

Why Do Households Trade So Much?*

JUHANI T. LINNAINMAA[†]

November 18, 2008

ABSTRACT

If households are uncertain about their trading abilities but learn by trading, some will trade even if they expect to lose. I show that a Bayesian trading model with frictions is consistent with three empirical regularities in household trading behavior: most households lose to the market, households' trading intensity depends on past performance, and poorly performing households stop trading altogether. Estimates from a hierarchical population model suggest that, first, moderate amounts of uncertainty can generate the regularities seen in the data; second, approximately 4% of investors could earn excess returns with some consistency; third, 18% of the investors who started active trading believed that they would lose to the market—the *possibility* of being skilled offset these investors' expected short-term losses; and fourth, because investors stop trading after poor performance, realized returns are significantly downwards biased measures of investors' true abilities.

*University of Chicago Booth School of Business. I would like to thank Tobias Adrian, Robert Battalio, Tony Bernardo, Alan Bester, Michael Brennan, John Cochrane, John Griffin, Mark Grinblatt, John Heaton, Matti Keloharju, Deborah Lucas, Kevin McCardle, Ľuboš Pástor, Mark Seasholes, Luke Taylor, Pietro Veronesi, and Yuzhao Zhang, and the AFA 2008-New Orleans Meetings, New York University, University of Chicago, and Notre Dame University seminar participants for their comments. This paper replaces an earlier version entitled “Learning from Experience.”

[†]**Correspondence Information:** Juhani Linnainmaa, University of Chicago Booth School of Business, 5807 S Woodlawn Ave, Chicago, IL 60637, tel: (773) 834-3176, fax: (773) 753-1052, <http://faculty.chicagogsb.edu/juhani.linnainmaa>, email: jlinnain@chicagogsb.edu.

Why Do Households Trade So Much?

ABSTRACT

If households are uncertain about their trading abilities but learn by trading, some will trade even if they expect to lose. I show that a Bayesian trading model with frictions is consistent with three empirical regularities in household trading behavior: most households lose to the market, households' trading intensity depends on past performance, and poorly performing households stop trading altogether. Estimates from a hierarchical population model suggest that, first, moderate amounts of uncertainty can generate the regularities seen in the data; second, approximately 4% of investors could earn excess returns with some consistency; third, 18% of the investors who started active trading believed that they would lose to the market—the *possibility* of being skilled offset these investors' expected short-term losses; and fourth, because investors stop trading after poor performance, realized returns are significantly downwards biased measures of investors' true abilities.

The average household trades frequently but loses to the market.¹ One explanation is that households derive utility from gambling in the stock market; another is that investors are overconfident about the precision of their information. Mahani and Bernhardt (2007), however, note that these explanations cannot account for other household trading patterns, such as the positive correlation between trading intensity and past performance. Their alternative is a model in which traders learn about their abilities. Whereas Mahani and Bernhardt examine theoretically how learning affects the market equilibrium, I study the quantitative effect of learning on investor behavior: can a learning model explain household behavior and, if so, what does the model tell us about households' beliefs?

I study the learning explanation by solving a dynamic trading model in which a Bayesian investor learns about his ability from each trade. Because the investor must trade to learn, trades have option value: an investor trades even if he expects to lose money if the value of another signal offsets the expected loss. This means that even those who think that they are unskilled may trade: the probability that they might nevertheless be skilled can give impetus to trade. Thus, the model delivers the stylized fact about the *average* investor trading “too much” given that he loses to the market. Such an investor also responds to gains and losses by varying trading intensity and quits altogether after a string of losses. The overconfidence and gambling explanations do not generate similar dynamics. For example, if overconfidence is a fixed investor characteristic, past performance is uncorrelated with future choices.

I use Finnish trading records on high-frequency (or “active”) traders to test the predictions of the learning model. These traders, who account for almost two-thirds of total household volume, behave in accordance with the learning model. First, traders often quit trading after experiencing a sequence of unsuccessful trades, suggesting that they learned that they were unskilled. Second, traders increase their trade sizes after successful trades and decrease trade sizes after unsuccessful trades. This is consistent with investors changing beliefs about their trading abilities after each trade. Both the exit and trade size effects are stronger early on in investors' careers, which is again consistent with the learning model: because the amount of skill uncertainty declines in trading experience, experienced investors' posterior beliefs place less weight on each new signal. Third, many traders initially execute very small trades that seem to be motivated only by their desire to learn more about their own trading skills. The finding about the changes in the speed of learning and the result about initial exploratory trades are, in particular, novel to the literature and unique

¹See, for example, Odean (1998, 1999), Barber and Odean (2000, 2001), Kumar (2007), Grinblatt and Keloharju (2008), and others for both empirical evidence and suggested explanations.

to the learning explanation.

The learning model can also address questions about the investor population: what distribution of beliefs in the population would generate the aggregate trading patterns seen in the data? I examine such questions by first making parsimonious distributional assumptions about investors' prior beliefs. Each investor's prior in this hierarchical model is drawn from the population-wide distributions and is assumed to be well-calibrated; i.e., investors are, on average, correct about their abilities. Because the resulting hierarchical model does not yield closed-form estimation equations, I use Simulated Method of Moments to estimate the structural parameters.

I use high frequency traders' conditional exit rates to identify the structural parameters—i.e., I count the number of investors who quit after different outcome sequences and then match these rates between the model and the data. Because investors can also avoid trading altogether, I use the initial entry decision as another moment condition. This takes into account the self-selection in the data: active traders are investors who *chose* to begin active trading while others *chose* to stay out. Because of the initial entry-moment condition, the estimates from the model are belief distributions for the entire population and not just for the high-frequency traders.

The learning model does a good job matching the initial selection step as well as the conditional exit rates in the data. The fact that this model can explain the initial selection is important: it indicates that active traders do not need to be a distinct group of investors with unreasonable beliefs or non-standard preferences. In contrast to such segregation, the investor population in the model is a continuum of investors; those who become active traders just happen to reside in particular corners of the belief parameter space.

The structural parameter estimates suggest that up to 18% of the investors who started active trading believed that they were not skilled: they traded because of the *possibility* that they might be. The average investor who begins trading assigns a probability of 0.72 to being skilled. The estimates also suggest that approximately 4% out of all investors have genuine trading skills, i.e., they can profit short-term price movements at least with some consistency. Although this skill may manifest itself as an ability to predict short-term shifts in fundamental prices, it may also reflect pseudo (or de facto) market making. In electronic limit order markets, individuals can try to profit from other investors' demand for immediacy by submitting limit orders.² The 4% estimate includes

²Barclay, Christie, Kandel, and Schultz (1999) and Barclay, Hendershott, and McCormick (2003) study the Nasdaq reform and the growth of electronic communications networks (ECNs), respectively, and emphasize the economic

both types of investors. This estimate is consistent with studies such as Coval, Hirshleifer, and Shumway (2005) and Grinblatt, Keloharju, and Linnainmaa (2008) who find persistent superior performance for a small fraction of individual investors. Finally, the model also implies that up to 11.1% of the genuinely skilled investors never trade on their abilities. These investors stay out because they deem the probability of being skilled to be too small.

The structural model results indicate that a reverse survivorship bias has ramifications for investor performance measurement. Because investors are more likely to quit trading after unsuccessful trades, realized performance is biased downwards relative to true, unobservable skill. This mechanism is best illustrated with a coin-flip example. Suppose that we start flipping a coin to measure its bias, but stop the experiment after we get the first tails. For a fair coin, the expected proportion of heads in such an experiment is $\lim_{n \rightarrow \infty} \left(\frac{1}{2}\right)^1 \frac{0}{1} + \left(\frac{1}{2}\right)^2 \frac{1}{2} + \dots + \left(\frac{1}{2}\right)^n \frac{n-1}{n} = \sum_{n=1}^{\infty} \left(\frac{1}{2}\right)^n \frac{n-1}{n} = 0.31$ —because the stopping rule is conditional on the tails-outcome, we oversample tails relative to the truth. This same mechanism affects investor performance measurement: if investors quit after poor performance, the observed data “oversamples” poor performance relative to true skill. This biases downwards the data-based estimates of investors’ abilities. The structural model gives an estimate of the economic magnitude of this bias: although high-frequency traders’ average true skill—expressed as a probability of a successful trade—is 0.512, the average *observed* skill is just 0.460.

Although the investors in the model have fixed trading abilities—i.e., they do not learn *how* to trade—trading experience is positively correlated with trading performance through a learning-by-survival mechanism. Because the skilled traders are more likely to survive, a higher fraction of the investors who remain in the market are skilled. We can quantify the economic significance of this effect by inserting the structural parameter estimates back into the model. In the sample of high-frequency traders, the fraction of unskilled investors decreases from 18.2% at the time of the first trade to just above 10% after two trades. Thus, the unskilled investors learn very fast that they are unskilled and quit the market. This makes the surviving pool of investors more skilled.

Investors in my model learn about their skills by trading. Empirical evidence overwhelmingly supports this assumption: numerous studies³ document that investors’ past performance significantly affects their future actions. What is not known is what generates this mechanism: why do not investors paper trade to learn about their abilities? Although it could be that paper trading sim-

importance that traders’ ability to compete with market makers has on liquidity.

³See the discussion in Section I.A.

ply fails to resolve *all* uncertainty—and that the residual uncertainty leads *some* investors to trade with real money—it could also arise from, e.g., bounded rationality: investors may learn more when they take part in the action instead of just paper trading.⁴ What is important is that my results themselves support the learning by trading assumption: if investors could indeed learn everything by paper trading, why does the learning-by-trading model fit the data so well?

This study is most closely related to the theoretical model of Mahani and Bernhardt (2007) that incorporates investor learning into a general equilibrium model of the markets. The model consists of risk-neutral agents who learn their types after the first period. Mahani and Bernhardt show that learning reduces the bid-ask spreads and the price impact of liquidity shocks. Although the paper is theoretical, Mahani and Bernhardt justify their setup by arguing that their model is qualitatively consistent with many empirical facts about financial speculators' behavior. This paper focuses on the empirical implications of a similar learning mechanism: can a learning model quantitatively explain investor behavior, and if so, what does the model tell us about households' beliefs and motivations for trading?

This study also complements papers that find evidence of investor learning in the data. The results suggest that this learning takes the form of a discovery: it is not that the investors who remain in the market the longest and perform the best learn *how* to trade—rather, it is that those who quit trading learn that they are unskilled. Seru, Shumway, and Stoffman (2007) reach a similar conclusion in their study of how the disposition effect weakens in investor experience: they find that the disposition effect only appears to weaken in investor experience because the high disposition effect investors are more likely to drop out of the sample.

Finally, structural models have not been widely used in finance to estimate learning processes. Only two recent papers, Sørensen (2007) and Taylor (2008), take this approach. Taylor (2008) estimates a structural model in which the board of directors learns about CEO skill; he finds that this learning can make the CEO firing decision very sensitive to firm performance. Sørensen (2007) uses a structural matching model to examine why those companies that are funded by more experienced VCs are more likely to go public. To the best of my knowledge, this study is the first one to use a structural model to estimate investors' learning processes from the data.

The paper is organized as follows. Section I motivates the study and discusses the learning-

⁴Section I.B discusses paper trading in greater detail.

by-trading assumption. Section II presents the model and examines its empirical implications. Section III describes the data. Section IV shows that investor behavior in the data is qualitatively consistent with the model. Section V estimates the hierarchical learning model from the data. Section VI concludes.

I. Additional Motivation

A. Stylized Facts about Household Behavior

The learning model of this study can reconcile a number of stylized factors about household behavior. I review some of these results to motivate the research. The trading records from (at least) Denmark, Finland, Norway, Taiwan, and the U.S. all have an average household that trades “too much”, loses to the market, and alters trading intensity in response to past returns. At the same time, some households do show superior performance.

Excessive trading and underperformance of the average household. Odean (1999), Barber and Odean (2000, 2001), Grinblatt and Keloharju (2000, 2008), and others find that the average household trades excessively: the poor after-transactions cost performance appears to suggest many would be better off by holding the market. Barber and Odean (2001a), Kumar (2007), and Grinblatt and Keloharju (2008) suggest that both overconfidence and sensation-seeking (or gambling) preferences may explain these results.

Performance Heterogeneity. Barber, Lee, Liu, and Odean (2004), Nicolosi, Peng, and Zhu (2004), Coval, Hirshleifer, and Shumway (2005), Bauer, Cosemans, and Eichholtz (2007), Seru, Shumway, and Stoffman (2007), and Goetzmann and Kumar (2008) all find that a small number of individual investors outperform both their peers and the market. Harris and Schultz (1998) find that some individual investors acting as SOES (Nasdaq’s Small Order Execution System) bandits are better than others, and that making money as a bandit takes skill. What makes these results important is the persistence in the superior performance: the best performing investors continue to outperform the worst performers from one period to the next. If so, these findings also suggest that not all high-intensity trading is necessarily excessive. For example, Grinblatt, Keloharju, and Linnainmaa (2008) document that investors with high intelligence scores trade more *and* outperform investors with low IQ scores.

Past performance and trading activity. Barber, Lee, Liu, and Odean (2004) find a positive correlation between the performance and subsequent trading activity among Taiwanese day traders: the best performing traders trade more while their fortunate peers trade less. Nicolosi, Peng, and Zhu (2004) find a similar relationship between past performance and trading activity among the customers of a U.S. discount broker. Seru, Shumway, and Stoffman (2007) find that “exits” is an important characteristic of the performance-trading activity results: the worst performing investors quit trading altogether. Mahani and Bernhardt (2007) also cite a number of older, unpublished manuscripts for similar evidence about the performance-trading activity relation.

B. Learning and Paper Trading

If a learning model is to explain investor behavior, investors need to learn from their own actions. Why could not individuals learn by just paper trading? This section discusses two (alternative) factors that may drive the learning-by-trading mechanism: the limits of paper trading and bounded rationality. Although I cannot distinguish between these explanations in the data, the existence of the mechanism itself appears undisputable: similar to Barber, Lee, Liu, and Odean (2004), Nicolosi, Peng, and Zhu (2004), and others, I also find that investors’ past performance greatly influences their future actions.

A learning by trading-mechanism arises if there are limits to what one can learn by paper trading: it is enough that investors cannot *fully* learn their type by paper trading. That is, investors may be left with some residual uncertainty no matter how much they paper trade. Some of the investors may then seek to resolve this residual uncertainty by experimenting with real money.

A number of practical difficulties with paper trading may impose limits of how much can be learned from it. First, trade execution quality is an important concern with short holding periods: whether an investor pays an average bid-ask spread of 0.1% or 1% amounts to a significant difference in realized returns. It is very difficult to learn about the execution component by paper trading. The first difficulty is that trade and quote databases are not readily available to the public. The second difficulty is that the data rarely show order-flow directions, making it difficult to assess *if* and *when* a hypothetical limit order would have executed. Lo, MacKinlay, and Zhang (2002, p. 31) find this to be an important issue: “hypothetical limit-order executions, constructed either theoretically from first-passage times or empirically from transactions data, are very poor proxies for actual limit-order executions.” Harris and Schultz (1998) find that good execution is vital to

individual investors acting as SOES bandits: bandits need to trade *within* quoted prices to make money.

Households' preference to trade small stocks—Barber and Odean (2001a, Table III), for example, show that the average household has an SMB loading of 0.776 in the Fama and French (1993) three-factor model—makes paper trading even more difficult. A household may not be atomistic in the smallest stocks: a trade or even a limit order submission may alter the behavior of other traders and change the stock's future price path. If an investor has non-negligible mass, paper trading cannot resolve all uncertainty: the investor does not know what *would have* happened had he submitted a real order.

Investors might get more out of paper trading by expending more time and money. For example, they might collect and study real-time market feeds to learn about execution. Such activities are, however, costly, creating a trade-off that may favor real trading: if the costs of paper trading (i.e., the actual monetary cost as well as the opportunity cost of time) are higher than what the investor expects to lose in the market by making small trades, real trading may dominate paper trading. These two considerations—i.e., the practical limits of paper trading and the relative cost-comparison—may be important reasons for why investors learn by trading.

An alternative mechanism is bounded rationality. Arrow (1962, p. 155), for example, motivates a learning by doing-process with an observation from psychology: “learning is the product of experience.” Although this mechanism could be rationalized (see, e.g., Grossman, Kihlstrom, and Mirman (1977) on drug use), it is often interpreted as reinforcement learning: individuals pay more attention to personal experiences than to generic observations. Kaustia and Knüpfer (2008), for example, study individual investors' IPO participation decisions and document that investors are more sensitive to their *own* successes and failures. Hence, it is possible that investors trade to learn *even though* they could learn the same things by paper trading.

The question in this paper does not require one to take a stance on *why* investors learn by doing; i.e., whether is because of the limits of paper trading, a relative cost comparison between paper and real trading, or because of bounded rationality. It is only necessary that, for whatever reason, investors do learn from the trades that they make. Thus, I treat the question “do investors learn by trading” as a hypothesis to be tested along with the implications of the learning model. In this respect, the results themselves validate the premise of the model: if we were to believe that investors

do not learn from their own trades, how come the learning model fits the data so well?

II. The Model

I solve a finite-horizon dynamic trading model in which a Bayesian investor learns about his ability from each trade. This uncertainty changes investor behavior significantly because the investor must take into account how the decisions made today affect the opportunities available in the future.

This model differs from other Bayesian life-cycle models (see, e.g., Brennan (1998) and Xia (2001)) in two ways. First, I add trading frictions to generate the learning by trading-mechanism. If an investor could trade infinitesimal amounts, he could effectively observe outcomes *without* trading. The second difference is the context of the model. Whereas investors in life-cycle models often have uncertainty about an objective parameter—e.g., the size of the equity premium or the predictive power of a state-variable—my investor is uncertain about his *own* trading skills.

A. Investor and Investment Opportunities

Assume that a single investor lives for T periods and maximizes power utility over terminal wealth,

$$U(W_T) = \frac{W_T^{1-\gamma}}{1-\gamma}, \tag{1}$$

where γ is the coefficient of relative risk-aversion. The investor can trade at dates $t = 1, 2, \dots, T-1$. The investor chooses an amount $x_t > 0$ to trade. It does not matter whether the investor attempts to profit from short-term price movements or from supplying liquidity to the market as a pseudo-market maker: x_t is just the dollar amount wagered in this strategy. The amount x_t is always non-negative because buying and selling actions are not considered separately; if there are N stocks, we can think of the investor as choosing from a universe of N potential purchases and N potential sales. The trade is a success with probability p and a failure with probability $1-p$. If the trade is a success, the payoff is $(1+\delta)x_t$ and if its a failure, the payoff is $(1-\delta)x_t$, where $\delta \leq 1$ is some constant.

The uncertainty in the model is about the parameter p , the probability that the trade is successful. I model this uncertainty by endowing the investor with a beta distributed prior about p . Beta distribution is very flexible, allowing for a very general description of an investor's beliefs

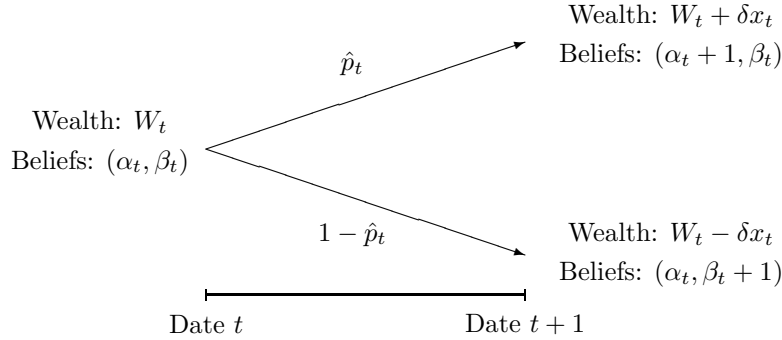


Figure 1: **The Evolution of Wealth and Beliefs in the Bayesian Trading Model.** A Bayesian investor has a wealth of W_t and a beta distributed prior with parameters (α_t, β_t) about his trading skill p at date t . This figure illustrates how the wealth and beliefs evolve if the investor bets an amount x_t on his view. The subjective probability of a good outcome is denoted by $\hat{p}_t = E_t[\tilde{p}] \equiv \frac{\alpha_t}{\alpha_t + \beta_t}$.

about his skills.⁵ Endowed with this prior, the investor believes that with probability $E(\tilde{p} > \frac{1}{2}) = \int_{\frac{1}{2}}^1 B(x; \alpha, \beta) dx$ he can make money from short-term trading with some consistency. ($B(x; \alpha, \beta)$ denotes the density of a beta distribution with parameters α and β .)

The investor learns about his skill p only by observing more outcomes. If the investor trades, he observes the outcome and updates his beliefs as a Bayesian. If the investor does not trade, he does not observe the outcome.

The investor updates his beliefs according to simple rules because of the beta prior/binomial outcome assumption. Starting from a prior $B(x; \alpha_1, \beta_1)$, the posterior would be $B(x; \alpha_1 + 1, \beta_1)$ after a successful trade and $B(x; \alpha_1, \beta_1 + 1)$ after a failed trade (see, e.g., DeGroot (1970, p. 160)). Hence, the two parameters of the beta distribution keep track of the number of positive and negative outcomes the investor experiences. Figure 1 illustrates how the investor's beliefs and wealth evolve in this model in one time-step.

I introduce a trading friction by setting a lower bound for permissible trade sizes, x_t . If an investor trades, the trade must be greater than some minimum amount \bar{x} . This assumption makes learning costly: if an investor believes that his p is below $\frac{1}{2}$, he needs to weigh the expected trading loss against the value of the information he would gain from one more trade.

The investor can also invest in a risk-free asset that pays no interest. I do not model the possibility that the investor could invest in the stock market instead of engaging in short-term trading. This

⁵See, e.g., Pratt, Raiffa, and Schlaifer (1995, p. 183) for an illustration of the various shapes that a Beta distribution can take.

restriction makes the problem tractable. Moreover, an addition of another asset should not alter the conclusions; because investors bet on both up- and down-movements of individual stocks, the return on the average trade can be nearly uncorrelated with the market. In the canonical dynamic portfolio choice model, an addition of an uncorrelated asset does not affect the signs of the original asset weights: the weights inherit their signs from the excess returns, $E[\tilde{r}_i] - r_f$ (Merton 1969, p. 256).

Because stock returns are continuous, the assumption that outcomes are binomial might be disconcerting. Although this assumption helps to retain tractability, it is also rooted in the performance evaluation literature: Henriksson and Merton (1981), Cumby and Modest (1987), and Hartzmark (1991) measure performance by asking how good traders are at forecasting price change directions. If an investor is uncertain about his trading abilities, his first step may well be to learn whether he can get the signs right. For example, suppose that an investor believes that a stock will go up by 10% the next day. If the stock falls by 5% instead, the binomial outcome-assumption says that the investor only cares about having gotten the sign wrong and ignores the magnitude of his error.

B. Investor's Optimization Problem

The investor chooses the optimal amount to trade at each date to maximize utility over terminal wealth. This task is complicated not only by the changes in \hat{p}_t but also by the minimum trade size requirement. The model becomes unsolvable in closed-form for $T > 3$ because of the two state variables, beliefs and wealth. I first write the investor's optimization problem as a dynamic programming problem and then translate it into a large-scale fixed point problem that is then solved numerically.

I first describe the investor's problem at date $T - 1$ when he makes his last decision and then move to the general problem at dates $t < T - 1$. I assume that $\gamma > 1$ to simplify notation; the same arguments hold for $\gamma < 1$.

B.1. Last Period Problem

Suppose the investor reaches date $T - 1$ with wealth W_{T-1} and beliefs \hat{p}_{T-1} . The investor has to choose from two alternatives: trade at least \bar{x} on the predicted stock price movement or quit and

consume wealth W_{T-1} . The date $T - 1$ optimization problem is then:

$$V_{T-1}(W_{T-1}, \hat{p}_{T-1}) = \max \left\{ \max_{x_{T-1} > \bar{x}} \left\{ \hat{p}_{T-1} \frac{(W_{T-1} + \delta x_{T-1})^{1-\gamma}}{1-\gamma} + (1 - \hat{p}_{T-1}) \frac{(W_{T-1} - \delta x_{T-1})^{1-\gamma}}{1-\gamma} \right\}, \frac{W_{T-1}^{1-\gamma}}{1-\gamma} \right\}, \quad (2)$$

where the inner maximization problem solves the trading problem subject to the minimum trade size constraint and the outer maximization problem compares this solution to the utility from quitting.

From the first-order condition, the optimal investment for the unconstrained investment problem in (2) is

$$x_{T-1}^* = \frac{1}{\delta} \frac{\hat{p}_{T-1}^{\frac{1}{\gamma}} - (1 - \hat{p}_{T-1})^{\frac{1}{\gamma}}}{\hat{p}_{T-1}^{\frac{1}{\gamma}} + (1 - \hat{p}_{T-1})^{\frac{1}{\gamma}}} W_{T-1}. \quad (3)$$

The inner maximization problem can then be replaced by its value $V_{T-1}^I(W_{T-1}, \hat{p}_{T-1})$:

$$V_{T-1}^I(W_{T-1}, \hat{p}_{T-1}) = \begin{cases} \frac{W_{T-1}^{1-\gamma}}{1-\gamma} 2^{1-\gamma} \left(\hat{p}_{T-1}^{\frac{1}{\gamma}} - (1 - \hat{p}_{T-1})^{\frac{1}{\gamma}} \right)^\gamma & \text{if } x_{T-1}^* \geq \bar{x}, \\ \frac{W_{T-1}^{1-\gamma}}{1-\gamma} \left(\hat{p}_{T-1} (1 + \delta \bar{\theta}_{T-1})^{1-\gamma} + (1 - \hat{p}_{T-1}) (1 - \delta \bar{\theta}_{T-1})^{1-\gamma} \right) & \text{if } x_{T-1}^* < \bar{x} \text{ and } W_{T-1} > \bar{x}, \\ -\infty & \text{otherwise,} \end{cases} \quad (4)$$

where $\bar{\theta}_{T-1} \equiv \frac{\bar{x}}{W_{T-1}}$ is the minimum trade size as a fraction of wealth. The third line evaluates the utility at the minimum allowed investment—i.e., because an unconstrained investor would invest $x_{T-1}^* < \bar{x}$, I evaluate the utility of this investor at the boundary, $x_{T-1}^* = \bar{x}$ —and the fourth line sets the utility to minus infinity if the investor cannot afford to trade the minimum amount.

The investor's date $T - 1$ indirect utility compares this utility from investing to the utility from quitting and consuming. This gives the utility function the following multiplicative form:

$$V_{T-1}(W_{T-1}, \hat{p}_{T-1}) = \max \left\{ V_{T-1}^I(W_{T-1}, \hat{p}_{T-1}), \frac{W_{T-1}^{1-\gamma}}{1-\gamma} \right\} = \frac{W_{T-1}^{1-\gamma}}{1-\gamma} k_{T-1}(W_{T-1}, \hat{p}_{T-1}), \quad (5)$$

where the implicitly defined $k_{T-1}(W_{T-1}, \hat{p}_{T-1})$ depends also on wealth because the minimum trade size constraint is not homogeneous in wealth.

B.2. General Problem

An investor who reaches date t with wealth W_t and beliefs \hat{p}_t evaluates two options: trade at least \bar{x} or quit and consume the wealth W_t . Let us conjecture that the investor's indirect utility function has the same functional form as in (5). The investor's date t optimization problem is then

$$V_t(W_t, \hat{p}_t) = \max \left\{ \max_{x_t > \bar{x}} \left\{ \hat{p}_t \frac{(W_t + \delta x_t)^{1-\gamma}}{1-\gamma} k_{t+1}^S + (1 - \hat{p}_t) \frac{(W_t - \delta x_t)^{1-\gamma}}{1-\gamma} k_{t+1}^F \right\}, \frac{W_t^{1-\gamma}}{1-\gamma} \right\},$$

where $k_{t+1}^S \equiv k_{t+1}(W_{t+1}, \hat{p}_{t+1})$ is the date $t+1$ multiplier in the indirect utility function if the trade is successful and k_{t+1}^F is the multiplier if the trade fails. Note that the date $t+1$ wealth and beliefs both depend on the date t outcome. Although coefficients k_{t+1}^S and k_{t+1}^F depend on the optimal x_t , let us assume that k_{t+1}^S and k_{t+1}^F are known. The numerical solution approach will make initial guesses about these values and then iterates the maximization problem to obtain the actual values. The optimal investment in the unconstrained inner maximization problem is

$$x_t^* = \frac{1}{\delta} \frac{[\hat{p}_t k_{t+1}^S]^{\frac{1}{\gamma}} - [(1 - \hat{p}_t) k_{t+1}^F]^{\frac{1}{\gamma}}}{[\hat{p}_t k_{t+1}^S]^{\frac{1}{\gamma}} + [(1 - \hat{p}_t) k_{t+1}^F]^{\frac{1}{\gamma}}} W_t. \quad (6)$$

The inner maximization problem in (6) simplifies to

$$V_t^I(W_t, \hat{p}_t) = \begin{cases} \frac{W_t^{1-\gamma}}{1-\gamma} 2^{1-\gamma} \left([\hat{p}_t k_{t+1}^S]^{\frac{1}{\gamma}} - [(1 - \hat{p}_t) k_{t+1}^F]^{\frac{1}{\gamma}} \right)^\gamma & \text{if } x_t^* \geq \bar{x}, \\ \frac{W_t^{1-\gamma}}{1-\gamma} \left(\hat{p}_t (1 + \delta \bar{\theta}_t)^{1-\gamma} k_{t+1}^S + (1 - \hat{p}_t) (1 - \delta \bar{\theta}_t)^{1-\gamma} k_{t+1}^F \right) & \text{if } x_t^* < \bar{x} \text{ and } W_t > \bar{x}, \\ -\infty & \text{otherwise,} \end{cases} \quad (7)$$

and the indirect date t utility function has the conjectured multiplicative form:

$$V_t(W_t, \hat{p}_t) = \frac{W_t^{1-\gamma}}{1-\gamma} k_t(W_t, \hat{p}_t). \quad (8)$$

C. Numerical Solution

The problem cannot be solved recursively by starting from the end, because the date t choice depends on future wealth and, at the same time, the future wealth depends on the date t choice. Nevertheless, the structure of the model facilitates exact numerical solution. My strategy is to cast the problem as a large fixed point problem. Because there are only two possible outcomes at each

date and because the changes in both wealth and beliefs are (conditionally) perfectly correlated with these outcomes, the problem has the shape of a non-recombining binomial tree. Figure 1 is one node in such a tree: starting from any wealth/belief pair (W_t, \hat{p}_t) , there are only two possible outcomes: $(W_t + x_t^*, \hat{p}_t^S)$ after a success and $(W_t - x_t^*, \hat{p}_t^F)$ after a failure, where x_t^* is the optimal trade size or zero, if the investor quits.

I first make initial guesses about each element in the investor's wealth process $\{W_t^j\}_{t=1, \dots, T-1}^{j=1, \dots, 2^{t-1}}$, where $j = 1, \dots, 2^{t-1}$ indexes all date t nodes of the tree. I can then solve the tree backwards by using (6), (7), and (8), because the initial wealth guesses determine the value function values in the subsequent nodes of the tree. The problem is then to find the wealth process $\{W_t^j\}_{t=1, \dots, T-1}^{j=1, \dots, 2^{t-1}}$ that supports the optimal choices in every node. If the optimal choice is to trade amount x_t^{j*} , the assumed wealth process supports this choice if

$$W_{t+1} = \begin{cases} W_t + x_t^{j*} & \text{if a success} \\ W_t - x_t^{j*} & \text{if a failure.} \end{cases} \quad (9)$$

If there are any wealth elements that do not support the optimal choices, I update wealth dynamics forward from date 1 to date $T - 1$ based on the new values for x_t^{j*} . I then pass again backwards through the tree and iterate until convergence.

The benefit of this solution methodology is that it is accurate: I do not need to approximate the value function either analytically or by numerically fitting splines. Moreover, because I use Simulated Method of Moments to estimate the model—which requires us to solve the model thousands of times—an important benefit of this solution approach is its speed: the solution usually converges in a matter of seconds. This technique does impose limits on the investment horizon T because the problem is exponential in T —the tree has $\sum_{i=1}^{T-1} 2^{i-1}$ nodes. Whereas a problem with $T = 15$ has just 16,383 decision nodes, a problem with $T = 20$ has already over 0.5 million nodes.

D. Empirical Implications

Three considerations in this model shape investor behavior and give rise to empirical implications. The first consideration is the usual myopic demand: given everything else, an investor with a higher prior about his trading ability p trades more. The empirical implication is that if two investors differ only in the levels of their priors, the optimistic investor trades more than the pessimistic investor.

The second consideration is the intertemporal hedging demand that arises because the returns are not i.i.d.: although the true p is fixed, an investor's beliefs about p change. The investor revises his beliefs upwards after a successful trade and downwards after an unsuccessful trade. Depending on how risk averse the agent is, this creates an additional positive or negative demand on top of the myopic demand. An investor with $\gamma > 1$ trades *less* because of this parameter uncertainty while an investor with $\gamma < 1$ trades more (see, e.g., Brennan (1998) for a proof and discussion). The empirical implication is that an investor's decision to continue trading and the choice of trade size depend on past outcomes.

The third consideration is the option value of trading. Trades have value beyond their expected payoffs because an investor observes the outcome only if he trades. Thus, even an investor with a belief $\hat{p} < \frac{1}{2}$ may choose to trade: ex ante, the (possibly small) probability that he is skilled may justify the expected loss from one more trade. Because an investor with $\hat{p} < \frac{1}{2}$ wants to minimize his expected loss, he trades the smallest permissible amount, \bar{x} . This mechanism runs in the opposite direction from the intertemporal hedging demand if $\gamma > 1$: an increase in trading skill uncertainty *lowers* the intertemporal hedging demand but *increases* the option value of trading. This option mechanism is similar to the one used in the sequential investment literature. In a sequential investment model, a firm may invest in early stages even though the NPV of the entire project is negative.⁶ The empirical implication is that an investor who is very uncertain about his trading abilities is more likely to begin trading than an investor with a very precise prior. Those investors who trade because of this option value mechanism trade smaller quantities than other investors.

Because I cannot solve the model analytically, I solve the model numerically for various types of investors to examine the relative strength of these three forces, myopic demand, intertemporal hedging demand, and the option value of trading. Figure 3 shows how the skill uncertainty affects investor behavior in the time-series. Although the investor believes that he is unskilled ($\hat{p}_1 = \frac{1}{2}$), he trades the minimum amount to observe another outcome. If the trade fails, the investor revises beliefs downwards to $\hat{p}_2 = \frac{1}{3}$, but still makes another minimum-sized trade. He would quit trading after one more loss. On the other hand, if the investor's first trade is successful, he revises beliefs upwards to $E\hat{p}_1 = \frac{2}{3}$) and trades more than the minimum amount. Throughout this tree, the changes in beliefs influence trade size choices considerably. This effect's strength depends on how

⁶For example, Roberts and Weitzman (1981), Weitzman, Newey, and Rabin (1981), and Dixit and Pindyck (1994, Chapter 10) discuss models where this mechanism is prominent.

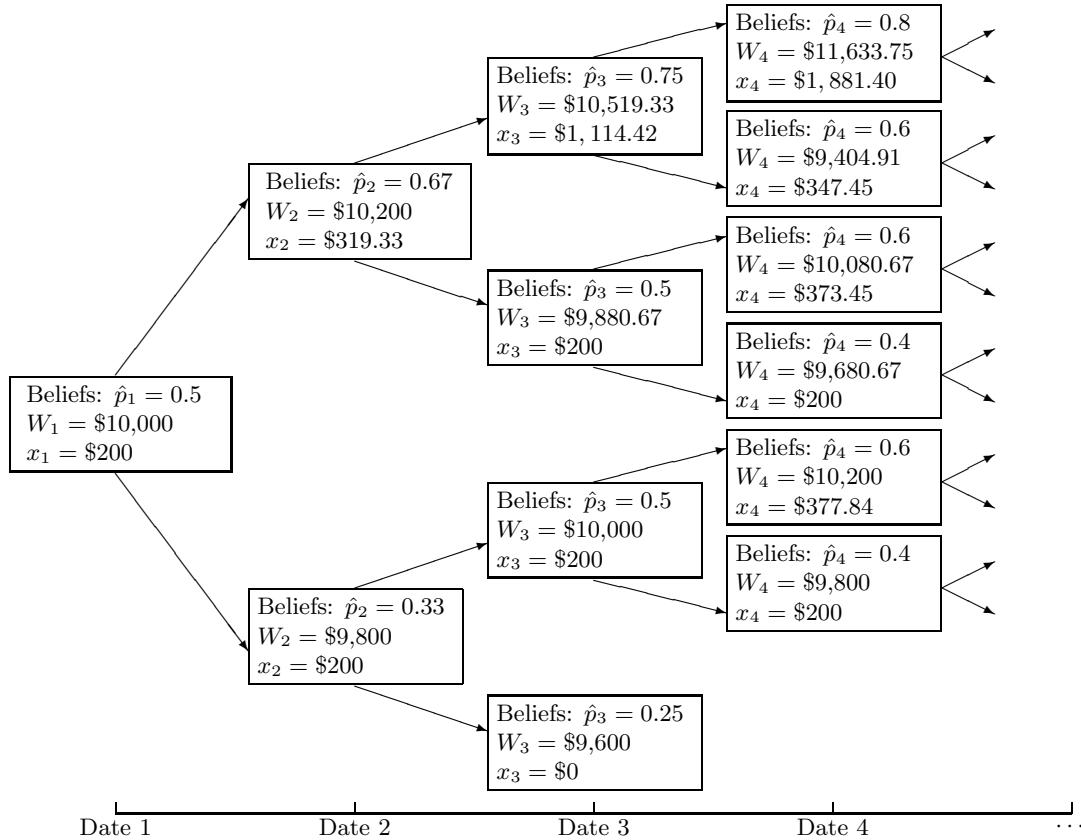


Figure 2: Optimal Investment Decisions with Trading Skill Uncertainty: An Example of Time-Series Changes in Behavior. This figure illustrates the optimal behavior of a risk-averse investor who is uncertain about his trading skills. The investor has an initial wealth of $W_0 = \$10,000$, a relative risk-aversion coefficient of $\gamma = 2$, and an investment horizon of $T = 16$. The minimum trade size is $\$200$. The investor in the model can trade amount $\$x$ at dates $t = 1, \dots, T - 1$. A trade is successful with a probability p and returns $2x$; it is a failure with a probability $1 - p$ and returns 0. The investor has a beta distributed prior about the value of p with parameters $\alpha_1 = 1$ and $\beta_1 = 1$. There is also a risk-free asset than pays no interest. The investor observes the outcome if he trades, in which case he updates his beliefs using the Bayes rule. All possible outcomes (with beliefs and wealth) are shown up to the investor's fourth trading decision.

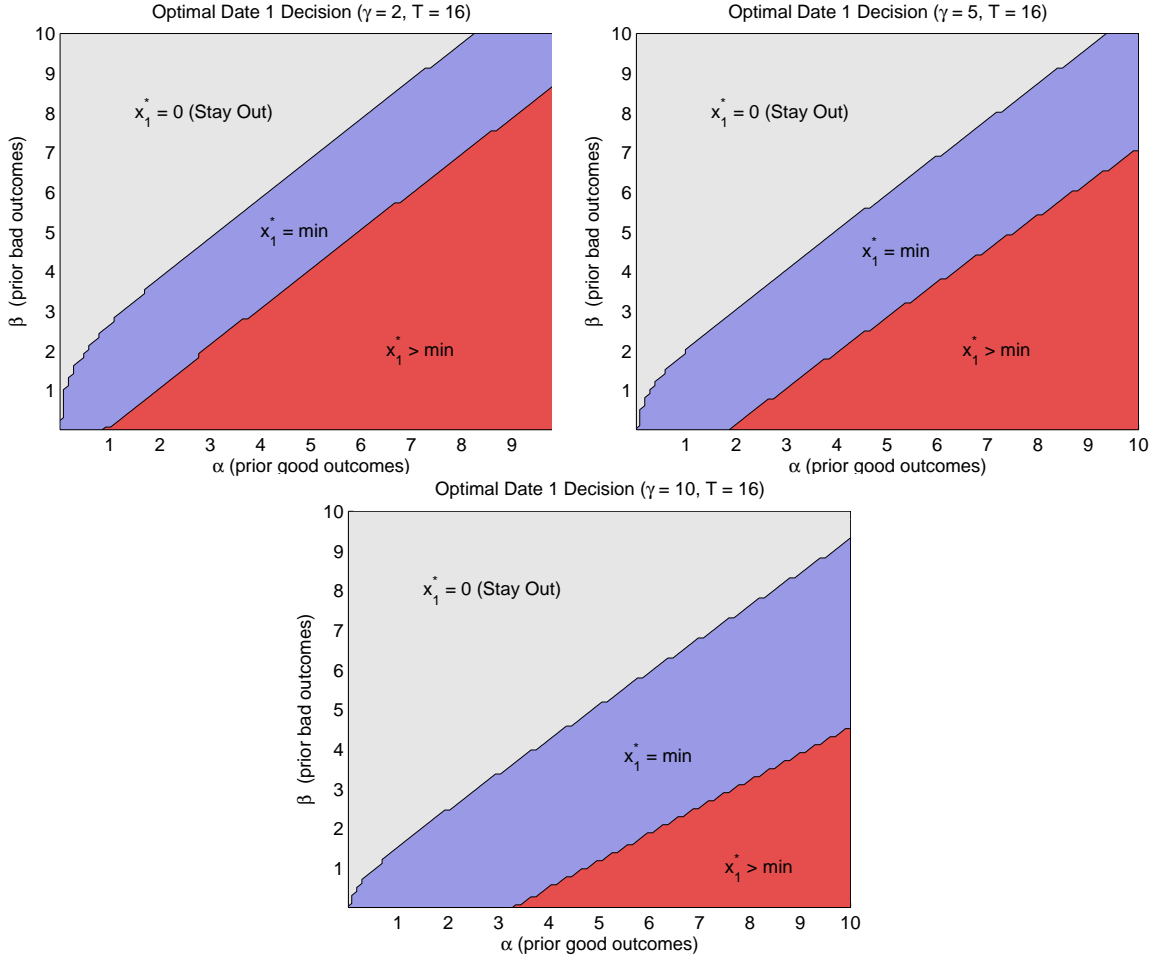


Figure 3: Optimal Initial Investment Decisions with Trading Skill Uncertainty. This figure illustrates the optimal behavior of a risk-averse investor who is uncertain about his trading skills. The investor has an initial wealth of $W_0 = \$10,000$ and an investment horizon of $T = 16$. The minimum trade size is \$200. The investor in the model can trade amount $\$x$ at dates $t = 1, \dots, T - 1$. A trade is successful with a probability p and returns $2x$; it is a failure with a probability $1 - p$ and returns 0. The investor has a beta distributed prior about the value of p with parameters $\alpha_1 = 1$ and $\beta_1 = 1$. There is also a risk-free asset than pays no interest. The investor observes the outcome if he trades, in which case he updates his beliefs using the Bayes rule. The three panels show investors' optimal decisions at date 1 as functions of the two parameters of the beta distributed prior, α_1 and β_1 . The mean and variance of the beta distribution are $\frac{\alpha}{\alpha+\beta}$ and $\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$, respectively; investors towards the right have more optimistic beliefs and investors close to the bottom-left corner have more dispersed priors than the investor close to the top-right corner. An investor above the diagonal believes that he is unskilled, $\hat{p}_1 < \frac{1}{2}$. The parameter space is divided into three regions: stay out, trade the minimum amount \bar{x} , or trade more than the minimum amount. The graphs are drawn for investors with relative risk aversion coefficients of 2, 5, and 10.

dispersed the prior is: if an investor's prior is very tight, one more observation affects beliefs very little, leading also to smaller changes in the trade size. This is a common characteristic of learning models: if belief changes induce changes in investor behavior, the changes should be stronger in the beginning when the prior is more dispersed.

Figure 2 illustrates how investors' initial choices depend on their prior beliefs and risk aversion. The graphs are drawn as functions of the two parameters of the beta distribution, α and β . (The mean of the prior increases in the ratio $\frac{\alpha}{\beta}$ while keeping $\alpha + \beta$ fixed, and the dispersion of the prior decreases in $\alpha + \beta$ while keeping the mean fixed.) The parameter space is divided into three regions: the first contains investors who never begin trading, the second contains investors whose first trade is the smallest possible, and the third contains investors whose first trade is larger than the minimum. Investors who lie under the 45 degree line believe that they are skilled ($\hat{p}_1 > \frac{1}{2}$) and usually trade more than the minimum amount. However, even some of these investors may stay out if they are risk averse enough: the negative intertemporal hedging demand offsets the positive myopic demand for these investors. Second, many mildly risk averse investors trade the minimum amount even if they are somewhat pessimistic about their skills ($\hat{p}_1 < \frac{1}{2}$). These are the investors who expect to lose money by trading, but who still trade because the value of another signal offsets their expected losses.

III. Data

A. Investor Trading Records and High-Frequency Traders

I use the Finnish Central Securities Depository data to test the learning model. This dataset records the portfolios and trading records from January 1, 1995 through November 29, 2002 of all household investors domiciled in Finland. The electronic records I use are exact duplicates of the official certificates of ownership and trades, and hence very reliable. Details on this dataset, which includes trades, holdings, and execution prices, are reported in Grinblatt and Keloharju (2000). The FCSD registry contains entries for approximately 1.1 million individual investors of whom 0.6 million traded during the eight-year sample period. In addition to stocks, the FCSD also records options and certain bond transactions. The investor file provides demographic data such as age, zip code, and primary language.

I limit the analysis to *high-frequency traders*. A high-frequency trader is a household investor who completes at least one intraday round-trip trade during the sample period. For example, if an investor buys 1,000 shares of a stock in the morning and sells some or all of these shares in the afternoon, the investor is classified as a high-frequency trader. Barber, Lee, Liu, and Odean (2004) call these investors “day traders” in their study of Taiwanese households. Although these traders are few in numbers—they represent 4.1% of all households who traded during the sample period—they are very active: 49.0% of all household trades and 63.8% of all household volume came from the investors.⁷

In addition to the round-trip trades, I also include high-frequency traders’ other short-term trades into the analysis. I identify these speculative transactions from a particular trading pattern: if an investor completes a round-trip trade in stock s on day t , I include all purchases from the same investor in the same stock s made during the next two weeks. Barber and Odean (1999, p. 49–50) use similar rules to separate speculative transactions from liquidity-motivated transactions.

High-frequency traders are well-suited for testing the learning model for a number of reasons. First, short-term speculative trades form a clean sequence of choices and outcomes: i.e., it is relatively transparent how the performance in past trades affects future behavior. An analysis of buy-and-hold investors would be more complicated because of the lack of natural sequencing; it would be difficult to say what the investor learns about his trading abilities and when. Second, by studying one-day returns, risk-adjustment is not an issue because intraday volatility swamps the *daily* risk premium. For example, an annual equity premium of 8% works out to a daily risk premium of just 0.03%. Thus, investors engaged in short-term trading try to profit from predicting price movements, or from acting as pseudo-market makers, and not from capturing the risk premium.

Third, high-frequency traders can learn about their abilities much faster than buy-and-hold investors. The difficulty with buy-and-hold investors is that these investors’ have unknown exposures to (possibly unknown) risk factors. When also these exposures and factors have to be estimated from the data, it becomes very difficult to disentangle luck from skill. Harris (2003, p. 453), for example, concludes that “[i]n practice, more than 20 years of returns data are typically required to obtain

⁷There are 128 high-frequency traders whose first intraday round-trip takes place before April 1995. I exclude these traders from this study for two reasons. First, I cannot measure intraday performance before April 1995 because these data are missing trade prices. Second, because these investors may have started short-term trading even before January 1995, I would not be able to control for their trading experience in the empirical analysis.

useful results for a given investment manager.”⁸ In contrast, the learning in short-term trading is akin to measuring whether a coin is biased towards the heads or not, with each trade representing a flip of a coin. Because this learning happens in *event* and not calendar time, a high-frequency trader resolves uncertainty much faster compared to a buy-and-hold investor. Harris and Schultz (1998) exploit the power of short-term trading in the same way and are able to draw inferences about the heterogeneity in SOES bandits trading abilities from just five trading days of data.

Finally, investors can bet almost symmetrically on one day up and down movements in stock prices. During the sample period, the Finnish brokers allowed customer to short shares with no additional contracts and low margin requirements: an investor could take a short position if he had 100% of the value of the short sale in either cash or portfolio holdings. This symmetry is useful: if short selling were difficult, I would not observe investors betting on negative signals. Fourth, these traders’ behavior holds the key to understanding “why do households trade so much” because they are responsible for almost two-thirds of all household trading volume.

Figure 4 shows information about the Finnish stock market and high-frequency traders’ round-trip trades. The upper figure is the level of the Finnish stock market index, the HEX portfolio index—it limits the weight of any one stock to 10% of the index. It returned an average of 12.4% per year during the sample period. This average masks the fact the market fluctuated considerably during this period: the sample includes not only the five years of the late-1990s bull market, but also more than two years of the post-Nasdaq crash period. The same graph also shows that the number of the round-trip trades is very small up to 1998 when several brokers started online services. The number of round-trip trades per month increased from fewer than 2,000 to over 10,000 in the early 2000.

Because the high-frequency traders try to profit from short-term price movements or from acting as pseudo-market makers, the second panel in Figure 4 shows how volatile individual stocks were during the sample period. I measure stock-specific volatility by first computing the average absolute *daily* log-return for each stock over each month. The solid line in the figure is the cross-sectional average of these stock-specific numbers. This average is around 2% for most of the period. The dashed line is the 90th percentile of the cross-sectional volatility distribution for each month. This measure may be more relevant to high-frequency traders who do not trade “average” stocks but the

⁸Fama and French (2008) take an extreme view of this problem and suggest that all results about mutual fund managers’ superior performance may be attributable to the difficulty to distinguishing between luck and skill. Looking at all mutual funds jointly, they find no evidence of managers trading on superior information.

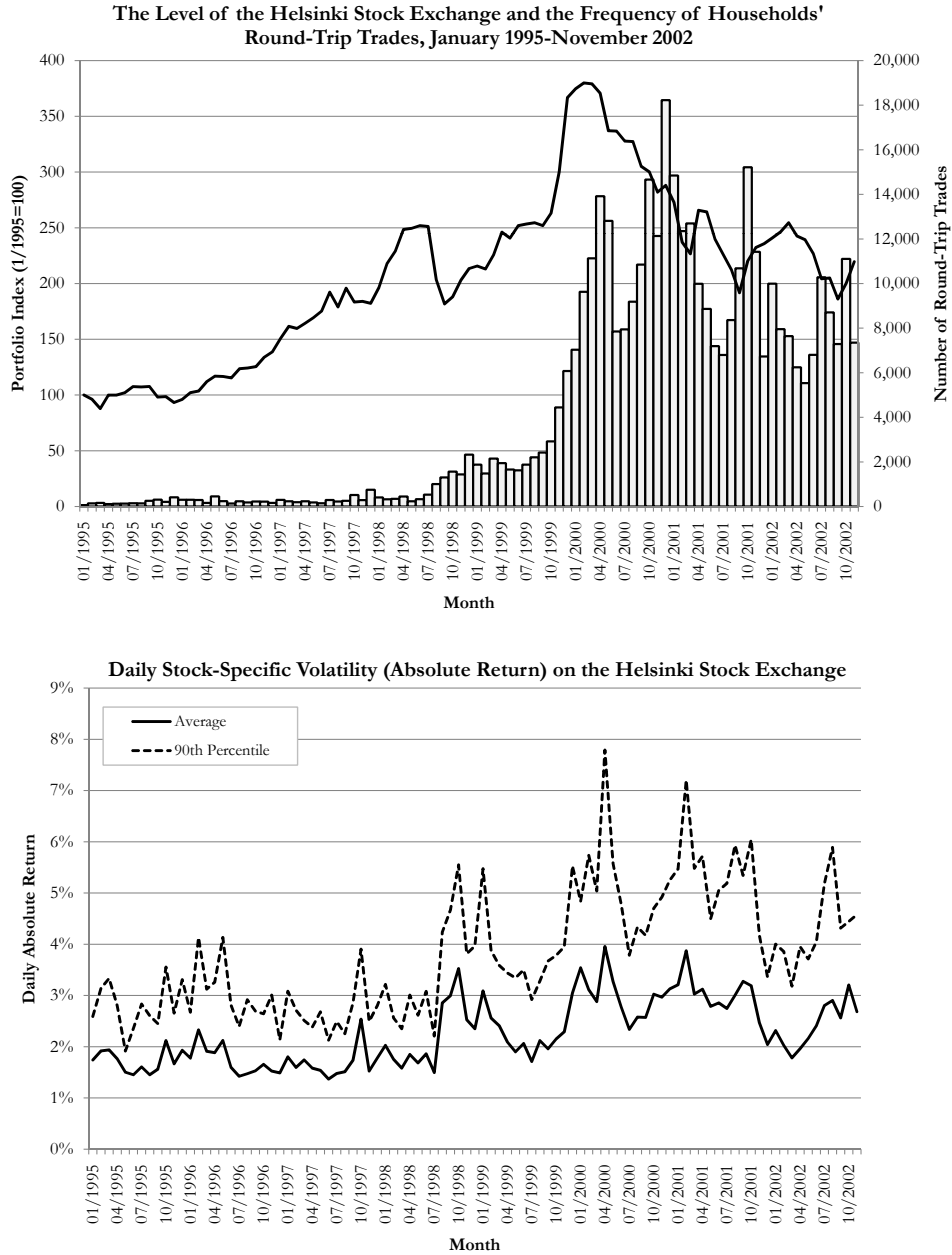


Figure 4: Market Returns, Daily Stock-Specific Volatility, and the Frequency of Households' Intraday Round-Trip Trades on the Helsinki Stock Exchange, January 1995—November 2002. This figure describes the behavior of the Finnish stock market during the sample period from January 1995 through November 2002. The line in the upper panel is the level of the Finnish Stock Market, HEX portfolio index. This is a value-weighted index that limits the weight of any one stock to 10% of the index. The bars show the monthly number of intraday round-trip trades from households. The lower panel shows how the stock-specific volatility evolved over the sample period. I measure stock-specific volatility by first computing the average absolute daily log-return for each stock over each month. The solid line is the cross-sectional average of these stock-specific numbers while the dashed line is the 90th percentile of the cross-sectional volatility distribution for each month. I remove the most thinly traded stock-days by requiring that a stock must have at least 10 trades per day. The average number of stocks in the monthly cross-sections is 76.3.

stocks that they expect to move significantly.

The peaks and troughs of the 90th percentile volatility series closely track the round-trip series from the upper panel. Although it is possible that households contributed to the increase in the stock-specific volatility—Campbell, Lettau, Malkiel, and Xu (2001) and Barber and Odean (2001b) suggest such possibility—the causality may run from higher volatility to higher trading activity.⁹ These volatility estimates suggest that investors *could* have profited from short-term trading: *if* an investor can predict price movements with some consistency, the magnitude of the average movement is large enough to cover transactions costs.

B. Variable Definitions

I measure the profitability of trade i on day t as

$$\text{gross profit}_{i,t} = \begin{cases} \sum_{s=1}^n (p_{i,s,t}^s - p_{i,s,t}^b) v_{i,s,t}^b & \text{if } v_{i,s,t}^b = v_{i,s,t}^s \\ \sum_{s=1}^n \left\{ (p_{i,s,t}^s - p_{i,s,t}^b) v_{i,s,t}^s + (p_{s,t}^c - p_{i,s,t}^b) (v_{i,s,t}^b - v_{i,s,t}^s) \right\} & \text{if } v_{i,s,t}^b > v_{i,s,t}^s \\ \sum_{s=1}^n \left\{ (p_{i,s,t}^s - p_{i,s,t}^b) v_{i,s,t}^b + (p_{i,s,t}^s - p_{s,t}^c) (v_{i,s,t}^s - v_{i,s,t}^b) \right\} & \text{if } v_{i,s,t}^b < v_{i,s,t}^s, \end{cases} \quad (10)$$

where n is the number of short-term trades from investor i on day t , $p_{i,s,t}^b$ and $p_{i,s,t}^s$ are the investor's average purchase and sale prices in stock s , $v_{i,s,t}^b$ and $v_{i,s,t}^s$ are the number of shares purchased and sold, and $p_{s,t}^c$ is the stock's closing price on the same day. The first line applies to the situation where the investors buys and sells the same number of shares the. The other two lines apply to the situations where an investor is left with a residual position; here, the remaining position $|x_{i,s,t}^s - x_{i,s,t}^b|$ is marked to market at the same day closing price.

I subtract commissions from gross profits to compute investors' net profits. Following Grinblatt and Keloharju (2008), I use the lowest available commission rate of 8.42 euros + 0.15% of trade value in these computations. The investor's day t return on short-term trading is then $r_{i,t} = \frac{\text{net profit}_{i,t}}{x_{i,t}}$,

⁹The comovement of the volatility series and the round-trip trading activity is statistically significant. A pooled regression of the absolute open-to-close price change on trading activity,

$$\left| \ln \frac{\text{closing price}_{i,t}}{\text{opening price}_{i,t}} \right| = a_i + b_1 \ln(\text{Number Of Trades}_{i,t} + 1) + b_2 \ln(\text{Number Of Round-Trip Trades}_{i,t} + 1) + \varepsilon_{i,t},$$

yields a slope of 0.0082 (s.e. = 0.0011) on the round-trip trading activity. (Subscripts i and t index stocks and dates, respectively; a_i 's are 211 stock fixed effects. The regression is estimated using 146,919 stock-day observations from time period January 1998 through November 2002. Errors are clustered by stock.) This slope estimate indicates that a move from no round-trip trades to 10 round-trip trades increases the daily absolute price movement by a factor of 1.09 even after controlling for other trading activity.

where $x_{i,t}$ is the size of the trade,

$$x_{i,t} = \begin{cases} \sum_{s=1}^n p_{i,s,t}^b v_{i,s,t}^b & \text{if } v_{i,s,t}^b \geq v_{i,s,t}^s \\ \sum_{s=1}^n p_{i,s,t}^s v_{i,s,t}^s & \text{if } v_{i,s,t}^s > v_{i,s,t}^b \end{cases} \quad (11)$$

This sets the trade size equal to the total value of shares purchased unless the investor's sales outnumber purchases.

C. Summary Statistics on High-Frequency Traders

Table I shows that high-frequency traders are quite different from other investors: they are younger and more often male than the rest of the investor population. These investors also trade more even after excluding round-trip trades. The median high-frequency trader trades 56 times while the median non-trader household only trades 2 times. The median number of round-trip trades is just 2, but there is significant variation across investors: while a quarter of the traders only complete one intra-day trade, the top 5% complete at least 63 round-trip trades.

High-frequency traders also concentrate their activity in high market beta stocks. The average betas are 1.53 and 1.42 for high-frequency traders' and other households' regular purchases, respectively. The stocks that are used for short-term trading have even higher average market beta of 1.80. The pairwise difference between the market betas of regular purchases and round-trip trades is 0.27 (s.e. = 0.01), indicating that investors choose riskier stocks for short-term trading. Although 39.8% of all the short-term trades take place in Nokia, many other stocks are also actively traded; for example, each of the 27 most popular stocks have at least 1,000 short-term trades in them.

High-frequency traders also trade options more often than other households; the participation rates in these securities are 14.8% and 1.7%, respectively. Unlike in the U.S., an investor in Finland does not need to sign additional contracts or demonstrate understanding of the financial markets to trade options; investors can trade options as easily as stocks.

The median high-frequency trader does not break even in short-term trading: he loses an average of just 20 euros per a round-trip trade. The profit distribution is, however, skewed to the right; because some traders earn substantial profits from short-term trading, the *average* investor's average profit is slightly positive at 18 euros per day.

Table I: Trading Activity and Demographics of High-Frequency Traders

A high-frequency trader is an investor who executes an intraday round-trip trade at least once during the sample. An intraday round-trip is a trade where an investor buys and sells (in any order) the same stock on the same day. The data are the complete trading records of all Finnish individuals from the Helsinki Stock Exchange for the period from January 1995 through November 2002. This table compares high-frequency traders' demographics and trading activity to other households. The number of trades, average trade sizes, and average betas are tabulated separately for regular purchases and round-trip trades. Each investor's average beta is computed using six months of daily data leading up to each purchase (or a round-trip trade), adjusted for nonsynchronous trading (Scholes and Williams 1977). *Warrants* is the fraction of investors who traded derivative securities at least once during the sample period. These securities include calls and puts as well as exchange-listed executive stock options.

Panel A: High-Frequency Household Traders, $N = 22,529$					
Variable	Mean	Std. dev	Percentiles		
			25	50	75
Age	39.5	13.8	29.0	38.0	49.0
Regular Trades					
Number of Trades	99.9	148.1	24.0	56.0	119.0
Average Trade Size, EUR	9,769.7	59,230.4	2,303.5	4,357.0	8,488.4
Average Beta	1.53	0.46	1.23	1.53	1.85
Intraday Round-Trip Trades					
Number of Trades	13.4	45.7	1.0	2.0	8.0
Average Trade Size, EUR	16,879.1	55,740.5	2,644.7	6,023.0	14,236.5
Average Beta	1.80	0.77	1.40	1.90	2.27
Average Profit, EUR	18.1	1,121.3	-81.5	-19.5	62.4
Men (%)	81.3%				
Warrants (%)	14.8%				
Panel B: Other Households, $N = 527,436$					
Variable	Mean	Std. dev	Percentiles		
			25	50	75
Age	46.2	18.6	33.0	47.0	59.0
Regular Trades					
Number of Trades	6.4	14.4	1.0	2.0	6.0
Average Trade Size, EUR	6,045.1	238,621.7	1,161.6	2,310.0	4,848.0
Average Beta	1.42	0.77	0.90	1.45	1.93
Men (%)	59.2%				
Warrants (%)	1.7%				

Overall, Table I indicates that high-frequency traders are quite different from other investors not only in terms of their demographics but also in their trading behavior. Such differences would arise in the learning model if these investor characteristics are correlated with either the means or the dispersions of priors, or with the risk-aversion coefficients. For example, the average high-frequency trader would be younger than the average non-trader if young investors are more uncertain about their abilities. Similarly, if women are more risk averse than men (see, e.g., Grinblatt and Keloharju (2008)), the learning model would predict that men are more likely to become high-frequency traders. The higher betas and the options trading could work through the same risk aversion channel: these less risk averse investors trade not only options and high beta stocks but also often become high-frequency traders.

IV. Testing the Implications of the Learning Model

The premise of the learning model is that some investors are skilled, some are not, and everyone is uncertain about his own abilities. Thus, I first test whether investors in the data exhibit persistent performance heterogeneity. I then study three trading patterns that are specific to the model: First, I measure whether investors are more likely to quit often after unsuccessful trades. If investors update their beliefs after each trade, negative outcomes always push investors towards more pessimistic beliefs. This leads some investors to quit trading. Second, I study whether investors increase trade sizes after successful trades and decrease them after unsuccessful trades. This prediction arises from the same mechanism as the first prediction: an investor's beliefs turn more positive after successful trades, making short-term trading a more attractive "asset," and increasing the demand for it. Third, I test whether investors initially make small exploratory trades. The intuition behind this prediction is the option value of trading. If there are investors who believe that they are unskilled and only trade to resolve uncertainty, they will trade the smallest possible amount to minimize expected losses. As time passes, investors' beliefs become more precise, weakening this motive for trading.

A. Performance Heterogeneity

If the learning mechanism is to explain investor behavior, some investors must believe that they *might* be skilled. If investors have reasonable (i.e., well-calibrated) priors, this means that some

investors *must* earn excess returns. If not, investors would have no incentives to trade to resolve uncertainty because they would know that they are unskilled.

Table II runs pooled performance persistence regressions using the data on high-frequency traders' short-term trades. The first specification in Panel A has the return on the j th trade as the dependent variable and the average career return (i.e., over trades $1, \dots, j - 1$) as the independent variable. The second specification in Panel B replaces returns with positive return-indicator variables, $\mathbf{1}(r_{i,t} > 0)$, and regresses this variable against the average "success rate" over the investor's earlier trades. If there are no differences in performance or if the differences are transitory, the slopes in these regressions would be zero. All regressions control for two types of fixed effects. First, I control for event-time effects by including dummies for trading dates—i.e., whether an investor trades for the second time, the third time, and so forth. Second, I control for calendar time effects by including monthly dummies.

(I estimate the model in the second specification as a linear probability model (LPM) because of the large number of fixed effects. A non-linear model such as a logit or probit would be subject to the incidental variables problem (see, e.g., Neyman and Scott (1948) and Hausman, Hall, and Griliches (1984)). Moreover, sources such as Wooldridge (2002, p. 456), advocate the use of linear probability models in situations such as this one where most RHS variables are indicator variables. Nevertheless, I confirm in unreported work that both probit and logit estimates are similar to those reported in the table.)

The estimates from both specifications indicate that some investors systematically outperform others. The coefficients for lagged performance are positive and significant, both statistically and economically. The full sample slope estimate of 0.205 (s.e. = 0.014) in the first specification indicates that an investor whose average career return is 1% expects to earn a return that is 0.205% higher than the return earned by an investor with a 0% career return. The results are similar in the second specification. For example, an investor whose every past trade has been a success has a 0.291 (s.e. = 0.008) higher probability of another success compared to an investor without any successes in the past.

The measurement errors in these regressions are heteroscedastic because the explanatory variables are sample averages with varying sample sizes. The last two columns in Table II split the sample in investors early ($t \leq 10$) and late ($t > 10$) trades to verify that this does not drive the

Table II: Heterogeneity and Persistence in High-Frequency Traders' Performance

This table reports estimates from regressions of traders' date t returns against their own lagged trading performance to measure performance persistence. Each observation is a single short-term trade, defined as a purchase and a sale of the same stock on the same day or a speculative purchase (see text). The data are the high-frequency traders' complete trading records from January 1995 through November 2002. Investors with any round-trip trades during the first three months of the sample are dropped (see text). Panel A regresses the return in trade t against the investor's average return from all previous trades. Panel B replaces continuous returns with positive-return indicator variables and regresses the success in the trade t against the investor's average success rate from all previous trades. Both regressions are estimated as pooled panels using OLS with calendar time and event-time fixed effects. The calendar time FEs are monthly dummies and the event-time FEs are dummy variables for trading dates $1, \dots, 50$; $t > 50$ trades are put in one category. The regressions are estimated in the full sample ($N = 280,549$), in a sample with investors' early trades ($t \leq 10$; $N = 74,593$), and in a sample with investors' late trades ($t > 10$; $N = 205,956$). Standard errors, clustered by investor, are in parentheses.

Panel A: OLS Panel Regression of Net Return (%) against Career Average Return			
	Full Sample	Early Career, $t \leq 10$	Late Career, $t > 10$
Career Average Return, $\frac{1}{t-1} \sum_{\tau=1}^{t-1} r_{\tau}$	0.205 (0.014)	0.167 (0.015)	0.386 (0.027)
Adjusted R^2	0.016	0.024	0.015
Std. dev. of Career Average Return	0.021	0.036	0.010
Panel B: OLS Panel Regression of Success-Dummy against Career Success Rate			
	Full Sample	Early Career, $t \leq 10$	Late Career, $t > 10$
Career Success Rate, $\frac{1}{t-1} \sum_{\tau=1}^{t-1} s_{\tau}$	0.291 (0.008)	0.179 (0.006)	0.559 (0.016)
Adjusted R^2	0.020	0.028	0.025
Std. dev. of Career Success Rate	0.198	0.324	0.124

results. The slopes are statistically significantly positive in both samples. Moreover, controlling for changes in the standard deviation of the RHS variable, the results are also economically similar in the two samples. For example, a one standard deviation shock to the average career return in Panel A increases today's return by 0.61% in the early trades-sample and by 0.38% in the late-trades sample. In Panel B, similar unit shocks increase success probabilities by 0.058 and 0.069, respectively.

B. Exits after Unsuccessful Trades

Table III examines whether investors are more likely to quit short-term trading after unsuccessful trades. The table reports estimates from LPM regressions of an exit-indicator variable against investor's performance in trade t as well as his average career performance. Because an investor who loses enough capital may be forced to quit even if there is no learning, the regressions also controls for the changes in investor's wealth. Specifically, I add the log-change in the investor's portfolio value from the time of the first trade to one day before the current trade to control for this alternative exit mechanism. The regressions also include event-time and calendar time dummies.

An important methodological issue is the definition of an "exit." The natural definition is the investor's last short-term trade. However, a concern with this definition arises from the right-censoring of the time-series data: an investor who takes a pause towards the end of the sample would be classified as having exited. I address this concern by ignoring all observations during the last three months of the sample from September 2002 through November 2002. Given that the average waiting time between an investor's two short-term trades is 15.4 days (the median is just 2 days), this three-month cutoff should eliminate many false exists. The calendar time fixed effects also help in this matter because they control for the (possible) mechanical increase in the exit rates towards the end of the sample. Importantly, any false exits would also favor the null hypothesis of no correlation between trade performance and exit decisions.

The regression estimates in Table III support the model's prediction that investors quit trading after negative outcomes. In the full sample, both the current trade performance as well as the average career performance are significant determinants of exit decisions. For example, a 5% return in Panel A lowers the quitting probability by 0.015 compared to the zero-return benchmark. The results are similar in Panel B that replaces returns with success dummies: a successful trade lowers the quitting probability by 0.030 relative to an unsuccessful trade.

Table III: A Pooled OLS Regressions of High-Frequency Traders' Exit Decisions against Performance

This table reports slope estimates from panel regressions of high-frequency traders' exit decisions against performance. Each observation is a single short-term trade, defined as a purchase and a sale of the same stock on the same day or a speculative purchase. The data are the investors' complete trading records from January 1995 through November 2002 from the Finnish FCSD registry. I drop investors with round-trip trades during the first three months of the sample and ignore trades during the last three months of the sample (see text). The dependent variable is set to one for each investor's last trade and to zero for all other trades. Panel A has the return from current trade as well as the average return from the earlier trades on the RHS. Panel B replaces continuous returns with success-indicator variables. Both regressions control for the change in investor wealth (as measured by portfolio values) from the first trade to one day before the current trade. Both regressions are estimated as pooled panels using OLS with calendar time and event-time fixed effects. The calendar time FEs are monthly dummies and the event-time FEs are dummy variables for trading dates $1, \dots, 50$; $t > 50$ trades are put in one category. The regressions are estimated in the full sample ($N = 281,618$), in a sample consisting of investors' early ($\tau \leq 10$) trades ($N = 92,255$), and in a sample consisting of investors' late trades ($N = 189,363$). Standard errors, clustered by investor, are in parentheses.

Panel A: Pooled OLS Regression of the Date t Exit Decision on Net Return (%)			
Explanatory Variables	Full Sample	Early Career, $t \leq 10$	Late Career, $t > 10$
Return in Trade t	-0.304 (0.015)	-0.474 (0.028)	-0.081 (0.012)
Career Average Return, $\frac{1}{t-1} \sum_{\tau=1}^{t-1} r_{\tau}$	-0.074 (0.032)	-0.078 (0.037)	0.051 (0.041)
$\ln(\text{Wealth}_t) - \ln(\text{Wealth}_0)$	0.000 (0.000)	-0.005 (0.001)	-0.001 (0.000)
Adjusted R^2	0.161	0.121	0.019
Panel B: Pooled OLS Regression of the Date t Exit Decision on Success Dummy			
Explanatory Variables	Full Sample	Early Career, $t \leq 10$	Late Career, $t > 10$
Success/Failure in Trade t	-0.030 (0.001)	-0.075 (0.002)	-0.007 (0.001)
Career Success Rate, $\frac{1}{t-1} \sum_{\tau=1}^{t-1} s_{\tau}$	-0.027 (0.003)	-0.034 (0.004)	0.003 (0.003)
$\ln(\text{Wealth}_t) - \ln(\text{Wealth}_0)$	0.000 (0.000)	-0.005 (0.001)	-0.001 (0.000)
Adjusted R^2	0.163	0.128	0.020

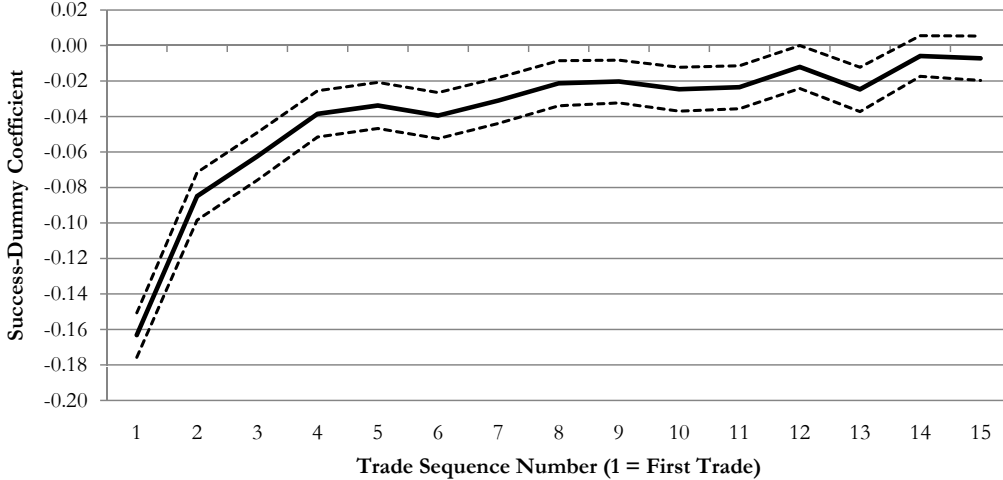


Figure 5: **Cross-Sectional Regressions of High-Frequency Traders’ Exit Decisions against a Success Indicator Variable.** This figure plots the slope estimates from regressions of investors’ exit decisions against a success-indicator variable. The regressions are run separately for each cross-section: investors trading for the first time, investors trading for the second time, and so forth. The regressions also control for the investor’s average career success rate and the change in wealth since the first trade and the regressions include monthly dummies for calendar time fixed effects. The dashed lines show the 95% confidence interval for the slope estimates.

Although the change in wealth is statistically significant in some regressions, the effect is economically small. For example, if an investor’s portfolio value increases (in log-terms) by 100%, the exit probability decreases by only 0.005. This suggests that unsuccessful trades influence exit decisions mainly through their effect on beliefs and not indirectly through their wealth effects.

The learning model predicts that investors’ early experiences should influence exit decisions more than (possible) later experiences. The longer an investor has traded, the sharper his posterior is and the less the investor pays attention to new data. For example, suppose that an investor in Section II’s model believes that $\hat{p}_t = 0.5$. Then, if the standard deviation of the prior is 0.2, one new positive observations leads the investor to update to $\hat{p}_{t+1} = 0.58$. However, if the standard deviation of the prior is 0.02 instead, a successful trade leads to a posterior of just $\hat{p}_{t+1} = 0.5008$.¹⁰ This deceleration of learning-result is a common feature of learning models; see, e.g., Harvey (1989) and Pratt, Raiffa, and Schlaifer (1995).

The early- and late-career estimates in Table III are consistent with this deceleration of learning-effect. For example, a positive outcome in the early sample lowers the exit probability by 0.075 but only by 0.007 in the late sample. The importance of the average career return variable also

¹⁰A beta distributed random variable with a mean of 0.5 has a standard deviation of $\sigma = \frac{1}{2\sqrt{2\alpha+1}}$. Thus, the parameters of the beta distribution are $\alpha_t = \beta_t = 2.625$ in the first case and $\alpha_t = \beta_t = 312.017$ in the second case.

disappears when moving from the early trades-sample to the late trades-sample. This is true for both the continuous return and indicator variable specifications.

Figure 5 further examines the changes in outcome sensitivity by plotting slope coefficients from separate cross-sectional regression for each $\tau = 1, \dots, 15$ (i.e., investors' first trades, second trades, and so forth). The coefficients increase rapidly at first and then tend more slowly towards zero. The concave pattern of the coefficients supports the deceleration of learning-argument. For example, the difference in the exit probabilities between successful and unsuccessful investors is 0.172 in the first trade but only 0.087 in the second trade. The changes in the slope estimates are statistically significant at a 5% level up to the fourth trade.

C. Trade Size Changes

Table IV reports slope estimates from regressions that measure how outcomes influence investor's trade size decisions. Because the investor in the learning model revises his beliefs after each trade, short-term trading becomes a more attractive investment opportunity after successful trades and less attractive after unsuccessful trades. The regressions are AR(1) type of models with the log-size of the trade t as the dependent variable and the log-size of trade $t - 1$ and the return in trade $t - 1$ as independent variables. I *condition* the regressions on trading: i.e., if an investor quits after trade t , I do not include quitting (zero trade size) as another observation. Because investors are more likely to quit after unsuccessful trades, this analysis understates the true effect that outcomes have on trade sizes. However, this is a desirable procedure because Table III already established the outcome-exit relation; the purpose of these regressions is to examine whether outcomes affect trade sizes even after ignoring exits.

The slope estimates indicate that past outcomes influence trade size choices significantly. Both the continuous return and indicator variable specifications support this conclusion. For example, the (unrestricted) full sample estimates that an investor with a successful trade increases the trade by $e^{0.084} - 1 = 8.8\%$ relative to an investor who experiences a failure. The results in Panel B are consistent with the deceleration of learning-argument: more experienced investors pay less attention to new data compared to the behavior of inexperienced investors. For example, an early sample investor who makes a successful trade places a trade that is 10.1% larger than the trade that he would make after a failure. This trade size effect is on top of the earlier exit decision-result. The difference in the slope estimates between the early and late trades-samples is statistically significant.

Table IV: A Pooled Regression of High-Frequency Traders' Trade Size Choices on Performance

This table runs pooled least squares regressions of high-frequency traders' log-trade sizes against two measures of past performance. Each observation is a single high-frequency trade, defined as a purchase and a sale of the same stock on the same day or a speculative purchase. The data are the investors' complete trading records from January 1995 through November 2002 from the Finnish FCS D registry. I drop investors with round-trip trades during the first three months of the sample. Panel A regresses the log-size of trade $t + 1$ against the size and return of trade t . Panel B replaces the continuous return with a success-indicator variable. The regressions also include the log-change in the investor's portfolio value from trade t to $t + 1$. An observation is dropped if the investor quits after trade t . The regressions are estimated as pooled panel regressions using OLS with calendar time and event-time fixed effects. The calendar time FEs are monthly dummies and the event-time FEs are dummy variables for trading dates $1, \dots, 50$; $t > 50$ trades are put in one category. The regressions are estimated in the full sample, in a sample consisting of investors' early trades, and in a sample consisting of investors' later trades. There are 280,867 observations in the full sample, 79,635 in the early trades-sample, and 201,232 in the late trades-sample. Standard errors, clustered by investor, are in parentheses.

Panel A: Pooled OLS Regression of Date $t + 1$ Log-Trade Size on Net Return (%)			
Explanatory Variables	Full Sample	Early Career, $t \leq 10$	Late Career, $t > 10$
Return in Trade t	0.703 (0.176)	0.531 (0.094)	0.840 (0.323)
Date t Log-Trade Size	0.691 (0.006)	0.656 (0.004)	0.702 (0.007)
Adjusted R^2	0.562	0.511	0.561
Std. dev. of Return in Trade t	0.040	0.049	0.036

Panel B: Pooled OLS Regression of Date $t + 1$ Log-Trade Size on Success Dummy			
Explanatory Variables	Full Sample	Early Career, $t \leq 10$	Late Career, $t > 10$
Success/Failure in Trade t	0.090 (0.004)	0.101 (0.006)	0.086 (0.005)
Date t Log-Trade Size	0.691 (0.006)	0.654 (0.004)	0.702 (0.007)
Adjusted R^2	0.562	0.512	0.562

Although the direction of the result in Panel A is the opposite—investors appear to become more, not less, sensitive to outcomes as they gain experience—the difference is insignificant because of the noisy coefficient estimate in the late trades-sample.

D. Exploratory Trades

An investor may trade in the learning model even if he expects to lose money because of the option value of trading: the value of another signal can offset the expected loss. The cross-sectional trade size distribution of high-frequency traders' first trades in Figure 6 suggests that this mechanism may also motivate trading in the actual investor data. Many of these initial trades are very small given the costs of short-term trading. To emphasize this observation, the figure overlays the implied break-even returns: i.e., the lower bound for the price change that lets the investor recover the commissions. For example, an investor who buys just EUR 1,000 worth of shares for short-term trading must sell at a 1.99% higher price to break-even after commissions.¹¹ This is a counter-intuitively large price movement: if an investor expects to predict correctly a price movement of more than 1.99%, why does not he trade more? A reasonable solution is that these very small trades are exploratory and not speculative trades. The histogram shows that a non-trivial number of the initial trades fit this description. For example, 13.2% of the initial trades are so small (\leq EUR 1,406) that the implied break-even return is at least 1.5%.

Panel B in Figure 6 shows how the cross-sectional average trade size changes in event time; i.e., what is the average (normalized) trade size for investors trading for the second time, for the third time, and so forth. The graph shows that the investors who continue trading increase their trade sizes significantly: an investor who goes on to trade for the second time increases the trade size, on average, by 82%. Two mechanisms in the learning model would generate this increase. First, those traders whose first trade is successful become more optimistic about their abilities, inducing them to trade more at date 2. Second, the traders whose first trade is unsuccessful become more pessimistic about their abilities. Some of these traders quit, shifting the date 2 trade size distribution to the right. Moreover, many of those who do not quit may still trade the same amount as they did at date 1 if the minimum trade size constraint was already binding.

Because it is the unsuccessful traders who quit from the left-tail of the trade size distribution,

¹¹This is based on the lowest available commission rate of EUR 8.42 + 0.15% of the value of the trade. If x is the initial investment, the implied break-even return, \hat{r} , is defined implicitly by $\hat{r} = \frac{2 \times 8.42 + 0.15\%x + 0.15\%(1+\hat{r})x}{x}$. The implied break-even return is then $\hat{r} = \frac{16.84 + 0.3\%x}{0.9985x}$.

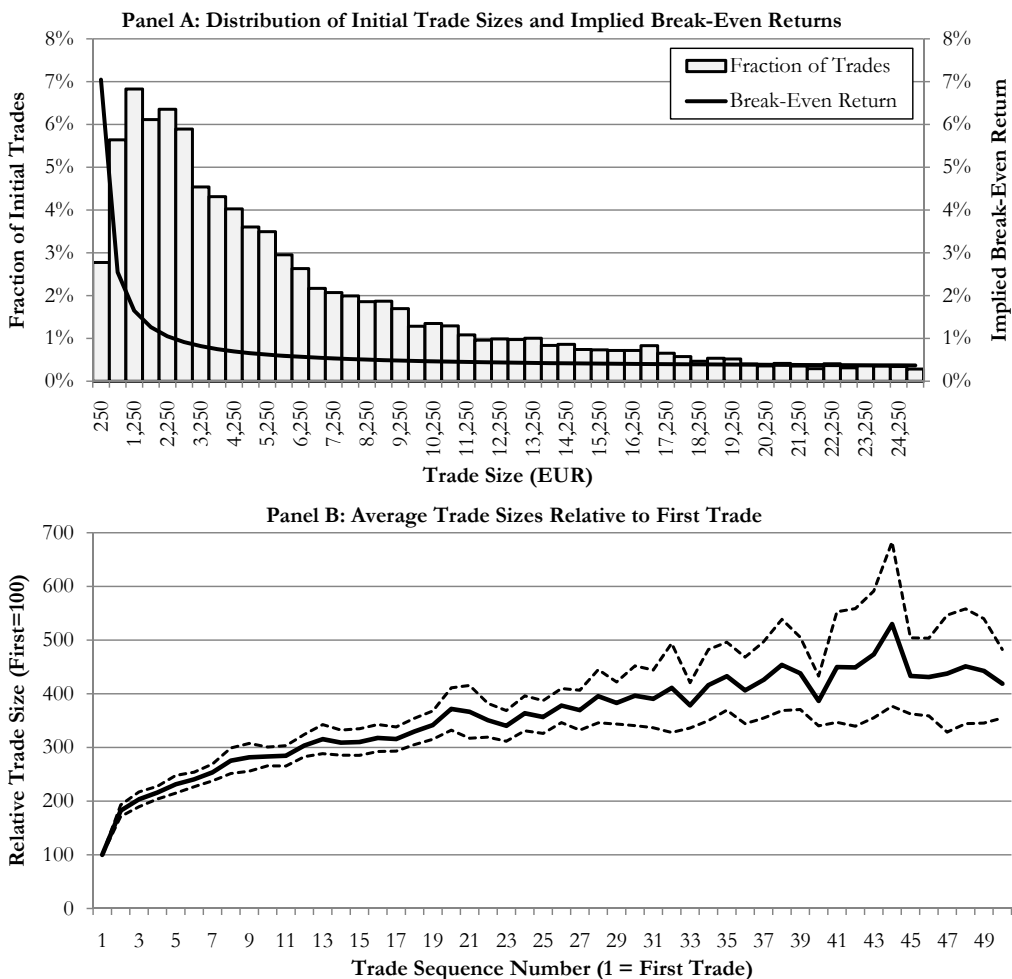


Figure 6: **Distribution of Initial Trade Sizes, Break-Even Returns, and Changes in Trade Sizes over Time.** Panel A shows the cross-sectional trade size distribution for high-frequency traders' first trades. Each category is EUR 500 wide with x -axis labels giving the midpoint of each category. The thick line is the implied break-even return: it is the percentage difference between the purchase and sale price is required for the investor to recover the EUR $8.42 + 0.15\%$ commission. This is computed as $\hat{r} = \frac{16.84 + 0.3\%x}{0.9985x}$, where x is the initial trade size. Panel B shows the cross-sectional average trade size conditional on the sequence number of the trade: investors trading for the first time, investors trading for the second time, and so forth. All trade sizes are normalized by the size of the same investor's first trade. The dashed lines denote the 95% confidence interval.

the trade size increases in the graph are conditional on survival; e.g., the date 5 trade size is the average trade size over those investors who survive up to date 5. This selection explains entirely the upward drift in Figure 6: the average *unconditional* trade size change over the entire panel ($N = 280,861$) is negative, -0.72% (s.e. = 0.06%). This indicates that it is indeed the surviving investors who increase their trade sizes; those investors who quit *decrease* their trade sizes before finally exiting.

V. Simulated Moments Parameter Estimation

This section tests whether the learning model can explain aggregate patterns in household behavior. I embed Section II’s trading model into a hierarchical model with population-wide distributions for investors’ prior beliefs. I use this model to estimate the beliefs for the entire population: I draw hypothetical investors from the belief distributions, put them through the trading model, and then compare their behavior against actual investor behavior. I estimate the model by requiring that (1) the fraction of investors who try high-frequency trading is close to the fraction in the data and that (2) the investors who try high-frequency trading respond to outcomes in the same way as they do in the data. This modeling approach assumes that there is a continuum of investors: those who become high-frequency traders are those who happen to inhabit particular corners of the prior beliefs-space.

A. Hierarchical Model of the Population

The first part of the hierarchical model is Section II’s trading model: an investor with a risk-aversion coefficient γ , initial wealth W_1 , trading horizon T , and prior beliefs (α_1, β_1) trades the optimal amount—or decides not to trade at all. The second part of the hierarchical model consists of distributions for investors’ prior beliefs. I fix investors’ other attributes (e.g., the initial wealth W_0) to reduce the number of parameters to estimate.

I describe population beliefs parsimoniously with four parameters. I draw the mean of each investor’s prior distribution from a beta distribution with parameters (α_μ, β_μ) . This distribution has the same $(0, 1)$ -support as the mean of the prior distribution itself. I draw the *relative* variance of each investor’s prior distribution from another beta distribution with parameters (α_v, β_v) independently of the mean. I model relative, not absolute variances because the support of the variance in beta distribution is a function of the mean: if the mean is \hat{p}_1 , the variance cannot be higher than

$\hat{p}_1(1 - \hat{p}_1)$. Relative variance as an index that runs from 0 to 100%; if an investor’s relative variance is v , the variance of the investor’s prior distribution is set to $v \times \hat{p}_1(1 - \hat{p}_1)$. Note that because a beta distributed variable has a variance of $\sigma^2 = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$ and a mean of $\mu = \frac{\alpha}{\alpha+\beta}$, the relative variance is $v = \frac{\sigma^2}{\mu(1-\mu)} = \frac{1}{\alpha+\beta+1}$; this is just the inverse of the “sample size”, $\alpha + \beta + 1$. The support of the beta distribution is again the same as the support of the relative variance.

B. Identifying Structural Parameters from Conditional Exit Rates

Because the hierarchical model does not yield closed-form estimation equations, I use Simulated Method of Moments for indirect inference. I draw hypothetical investors from the trial belief distribution, put them through the trading model, and finally compare model-implied behavior against the actual behavior in the data.

I identify the structural parameters from investors’ conditional exit rates. These rates measure how many investors quit high-frequency trading after a particular sequence of outcomes. These rates facilitate estimation of the belief parameters because they depend on both the mean and the variance of the prior distribution. An increase in the mean of the prior, given everything else, increases myopic demand, which in turn decreases the exit propensity. Similarly, an increase in the variance of the prior decreases exit rates by increasing the option value of trading. An increase in the variance also makes the posterior place less weight on the prior, leading to more extreme belief shifts. For example, an investor with a very uninformative prior (α, β close to zero) might trade at date 0 because of the high option value of trading, but would almost certainly quit after just one unsuccessful trade because the posterior would be close zero with high precision.

I use the conditional exit rates in the following five binomial tree nodes as moment conditions: 0, d , dd , ddd , and $dddd$. Node 0 is the initial node: it is set to zero for investors who make at least one high-frequency trade and to one otherwise. I use these five moments to estimate the four parameters ($\alpha_\mu, \beta_\mu, \alpha_v, \beta_v$) to have an overidentified model. I also report the estimation results separately for a case where I add a sixth moment condition, average investor skill. Each investor’s skill is estimated from the data as the average success rate over the investor’s all short-term trades.

Two considerations motivate these moment conditions. First, because I fix each investor’s initial wealth and the minimum trade size, I want to avoid moment conditions written on levels. For example, because CRRA investors trade a fraction of their wealth, any moment conditions written

on trade sizes would require us to estimate each investor’s wealth from the data. Second, because I solve a finite horizon model, I use investors’ early decisions for identification. I assume that by having a horizon T large enough relative to the timing of the last moment condition (dated at date 4), a turnpike result holds: i.e., that the investor behavior in the early nodes is the same as what it would be for any $T' > T$ (see, e.g., McCardle and Winkler (1989) for a turnpike result in a learning model).

I use Simulated Method of Moments (SMM; see, e.g., McFadden (1989) and Pakes and Pollard (1989)) to estimate the four parameters describing the investor population. Suppose that an investor with an attribute pair (μ, v) generates a 5×1 vector of moments denoted by $H(\mu, v)$. Then, the model implied k^{th} moment can be computed by integrating over the two distributions:

$$m_k(\theta) = \frac{\int_0^1 \int_0^1 B(\mu; \alpha_\mu, \beta_\mu) B(v; \alpha_v, \beta_v) \Pr(k|\mu, v) H_k(\mu, v) d\mu dv}{\int_0^1 \int_0^1 B(\mu; \alpha_\mu, \beta_\mu) B(v; \alpha_v, \beta_v) \Pr(k|\mu, v) d\mu dv}, \quad (12)$$

where $B(\cdot)$ is the probability density for the beta distribution and $\Pr(k|\mu, v)$ is the probability that an investor with beliefs (μ, v) reaches the node described by the k^{th} moment condition. Except for the first moment, the expectations in (12) are conditional on the investor reaching the relevant node in the binomial tree. For example, all investors reach the zero node (the initial “enter or do not enter” node) with probability one, so $\Pr(1|\cdot) \equiv 1$ for all investors. I set $\Pr(k|\mu', v') \equiv 0$ for all moment conditions that an investor with beliefs (μ', v') cannot reach. For example, investors who do not begin trading at all cannot reach node d , so $\Pr(2|\cdot) = 0$ for such investors. Because I assume that investor’s prior beliefs are well-calibrated, the probability that a (surviving) investor reaches, say, node dd is $(1 - \mu)^2$.

Because I cannot compute the elements of $H(\mu, v)$ in closed form, I compute both $H(\mu, v)$ and $\Pr(\mu|m, v)$ over a 500×500 grid and evaluate the integrals in (12) numerically using Simpson’s Rule:

$$m_k(\theta) \approx \frac{\sum_{i=1}^N \sum_{j=1}^N V(i, j) B(\mu_i; \alpha_\mu, \beta_\mu) B(v_j; \alpha_v, \beta_v) \Pr(k|\mu_i, v_j) H_k(\mu_i, v_j)}{\sum_{i=1}^N \sum_{j=1}^N V(i, j) B(\mu_i; \alpha_\mu, \beta_\mu) B(v_j; \alpha_v, \beta_v) \Pr(k|\mu_i, v_j)}, \quad (13)$$

where $V(i, j)$ is a multiplier specific to the Simpson rule.¹²

¹²Simpson’s rule (see, e.g., Judd (1998, p. 253–254)) approximates the integral over the interval $[a, b]$ by

$$\int_a^b f(x) dx \approx \frac{b-a}{6} \left(f(a) + 4f\left(\frac{a+b}{2}\right) + f(b) \right)$$

A composite Simpson’s rule divides the main interval (or in this case, area) into subintervals and uses the basic rule

I find the structural parameters θ by maximizing a GMM-type objective function,

$$\hat{\theta} = \arg \min_{\theta} \left(\hat{M} - m(\theta) \right)' W \left(\hat{M} - m(\theta) \right), \quad (14)$$

where \hat{M} is the vector of estimated moments from the data, $m(\theta)$ are the model implied moments from (12), and W is an arbitrary positive-definite weighting matrix. I use the optimal weighting matrix for estimation.

I adjust the four exit-related sample moments in \hat{M} to account for spurious exits; i.e., for exits that are driven by reasons outside the model. Although an investor in the theoretical model would never quit after a good outcome, some investors in the data do stop after successes. For example, 15.9% of investors with a successful second trade stop trading. The exit rate among investors losing for the second time in a row is 30.7%. I adjust all exit rates by subtracting the average exit rate of the *successful* investors who quit on the same date. In this example, the adjusted exit rate for the *dd* node is $30.7\% - 15.9\% = 14.8\%$. These adjustment rates are first-stage estimates of the spurious, time-dependent exit rates. Within the context of the model, this method of estimating the spurious exit rates is valid because the successful investors *must* have quit for reasons outside the model.

I set the non-belief related parameters of the model as follows: (1) the trading horizon is set to $T = 16$, (2) the initial wealth is set to $W_0 = \$20,000$, (3) the relative risk-aversion coefficient is set to $\gamma = 2$, and (4) the minimum trade size set to $\bar{x} = \$300$. The values for initial wealth and minimum trade size are motivated by the data: the median portfolio value of an investor trading for the first time is EUR 17,100 and the average loss for these investors is EUR 327. (This is the average computed over investors who lost money in their first trade. Because the trading outcome in the model is binomial—i.e., the investor gains or loses x_t —I interpret the minimum trade size as the minimum loss conditional on an unsuccessful outcome.)

C. Results

Table V compares the moment conditions between the data and the simulation. The learning model does a good job matching the conditional exit rates in the data. For example, 20.50% of the investors

separately for each of the subintervals. I construct the grid in an adaptive way, making the μ -grid denser around $\mu = \frac{1}{2}$ and the v -grid denser around $v = 0$. $V(i, j)$ in (13) is the area of the parameter space around grid point (i, j) multiplied by the appropriate factors (i.e., 1, 2, or 4) from the univariate composite Simpson rule that depend on the indexes i and j .

Table V: Moment Conditions and Structural Parameter Estimates of the Hierarchical learning Model

This table compares moments from the data to the moments simulated from a hierarchical learning model (Panel A) and shows the structural parameter estimates (Panel B). The first level of the hierarchical model is a trading model in which a risk-averse investor updates beliefs about his trading abilities after each trade. The second level of the model consists of distributions for population beliefs. Each investor’s prior distribution is constructed by drawing the mean from a beta distribution with parameters (α_μ, β_μ) and the relative variance (= variance divided by the maximum possible variance) from a beta distribution with parameters (α_v, β_v) . All investors have the same initial wealth ($W_1 = \$20,000$), relative risk-aversion coefficient ($\gamma = 2$), and trading horizon ($T = 16$); the minimum trade size is set to \$300. The model is estimated using Simulated Method of Moments. The moment conditions in the first specification are the conditional exit rates in the following five binomial tree nodes: 0, d , dd , ddd , and $dddd$. Node 0 is the initial node: this moment is set to zero for investors who make at least one high-frequency trade and to one otherwise. The second specification adds the average skill as a 6th moment condition. Panel A compares the actual moments to the simulated moments. Panel B reports the four structural parameters with standard errors in parentheses. χ^2 is the test statistic from the test of overidentifying restrictions; its p -value is in parentheses. Panel C tabulates some investor characteristics from the estimated model. The statistics are reported for high-frequency traders (*traders*), other investors (*non-traders*), and for both groups combined (*everyone*). $100 \times E[v]$ is average of the relative variance; average $\Pr(p > 0.5)$ is the average probability that the investor assigns to being skilled; % of skilled investors is the percentage of investors with $p > 0.5$; and average skill is the average performance that would be estimated in the data—this is lower than the average prior (true skill) in the first column because of the reverse survival bias (see text). Each investor’s average skill is estimated as the average success rate up to the fifth trade.

Panel A: Comparison of Actual Moments and Simulated Moments							
Specification		Moment Conditions					Average Skill
		Conditional Exit Rates					
		Node 0	Node d	Node dd	Node ddd	Node $dddd$	
I	Actual	0.9590	0.2050	0.1554	0.1311	0.1001	
	Simulated	0.9593	0.2015	0.1577	0.1366	0.0952	
II	Actual	0.9590	0.2050	0.1554	0.1311	0.1001	0.4367
	Simulated	0.9588	0.2058	0.1438	0.1372	0.0879	0.4599

Panel B: Structural Parameter Estimates						
Specification		Distribution Parameters				χ^2
		α_μ	β_μ	α_v	β_v	
I	Estimate	17.179	29.442	0.470	51.848	8.44
	<i>s.e.</i>	(7.003)	(9.106)	(0.016)	(9.309)	(0.004)
II	Estimate	68.186	90.276	0.504	93.358	124.72
	<i>s.e.</i>	(14.095)	(16.047)	(0.014)	(9.290)	(0.000)

Panel C: Implied Distributions of Prior Beliefs						
Specification	Population Subset	Average Prior about p	$100 \times E[v]$	Average $\Pr(p > 0.5)$	% of Skilled Investors	Average Skill
I	Everyone	0.3685	0.8103	0.0572	3.36	
	Traders	0.5182	1.6523	0.7243	78.16	0.4687
	Non-Traders	0.3621	0.7746	0.0289	0.19	
II	Everyone	0.4303	0.4955	0.0831	3.80	
	Traders	0.5120	0.9754	0.7186	81.82	0.4599
	Non-Traders	0.4268	0.4749	0.0558	0.44	

in the data quit after the first unsuccessful trade while the model delivers a slightly lower exit rate of 20.15%. Although the differences between the actual and simulated moments are economically small, the test of overidentifying restrictions rejects the model with a p -value of 0.004.

The model encounters more difficulties in matching investor's average success rate in the second specification. The average success rate in the data is 43.7%; i.e., this is the average investor's unconditional probability of a successful trade. The model, however, delivers a somewhat higher average success rate of 46.0%. This leads to a stronger rejection of the model in the second specification. Although the first specification has no moment conditions to directly constrain average skill, this specification leads to an average skill that is quite close to the one obtained in the second specification; the last column in Panel C shows that the average skill is 46.9% in the first specification. Although the model does not precisely match the average skill of high-frequency traders, it errs to the conservative side: the model slightly overestimates traders' abilities. Because the second specification uses average skill as a moment condition—which is very informative about the prior distribution in the model—it yields considerably sharper estimates of α_μ and β_μ that control the means of investors' prior distributions,

Panel C in Table V tabulates investor characteristics based on the estimated model. These characteristics are computed for the entire population as well as conditional on the entry decision: one set of attributes for investors who begin trading and another set of attributes for those who stay out. Although the population is a continuum of investors, the conditional distributions are very different from one another because of the self-selection; an investor who becomes a high-frequency trader must either have a high prior about his skill \hat{p}_1 or he must be very uncertain about this skill. Conversely, an investor who stays out must either be pessimistic about his abilities (i.e., a low \hat{p}_1) or have a relatively tight prior.

A comparison of the conditional distributions indicates that both the mean and the variance of the prior play roles in turning some investors into high-frequency traders. Because of the similarities between the distributions in the two specifications, I use the parameter estimates from the second specification in the following discussion. The average high-frequency traders' prior about his skill is $\hat{p}_1 = 0.512$ in the simulations. Although above 0.5, this is a relatively low number: a trader with this prior expects to be correct about the direction of the price movement approximately 51 times out of 100 trades. The average belief among those who do not trade is considerably lower at $\hat{p}_1 = 0.362$. High-frequency traders also have more highly dispersed beliefs: the average relative

variance is 2.1 times higher for high-frequency traders than what it is for those who stay out. Given the statistics in Panel C, the *average* high-frequency trader believes that his true trading ability lies in the $[0.415, 0.608]$ interval with a 95% probability.¹³ The average investor who stays out, on the other hand, believes that his trading skill lies somewhere between 0.361 and 0.494 with a 95% probability.

The investors in the tails of the distributions are more interesting than the average investors of the two groups. For example, many high-frequency traders thought that they would lose money ($\hat{p}_1 < \frac{1}{2}$) but still decided to trade. Similarly, many investors who decided to stay out did so even though they were optimistic about their abilities ($\hat{p}_1 > \frac{1}{2}$). Given the assumption about unbiased priors, the belief distributions coincide with the true skill distributions in the population: i.e., if 10% of investors believe that they can make money by trading, exactly 10% of investors in a large population can make money by trading. (Of course, the genuinely skilled investors do not need to be those who *believe* that they are skilled.) Panel C in Table V reports two sets of values that investigate investors beliefs from the viewpoint: (1) column “Average $\Pr(p > 0.5)$ ” reports the average probability that investors assign to being skilled; (2) column “% of Skilled Investors” reports the percentage of investors whose mean of the prior is above 0.5.¹⁴

The estimates indicate that 81.8% of those who tried high-frequency trading were skilled; or, conversely, as many as 18.2% of high-frequency traders believed that they were not skilled but traded nevertheless because of the chance that they might be wrong. These numbers are based on the mean of the prior distribution; the average high-frequency trader believed that he was skilled with a probability of 0.72. Only 0.4% of the investors who stayed out were skilled—and the average investor in this group assigned a probability of just 0.06 to being skilled. However, because of the

¹³Given the mean of the prior $\hat{p}_1 = 0.5120$ and the relative variance of $v = 0.9754/100$, this investor’s prior is distributed $B(52.0, 49.5)$.

¹⁴Here are the computational details. If an investor’s prior distribution is $\text{Beta}(\alpha, \beta)$, the investor believes that he is skilled with probability $\int_{\frac{1}{2}}^1 B(x; \alpha, \beta) dx$. Because the structural model parameters are distributions for the means and relative variances (instead of α ’s and β ’s directly), I first compute the parameters α and β from the means and variances. An investor with a prior of \hat{p} and a relative variance of prior of v has a prior with parameters

$$A = \hat{p} \left(\frac{\hat{p}(1 - \hat{p})}{\sigma^2} - 1 \right), \quad B = (1 - \hat{p}) \left(\frac{\hat{p}(1 - \hat{p})}{\sigma^2} - 1 \right), \quad \text{where } \sigma^2 \equiv v \hat{p} (1 - \hat{p}).$$

The average amount of mass in each investor’s skill distribution above the $\hat{p} = \frac{1}{2}$ cutoff point is then

$$E[\Pr(p > 0.5)] = \int_0^1 \int_0^1 \int_{\frac{1}{2}}^1 B(\mu; \alpha_\mu, \beta_\mu) B(v; \alpha_v, \beta_v) B(x; A, B) d\mu dv dx, \quad (15)$$

where $\alpha_\mu, \beta_\mu, \alpha_v$, and β_v are the model parameter estimates.

relative sizes of the two groups (i.e., those who trade and those who do not), these figures imply that $\frac{(0.44\%)(95.88\%)}{(0.44\%)(95.88\%)+(81.82\%)(1-95.88\%)} = 11.12\%$ of all skilled investors in the population never traded to profit from their abilities.

The overall fraction of skilled investors in the population is 3.8%. Importantly, this number is well within the bounds implied by, e.g., Coval, Hirshleifer, and Shumway (2005) who find some persistent superior performance among the investors in the top decile. It is also consistent with Grinblatt, Keloharju, and Linnainmaa (2008) that finds superior performance among individual investors in the highest IQ stanine category (i.e., the top 4%).

An interesting application of the structural model is to use it to examine the learning by survival-mechanism. In the trading model, investors do not learn how to trade: their true skill p is fixed and they only *discover* their true skill as they trade repeatedly. However, because (1) investors with low p 's are more likely to be unsuccessful and because (2) unsuccessful investors are more likely to quit, the investors who survive are better than what the average investor is in the beginning.

I can examine directly how this selection process works over time by inserting the structural parameter estimates back into the model. An unskilled investor (i.e., with a true skill $p < 0.5$) who begins trading quits after the first trade with probability 0.418. The probability that such an investor quits after the first *or* the second trade is 0.516. Thus, over half of the unskilled investors reach the correct conclusion by the second trade. In contrast, these cumulative exit rates are only 3.0% and 5.4% for the skilled investors; i.e., 5.4% of the skilled investors receive bad draws of luck and exit despite being (unbeknownst to them) skilled. The differences in these exit rates alter the composition of the trader pool, increasing the skill of the average investor: whereas 18.2% of the investors trading for the first time are unskilled (Table V, Panel C), this fraction decreases to $\frac{18.2\% \times (1-41.8\%)}{18.2\% \times (1-41.8\%) + (1-18.2\%) \times (1-3.0\%)} = 11.8\%$ after the first trade and to 10.2% after the second trade.

Because the skilled traders are more likely to survive, a regression of trade performance on trading experience would produce a positive coefficient despite the fact that everyone's skill is fixed. Similarly, a regression of trade size changes on performance would yield a positive coefficient even when shutting down the survival mechanism. For example, the average investor who reaches the *uu* node places a trade that is 11.3% larger than his first trade; the average investor who reaches the *dd* node and keeps on trading places a trade that is 20.8% smaller than his first trade.

The fact that investors quit after unsuccessful trades leads to a reverse survivorship bias when

measuring investors' abilities from the data. If investors never quit, the average skill measured from the data would be the same as the mean of the prior given the unbiased priors-assumption. However, because investors quit after unsuccessful trades, the skill estimates from the data are biased downwards: the observed data oversamples poor performance, yielding too low estimates of investors' true skill. This statement does not violate optional stopping-time theorems:¹⁵ even if $\tilde{\varepsilon}$ is mean zero and a martingale, and thus $E \left[\sum_{t=1}^{\tilde{T}} \tilde{\varepsilon} \right] = 0$, the *average* $\tilde{\varepsilon}$ is different from zero if $\tilde{\varepsilon}$ is correlated with the stopping time \tilde{T} : $E \left[\frac{1}{\tilde{T}} \sum_{t=1}^{\tilde{T}} \tilde{\varepsilon} \right] = \text{cov} \left(\frac{1}{\tilde{T}}, \sum_{t=1}^{\tilde{T}} \tilde{\varepsilon} \right)$.

The gap between the true and observed skill is economically significant: although the average high-frequency trader's skill (i.e., the mean of the prior) is 0.512, the average *observed* skill is just 0.460 in Table V, Panel C. This reverse survivorship bias thus appears to be a highly important consideration when the task is to assess individual investors' abilities: if investors stop trading after suffering losses, *realized* returns are significantly downwards biased measures of investors' true, unobservable trading abilities.

VI. Summary and Conclusions

I show that a model in which investors learn about their trading skills over time is consistent with several empirical regularities in household trading behavior: households appear to trade excessively given that their returns do not cover trading costs, good past performance increases trading intensity, and some households quit trading altogether after experiencing poor performance. The key mechanism in the Bayesian learning model with frictions is the option value of trading; because investors learn by trading, investors have an incentive to trade even if they think that they are unskilled—the (possibly small) probability of being skilled may outweigh the expected trading losses.

The estimates from a hierarchical learning model suggest that, first, moderate dispersion in beliefs can generate the trading patterns observed in the data; second, up to 18% of the investors who started active trading believed that they would lose to the market; third, approximately 4% of the investors in the population can either predict short-term price movements with some consistency or, alternatively, earn excess returns by acting as pseudo-market makers; fourth, because unskilled investors are more likely drop out over time, a regression of performance against experience generates a positive coefficient; and, fifth, realized returns are significantly downwards biased measures of

¹⁵See, e.g., Williams (1991).

investors' true abilities—because investors are more likely to quit after unsuccessful trades, realized investor performance “oversamples” unsuccessful trades relative to true skill.

Although I assume that investors are rational and have properly calibrated beliefs about their abilities, the results do not imply that investors must be rational in all respects. For example, I do not know if investors learn by trading because there are limits to paper trading, or because real trading is more cost-effective than paper trading, or because investors learn more when they take part in the action than when they merely observe as outsiders (reinforcement learning). Similarly, because I do not impose any constraints on investors' prior beliefs, investors' priors may be looser or tighter than what they should be given the investors' full histories. The main conclusion, however, is immune to these caveats: investors learn from their own experiences, and the evolution of beliefs that results from this appears to be an important driver of investor behavior.

Because a hierarchical learning model appears to be well-suited for describing investor behavior, this family of models may very well have other useful applications. For example, the learning-argument applies to all investors and not just to the high-frequency traders that I examine; I study these fast traders because they offer the cleanest test of the learning hypothesis. An analysis of casual household investors could, for example, extract investors' beliefs not only about their stock-picking abilities but also about their market timing abilities. If so, it would be interesting to study how large market movements alter investor beliefs, and measure whether these changes track the in- and outflow of money from index funds. If there is a market event that turns many investors simultaneously pessimistic about their stock-picking abilities, this event should be followed by a net inflow of money into index funds. Another direction for future research is to examine more general models. For example, a learning model in which the profitability of short-term trading is (possibly) correlated with the aggregate stock market could yield richer predictions about investors' trading decisions and portfolio choices.

References

- Arrow, K. J. (1962). The economic implications of learning by doing. *Review of Economics and Statistics* 29(3), 155–173.
- Barber, B. M., Y.-T. Lee, Y.-J. Liu, and T. Odean (2004). Do individual day traders make money? Evidence from Taiwan. University of California, Berkeley working paper.
- Barber, B. M. and T. Odean (1999). The courage of misguided convictions: The trading behavior of individual investors. *Financial Analysts Journal* Nov/Dec 1999, 41–55.
- Barber, B. M. and T. Odean (2000). Trading is hazardous to your wealth: The common stock investment performance of individual investors. *Journal of Finance* 55(2), 773–806.
- Barber, B. M. and T. Odean (2001a). Boys will be boys: Gender, overconfidence, and common stock investment. *Quarterly Journal of Economics* 116(1), 261–292.
- Barber, B. M. and T. Odean (2001b). The Internet and the investor. *Journal of Economic Perspectives* 15(1), 41–54.
- Barclay, M. J., W. G. Christie, E. Kandel, and P. H. Schultz (1999). Effects of market reform on the trading costs and depths of Nasdaq stocks. *Journal of Finance* 54, 1–34.
- Barclay, M. J., T. Hendershott, and D. T. McCormick (2003). Competition among trading venues: Information and trading on electronic communications networks. *Journal of Finance* 58, 2637–2665.
- Bauer, R., M. Cosemans, and P. Eichholtz (2007). The performance and persistence of individual investors: Rational agents or tulip maniacs? Maastricht University working paper.
- Brennan, M. J. (1998). The role of learning in dynamic portfolio decisions. *European Economic Review* 1(3), 295–306.
- Campbell, J. Y., M. Lettau, B. G. Malkiel, and Y. Xu (2001). Have individual stocks become more volatile? An empirical exploration of idiosyncratic risk. *Journal of Finance* 56(1), 1–43.
- Coval, J. D., D. Hirshleifer, and T. Shumway (2005). Can individual investors beat the market? Working paper.
- Cumby, R. E. and D. M. Modest (1987). Testing for market timing ability: A framework for forecast evaluation. *Journal of Financial Economics* 19(1), 169–189.
- DeGroot, M. H. (1970). *Optimal Statistical Decisions*. New York: McGraw-Hill College.
- Dixit, A. K. and R. S. Pindyck (1994). *Investment under Uncertainty*. Princeton: Princeton University Press.
- Fama, E. J. and K. French (1993). Common risk factors in the returns of stocks and bonds. *Journal of Financial Economics* 33(1), 3–56.
- Fama, E. J. and K. French (2008). Mutual fund performance. University of Chicago working paper.
- Goetzmann, W. N. and A. Kumar (2008). Equity portfolio diversification. *Review of Finance* 12, 433–463.
- Grinblatt, M. and M. Keloharju (2000). The investment behavior and performance of various investor-types: A study of Finland’s unique data set. *Journal of Financial Economics* 55(1), 43–67.
- Grinblatt, M. and M. Keloharju (2008). Sensation seeking, overconfidence, and trading activity. *Journal of Finance*. Forthcoming.

- Grinblatt, M., M. Keloharju, and J. Linnainmaa (2008). Do smart investors outperform dumb investors? University of California, Los Angeles working paper.
- Grossman, S. J., R. E. Kihlstrom, and L. J. Mirman (1977). A Bayesian approach to the production of information and learning by doing. *Review of Economics and Statistics* 44(3), 533–547.
- Harris, J. H. and P. H. Schultz (1998). The trading profits of SOES bandits. *Journal of Financial Economics* 50(1), 39–62.
- Harris, L. (2003). *Trading and Exchanges: Market Microstructure for Practitioners*. Oxford University Press.
- Hartzmark, M. L. (1991). Luck versus forecast ability: Determinants of trader performance in futures markets. *Journal of Business* 64(1), 49–74.
- Harvey, A. C. (1989). *Forecasting, Structural Time Series Models and The Kalman Filter*. Cambridge: Cambridge University Press.
- Hausman, J., B. H. Hall, and Z. Griliches (1984). Econometric models for count data with an application to the patents-R&D relationship. *Econometrica* 52(4), 909–938.
- Henriksson, R. D. and R. C. Merton (1981). On market timing and investment performance II – Statistical procedures for evaluating forecasting skills. *Journal of Business* 54(4), 513–533.
- Judd, K. L. (1998). *Numerical Methods in Economics*. Cambridge, Massachusetts: The MIT Press.
- Kaustia, M. and S. Knüpfer (2008). Do investors overweight personal experience? Evidence from IPO subscriptions. *Journal of Finance*. Forthcoming.
- Kumar, A. (2007). Who gambles in the stock market? *Journal of Finance*. Forthcoming.
- Lo, A. W., A. C. MacKinlay, and J. Zhang (2002). Econometric models of limit-order executions. *Journal of Financial Economics* 65(1), 31–71.
- Mahani, R. S. and D. Bernhardt (2007). Financial speculators’ underperformance: Learning, self-selection, and endogenous liquidity. *Journal of Finance* 62(3), 1313–1340.
- McCardle, K. F. and R. L. Winkler (1989). All roads lead to risk preference: A turnpike theorem for conditionally independent returns. *Journal of Financial and Quantitative Analysis* 24(1), 13–27.
- McFadden, D. (1989). A method of simulated moments for estimation of discrete response models without numerical integration. *Econometrica* 57(5), 995–1026.
- Merton, R. C. (1969). Lifetime portfolio selection under uncertainty: The continuous-time case. *Review of Economics and Statistics* 51(3), 247–257.
- Neyman, J. and E. L. Scott (1948). Consistent estimates based on partially consistent observations. *Econometrica* 16(1), 1–32.
- Nicolosi, G., L. Peng, and N. Zhu (2004). Do individual investors learn from their trading experience? University of Cincinnati working paper.
- Odean, T. (1998). Are investors reluctant to realize their losses? *Journal of Finance* 53(5), 1775–1798.
- Odean, T. (1999). Do investors trade too much? *American Economic Review* 89(5), 1279–1298.
- Pakes, A. and D. Pollard (1989). Simulation and the asymptotics of optimization estimators. *Econometrica* 57(5), 1027–1057.

- Pratt, J. W., H. Raiffa, and R. Schlaifer (1995). *Introduction to Statistical Decision Theory*. Cambridge, Massachusetts: The MIT Press.
- Roberts, K. and M. L. Weitzman (1981). Funding criteria for research, development, and exploration projects. *Econometrica* 49(5), 1261–1288.
- Scholes, M. and J. Williams (1977). Estimating betas from nonsynchronous data. *Journal of Financial Economics* 5(3), 309–327.
- Seru, A., T. Shumway, and N. Stoffman (2007). Learning by trading. University of Michigan working paper.
- Sørensen, M. (2007). How smart is smart money: A two-sided matching model of venture capital. *Journal of Finance* 62, 2725–2762.
- Taylor, L. A. (2008). Why are CEOs rarely fired? Evidence from structural estimation. Wharton School, working paper.
- Weitzman, M. L., W. Newey, and M. Rabin (1981). Sequential R&D strategy for synfuels. *Bell Journal of Economics* 12(2), 574–590.
- Williams, D. (1991). *Probability with Martingales*. Cambridge, Massachusetts: Cambridge University Press.
- Wooldridge, J. M. (2002). *Econometric Analysis of Cross Section and Panel Data*. Cambridge, Massachusetts: The MIT Press.
- Xia, Y. (2001). Learning about predictability: The effects of parameter uncertainty on dynamic asset allocation. *Journal of Finance* 56(1), 205–246.