} - c - \bar{r} \right) + \frac{1}{2}\alpha_t \left(1 - \alpha_t \right) \sigma^2, \qquad (32) \end{aligned}$$

where the last line substitutes the definition $r_{t+1} - m_t = \eta_{t+1}$. This expression translates to our recursive formulation in the paper. Indeed, Equation (31) holds exactly as long as α_t is constant between t and $t+1$, which requires continuous rebalancing (e.g., after a positive excess return on the algorithm, some of the portfolio needs to be sold to keep α_t constant). Hence, in a discrete time environment where portfolios are not rebalanced within the time interval, this equation holds only approximately, and the approximation is around the state of the world where excess log returns on the algorithm are zero.

Electronic copy available at: https://ssrn.com/abstract=4301514

Online Appendix A.2 Validating the Quasi-Random Assignment of Clients to Advisors

The institutional details outlined above indicate that the assignment of new clients to advisors during the sample period was random with respect to advisor type. Assignment appeared to depend on advisor capacity and availability but not on advisors' skills or characteristics. To be precise, the probability that a new investor is assigned to an advisor at any point in time negatively depends on the load of the advisor at that time, but not on advisor type.

Put differently, the contention is that assignment of advisors to clients did not depend on advisor-specific retention rates (our empirical proxy for advisor type). To test whether this institutional feature is indeed reflected in the data, we first report in Figure Online I the distribution of Log Assets (Subfigure (a)), Equity Share (Subfigure (b)), and Age (Subfigure (c)) for users matched with high-retention advisors (in blue) and low-retention advisors (in red). In all cases, user characteristics are virtually identical, irrespective of the type of advisor they are matched with. We then estimate the following baseline specification at the daily frequency:

$$Clients_Added_{i,t} = \alpha_M + \beta_1\ Clients_Under_Management_{i,t-1} + \beta_2\ Advisor_Retention_{i,t-1} + \epsilon_{i,t}, \quad (33)$$

where $Clients_Added_{i,t}$ denotes the number of clients assigned to advisor i on day t, the number of clients under management by advisor i as of day $t-1$ is denoted by $Clients_Under_Management_{i,t-1}$, and $Advisor_Retention_{i,t-1}$ denotes the fraction of clients assigned to advisor i that have stayed in the service. Note that this last quantity is—crucially—computed only using information up to day $t-1$. Finally, α_M denotes monthly fixed effects. The description of the institutional setting reported above implies a negative and significant coefficient on β_1 and an insignificant coefficient on β_2.

Table Online I Panel A reports the results from this exercise. The first specification (column one) removes those advisors in the bottom 8% of the distribution in terms of clients under management at time $t-1$ to avoid including advisors who are either in the initial or final phases of their career as advisors.[12] We also only include instances where two or fewer clients are added to an advisor's set, to

[12] In the first case, we would expect the advisor manager to limit the number of new clients added to the advisor during an initial training period. In the second, we would expect the advisor manager to progressively reduce the number of clients advised by each advisor until they reach zero.

Electronic copy available at: https://ssrn.com/abstract=4301514

focus on the marginal client additions to the set of clients associated with each advisor.

The β_1 coefficient is negative and significant, suggesting that those advisors with a lower number of clients under management have a higher likelihood of being assigned a new client. The coefficient (multiplied by 100) is -0.058, meaning that an advisor with 10 clients more than the average load has a 0.58% lower probability of being assigned a new client on a given day. The β_2 coefficient, on the other hand, is statistically insignificant. Consistent with the assignment rule described above, it is not the case that advisors with different histories of retention are systematically more or less likely to be assigned new clients once we control for advisors' current number of clients under management.

Columns two through three repeat the analysis eliminating advisors in the bottom 6% and 4% of the distribution in terms of investors under management. Finally, the last column includes all advisors in the analysis. These specifications result in economically smaller (less negative) and less statistically significant coefficients for $Clients_Under_Management_{i,t-1}$: -0.052 (t-stat$=-8.1$), -0.044 (t-stat$=-6.59$), and -0.015 (t-stat$=-2.52$), respectively. At the same time, the coefficients on $Advisor_Retention_{i,t-1}$ become economically larger and become statistically significant in the last specification: 0.023 (t-stat$=0.72$), 0.049 (t-stat$=1.23$), and 0.127 (t-stat$=3.74$), respectively.

The coefficient estimates vary across the four specifications, even though the number of observations is rather similar. This suggests that client allocation rates are different for advisors with few clients. These advisors may also have lower retention rates, since they are at the beginning or end of their careers at the firm, compared with the rest of the advisor population. It may also, of course, suggest that despite the institutional framework, there is some discretion to allocate clients to specific advisors, especially for advisors that are at the beginning or end of their stints at the firm, violating the orthogonality condition.

To figure out which of these mechanisms is at work, we estimate an extension of Equation 33 where we estimate our model on the full dataset but allow for different sensitivities of new additions for advisors in the bottom 5% or top 95% of the distribution in terms of clients under management.

Electronic copy available at: https://ssrn.com/abstract=4301514

That is, we estimate:

$$\begin{aligned} Clients_Added_{i,t} = \alpha_M &+ \beta_1\ Clients_Under_Management_{i,t-1} \\ &+ \beta_1\ Advisor_Retention_{i,t-1} + \gamma_1\ Few_Clients_{i,t-1} \\ &+ \gamma_2\ Few_Clients_{i,t-1} \times Clients_Under_Management_{i,t-1} + \\ &+ \gamma_2\ Few_Clients_{i,t-1} \times Advisor_Retention_{i,t-1} + \epsilon_{i,t}, \end{aligned} \tag{34}$$

where all regressors are defined as in Equation 33 and $Few_Clients_{i,t-1}$ is a dummy variable that equals 1 if advisor i at time $t-1$ is in the bottom 5% in terms of clients under management.

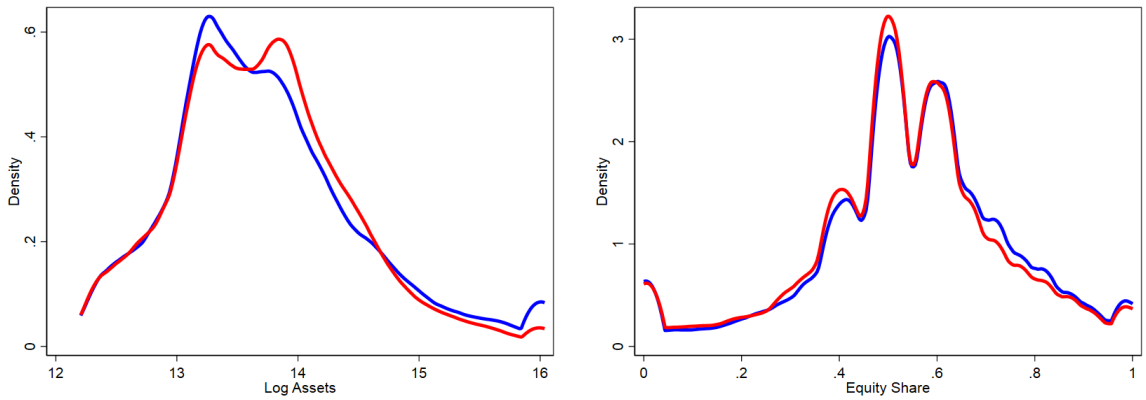
The first column in Panel B of Table Online I confirms that there are indeed different allocation rules to advisors in the bottom 5% in terms of clients under management compared to the remainder of the advisor population. For advisors in the top 95% of the distribution, clients under management come in with a coefficient (multiplied by 100) of -0.043 and a t-statistic of -6.44. As expected, both the coefficient estimate and the t-statistic are very close to the ones in column 3 of Panel A. In contrast, the coefficients on the dummy "Few Clients" and its interactions with the other regressors show that 1) advisors in the bottom 5% of the distribution are less likely to be assigned a new client, and 2) there is a positive relation between clients under management and the probability of being assigned a new client for these advisors, which is consistent with slower assignment to newly-joining financial advisors. Finally, once this difference in the speed of adjustment is accounted for, advisor retention is not significant for both groups. This is consistent with the results reported in Panel A and confirms the orthogonality of client assignment to advisor past retention once the number of current clients under management is accounted for.

The results reported so far use only the marginal additions of new clients to advisors and therefore only include values of $Clients_Added_{i,t}$ that range between zero and two. In the following three columns, we allow the variable $Clients_Added_{i,t}$ to range between zero and three in column 2 and between zero and 10 in column 3 to include the roughly 4% of days in which more clients are added. Finally, the last column changes the definition of the dependent variable into a dummy variable equal to one when advisor i is allocated one or more clients in period t and equal to zero otherwise, i.e., a

Electronic copy available at: https://ssrn.com/abstract=4301514

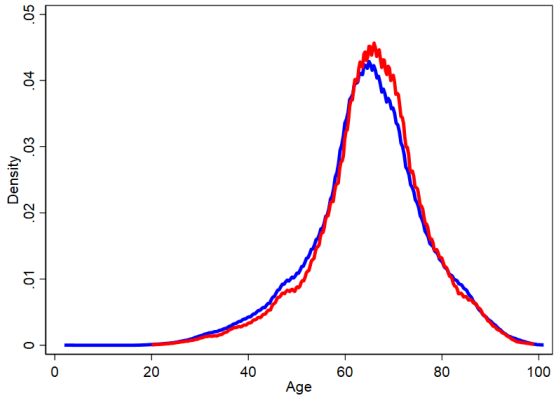
linear probability model. In all cases, we find that the number of clients under management predicts client assignment to advisors while advisor retention does not.

Electronic copy available at: https://ssrn.com/abstract=4301514



(a) Log Assets

(b) Equity Share



(c) Age

Figure Online I: This figure reports the distributions of Log Assets (Subfigure (a)), Equity Share (Subfigure (b)), and Age (Subfigure (c)) for users matched with high-retention advisors (in blue) and low-retention advisors (in red).

Electronic copy available at: https://ssrn.com/abstract=4301514

Table Online I. Assignment of New Investors to Advisors

Panel A. Baseline Specifications

	$> 8\%$	$> 6\%$	$> 4\%$	All
Clients Under Management (×100)	-0.058***	-0.052***	-0.044***	-0.015**
	(-10.19)	(-8.10)	(-6.59)	(-2.52)
Advisor Retention	0.011	0.023	0.049	0.127***
	(0.35)	(0.72)	(1.23)	(3.74)
Constant	0.213***	0.206***	0.196***	0.165***
	(24.65)	(22.59)	(21.12)	(17.18)
Fixed Effects	Month	Month	Month	Month
R-square	0.016	0.014	0.013	0.008
N	291,241	300,747	311,120	333,738

Panel B. Specifications with Interactions

	≤ 2	≤ 3	≤ 10	Lin. Prob.
Clients Under Management (×100)	-0.043***	-0.072***	-0.105***	-0.063***
	(-6.44)	(-6.87)	(-6.46)	(-6.93)
Advisor Retention	0.049	0.075	0.154	0.072
	(1.65)	(1.46)	(1.63)	(1.60)
Few Clients	-0.208***	-0.315***	-0.443***	-0.284***
	(-18.07)	(-16.69)	(-12.98)	(-17.15)
Few Clients × Clients Under Management	0.026***	0.039***	0.050***	0.035***
	(7.13)	(7.11)	(7.17)	(7.49)
Few Clients × Advisor Retention	0.008	0.016	-0.085	-0.005
	(0.24)	(0.29)	(-0.91)	(-0.11)
Constant	0.196***	0.293***	0.401***	0.263***
	(21.07)	(19.13)	(15.41)	(19.74)
Fixed Effects	Month	Month	Month	Month
R-square	0.019	0.025	0.024	0.029
N	333,738	344,836	350,462	350,462

This table reports regression-based tests to validate the quasi-random assignment of investors to advisors. As detailed in Section Online Appendix A.2, Panels A and B report estimates for different, but complementary, specifications. Panel A restricts the sample of data used in our estimates, while Panel B adopts flexible specifications relating new clients' assignment and current advisor capacity.

Electronic copy available at: https://ssrn.com/abstract=4301514