Why Do Households Trade So Much?*

JUHANI T. LINNAINMAA†

October 16, 2009

ABSTRACT

When agents can learn about their abilities as active investors, they rationally “trade to learn” even if they expect to lose from active investing. The model used to develop this insight draws conclusions that are consistent with empirical study of household trading behavior: Households’ portfolios underperform passive investments; their trading intensity depends on past performance, and they begin by trading small sums of money. Using household data from Finland, the paper estimates a structural model of learning and trading. The estimated model shows that investors trade to learn even if they are pessimistic about their abilities as traders. It also demonstrates that realized returns are downward-biased measures of investors’ true abilities. This bias is large.

Keywords: Learning, individual investor behavior, individual investor performance
JEL classification: D10, G11

*I would like to thank Tobias Adrian, Robert Battalio, Tony Bernardo, Alan Bester, Michael Brennan, John Cochrane, Gene Fama, John Griffin, Mark Grinblatt, John Heaton, Matti Keloharju, Deborah Lucas, Kevin McCardle, Luboš Pástor, Josh Pollet, Mark Seasholes, Luke Taylor, Pietro Veronesi, Yuzhao Zhang, and seminar participants at the 2008 AFA meetings, New York University, University of Chicago, Notre Dame University, and Michigan State University. This paper replaces an earlier version entitled “Learning from Experience.” Financial support from the University of Chicago Booth School of Business is gratefully acknowledged. Sandra Sizer served as copy editor.

†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.chicagobooth.edu/juhani.linnainmaa, email: jlinnain@chicagobooth.edu.
Many households trade frequently but underperform passive investments. I ask if a model in which investors learn from experience can explain this behavior, and, if so, what the model tells us about households’ beliefs.

In my model, investors are uncertain about their abilities. They learn about their abilities as they trade. If the value of another signal is high, then an investor trades even if he expects to lose money, thus apparently trading “too much.” If a trade is successful, the investor infers skill, and subsequently trades more. If an investor loses money, he will infer less skill and subsequently trade less. After enough losses, he stops trading altogether. Investors who are especially uncertain about their abilities trade small amounts early in their careers until they infer skill or leave. Finally, new signals influence subsequent behavior more early in an investor’s career, when the investors’ prior is diffuse.

I use Finnish trading records on active traders to test the predictions of the learning model. These traders increase their trade sizes after successful trades and decrease trade sizes or quit after unsuccessful trades. Both the exit and trade size effects are stronger early on in investors’ careers. Some traders initially execute very small trades that seem to be motivated, in part, by their desire to learn more about their own trading skills.

I add belief distributions on top of a general trading model with non-tradable income to create a structural model of an investor population. The parameters in this model control investors’ prior belief distributions. I extract from this model the distribution of beliefs that comes closest to matching the aggregate trading patterns in the data. I identify the structural parameters by matching, between the model and the data, the number of investors who quit after each trading date and how sensitive investors’ exit decisions are to returns. Because investors can also self-select and not trade at all, I also match the number of investors who trade at least once. Because of this entry moment condition, the model parameters are belief distributions for the entire population, not just for the active traders.

The learning model matches the initial selection step, the conditional exit rates, and the
performance-exit sensitivity in the data. The fact that this model can explain the initial selection indicates that active traders do not need to be unlike those investors who choose not to trade. Those who become active traders just happen to reside in particular corners of the belief parameter space.

The parameter estimates suggest that many investors start active trading without believing that they are skilled. They start trading because they might be. More than a quarter of all traders begin with a belief that there is less than one-in-seventy chance of being skilled. Approximately 2.3% of all investors have genuine trading skills. Although this skill may manifest itself as an ability to predict short-term price changes, an analysis of active traders’ order choices indicates that many profit (or try to profit) by supplying liquidity to other investors. The 2.3% estimate, which includes both types of investors, is consistent with studies such as Coval, Hirshleifer, and Shumway (2005) and Grinblatt, Keloharju, and Linnainmaa (2009a).

The structural model yields an estimate of the size of a reverse survivorship bias that affects investor performance measurement. Because investors are more likely to quit after a series of unsuccessful trades, their realized performance is biased downwards relative to true, unobservable skill. This mechanism is best illustrated with a coin-flip example. Suppose that I start flipping a coin to measure its bias, but stop the experiment after I get the first tails. For a fair coin, the expected proportion of heads in such an experiment is

$$\lim_{n \to \infty} \left( \frac{1}{2} \right)^{n} + \cdots + \left( \frac{1}{2} \right)^{n} = 0.31.$$

Because the stopping rule is conditional on the tails-outcome, I oversample tails relative to the truth. Similarly, if investors quit after poor performance, then the observed data oversample poor performance relative to true skill and the estimates of investors’ abilities become downwards-biased. Although active traders’ average true skill, expressed as a probability of a successful trade, is 0.344 in the estimated structural model, their average observed skill is just 0.267.

Although the traders in the model have constant abilities, more experienced traders are more skilled due to a learning-by-survival mechanism. Because a skilled trader is more likely to survive, a larger fraction of the investors who remain in the market are skilled. By inserting the structural
parameter estimates back into the model, I find that the fraction of skilled traders increases from 36% at the time of the first trade to 45% after ten trades. Thus, the unskilled traders learn quickly that they are unskilled and quit.

My study is most closely related to Mahani and Bernhardt (2007), who incorporate investor learning into a general equilibrium model of the markets. They show that learning reduces the bid-ask spreads and the price impact of liquidity shocks. Mahani and Bernhardt motivate their setup by noting that their model is qualitatively consistent with many empirical facts about investor behavior while a model built on overconfidence alone is not.

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. Sørensen (2007) uses a structural matching model to examine why companies funded by more experienced VCs are more likely to go public. Taylor (2008) estimates a structural model in which the board of directors learns about CEO skill.

The paper is organized as follows. Section I provides the background for the study and discusses the learning-by-trading assumption. Section II presents a stylized trading 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 learning mechanism. Section V formulates and estimates a structural model of learning and trading. Section VI concludes.

I. Background

A. Stylized Facts about Household Behavior

My study’s learning model reconciles several stylized facts about household behavior. The trading records from (at least) Denmark, Finland, Norway, Taiwan, and the U.S. indicate that all these countries 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 exhibit superior performance.
**Excessive trading and underperformance of the average household.** Studies such as those by Odean (1999), Barber and Odean (2000, 2001), Grinblatt and Keloharju (2000, 2009), and Barber, Lee, Liu, and Odean (2009) find that the average household trades excessively. The poor performance after trading costs appears to suggest many would be better off by holding the market. Barber and Odean (2001), Grinblatt and Keloharju (2009), French (2008), and Kumar (2009) suggest that both overconfidence and the desire to gamble contribute to these results.

**Performance heterogeneity.** Studies by Nicolosi, Peng, and Zhu (2004), Barber, Lee, Liu, and Odean (2005), Coval, Hirshleifer, and Shumway (2005), Bauer, Cosemans, and Eichholtz (2007), Goetzmann and Kumar (2008), Ivković, Sialm, and Weisbenner (2008), Seru, Shumway, and Stoffman (2009), and Grinblatt, Keloharju, and Linnainmaa (2009a) find that a small number of individual investors outperform the market or their peers. Harris and Schultz (1998) find that some individuals acting as Nasdaq’s Small Order Execution System (SOES) “bandits” are better than others. In many of these studies, the best performing investors continue to outperform the worst performers from one period to the next.


**B. Learning and Paper Trading**

If a learning model is to explain investor behavior, investors must learn from their own actions, not just paper trading. Here, I discuss factors that may drive the learning-by-trading mechanism. Although I cannot distinguish between these explanations in the data, the existence of the
mechanism itself is undisputable. Like Barber, Lee, Liu, and Odean (2005), Nicolosi, Peng, and Zhu (2004), and others, also I find that investors’ past performance greatly influences their future actions.

A learning-by-trading mechanism kicks in if there are limits to what an individual can learn by paper trading. If there are such limits to paper trading, investors may be left with some residual uncertainty no matter how much they paper trade. Some investors may then seek to resolve this residual uncertainty by experimenting with real money.

Several practical difficulties with paper trading may impose limits on how much can be learned from it. First, trade execution quality is an important concern with short holding periods. It is difficult to learn about the execution component by paper trading, first, because trade and quote databases are not readily available to the public, and second, because the data rarely show order-flow directions. Moreover, Lo, MacKinlay, and Zhang (2002) find that hypothetical limit-order executions, “constructed either theoretically from first-passage times or empirically from transactions data,” are poor proxies for actual limit-order executions. (The traders that I analyze in this study rely extensively on limit orders in their execution strategies.) Harris and Schultz (1998) find that good execution is vital to SOES bandits, because these bandits need to trade within quoted prices to make money.

Households’ preference for small stocks increases the difficulty of paper trading. (Barber and Odean (2001) show that the average household has an SMB loading of 0.78 in the Fama and French (1993) three-factor model.) A trade or a limit order submission in a small stock may alter the behavior of other traders and change the stock’s future price path. If an investor has non-negligible mass, then 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. This, however, creates a trade-off that may favor real trading. If the costs of paper trading are higher than what the investor expects to lose in the market by making small trades, then the investor may
replace paper trading with real trading. These two considerations, the 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) motivates a learning-by-doing process with an observation from psychology, “learning is the product of experience.” Kaustia and Knüpfer (2008), for example, find that individual investors are more sensitive to their own successes and failures in IPO subscriptions than they are to other observations. Hence, investors may trade to learn even though they also could learn by paper trading.

II. A Simplified Trading Model

I use a simplified binomial trading model to demonstrate the consequences skill uncertainty on investor behavior. This uncertainty alters investor behavior because the investor must consider how actions taken today affect the options available in the future. I can describe an investor’s optimal behavior in a binomial tree because each trade has only two possible outcomes. I can then illustrate how the investor conditions on past outcomes and how the investor’s beliefs evolve after each trade.

This model differs from other Bayesian portfolio choice models\(^1\) in two ways. First, I add a friction to generate the learning-by-trading mechanism. If an investor could trade infinitesimal amounts without a cost, then he could effectively observe outcomes without trading. The second difference is the context of the model. Investors in other models are often uncertain about some objective parameter, such as the size of the equity premium or the predictive power of a state-variable. The investor in my model is uncertain about his own trading skill.

\(^1\)See, for example, Detemple (1986), Brennan (1998), and Xia (2001). Pástor and Veronesi (2009) survey this literature.
A. Investor and Investment Opportunities

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

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

where $\gamma$ is the coefficient of relative risk aversion. The investor can trade at dates $t = 1, 2, \ldots, T-1$. Each day, the investor receives a positive or negative signal about a single stock. (This signal can be about a short-term price movement or about an opportunity to supply liquidity as a pseudo market-maker.) The investor chooses an amount $x_t > 0$ to invest in this signal. The signal is genuine with probability $p$, in which case the payoff is $2x_t$. The signal is useless with probability $1-p$, in which case the payoff is zero.\(^2\) There is also a risk-free asset that pays no interest.

The investor has a beta-distributed prior belief about $p$ and learns about it only by observing more outcomes. If the investor trades, he observes the outcome and updates his beliefs using Bayes’ rule. If the investor does not trade, he does not observe the outcome. Starting from a prior distribution $B(\alpha_1, \beta_1)$, the posterior distribution is $B(\alpha_1 + 1, \beta_1)$ after a successful trade and $B(\alpha_1, \beta_1 + 1)$ after a failed trade.\(^3\) Hence, $\alpha$ and $\beta$ 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 one time-step.

I introduce a trading friction by setting a minimum trade size, $\bar{x}$. If an investor trades, he must choose $x_t \geq \bar{x}$. This friction makes learning costly. If an investor believes that $p < \frac{1}{2}$, then he needs to weigh the expected trading loss against the value of the information he would gain from one more trade. The Appendix describes the investor’s optimization problem and provides a numerical solution method.

\(^2\)The binomial-outcome assumption bears similarities to the tests 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 directions of price changes.

\(^3\)See, for example, DeGroot (1970).
B. Empirical Implications

Three considerations in this model shape investor behavior. The first factor is the myopic demand: given everything else, an investor with a higher prior belief about his trading ability trades more. The implication is that if two investors differ only in the levels of their prior beliefs, then the optimistic investor trades more than the pessimistic investor.

The second factor is the intertemporal hedging demand. Although the true \( p \) is fixed, an investor revises his beliefs upwards after a successful trade and downwards after an unsuccessful trade. This hedging demand component is negative for an investor with \( \gamma > 1 \) and positive for an investor with \( \gamma < 1 \). The empirical implication is that this hedging-demand component dissipates over time as investors resolve uncertainty.

The third factor is the option value of trading. Trades have value beyond their expected payoffs because investors observe outcomes only if they trade. The (possibly small) probability of being 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, \( \bar{x} \). If \( \gamma > 1 \), an increase in skill uncertainty lowers the intertemporal hedging demand but increases the option value of trading. The empirical implication of this mechanism is that an investor with a very diffuse prior distribution is more likely to trade than an investor with a very precise prior distribution. Moreover, an investor who trades only to learn, makes small trades.

[Figure 2 about here.]

Figure 2 shows how the skill uncertainty affects investor behavior in the time series. The investor’s prior belief about his skill is \( \hat{p}_1 = \frac{1}{2} \). At date 1, he trades the minimum amount to observe another outcome. If the trade fails, then the investor revises his beliefs downwards to \( \hat{p}_2 = \frac{1}{3} \), but he still makes another minimum-size trade. He quits trading after one more loss. If the investor’s first trade is successful, he revises his beliefs upwards to \( \hat{p}_2 = \frac{2}{3} \) and trades more than the minimum amount. The changes in beliefs have a considerable influence on trade size choices.
This effect’s strength depends on the how dispersed the prior distribution is: if the distribution is very tight, one more observation matters very little, which also leads to smaller changes in the trade size. This is a common characteristic of learning models: if belief changes induce changes in investor behavior, then the changes should be stronger in the beginning when the prior distribution is more dispersed.

[Figure 3 about here.]

Figure 3 illustrates how investors’ initial choices depend on their prior beliefs and risk aversion. The graphs show functions of the two parameters of the beta distribution, $\alpha$ and $\beta$. I divide the parameter space 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 below 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 stay out when they are risk-averse enough. For these investors the negative intertemporal hedging demand offsets the other demand components. By contrast, many mildly risk-averse investors trade the minimum amount even if they are 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 Active Traders

I use the Finnish Central Securities Depository data to test the learning model. This data set records the portfolios and trades from January 1, 1995 through November 29, 2002 of all individual investors in Finland. The electronic records I use are exact duplicates of the official certificates of ownership, trades, and demographic information and hence very reliable.

I study the behavior of active traders. I define an active trader as an individual investor who
completes at least one intraday round-trip trade during the sample period. If an investor buys, say, 1,000 shares of a stock in the morning and sells some or all of these shares in the afternoon, I classify that investor as an active trader. (Barber, Lee, Liu, and Odean (2005) call these investors “day traders” in their study of Taiwanese households.) Although these traders represent only 4.1% of all households who traded during the sample period, they account for 64% of all household trading volume.

I use intraday round-trip trades to identify the point in time when an investor begins to trade very frequently. After an investor makes a round-trip trade, I also include in the sample the same investor’s all short-term trades. I identify these speculative transactions as all purchases by the same investor in any stock during the next two weeks following either a round-trip trade or a speculative purchase. I stop tracking an investor when he no longer makes any round-trip trades and also stops making purchases for at least two weeks. Thus, these rules identify investors who start trading very actively and then follows them for as long as they remain active.

Active traders are well-suited for testing the learning model because their trades form a clean sequence of choices and outcomes. By contrast, an analysis of buy-and-hold investors would be complicated by the lack of natural sequencing. It is difficult to say what these investors learn about their abilities and when. Also, risk-adjustment is not an issue with short-term trades because intraday volatility swamps the daily risk premium. (The daily risk premium is less than 5 basis points even if the annual risk premium is 10%.) Thus, short-term traders try to profit from predicting price movements, or from supplying liquidity as pseudo-market makers, and not from capturing the risk premium.

Active traders can also learn about their abilities much faster than buy-and-hold investors. The difficulty with buy-and-hold investors is that we do not know their exposures to the (possibly unknown) risk factors. When these exposures and factors have to be estimated from the data, it becomes difficult to disentangle luck from skill. Harris (2003) advises that “[i]n practice, more than 20 years of returns data are typically required to obtain useful results for a given investment manager.” Fama and French (2008) suggest that all results about mutual fund managers’ superior
performance may be attributable to the difficulty in distinguishing luck from skill. In contrast, the learning in short-term trading is akin to measuring the bias of a coin, with each trade representing a flip of a coin. Because active traders learn in event-time rather than in calendar-time, these traders learn much faster than buy-and-hold investors. Harris and Schultz (1998), for example, exploit the power of short-term trading and draw inferences about traders’ abilities from just five days of trading data.

Yet another benefit of studying active traders is that they can bet almost symmetrically on one-day up and down movements in stock prices. During the sample period, Finnish brokers allowed their customers to sell shares short with no additional contracts. Moreover, an investor did not need to borrow the shares that he sold as long as he covered the position by the end of the day. An investor could enter 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.\(^4\)

Finally, these traders’ behavior holds the key to understanding why households trade so much, because they are responsible for almost two thirds of all household trading volume.\(^5\)

\textit{B. Variable Definitions}

I measure the profitability of a short-term trade \(i\) on day \(t\) as

\[
gross \text{ profit}_{i,t} = \begin{cases} 
\sum_{s=1}^{n} (p_{i,s,t} - p_{b,i,s,t})v_{b,i,s,t} & \text{if } v_{b,i,s,t} = v_{s,i,s,t} \\
\sum_{s=1}^{n} (p_{i,s,t} - p_{b,i,s,t})v_{s,i,s,t} + (p_{b,i,s,t} - p_{s,i,s,t})(v_{b,i,s,t} - v_{s,i,s,t}) & \text{if } v_{b,i,s,t} > v_{s,i,s,t} \\
\sum_{s=1}^{n} (p_{s,i,s,t} - p_{b,i,s,t})v_{b,i,s,t} + (p_{s,i,s,t} - p_{b,i,s,t})(v_{b,i,s,t} - v_{s,i,s,t}) & \text{if } v_{b,i,s,t} < v_{s,i,s,t}, 
\end{cases}
\]

\(^4\)Although investors can bet symmetrically on positive and negative signals, investors may be reluctant to do so for psychological reasons. (In the data active traders start 25.1% of their intraday round-trip trades with a naked short sale.) However, even if investors do not trade on some of their signals, there is no reason to expect that these unrealized trades influence the data to favor the learning story.

\(^5\)I note that an investor who stops “active trading” may still engage in stock picking. However, if an investor learns from his short-term trading abilities, the same investor also should (eventually) learn about his stock-picking abilities. For the speed-of-learning reasons detailed above, I focus my analysis on the period of time when investors trade extremely frequently.
where \( n \) is the number of trades by investor \( i \) on day \( t \), \( p_{b,i,s,t}^b \) and \( p_{s,i,s,t}^s \) 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 \( p_{c,t}^s \) is the stock’s closing price on the same day. If an investor buys and sells different amounts, 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 (2009) I use the lowest available commission rate of 8.42 euro plus 0.15% of trade value. The investor’s day \( t \) return on short-term trading is then \( r_{i,t} = \frac{\text{net profit}_{i,t}}{x_{i,t}} \), where \( x_{i,t} \) is the total size of the trade. The trade size is the maximum of the number of shares bought and sold multiplied by the volume-weighted average transaction price across all trades:

\[
x_{i,t} = \sum_{s=1}^{n} \left( \text{max}(v_{i,s,t}^{b}, v_{i,s,t}^{s}) \frac{p_{b,i,s,t}^{b} v_{i,s,t}^{b} + p_{s,i,s,t}^{s} v_{i,s,t}^{s}}{v_{i,s,t}^{b} + v_{i,s,t}^{s}} \right). \tag{3}
\]

C. Summary Statistics on Active Traders

Table I shows that the average active trader (Panel A) is younger and more often male than the average non-active investor (Panel B). Active traders also trade more frequently than other investors even after excluding their short-term trades. While the average active trader trades 113.3 times, the average non-active trader completes just 6.4 trades. The median number of (identified) short-term trades is just four, but this number varies considerably across investors. One quarter of traders complete only two short-term trades, but the top quarter of traders complete at least 15 short-term trades. The top 5% of traders complete at least 112 short-term trades.

[Table I about here.]

The stringent rules that I use to identify “short-term trades” may miss out on some trades that should be included in the analysis. Nevertheless, these rules identify as short-term trades most trades that investors make during their “active phase.” (An investor’s active phase is the period of time between the investor’s first intraday round-trip trade and the last trade that is identified as a short-term trade.) The fraction-of-purchases-captured row in Table I indicates that over 80% of the
average trader’s active-phase trades are identified as short-term trades. All trades are identified as short-term trades for 55.5% of active traders. These statistics suggest that I successfully identify a (possibly short) period of time when an investor trades very frequently. Even if these rules exclude some genuine short-term trades that took place before or after the active phase, such omissions would only add noise to the resultant sample and mask the effects of learning.

Active traders concentrate their activity in high beta stocks. The average market beta is 1.73 for active traders’ short-term trades and 1.51 for their other trades. Although the mobile communications company Nokia accounts for 32.1% of all the short-term trades, each of the 48 most popular stocks record at least 1,000 short-term trades over the sample period.

The median active trader does not break even in short-term trading, losing an average of just 25 euro per a short-term trade (and the opportunity cost of capital). However, the profit distribution is skewed to the right. Since substantial short-term profits are more frequent than substantial short-term losses, the average investor’s (average) loss is smaller at 7 euro per day. These profits estimates may, however, be upwards biased if some investors in the data exhibit severe disposition effect. Such an investor may sell shares purchased earlier on the same day if the price increases significantly. Although I cannot clean the data of such disposition-effect induced trades, their presence would not affect my analyses. Even if the average return is upwards-biased, the covariances between returns, exit decisions, and trade size changes are unaffected.

Limit orders play a prominent role in active traders’ execution strategies. Whereas 47.2% of other individual investors’ trades originate from limit orders, this proportion is 52.5% for active traders. This difference suggests that many active traders may try to profit by supplying liquidity to other (impatient) investors, not from predicting short-term price movements. Moreover, when active traders are assigned into number-of-short-term-trades deciles, the proportion of limit-order

---

6 I measure market betas against a constrained market index because Nokia dominates the unconstrained index for most of the sample period. The constrained market index caps the weight of each individual stock to 10%. The typical stock bought by an individual investor has a market beta of 0.81 when betas are measured against the unconstrained market index. This average is 1.43 for the constrained market index.

7 See, for example, Shefrin and Statman (1985), Odean (1998), and Barberis and Xiong (2009).

8 Kaniel, Saar, and Titman (2008) find that individuals trading on the NYSE profit by supplying liquidity to institutions.
trades increases monotonically from 50.0% (decile 1) to 56.0% (decile 10). Thus, those active traders who remain in the market the longest use more limit orders in their execution strategies. If longevity correlates with success, these numbers indicate that liquidity-providing individuals are more successful than those who try to use market orders to profit from short-term price movements.

Active traders demonstrate better-than-average cognitive abilities in the Finnish Armed Forces intelligence test. The score from this IQ test, which is administered to Finnish males in a 20-year age range upon induction into Finland’s mandatory military service, is standardized to follow the stanine distribution (integers 1 through 9 with 9 being most intelligent).\(^9\) Using the same data as Grinblatt, Keloharju, and Linnainmaa (2009a, 2009b), I find an average IQ score of 6.12 (standard error = 0.02) for those 7,229 male active traders who entered military between 1982 and 2001. The average test score over all inductees is just 4.83. Thus, active traders’ scores are over half a standard deviation above the average. Active traders also earn higher IQ scores than other individual investors who receive an average IQ score of 5.99 (Grinblatt et al. 2009b). The 0.13 point difference between active traders and other individual investors is statistically significant with a \(p\)-value < 0.01. Grinblatt et al. (2009a) find that high IQ score individuals’ investments outperform low IQ individuals’ investments. Thus, active traders’ better-than-average cognitive abilities indicate that some active traders may rightfully believe that they could be skilled.

IV. Testing the Implications of the Learning Model

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 prior beliefs, this requires that some investors are genuinely skilled. If not, investors would know that they must be unskilled and they would not trade.

\(^9\)A stanine distribution divides a normal distribution into nine intervals, each of which has a width of one half of a standard deviation except for the outermost categories. Stanine 5 represents the mean. I thank Matti Keloharju for providing me these IQ score distributions from the Finnish Armed Forces database.
Table II reports on performance persistence regressions that use data on active traders’ short-term trades. The first specification has the return on the $j^{th}$ trade as the dependent variable and the average career return over trades $1, \ldots, j-1$ as the regressor. The second specification replaces returns with positive return-indicator variables, $1(r_{i,t} > 0)$, and uses the average “success rate” as the regressor. If there are no performance differences or if such differences are transitory, then the slopes in these regressions will be zero. All regressions include event-time and calendar-time fixed effects.

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 first specification indicates that an investor whose average career return is 1% expects to earn a return that is 0.166% higher than the return earned by an investor with a career return of zero. The results are similar in the second specification. An investor whose every past trade has been a success has a 0.283 higher probability of another success compared to an investor with no past successes.

The last two columns in Table II split the sample in investors’ early ($t \leq 10$) and late ($t > 10$) trades. The slopes are positive and statistically significant in both samples. A one standard deviation shock to the average career return in Panel A increases today’s return by 0.35% in the early-trades sample and by 0.34% in the late-trades sample. In Panel B, such unit shocks increase success probabilities by 0.046 and 0.067, respectively.

B. Exits after Unsuccessful Trades

In Table III I examine whether investors are more likely to quit short-term trading after unsuccessful trades. The table reports on linear probability model (LPM) regressions in which the dependent variable is an exit-indicator variable. The main regressors are the investor’s performance in trade $t$ as well as his average career performance. Because investors who lose enough capital may be
forced to quit even in absence of learning, I also include the log-change in the investor’s portfolio value from the time of the first trade to day \( t - 1 \). The regressions also include event-time and calendar-time dummies.

I define “exit” as each investor’s last short-term trade. A concern with this definition is that an investor who pauses 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. Given that the average waiting time between an investor’s two short-term trades is 11.6 days (the median is two days), this three-month cutoff should eliminate most false exits. Moreover, any remaining false exits would favor the null hypothesis of no correlation between trade performance and exit decisions.

The regression estimates in Table III indicate that investors often exit after negative outcomes. In the full sample, both the current trade performance and the average career performance are significant determinants of the exit decision. A successful trade on day \( t \) lowers the quitting probability by 0.02 compared to the unsuccessful trade benchmark.

Although the change in wealth is statistically significant in some regressions, the effect is economically small. Even if an investor’s portfolio doubles in value, the exit probability (in Panel B) decreases by only 0.0003. Thus, unsuccessful trades appear to influence exit decisions mainly through their effect on beliefs, not indirectly through their wealth effects.

The learning model predicts that investors’ early experiences should influence exit decisions more than (possible) later experiences. An investor who has traded for a long time has a sharp posterior belief and pays less attention to new data. For example, suppose that an investor in the binomial model has a prior mean of \( \hat{p}_t = 0.5 \). If the standard deviation of the prior distribution is 0.2, one successful trade leads to a posterior mean of \( \hat{p}_{t+1} = 0.58 \). If, on the other hand, the standard deviation of the prior is 0.02, one successful trade leads to a posterior belief of just \( \hat{p}_{t+1} = 0.5008 \).
This deceleration-of-learning result is a common feature of learning models\textsuperscript{10}, and the early- and late-career estimates in Table III are consistent with it. A positive outcome in the early sample lowers the exit probability by 0.061 but only by 0.007 in the late sample. The average career return also becomes insignificant when moving from the early-trades sample to the late-trades sample.

In Figure 4 I illustrate the changes in outcome sensitivity by plotting slope coefficients from separate cross-sectional regression for investors’ first trades, second trades, and so forth. The coefficients increase rapidly at first and then tend more slowly towards zero. This concave coefficient pattern indicates that the speed of learning decreases from one trade to the next. These changes are economically significant. Whereas the difference in the exit probabilities between successful and unsuccessful investors is 0.16 in the first trade, the difference is only 0.05 in the second trade.

C. Trade Size Changes

Table IV reports on regressions that measure how outcomes influence investors’ trade size decisions. Because the investor in the learning model revises his beliefs after each trade, short-term trading becomes more attractive 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 and the return in trade $t - 1$ as regressors. I ignore investors who exit after day $t - 1$. Because investors are more likely to quit after unsuccessful trades, this understates the true effect that outcomes have on trade sizes. However, this is a desirable restriction because Table III already establishes the outcome-exit relation. These regressions measure whether outcomes affect trade sizes even after ignoring exits.

\textsuperscript{10}See, for example, Harvey (1989) and Pratt, Raiffa, and Schlaifer (1995).
The estimates from both specification indicate that past outcomes have a significant influence on trade size choices. For example, the full sample estimates show that an investor with a successful trade increases the trade size by $e^{0.09} - 1 = 9.4\%$ relative to an investor who experiences a failure. The results in Panel B also support the deceleration-of-learning argument: more experienced investors pay less attention to new data compared to the inexperienced investors. Although the direction of the result in Panel A is the opposite, the difference is not statistically significantly different from zero because of the noisy coefficient estimate in the late-trades sample.

D. Exploratory Trades

In the model, an investor trades even if he expects to lose money when the value of another signal is high. Such an investor trades only the smallest permissible amount. The distribution of trade sizes (for investors trading for the first time) in Figure 5 suggests that this mechanism may also be present in the data. Many of these initial trades are quite small considering the costs of short-term trading. To emphasize the significance of these costs, the figure overlays the implied break-even returns. An investor has to earn at least this high a return to recover the commissions. For example, an investor who buys shares for 1,000 euro must sell them at a 1.99\% higher price to break even after 8.42 euro + 0.15\% commissions.

[Figure 5 about here.]

The implied break-even return is often counter-intuitively high in Figure 5 if there is no learning involved. If an investor expects to correctly predict a one-day price movement of more than 1.99\%, why does he not trade more? By contrast, these high break-even returns are less of a puzzle if some investors trade just to learn. These investors are willing to lose a small sum of money to learn about their abilities. The histogram shows that a nontrivial number of the initial trades fit this description. For example, 11.6\% of the initial trades are so small (< 1,406 euro) that the implied break-even return is at least 1.5\%.
Consistent with the interpretation that some initial trades are exploratory, investors who continue trading increase their trade sizes significantly. Panel B shows that the average investor who goes on to trade for the second time increases the trade size by 46%. Investors trade size changes and exit decisions interact as well. For example, the average investor still alive at date 25 enters a trade that is 2.2 times as large as the same investor’s date-1 trade. By day 50 the average trade is 2.6 times as large as the first trade. This increasing trade size pattern, however, holds only for successful investors: the average day-to-day trade size change across all observations is negative (−1.3%). This surviving investors-versus-all investors difference indicates that unsuccessful investors decrease their trade sizes until they ultimately quit.

V. Simulated Moments Parameter Estimation

A. A General Trading Model

I solve a general trading model that relaxes the simplifying assumptions of the binomial model. I then add belief distributions on top of this trading model to create a structural model of investor population that can be estimated from the data. I assume that an infinite-horizon investor maximizes log-utility over consumption:

$$\sum_{t=0}^{\infty} \beta^t \log c_t,$$

where $\beta$ is the subjective discount rate. The wealth dynamics are given by

$$W_{t+1} = (W_t - c_t)(1 + \tilde{y} + \theta_t(\tilde{r}_t - \tilde{y})),$$

where $\theta_t$ is the proportion of wealth invested in a risky trading opportunity, $\tilde{r}$ is the return on this trading opportunity, and $\tilde{y}$ is the uncertain stream of income from other sources. This other-sources income, which can present both labor income and risky savings, is proportional to the after-consumption wealth and net of the resources allocated to the risky trading opportunity. The density $f(y)$ is such that $f(y) = 0$ for all $y \leq -1$. The investor knows $f(y)$. 
The returns on the trading opportunity are drawn from a left-truncated normal distribution with truncation at \( x = -1 \). The underlying untruncated normal distribution has a mean of \( \mu \) and a variance of \( \sigma^2 \), where \( \mu \) represents the investor’s trading ability. While \( \sigma^2 \) is fixed and known, the investor is uncertain about his ability \( \mu \). The investor’s date-\( t \) prior distribution about \( \mu \) is normal with a mean of \( m_t \) and a variance of \( \sigma^2_t \). The investor updates his beliefs based on each return realization \( r_t \). As a technical matter, the investor backs out from each \( r_t \) the corresponding realization \( r_t' \) for the untruncated normal distribution and uses this value to update from prior \( m_{t-1} \) to posterior \( m_t \). I compute this transformation by first looking up the cumulative density at \( r_t \) and then finding the value \( r_t' \) for the untruncated distribution with the same cumulative density.\(^{11}\)

The investor’s actions are constrained by a minimum trade size, \( \theta \geq \bar{\theta} \). If the investor trades, he observes the realization \( r_t \), backs out \( r_t' \), and updates his beliefs about \( \mu \) accordingly. The dynamics of the mean and variance of the belief distribution are given by

\[
\begin{align*}
    m_{t+1} & = \frac{m_t + r_t'}{\frac{1}{\sigma_t^2} + \frac{1}{\sigma^2}}, \\
    \sigma^2_{t+1} & = \frac{1}{\frac{1}{\sigma_t^2} + \frac{1}{\sigma^2}}.
\end{align*}
\]

(6)  (7)

If the investor does not trade, he does not observe the return and his beliefs remain unchanged.

The investor’s indirect utility is a function of three state variables: the current wealth \( W_t \), the mean of the prior distribution \( m_t \), and the variance of the prior distribution, \( \sigma^2_t \). Because the investor maximizes log-utility, the wealth and belief terms are additive in the indirect utility.

\(^{11}\)The use of a truncated normal distribution gives the trading opportunity limited liability. Without this assumption, the log-utility investor would never invest in the trading opportunity because of the fear of losing more than everything. This same transformation trick could be used for any (continuous) return distribution that can be mapped one-to-one to the underlying normal distribution. For example, if the returns were log-normally distributed, an investor could learn about the mean of the normal variate \( \tilde{x} \) in \( \tilde{r} \equiv e^{\tilde{x}} - 1 \). Although elegant, this log-normal assumption is problematic because also the variance of \( \tilde{r} \) changes as the investor updates his beliefs about the mean of \( \tilde{x} \). By contrast, the truncation of the normal distribution at \( x = -1 \) is an innocuous return-distribution twist because, for reasonable return volatilities, the amount of mass that gets truncated is close to zero. See Johnson, Kotz, and Balakrishnan (1994) for details on truncated normal distributions.
function. I conjecture that

$$V(W_t, m_t, \sigma_t^2) = A + B \log W_t + g(m_t, \sigma_t^2),$$  \hspace{1cm} (8)$$

where $A$ and $B$ are constants and $g(\cdot)$ is a function of the investor’s date-$t$ beliefs. The investor’s optimization problem with this conjecture is

$$V(W_t, m_t, \sigma_t^2) = \max_{c_t} \left\{ \log c_t + \beta \max \left\{ A + B E \left[ \log ((W_t - c_t)(1 + \tilde{y})) \right] + g(m_t, \sigma_t^2), \right. \\
\left. \max_{\theta_t \geq \bar{\theta}} \left\{ E \left[ A + B \log ((W_t - c_t)(1 + \tilde{y} + \theta(\tilde{r}_t - \tilde{y}))) + g(\tilde{m}_{t+1}, \sigma_t^2) \right] \right\} \right\}. \hspace{1cm} (9)$$

The optimal consumption from the first-order condition is $c_t^* = \frac{1}{1 + \beta B} W_t$. Constants $A$ and $B$ can be solved by inserting the optimal consumption back into expression (9) and matching coefficients against the value-function conjecture in expression (8). The value function then simplifies to

$$V(W_t, m_t, \sigma_t^2) = \frac{\beta \log \beta + (1 - \beta) \log (1 - \beta)}{(1 - \beta)^2} + \frac{1}{1 - \beta} \log W_t + g(m_t, \sigma_t^2),$$  \hspace{1cm} (10)$$

where $g(m_t, \sigma_t^2)$ solves the following functional equation:

$$g(m_t, \sigma_t^2) = \beta \max \left\{ \frac{1}{1 - \beta} E \left[ \log (1 + \tilde{y}) \right] + g(m_t, \sigma_t^2), \frac{1}{1 - \beta} \max_{\theta_t \geq \bar{\theta}} \left\{ E \left[ \log (1 + \tilde{y} + \theta(\tilde{r}_t - \tilde{y})) \right] + E \left[ g(\tilde{m}_{t+1}, \sigma_t^2) \right] \right\} \right\}. \hspace{1cm} (11)$$

Expression (10) verifies the conjecture about the form of the value function. I note that while the posterior mean is uncertain at time $t$, the posterior variance $\sigma_{t+1}$ is deterministic. The solution to expression (11), which does not depend on the investor’s wealth, determines the investor’s optimal trading behavior in the model. The first argument in the outer maximization problem is the payoff that the investor receives if he does not trade. In this case the investor only receives the income stream $\tilde{y}$ and, because he does not update his beliefs, he must forever remain “stuck” in the
same belief state \((m_t, \sigma_t^2)\). Thus, \(g(m_t, \sigma_t^2)\) in these states solves a stochastic consumption-savings problem. The solution to this problem is

\[
g(m_t, \sigma_t^2) = \frac{\beta}{(1-\beta)^2} E[\log(1 + \tilde{y})]. \tag{12}
\]

Although this value is independent of the prior distribution, I note that the solution is conditional on \(m_t\) and \(\sigma_t^2\) being such that the investor refrains from trading. Using this result, the functional equation for \(g(m_t, \sigma_t^2)\) simplifies to

\[
g(m_t, \sigma_t^2) = \max \left\{ \frac{\beta}{(1-\beta)^2} E[\log(1 + \tilde{y})], \frac{\beta}{1-\beta} \max_{\theta_t \geq \tilde{\theta}} \left\{ E[\log(1 + \tilde{y} + \theta(\tilde{r}_t - \tilde{y}))] \right\} + \beta E[g(\tilde{m}_{t+1}, \sigma_{t+1}^2)] \right\}. \tag{13}
\]

Function \(g(m_t, \sigma_t^2)\) can also be computed explicitly for the boundary states in which the investor has resolved all uncertainty about his abilities (i.e., \(\sigma_t^2 = 0\)). If the investor’s optimal decision in such a boundary state is to trade, then

\[
g(m_t, 0) = \frac{\beta}{(1-\beta)^2} \max_{\theta_t \geq \tilde{\theta}} \left\{ E[\log(1 + \tilde{y} + \theta(\tilde{r}_t - \tilde{y}))] \right\}. \tag{14}
\]

The static portfolio choice problem in expression (14), which only depends on the mean of the investor’s prior distribution, must be solved numerically. Combining expressions (12) and (14) at the \(\sigma_t^2 = 0\) boundary, the general solution for \(g(m_t, 0)\) is

\[
g(m_t, 0) = \max \left\{ \frac{\beta}{(1-\beta)^2} E[\log(1 + \tilde{y})], \frac{\beta}{(1-\beta)^2} \max_{\theta_t \geq \tilde{\theta}} \left\{ E[\log(1 + \tilde{y} + \theta(\tilde{r}_t - \tilde{y}))] \right\} \right\}. \tag{15}
\]

All values of \(g(m_t, \sigma_t^2)\) can now be computed recursively on a discrete mean-variance grid by using expression (13). We start from \(\sigma_t^2 = 0\) for which expression (15) gives the solution. We then increase the variance of beliefs slightly. The key observation is that if an investor stands in grid point \((m^{(i)}, \sigma^{(k)})\), the current value \(g(m^{(i)}, \sigma^{(k)})\) only depends on the “next-period” values.
$g(m^{(j')}, \sigma^{(k+1)})$. There is no feedback from $k$ to $k + 1$ because the investor always transitions to a lower variance state, $k \rightarrow k + 1$. Hence, if we know $g(m^{(j')}, \sigma^{(k+1)})$ for all posterior mean values $m^{(j')}$, then the expectation $E[g(\bar{m}_{t+1}, \sigma^2_{t+1})]$ in expression (13) is approximately

$$E[g(\bar{m}_{t+1}, \sigma^2_{t+1})] \approx \sum_{j'=1}^{N_\mu} \Pr(m^{(j')}|m^{(j)}, \sigma^{(k)}) \ast g(m^{(j')}, \sigma^{(k+1)}),$$

(16)

where $N_\mu$ is the number of grid points for the mean. The transition probability $Pr(m^{(j')}|m^{(j)}, \sigma^{(k)})$ from a prior-mean state $m^{(j)}$ to a posterior-mean state $m^{(j')}$ is given by the investor’s prior distribution together with the belief dynamics from expression (6). The value of $g(m^{(j)}, \sigma^{(k)})$ can then be determined by evaluating the outer maximization in expression (13). The benefit of this log-utility model is thus its computational efficiency. Because each grid point needs to be visited only once (i.e., without value-function iterations), the entire problem can be solved in a reasonable amount of time even for large grids. I use a grid with one thousand points in both the mean and variance dimensions for a grid with one million elements.

The form of function $g(m_t, \sigma^2_t)$ in expression (13) is instructive about the investor’s behavior in this model. An investor’s decision to trade depends on the future evolution of beliefs which, in turn, depend crucially on the variance of beliefs, $\sigma^2_t$. However, if the investor decides to trade, the precise amount invested, $\theta_t$, is independent of $\sigma^2_t$. This disconnect arises because a log-utility investor’s intertemporal hedging demand is zero. Thus, this model cleanly isolates the option-value-of-trading mechanism: the investor will trade, even if he expects to lose, if by doing so he learns about his own abilities as a trader.

While the binomial-model investor never stops trading after a positive outcome, a continuous-return investor may do so. Suppose that $m_t < 0$, so that the investor is pessimistic about his trading abilities. Although a positive outcome is an unequivocally good signal in that $m_{t+1} > m_t$, the new observation also lowers the variance of beliefs. This decrease in variance may discourage the trader from trading again. Despite having observed a positive outcome, the investor can be more confident in that his true ability is below the critical threshold determined by the minimum
trade size and income from other sources.

B. Structural Model of the Population

I add belief distributions on top of the trading model to create a structural model of investor population that can be estimated from the data. The first part of the model is the life-cycle model detailed above. The second part consists of distributions for investors’ prior beliefs.

I draw the mean of each investor’s prior distribution from a normal distribution $N(\mu^*, \sigma^*)$. I then (independently) draw the variance of the prior distribution from a gamma distribution with a shape parameter $\alpha^*$ and a location parameter $\beta^*$. (I use asterisks to emphasize that these parameters are specific to the investor population, not to any one individual.) The supports of these distributions coincide with the supports of both the mean ($m_t$) and the variance ($\sigma_t^2$) of each investor’s prior distribution.

I assume that fraction $\eta$ of investors exit each period for reasons that fall outside the scope of the trading model. These unmodeled reasons could include exits driven by unemployment, serious illness, death, (independent) negative signals about one’s own abilities, as well as exits by disposition-effect investors. (A disposition-effect investor may “unintentionally” complete a short-term trade if the stock price shoots up on the purchase day.) I estimate the random exit frequency $\eta$ from the data along with the four population-distribution parameters. The estimate of this parameter measures the importance that the learning mechanism plays in explaining these investors’ actions.

C. Identifying Structural Parameters from Conditional Exit Rates and Performance-Exit Sensitivity

Because the structural model does not yield closed-form estimation equations, I use the simulated method of moments (SMM) for indirect inference. I draw hypothetical investors from a trial belief distribution, put them through the trading model, and 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 trading after date $t$ conditional on still trading at date $t$. These rates identify the structural parameters, because they depend on both the mean and variance of the prior distribution. An increase in the mean of the prior belief, given everything else, increases myopic demand, which in turn decreases the exit propensity. An increase in the variance of the prior belief increases the option value of trading and thus increases the number of traders who trade at least once. An increase in the variance also leads to more extreme belief shifts. For example, an investor with a diffuse prior distribution might trade initially because of the high option value of trading, but he is also very likely to quit because each new observation influences his beliefs dramatically.

I use the conditional exit rates over the first ten trading events as moment conditions. I also include the date-0 enter/stay out decision that takes the value of one for investors who trade at least once and the value of zero for those who never trade. The last moment condition is the slope coefficient from a regression of date-1 exit decision on a constant and the realized return. By using the slope coefficient from an exit regression (instead of the average realized return), I allow for the possibility that disposition-effect induced round-trip trades bias realized returns upwards. This slope-coefficient moment condition forces the model to identify the structural parameters from the covariance between returns and exit decisions. I use these twelve moments to estimate the five parameters ($\mu^*, \sigma^*, \alpha^*, \beta^*, \eta$) to have an overidentified model.

I use SMM to estimate the five parameters that describe the investor population. For an investor with prior beliefs $(m_0, \sigma^2_0)$, I let $H(m_0, \sigma^2_0, \eta)$ denote an $12 \times 1$ vector of the conditional exit rates moments and the exit-regression slope coefficient. The first element of this moment vector is either zero (do not trade) or one (trade at least once); the second element is the probability that an investor who trades at least once exits after the first trade; the third element is the probability that an investor who trades at least twice exits after the first trade; and so forth. The model implied
The $k^{th}$ moment is then

\[
M_k(\theta) = \frac{\int_0^\infty \phi(m_0; \mu^*, \sigma^*) \int_0^\infty \varphi(m_0; \mu^*, \sigma^*) \Pr(1_k|m_0, \sigma_0^2) H_k(m_0, \sigma_0^2, \eta) \, dm_0 \, d\sigma_0^2}{\int_0^\infty \phi(m_0; \mu^*, \sigma^*) \int_0^\infty \varphi(m_0; \mu^*, \sigma^*) \Pr(1_k|m_0, \sigma_0^2) \, dm_0 \, d\sigma_0^2},
\]

where $\theta$ is a vector of the structural parameters $\mu^*$, $\sigma^*$, $\alpha^*$, $\beta^*$, and $\eta$, $\varphi(\cdot)$ is the normal density, $g(\cdot)$ is the gamma density, $\Pr(1_k|m_0, \sigma_0^2)$ is the probability that an investor with initial beliefs $(m_0, \sigma_0^2)$ reaches the $k^{th}$ moment condition, and $k$ indexes moment conditions from 0 to 11. The inclusion of $\Pr(1_k|m_0, \sigma_0^2)$ turns each exit rate conditional: $M_k(\theta)$ is the date-$k$ exit rate for investors still alive at date $k - 1$. ($\Pr(1_{11}|m_0, \sigma_0^2)$ is the same as $\Pr(1_1|m_0, \sigma_0^2)$, because the slope-coefficient moment condition applies to all investors who trade at least once.) I compute the exit rates in the simulations by assuming that investors’ prior beliefs are well calibrated. For example, if an investor’s posterior mean crosses the quit-trading threshold for returns $r_t < -5\%$, the investor’s exit rate is equal to the probability of drawing a return less than $-5\%$ given the prior distribution.

The use of the mean-variance belief grid in the life-cycle model solution significantly simplifies the estimation of the full structural model. I first compute and record the twelve elements of $H(m_0, \sigma_0^2, \eta)$ for each grid point: the date-0 entry decision, the conditional exit rates for dates 1, ..., 10, and the expectations required for the exit regression. The population-distribution parameters in vector $\theta$ determine how likely each of these grid points is in the current trial population: what is the probability of drawing an investor with prior beliefs close to some grid point $(m^{(j)}, \sigma^{(k)})$? Expression (17) can thus be evaluated by applying appropriate weights to each of the grid points. The SMM-estimator $\hat{\theta}$ is then

\[
\hat{\theta} = \arg \min_\theta \left( \hat{M} - M(\theta) \right) W \left( \hat{M} - M(\theta) \right),
\]

where $\hat{M}$ is the vector of estimated moments from the data, $M(\theta)$ is the vector of model-implied moments from Equation (17), and $W$ is an arbitrary positive-definite weighting matrix. I use the optimal weighting matrix for estimation. Because I use conditional exit rates as moment conditions, the covariance matrix for the first eleven moment conditions is diagonal: for each investor who is
still alive at date \( t \), the date \( t - 1 \) exit realization must have been zero.

I set the other parameters of the model as follows. First, I set the subjective discount rate to \( \beta = 0.997 \), the mean income yield to \( \mathbb{E}(\tilde{y}) = 0.002 \), and the standard deviation of the income yield to \( \sigma(\tilde{y}) = 0.027 \). I assume, for simplicity, that \( \tilde{y} \) has a binomial distribution. I note that because the income yield is independent of the return on the trading opportunity, \( \tilde{y} \) acts as background risk that increases the investor’s effective risk-aversion coefficient.\(^{12}\) I choose \( \beta \), \( \mathbb{E}(\tilde{y}) \), and \( \sigma(\tilde{y}) \) based on the estimates in studies as those by Krebs (2003) and Storesletten, Telmer, and Yaron (2007). I assume that each period in the model has a length of one week and scale these parameters accordingly. I set the trading-opportunity return volatility to \( \hat{\sigma}_\varepsilon = 0.08 \), which is the volatility estimate in the data for first-time traders.\(^{13}\) Lastly, I set the minimum trade size to \( \bar{\theta} = 2\% \). With this assumption, an investor with a wealth of $25,000 needs to commit at least $500 to the trading opportunity.

\textit{D. Results}

Table V Panel A shows that the discrepancies between the model moments and data moments are both statistically and economically small. The model matches both the initial entry decision and performance-exit sensitivity almost perfectly. Similarly, both in the model and in the data over a quarter of investors exit after just one trade. The model can deliver this high exit rate because investors find the first-day signal very valuable. Even if an investor is very uncertain about his trading abilities, the variance of returns constitutes an upper bound for the posterior variance. (Expression (7) shows that \( \lim_{\sigma^2_0 \to \infty} \sigma^2_I = \sigma^2_\varepsilon \).) Thus, no matter how uncertain an investor is about his beliefs, the standard deviation of beliefs is always below 8% after the first trade, below \( \frac{8\%}{\sqrt{2}} = 5.7\% \) after the second trade, and so forth. The largest conditional exit rate discrepancy between the data and the simulations is the date-7 exit rate. Here, the model delivers an exit rate of 6.63% while the actual exit rate in the data is 6.32%. The test of over-identifying restrictions (reported on in Panel B) fails to reject the model at a 0.05 significance level.

\(^{12}\)See, for example, Heaton and Lucas (2000).

\(^{13}\)The volatility estimate over all short-term trades is 0.05. Both of these estimates are very high because individuals concentrate their short-term trades (both round-trip and other trades) in both high-volatility stocks and days. The structural-model estimates are essentially the same for both of these volatility figures.
Panel B reports on the structural parameter estimates. (Instead of describing the gamma distribution in terms of its primitive parameters, I report on the distribution’s mean and variance.) The estimate of the random exit frequency $\eta$ shows that almost all exits in the data fall within the scope of the life-cycle model. Only 1.26% of investors quit each date for non-learning reasons.

The mean-distribution estimates suggest that most individuals are very pessimistic about their trading abilities. The mean of each trader’s prior distribution is drawn from a normal distribution with a mean of $-0.509$ and a standard deviation of $0.189$. This suggests that just 0.36% of all individuals have a positive prior mean about their abilities. However, because investors in the model always consider what might be, the level of the prior is largely inconsequential for these investors. It is the upper-tail probability $\Pr(\mu > 0)$ that plays a crucial role in determining which investors trade and which do not. The low mean in the population-wide distribution is the model’s way of ensuring that this probability is so low for most individuals that they have no incentive to trade.

Panel C report on traders’ abilities based on the estimated model. I report both the probability that an investor is skilled and the probability that an investor receives a positive first-day return. (For non-active investors, the first-day return computation applies to a hypothetical trade.) The average trader assigns a probability of 0.359 to being skilled and a probability 0.344 to obtaining positive first-day return. These numbers do not coincide because returns reflect two sources of uncertainty: the variance in ability and the variance in returns. For example, although 5% of active traders know that they are skilled, the probability of a positive one-day return is still just 0.727 because of the randomness in returns. Thus, even these traders make money in only from seven out of ten trades.
Panel D shows the traders’ 95% confidence intervals about their trading abilities. For this computation I partition traders into three groups based on how precise their prior distributions are. For example, the top 5% of investors, who know that they are skilled, believe that their expected one-day return from trading is just 0.14%. Thus, although I impose no constraints on individuals’ prior distributions, the model yields reasonable levels of beliefs for those traders who believe that they can profit by trading. As the dispersion in beliefs increases, the traders in the model become more pessimistic about their trading abilities. For example, the 95% confidence interval for the top half of the trader population is just $-0.2475 < \mu < -0.0126$. These investors trade because of the (very) small chance that the true $\mu$ is greater than zero. A number of outliers influence the confidence-interval estimates for all traders on the last row of Panel D. These outliers represent those investors who learn the most from the first signals and often quit after just a few trades.

Most active traders think that they will lose money but still trade. At the same time, some investors stay out even though they might be skilled. If we assume that investors’ beliefs are unbiased, then the belief distributions coincide with the true skill distributions in the population. Thus, the ability distributions in Panel C also imply that 36% of those who try active trading are skilled. The remaining two thirds are unskilled but they trade because of the chance that they might be wrong. By contrast, only 0.9% of those who stay out are skilled. However, given the relative sizes of the trader and non-trader groups, these numbers imply that more than a third (36.6%) of all skilled investors never trade actively.

The total amount of (genuinely) skilled investors is 2.3% in the calibrated model. This figure is within the estimate in Coval, Hirshleifer, and Shumway (2005), who find persistent superior performance among the investors in the top decile. This estimate is also in line with Grinblatt, Keloharju, and Linmainmaa (2009a), who find superior performance among the 4% of individuals who receive the highest possible score in the Finnish Armed Forces intelligence test.

I compute these intervals for the left-truncated normal distribution and not for the underlying untruncated distribution. Thus, these confidence intervals correspond to investors’ expected one-day returns from the trading opportunity.
E. Applications

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. Instead, investors only learn about their true skill $\mu$ as they trade. However, because investors with low abilities are more likely to be unsuccessful, and because unsuccessful investors are more likely to quit, the investors who survive are better than the average investor.

I study this selection process by inserting the structural parameter estimates back into the model. Table V Panel E tabulates the cumulative exit rates for the skilled and unskilled investors and shows how the average ability of the trader pool changes over the first ten trades. The calibrated model shows that an unskilled trader quits after the first trade with probability 0.26. The probability that he quits after the first or the second trade is 0.42. Thus, two fifths of the unskilled investors reach the correct conclusion about their abilities already by the second trade. The cumulative exit rates are lower (0.20 and 0.34, respectively) for the skilled investors. Thus, one third of the skilled investors are unlucky and exit after the second trade despite being, unbeknownst to them, skilled. These exit rate differences change the composition of the trader pool, increasing the skill of the average investor monotonically over time. While 35.9% of the investors trading for the first time are skilled, this fraction increases to 42.1% by the fifth trade and to 45.2% by the tenth trade. Because of this survival mechanism, trade performance correlates positively with trading experience despite the fact that investors’ abilities are fixed.

Another application of the model is to measure a reverse survivorship bias. This bias arises because investors quit after unsuccessful trades. If an investor never quits, then the average skill measured from the data would essentially be a draw from the investor’s prior distribution. However, because investors quit after unsuccessful trades, the skill estimates from the data are biased downwards. The skill estimates are too low relative to investors’ true skills because the data oversample poor performance.

This observation follows from optional stopping-time theorems. If $\xi$ is mean zero and a mar-
tingale, 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 $\tilde{T}$:

$$E \left[ \frac{1}{\tilde{T}} \sum_{t=1}^{\tilde{T}} \tilde{\varepsilon} \right] = \text{cov} \left( \frac{1}{\tilde{T}}, \sum_{t=1}^{\tilde{T}} \tilde{\varepsilon} \right).$$

(19)

The gap between the true and observed skill is economically significant. Although the average active trader obtains a positive first-day return with a probability of 0.344, the probability of success (estimated by using data simulated from the model) is just 0.267. Thus, this reverse survivorship bias is an important consideration when measuring traders’ abilities. Expression (19) shows that the magnitude of this bias depends on how sensitive investors’ exit decisions are to poor performance. Because active traders are very sensitive to poor performance, the resultant bias in their performance estimates is large.

**F. Robustness Tests**

In addition to the tests detailed above, I also completed several robustness checks of the results. First, instead of using the infinite-horizon life-cycle model as the basis of the structural model, an earlier version of this paper used the simplified binomial model from Section II. The estimates from this simplified model are consistent with the main results presented in this paper. For example, these alternative-model estimates suggest that approximately 4% of all individual investors are skilled. The binomial assumption, however, lets investors to learn too quickly about their abilities. As a consequence, almost all unskilled investors quit by the tenth trade and the model encounters difficulties in fitting the exit rates to the data.

I also constructed an alternative sample of traders by extending the analysis to all 75,963 individual investors who make any speculative purchases during the sample period. (I follow Barber and Odean (2002) and define a speculative purchase as a purchase that follows a complete sale of other position within three weeks.) I further distinguish this alternative sample from the main sample by, first, excluding all round-trip trades and, second, by analyzing traders’ one-month market-adjusted returns. Both the qualitative tests (Section IV) and the structural model estimates
support the learning story in this alternative sample. For example, an investor whose purchase underperforms the market is 4.9% more likely to stop trading than an investor who outperforms the market.

VI. Conclusion

I show that a Bayesian 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 poor performance. The key mechanism in the learning model is the option value of trading. Because investors learn by trading, they have an incentive to trade even if they think that they are unskilled. Even a small probability of being skilled can outweigh the expected trading losses.

A structural trading model matches many of the salient features of investor trading records. This model gets right the number of investors who begin active trading and fits the decreasing pattern in the conditional exit rates. The performance-exit sensitivity also is the same in the estimated model and in the data. The estimates of this model suggest that, first, most traders are pessimistic about their abilities; second, that up to two thirds of the investors who start active trading are unskilled; third, that over a quarter of all traders believe that their chance of being skilled is less than one-in-seventy; fourth, that approximately 2.3% of all investors (including non-active traders) are genuinely skilled; fifth, that because unskilled investors are more likely to stop trading, performance correlates positively with experience; and sixth, that realized returns are significantly downwards-biased measures of investors’ true abilities.

Although I assume that investors are rational and have well-calibrated prior distributions, the results neither prove nor disprove these assumptions. For example, I do not know if investors learn by trading because there are limits to paper trading or if there is some other reason for this mechanism. Also, investors’ prior distributions may be looser (or tighter) than what they should be given the sum of investors’ all knowledge. However, my main conclusions are robust to these
reservations. Investors learn from their own experiences, and the evolution of beliefs that results from this is an important driver of investor behavior.
This Appendix presents the solution to Section II’s trading model. I first formulate the investor’s optimization problem and then describe a numerical solution method.

A. Investor’s Optimization Problem

To maximize utility over terminal wealth, the investor chooses the optimal amount to trade at each date. This task is complicated not only by the changes in \( \hat{p}_t \), but also by the minimum trade size requirement. I write the investor’s optimization problem as a dynamic programming problem and then solve it numerically as a large-scale fixed point 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 \( \bar{x} \) or quit and consume wealth \( W_{T-1} \). The date \( T-1 \) optimization problem is then:

\[
V_{T-1}(W_{T-1}, \hat{p}_{T-1}) = \max \left\{ \max_{x_{T-1} > \bar{x}} \left\{ \frac{\hat{p}_{T-1} (W_{T-1} + x_{T-1})^{1-\gamma}}{1-\gamma} + (1 - \hat{p}_{T-1}) \frac{(W_{T-1} - x_{T-1})^{1-\gamma}}{1-\gamma} \right\}, W_{T-1}^{1-\gamma} \right\}, \quad (20)
\]

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. The optimal investment for an unconstrained investment problem in expression (20) is

\[
x_{T-1}^* = \frac{\frac{1}{\hat{p}_{T-1}} - (1 - \hat{p}_{T-1})^{\frac{1}{\gamma}}}{\frac{1}{\hat{p}_{T-1}} + (1 - \hat{p}_{T-1})^{\frac{1}{\gamma}}} W_{T-1}. \quad (21)
\]
I can then replace the inner maximization problem by its value \( V_{T-1}(W_{T-1}, \hat{p}_{T-1}) \):

\[
V_{T-1}(W_{T-1}, \hat{p}_{T-1}) =
\begin{cases}
\frac{W_{T-1}^{1-\gamma}}{1-\gamma} \phi^{1-\gamma} \left( \frac{1}{\hat{p}_{T-1}} - (1 - \hat{p}_{T-1}) \right)^{\frac{1}{\gamma}} & \text{if } x^*_{T-1} \geq \bar{x}, \\
\frac{W_{T-1}^{1-\gamma}}{1-\gamma} \left( \hat{p}_{T-1}(1 + \bar{T}_{T-1})^{1-\gamma} + (1 - \hat{p}_{T-1})(1 - \bar{T}_{T-1})^{1-\gamma} \right) & \text{if } x^*_{T-1} < \bar{x} \text{ and } W_{T-1} > \bar{x}, \\
-\infty & \text{otherwise},
\end{cases}
\]

(22)

where \( \bar{T}_{T-1} \equiv \frac{\bar{x}}{W_{T-1}} \) is the minimum trade size as a fraction of wealth. If an unconstrained investor would invest less than the minimum amount, \( x^*_{T-1} < \bar{x} \), I evaluate the utility of this investor at the boundary, \( x^*_{T-1} = \bar{x} \). If the investor cannot afford to trade the minimum amount, I set the utility to minus infinity.

The investor’s date \( T - 1 \) indirect utility compares expression (22) to the utility from quitting and consuming. The utility function thus has the following multiplicative form:

\[
V_{T-1}(W_{T-1}, \hat{p}_{T-1}) = \max \left\{ V_{T-1}(W_{T-1}, \hat{p}_{T-1}), \frac{W_{T-1}^{1-\gamma}}{1-\gamma} \right\} = \frac{W_{T-1}^{1-\gamma}}{1-\gamma} k_{T-1}(W_{T-1}, \hat{p}_{T-1}),
\]

(23)

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

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 \( W_t \). I assume that the investor’s indirect utility function has the same functional form as in expression (23). The investor’s date \( t \) optimization problem is then

\[
V_t(W_t, \hat{p}_t) = \max \left\{ \max_{x_t > \bar{x}} \left\{ \hat{p}_t \left( \frac{W_t + x_t}{1 - \gamma} \right) k_t^{S_{t+1}} + (1 - \hat{p}_t) \left( \frac{W_t - x_t}{1 - \gamma} \right) k_t^{F_{t+1}} \right\}, \frac{W_t^{1-\gamma}}{1-\gamma} \right\},
\]

(24)

where \( k_t^{S_{t+1}} \equiv k_{t+1}(W_{t+1}, \hat{p}_{t+1}) \) is the date \( t + 1 \) multiplier in the indirect utility function if the trade is successful, and \( k_t^{F_{t+1}} \) is the multiplier if the trade fails. Although coefficients \( k_t^{S_{t+1}} \) and \( k_t^{F_{t+1}} \) depend on the optimal \( x_t \), I assume here that \( k_t^{S_{t+1}} \) and \( k_t^{F_{t+1}} \) are known. The numerical solution
approach makes initial guesses about these values and then iterates the maximization problem to obtain the actual values. Solving for the optimal investment, the inner maximization problem in expression (24) simplifies to

\[
V_t^I(W_t, \hat{p}_t) = \begin{cases} 
W_t^{1-\gamma} & \text{if } x_t^* \geq \bar{x}, \\
W_t^{1-\gamma} \left( \hat{p}_t (1 + \theta_t) - (1 - \hat{p}_t) (1 - \theta_t) \right) & \text{if } x_t^* < \bar{x} \text{ and } W_t > \bar{x}, \\
-\infty & \text{otherwise},
\end{cases}
\]

and the indirect date \( t \) utility function has the conjectured form:

\[
V_t(W_t, \hat{p}_t) = \frac{W_t^{1-\gamma}}{1-\gamma} k_t(W_t, \hat{p}_t).
\]

**B. Numerical Solution**

The problem cannot be solved recursively, because the date-\( t \) choice depends on future wealth and, at the same time, the future wealth depends on the date-\( t \) choice. However, because there are only two possible outcomes at each date, the problem has the shape of a non-recombining binomial tree. Starting from any wealth-belief pair \((W_t, \hat{p}_t)\), the next node is either \((W_t + x_t^*, \hat{p}_t^S)\) or \((W_t - x_t^*, \hat{p}_t^F)\), where \( x_t^* \) is the optimal trade size or zero, if the investor quits.

I make initial guesses about each element in the investor’s wealth process \(\{W_t^j\}_{t=1,...,T-1}^{j=1,...,2^{t-1}}\), where \( j = 1, \ldots, 2^{t-1} \) indexes all date-\( t \) nodes of the tree. I can then solve the tree backwards, because the initial wealth guesses determine the value function values in the subsequent nodes of the tree. The objective is to find \(\{W_t^j\}_{t=1,...,T-1}^{j=1,...,2^{t-1}}\) 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

\[
W_{t+1} = \begin{cases} 
W_t + x_t^{j*} & \text{if a success} \\
W_t - x_t^{j*} & \text{if a failure.}
\end{cases}
\]

36
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 $x_i^*$. I iterate until convergence.

This solution method is fast and accurate. However, this method does impose limits on the investment horizon $T$, because the problem size is exponential in $T$. While a problem with $T = 15$ trading periods has just 16,383 decision nodes, a problem with $T = 20$ has over half a million nodes.
REFERENCES


Figure 1. The Evolution of Wealth and Beliefs in the Trading Model. A Bayesian investor has a wealth of $W_t$ and a beta distributed prior with parameters $(\alpha_t, \beta_t)$ 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. I denote the probability of a good outcome by $\hat{p}_t \equiv \frac{\alpha_t}{\alpha_t + \beta_t}$. 
Figure 2. Changes in Optimal Behavior in the Trading Model. 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 $W_1 = 10,000$, a relative risk aversion coefficient of $\gamma = 2$, and an investment horizon of $T = 16$. The minimum trade size is $200$. The investor can trade an amount $x_t$ at dates $t = 1, \ldots, T - 1$. A trade is successful with a probability $p$ and returns $2x$. It is a failure with a probability $1 - p$ and returns zero. 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 that pays no interest. The investor observes the outcome if he trades, in which case he updates his beliefs using Bayes rule. All possible outcomes (with beliefs and wealth) are shown up to the investor’s fourth trading decision.
Figure 3. Optimal Initial Behavior in the Trading Model. 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 \( t = 1, \ldots, T - 1 \). A trade is successful with a probability \( p \) and returns \( 2x \). It is a failure with a probability \( 1 - p \) and returns zero. 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 that pays no interest. The investor observes the outcome if he trades, in which case he updates his beliefs using Bayes rule. This figure shows investors' optimal decisions at date 1 as functions of the two parameters of the beta distributed prior, \( \alpha_1 \) and \( \beta_1 \). (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-hand side 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 three panels represent investors with relative risk aversion coefficients of two, five, and ten.
Figure 4. Cross-Sectional Regressions of Active Traders’ Exit Decisions against a Success Dummy Variable. This figure plots the slope coefficient estimates from regressions of active traders’ exit decisions against a success dummy variable. I estimate separate regressions for investors trading for the first time, investors trading for the second time, and so forth. The regressions contain controls for each investor’s average career success rate, the change in wealth since the first trade, and calendar-time fixed effects. The dashed lines show the 95% confidence interval for the slope estimates. The data are the complete trading records of all active traders in Finland from January 1995 through November 2002.
Panel A: Distribution of Initial Trade Sizes

Panel B: Average Trade Sizes Relative to First Trade

Figure 5. Distribution of Initial Trade Sizes, Break-Even Returns, and the Increase in Surviving Traders’ Trade Sizes. Panel A shows the distribution of trade sizes for active traders who are trading for the first time. Each category is 500 euro wide with x-axis labels giving the midpoint of each category. The line shows the implied break-even return, which is the percentage difference between the purchase and sale price that is required for the investor to recover the 8.42 euro + 0.15% commission. The bars show the fraction of trades that fall into each trade size category. Panel B shows the average trade size of an investor still alive at date t. Trade sizes are standardized by the size of the first trade. The dashed lines show the 95% confidence interval for the average. The data are the complete trading records of all active traders in Finland from January 1995 through November 2002.
Table I
Trading Activity and Demographics of Active Traders

In this table I compare active traders’ demographics and trading activity to other individual investors. I define an active trader as an individual investor who completes at least one intraday round-trip trade during the sample period. An investor completes an intraday round-trip trade if he 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 January 1995 through November 2002. I drop investors with round-trip trades during the first three months of the sample. I tabulate separately the number of trades, average trade sizes, and average betas for short-term trades (i.e., round-trip trades plus other speculative trades) and other trades. I compute Scholes and Williams (1977) betas using six months of daily data leading up to each trade. I use a constrained market index, which caps the weight of individual stocks in the index to 10%, to compute betas. Fraction-of-purchases-captured statistic is the number of days with short-term trades divided by the total number of stock-purchase days during an investor’s active phase. The active phase is the period of time between the investor’s first and last short-term trade.

Panel A: Active Traders, \( N = 22,529 \)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>SD</th>
<th>25</th>
<th>50</th>
<th>75</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>39.5</td>
<td>13.8</td>
<td>29.0</td>
<td>38.0</td>
<td>49.0</td>
</tr>
<tr>
<td>Short-Term Trades</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of Trades</td>
<td>24.8</td>
<td>78.3</td>
<td>2.0</td>
<td>4.0</td>
<td>15.0</td>
</tr>
<tr>
<td>Average Trade Size, EUR</td>
<td>13,231.0</td>
<td>49,228.2</td>
<td>2,074.9</td>
<td>4,493.1</td>
<td>10,553.8</td>
</tr>
<tr>
<td>Average Beta</td>
<td>1.73</td>
<td>0.71</td>
<td>1.32</td>
<td>1.80</td>
<td>2.17</td>
</tr>
<tr>
<td>Average Profit, EUR</td>
<td>-7.1</td>
<td>968.7</td>
<td>-82.6</td>
<td>-25.3</td>
<td>26.9</td>
</tr>
<tr>
<td>Fraction of Purchases Captured</td>
<td>0.823</td>
<td>0.258</td>
<td>0.667</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>Other Trades</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of Trades</td>
<td>88.5</td>
<td>119.3</td>
<td>23.0</td>
<td>52.0</td>
<td>110.0</td>
</tr>
<tr>
<td>Average Trade Size, EUR</td>
<td>9,668.2</td>
<td>59,138.9</td>
<td>2,285.3</td>
<td>4,318.2</td>
<td>8,411.7</td>
</tr>
<tr>
<td>Average Beta</td>
<td>1.51</td>
<td>0.47</td>
<td>1.20</td>
<td>1.50</td>
<td>1.82</td>
</tr>
<tr>
<td>Male</td>
<td>81.3%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Panel B: Other Individual Investors, \( N = 527,436 \)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>SD</th>
<th>25</th>
<th>50</th>
<th>75</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>46.2</td>
<td>18.6</td>
<td>33.0</td>
<td>47.0</td>
<td>59.0</td>
</tr>
<tr>
<td>Regular Trades</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of Trades</td>
<td>6.4</td>
<td>14.4</td>
<td>1.0</td>
<td>2.0</td>
<td>6.0</td>
</tr>
<tr>
<td>Average Trade Size, EUR</td>
<td>6,045.0</td>
<td>238,620.8</td>
<td>1,161.6</td>
<td>2,310.0</td>
<td>4,848.1</td>
</tr>
<tr>
<td>Average Beta</td>
<td>1.41</td>
<td>0.76</td>
<td>0.90</td>
<td>1.44</td>
<td>1.92</td>
</tr>
<tr>
<td>Male</td>
<td>59.2%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

46
Heterogeneity and Persistence in Active Traders' Performance

This table reports on regressions that measure active traders’ performance persistence. Each observation is a single short-term trade. A short-term trade is an intraday round-trip trade or a purchase that follows a round-trip trade within two weeks. The data are the complete trading records of all active traders in Finland from January 1995 through November 2002. I drop investors with round-trip trades during the first three months of the sample. In Panel A I regress the return in trade \( t \) against the investor’s average return from all previous trades. In Panel B I regress the success-dummy in trade \( t \) against the investor’s average success rate from all previous trades. Both regressions include calendar-time and event-time fixed effects. The calendar-time FE are monthly dummies and the event-time FE are dummy variables for trading dates \( 1, \ldots, 50 \). Trades \( t > 50 \) are the omitted category. The regressions are estimated in the full sample (\( N = 449,362 \)), in a sample with investors’ early trades (\( N = 96,836 \)), and in a sample with investors’ late trades (\( N = 352,526 \)). Standard errors, clustered by investor, appear in parentheses.

### Panel A: Regression of Date-\( t \) Return against Career Average Return

<table>
<thead>
<tr>
<th></th>
<th>Full Sample</th>
<th>Early Career, ( t \leq 10 )</th>
<th>Late Career, ( t &gt; 10 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Career Average Return, ( \frac{1}{t-1} \sum_{\tau=1}^{t-1} r_\tau )</td>
<td>0.166</td>
<td>0.107</td>
<td>0.338</td>
</tr>
<tr>
<td></td>
<td>(0.012)</td>
<td>(0.009)</td>
<td>(0.035)</td>
</tr>
<tr>
<td>Adjusted ( R^2 )</td>
<td>0.011</td>
<td>0.014</td>
<td>0.014</td>
</tr>
<tr>
<td>SD(Career Average Return)</td>
<td>0.018</td>
<td>0.033</td>
<td>0.010</td>
</tr>
</tbody>
</table>

### Panel B: Regression of Date-\( t \) Success Dummy against Career Success Rate

<table>
<thead>
<tr>
<th></th>
<th>Full Sample</th>
<th>Early Career, ( t \leq 10 )</th>
<th>Late Career, ( t &gt; 10 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Career Success Rate, ( \frac{1}{t-1} \sum_{\tau=1}^{t-1} s_\tau )</td>
<td>0.283</td>
<td>0.141</td>
<td>0.543</td>
</tr>
<tr>
<td></td>
<td>(0.008)</td>
<td>(0.005)</td>
<td>(0.013)</td>
</tr>
<tr>
<td>Adjusted ( R^2 )</td>
<td>0.018</td>
<td>0.018</td>
<td>0.024</td>
</tr>
<tr>
<td>SD(Career Success Rate)</td>
<td>0.193</td>
<td>0.327</td>
<td>0.123</td>
</tr>
</tbody>
</table>
This table reports on regressions that measure how sensitive active traders’ exit decisions are on performance. Each observation is a single short-term trade. A short-term trade is an intraday round-trip trade or a purchase that follows a round-trip trade within two weeks. The data are the complete trading records of all active traders in Finland from January 1995 through November 2002. 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. I set the dependent variable to one for each investor’s last trade, and to zero for all other trades. In Panel A I regress the exit decision against the investor’s return from current trade, the average return from all previous trades, and the change in the investor’s wealth from the first trade to one day before the current trade. In Panel B I replace continuous returns with success dummy variables. I estimate both 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, . . . , 50. Trades $t > 50$ are the omitted category. I estimate the regressions in the full sample ($N = 441,662$), in a sample consisting of investors’ early trades ($N = 113,944$), and in a sample consisting of investors’ late trades ($N = 327,718$). Standard errors, clustered by investor, appear in parentheses.

<table>
<thead>
<tr>
<th>Panel A: Regression of Date-$t$ Exit Decision on Date-$t$ Return</th>
<th>Full Sample</th>
<th>Early Career, $t \leq 10$</th>
<th>Late Career, $t &gt; 10$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Explanatory Variables</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Date-$t$ Return</td>
<td>$-0.216$</td>
<td>$-0.423$</td>
<td>$-0.085$</td>
</tr>
<tr>
<td>(0.012)</td>
<td>(0.026)</td>
<td>(0.009)</td>
<td></td>
</tr>
<tr>
<td>Career Average Return, $\frac{1}{t-1} \sum_{\tau=1}^{t-1} r_\tau$</td>
<td>$-0.155$</td>
<td>$-0.221$</td>
<td>0.111</td>
</tr>
<tr>
<td>(0.027)</td>
<td>(0.034)</td>
<td>(0.029)</td>
<td></td>
</tr>
<tr>
<td>$(\ln(\text{Wealth}_t) - \ln(\text{Wealth}_0))/100$</td>
<td>0.007</td>
<td>$-0.086$</td>
<td>$-0.105$</td>
</tr>
<tr>
<td>(0.025)</td>
<td>(0.112)</td>
<td>(0.021)</td>
<td></td>
</tr>
<tr>
<td>Adjusted $R^2$</td>
<td>0.098</td>
<td>0.075</td>
<td>0.018</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B: Regression of Date-$t$ Exit Decision on Date-$t$ Success Dummy</th>
<th>Full Sample</th>
<th>Early Career, $t \leq 10$</th>
<th>Late Career, $t &gt; 10$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Explanatory Variables</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Date-$t$ Success Dummy</td>
<td>$-0.019$</td>
<td>$-0.056$</td>
<td>$-0.006$</td>
</tr>
<tr>
<td>(0.001)</td>
<td>(0.002)</td>
<td>(0.000)</td>
<td></td>
</tr>
<tr>
<td>Career Success Rate, $\frac{1}{t-1} \sum_{\tau=1}^{t-1} s_\tau$</td>
<td>$-0.035$</td>
<td>$-0.047$</td>
<td>$-0.001$</td>
</tr>
<tr>
<td>(0.002)</td>
<td>(0.003)</td>
<td>(0.002)</td>
<td></td>
</tr>
<tr>
<td>$(\ln(\text{Wealth}_t) - \ln(\text{Wealth}_0))/100$</td>
<td>$-0.002$</td>
<td>$-0.117$</td>
<td>$-0.101$</td>
</tr>
<tr>
<td>(0.026)</td>
<td>(0.112)</td>
<td>(0.021)</td>
<td></td>
</tr>
<tr>
<td>Adjusted $R^2$</td>
<td>0.099</td>
<td>0.080</td>
<td>0.018</td>
</tr>
</tbody>
</table>
Table IV
Pooled Regressions of Active Traders’ Trade Size Changes on Performance

This table reports on regressions that measure how sensitive active traders’ trade size choices are on performance. Each observation is a single short-term trade. A short-term trade is an intraday round-trip trade or a purchase that follows a round-trip trade within two weeks. The data are the complete trading records of all active traders in Finland from January 1995 through November 2002. I drop investors with round-trip trades during the first three months of the sample. In Panel A I regress the log-size of trade $t + 1$ against the log-size and return of trade $t$. In Panel B I replace the continuous return with a success dummy variable. Both regressions also include the change in the investor’s portfolio value from one day before trade $t$ to one day before trade $t + 1$. I drop an observation if the investor quits after trade $t$. I estimate both 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, . . . , 50. Trades $t > 50$ are the omitted category. I estimate the regressions in the full sample ($N = 447,414$), in a sample consisting of investors’ early trades ($N = 103,370$), and in a sample consisting of investors’ late trades ($N = 344,044$). Standard errors, clustered by investor, appear in parentheses.

### Panel A: Regression of Date-$t + 1$ Log-Trade Size on Date-$t$ Return

<table>
<thead>
<tr>
<th>Explanatory Variables</th>
<th>Full Sample</th>
<th>Early Career, $t \leq 10$</th>
<th>Late Career, $t &gt; 10$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Date-$t$ Return</td>
<td>0.661</td>
<td>0.515</td>
<td>0.773</td>
</tr>
<tr>
<td></td>
<td>(0.090)</td>
<td>(0.109)</td>
<td>(0.131)</td>
</tr>
<tr>
<td>Date-$t$ Log-Trade Size</td>
<td>0.625</td>
<td>0.583</td>
<td>0.637</td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td>(0.005)</td>
<td>(0.008)</td>
</tr>
<tr>
<td>Adjusted $R^2$</td>
<td>0.501</td>
<td>0.450</td>
<td>0.500</td>
</tr>
<tr>
<td>SD(Date-$t$ Return)</td>
<td>0.037</td>
<td>0.045</td>
<td>0.033</td>
</tr>
</tbody>
</table>

### Panel B: Regression of Date-$t + 1$ Log-Trade Size on Date-$t$ Success Dummy

<table>
<thead>
<tr>
<th>Explanatory Variables</th>
<th>Full Sample</th>
<th>Early Career, $t \leq 10$</th>
<th>Late Career, $t &gt; 10$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Date-$t$ Success Dummy</td>
<td>0.092</td>
<td>0.116</td>
<td>0.085</td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
<td>(0.006)</td>
<td>(0.005)</td>
</tr>
<tr>
<td>Date-$t$ Log-Trade Size</td>
<td>0.624</td>
<td>0.580</td>
<td>0.636</td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td>(0.005)</td>
<td>(0.008)</td>
</tr>
<tr>
<td>Adjusted $R^2$</td>
<td>0.502</td>
<td>0.452</td>
<td>0.501</td>
</tr>
<tr>
<td>SD(Date-$t$ Success Dummy)</td>
<td>0.497</td>
<td>0.497</td>
<td>0.498</td>
</tr>
</tbody>
</table>
Table V
Moment Conditions and Parameter Estimates of the Structural Learning Model

This table reports on the estimates of the structural learning model. The model is estimated using data on all active traders in Finland. An active trader is an individual who completes at least one intraday round-trip trade during the sample period. Panel A compares simulated moments to sample moments, Panel B reports on the structural parameter estimates, Panel C tabulates the probabilities that traders assign to being skilled and to realizing a positive first-day return, Panel D reports on the 95% confidence intervals that traders have about their trading abilities, and Panel E reports on skilled and unskilled traders’ cumulative exit rates. The first part of the structural model is a trading model in which a log-utility investor (with risky income from a non-tradable asset) updates beliefs about his abilities as he trades. The second part consists of distributions for population beliefs. I construct each investor’s prior distribution by drawing the mean from a normal distribution and the variance from a gamma distribution. A trader stops trading on day $t$ for non-learning reasons with probability $\eta$ (random exits). The subjective discount rate is $\beta = 0.997$, the mean income yield is $E(\tilde{y}) = 0.002$, the standard deviation of the income yield is $\sigma(\tilde{y}) = 0.027$, the volatility of returns is $\tilde{\sigma}_\epsilon = 0.08$, and the minimum trade size is $\bar{\theta} = 2\%$ of wealth. I estimate the model by using the Simulated Method of Moments with the optimal weighting matrix. The twelve moment conditions are the enter/do not enter decision at date 0, the conditional exit rates for trading dates 1, \ldots, 10, and the slope coefficient from a date-1 OLS regression of an exit dummy variable on realized returns. The rightmost column in Panel B tests for the over-identifying restrictions. Its $p$-value appears in parentheses.

<table>
<thead>
<tr>
<th>Moment Condition</th>
<th>Data</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Date-0 Entry</td>
<td>4.10%</td>
<td>4.10%</td>
</tr>
<tr>
<td>Conditional Exit Rates</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Date 1</td>
<td>25.19%</td>
<td>25.17%</td>
</tr>
<tr>
<td>Date 2</td>
<td>16.50%</td>
<td>16.50%</td>
</tr>
<tr>
<td>Date 3</td>
<td>12.01%</td>
<td>12.27%</td>
</tr>
<tr>
<td>Date 4</td>
<td>10.39%</td>
<td>10.13%</td>
</tr>
<tr>
<td>Date 5</td>
<td>8.69%</td>
<td>8.53%</td>
</tr>
<tr>
<td>Date 6</td>
<td>7.45%</td>
<td>7.46%</td>
</tr>
<tr>
<td>Date 7</td>
<td>6.32%</td>
<td>6.63%</td>
</tr>
<tr>
<td>Date 8</td>
<td>6.03%</td>
<td>5.95%</td>
</tr>
<tr>
<td>Date 9</td>
<td>5.74%</td>
<td>5.54%</td>
</tr>
<tr>
<td>Date 10</td>
<td>5.01%</td>
<td>5.13%</td>
</tr>
<tr>
<td>Exit-Regression Slope Coefficient</td>
<td>$-0.966$</td>
<td>$-0.969$</td>
</tr>
</tbody>
</table>
### Panel B: Structural Parameter Estimates

<table>
<thead>
<tr>
<th>Prior Distribution Parameters</th>
<th>Mean (Normal)</th>
<th>Variance (Gamma)</th>
<th>Random Exits</th>
<th>( J )-test, ( \chi^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>-0.5086</td>
<td>0.0019</td>
<td>0.0126</td>
<td>10.088</td>
</tr>
<tr>
<td>Variance</td>
<td>0.0358</td>
<td>0.2211</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SE</td>
<td>(0.0289)</td>
<td>(0.0053)</td>
<td>(0.0026)</td>
<td>(0.184)</td>
</tr>
</tbody>
</table>

### Panel C: Trader and Non-Trader Abilities (%)

<table>
<thead>
<tr>
<th>Population Subset</th>
<th>Pr(Skilled)</th>
<th>Pr(( r_1 &gt; 0 ))</th>
<th>5%</th>
<th>25%</th>
<th>50%</th>
<th>75%</th>
<th>95%</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>2.327</td>
<td>2.631</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>2.661</td>
</tr>
<tr>
<td></td>
<td>35.940</td>
<td>34.448</td>
<td>0.000</td>
<td>1.472</td>
<td>31.180</td>
<td>66.908</td>
<td>100.000</td>
</tr>
<tr>
<td></td>
<td>0.885</td>
<td>1.266</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>

### Panel D: 95% Confidence Intervals of Ability by the Precision of Prior Beliefs

<table>
<thead>
<tr>
<th>Precision</th>
<th>Lower</th>
<th>Upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>Top-5%</td>
<td>0.14%</td>
<td>0.14%</td>
</tr>
<tr>
<td>Top-25%</td>
<td>-6.04%</td>
<td>-2.39%</td>
</tr>
<tr>
<td>Top-50%</td>
<td>-24.70%</td>
<td>-1.21%</td>
</tr>
<tr>
<td>All</td>
<td>-74.57%</td>
<td>113.29%</td>
</tr>
</tbody>
</table>

### Panel E: Cumulative Exit Rates for Skilled and Unskilled Traders

<table>
<thead>
<tr>
<th>Trading Date</th>
<th>Cumulative Exit Rates (%)</th>
<th>% of Skilled Investors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Date</td>
<td>Skilled</td>
<td>Unskilled</td>
</tr>
<tr>
<td>0</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>1</td>
<td>20.21</td>
<td>25.99</td>
</tr>
<tr>
<td>2</td>
<td>34.09</td>
<td>42.45</td>
</tr>
<tr>
<td>3</td>
<td>42.98</td>
<td>52.54</td>
</tr>
<tr>
<td>4</td>
<td>49.33</td>
<td>59.52</td>
</tr>
<tr>
<td>5</td>
<td>54.01</td>
<td>64.54</td>
</tr>
<tr>
<td>6</td>
<td>57.67</td>
<td>68.36</td>
</tr>
<tr>
<td>7</td>
<td>60.61</td>
<td>71.36</td>
</tr>
<tr>
<td>8</td>
<td>63.00</td>
<td>73.75</td>
</tr>
<tr>
<td>9</td>
<td>65.06</td>
<td>75.77</td>
</tr>
<tr>
<td>10</td>
<td>66.86</td>
<td>77.47</td>
</tr>
</tbody>
</table>