FUNDAÇÃO GETULIO VARGAS
ESCOLA DE ECONOMIA DE SÃO PAULO

BERNARDO LEMBI RAMALHO MACIEL

LEARNING IN BRAZILIAN DAY TRADERS

São Paulo

2022

BERNARDO LEMBI RAMALHO MACIEL

LEARNING IN BRAZILIAN DAY TRADERS

Dissertação apresentada à Escola de Economia de São Paulo como pré-requisito à obtenção de título de mestre em Economia de Empresas.

Orientador: Bruno Cara Giovannetti.

Coorientadores: Bernardo de Vasconcellos Guimarães e Fernando Daniel Chague.

São Paulo

2022

Maciel, Bernardo Lembi Ramalho.
Learning in Brazilian day traders / Bernardo Lembi Ramalho Maciel. - 2022.
56 f.

Orientador: Bruno Cara Giovannetti.
Co-orientador: Bernardo de Vasconcellos Guimarães.
Co-orientador: Fernando Daniel Chague.
Dissertação (mestrado CMEE) – Fundação Getulio Vargas, Escola de Economia de São Paulo.

1. Especulação diária (Finanças). 2. Investimentos - Aspectos psicológicos. 3. Investidores (Finanças). 4. Processo decisório. I. Giovannetti, Bruno Cara. II. Guimarães, Bernardo de Vasconcellos. III. Chague, Fernando Daniel. IV. Dissertação (mestrado CMEE) – Escola de Economia de São Paulo. V. Fundação Getulio Vargas. VI. Título.

CDU 336.767

Ficha Catalográfica elaborada por: Isabele Oliveira dos Santos Garcia CRB SP-010191/O
Biblioteca Karl A. Boedecker da Fundação Getulio Vargas - SP

BERNARDO LEMBI RAMALHO MACIEL

LEARNING IN BRAZILIAN DAY TRADERS

Dissertação apresentada à Escola de Economia de São Paulo como pré-requisito à obtenção de título de mestre em Economia de Empresas.

Data de aprovação: 24/05/2022

Banca examinadora:

Prof. Bruno Cara Giovannetti (Orientador)
FGV-EESP

Prof. Bernardo de Vasconcellos Guimarães
FGV-EESP

Prof. Fernando Daniel Chague
FGV-EESP

Prof. Marco Antonio Cesar Bonomo
Insper

Agradecimentos

Agradeço à minha família e minha namorada, Ana, pelo apoio imenso durante os últimos dois anos. Agradeço a meus colegas e amigos do mestrado pelo companheirismo e ajuda que, mesmo dificultados pela distância, foram essenciais. Agradeço aos professores da EESP por terem proporcionado um aprendizado inigualável durante o período. Agradeço principalmente a meus orientadores, Bruno Giovannetti, Bernardo Guimarães e Fernando Chague por se disponibilizarem a me orientar e pelos ensinamentos transmitidos. Finalmente, agradeço aos professores e colegas que contribuíram para este trabalho com comentários durante o seminário de tese e em outros momentos.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.

“If you don’t know, the thing to do is not to get scared, but to learn.“

(Ayn Rand, "Atlas Shrugged")

Resumo

É bem documentado que a maioria das pessoas perde dinheiro com *day trade* (mesmo sem considerar custos de transação). No entanto, como mostramos neste trabalho, a probabililidade de uma pessoa lucrar em um dado dia de *day trade* é maior do que 50%, o que pode iludi-la quanto à sua verdadeira habilidade. Tentamos conciliar esses dois fatos empíricos aparentemente contraditórios. Nós mostramos que o *dispositon effect* apresentado pelos day traders produz a medida ilusória de habilidade.

Palavras-chave: *day trading.* vieses comportamentais. aprendizado de investidor. *disposition effect.*

Abstract

It is well-documented that most individuals lose by day trading (even before transaction costs). However, as we document in this study, the odds of individuals profiting from day trading on a given day are above 50%, which may delude them about their actual day trading skills. We try to conciliate these two apparently contradictory empirical facts. We show that the disposition effect displayed by day traders generates the polluted measure of day trading skill.

Keywords: day trading. behavioral biases. investor learning. disposition effect.

JEL Classification: G40

List of Figures

List of Tables

Contents

1 Introduction

Day trading is booming among individuals. The popularization of technologies that allow regular people to trade with a few clicks has been followed by online courses and financial commentators teaching trading techniques and strategies that promise large profits. However, it is well documented that most day traders lose money (Chague et al., 2019; Barber et al., 2019), even before transaction costs. Yet, many of them keep day trading for a long time. Why does this happen?

In this paper, we first document that even though most individuals lose by day trading, the unconditional probability of individuals profiting in the activity on a given day is above 50%. This could induce them to believe they are skilled day traders. After documenting this fact, we show that the disposition effect displayed by day traders (Heimer, 2016; Jordan and Diltz, 2004; Linnainmaa, 2005) is what generates the polluted measure of day trading skill.

We can understand this with a simple example. Imagine an unskilled individual who day trades as follows. Every day, at 10 am, he decides to buy or sell an asset. Since he has no skill, the odds of a good decision are 50%. At noon, he looks again at the market. If he is profiting from his position, he closes the position and quits day trading for the day. However, if he is losing, he doubles the bet and waits for another couple of hours. Again, the odds of a future profit are 50%. At 2 pm, he finally closes his position, whatever the result is. An individual with this simple behavior will profit in three out of four days, even though he has no skill. That is, if he looks at the proportion of days he profits to infer his day trading skill, he will have a very polluted (upward biased) measure.

We first document that the daily day trading results are significantly skewed to the left, consistent with the previous example. That is, losses tend to be larger than profits. We then propose a model that allows us to estimate, from the daily day trading results, the disposition effect of day traders (i.e., the willingness to increase the position when facing a loss) and their real day trading skill (i.e., the likelihood of profiting in the first trade session of the day). The model is described and analyzed in Chapter 2. We show that individuals do display the disposition effect and that the real odds of profiting in the first trading session of the day are less than 50%.

In Chapter 3, we then evaluate whether day trading skill improves with experience. We show that only the polluted measure of skill (the odds of a daily profit) improves over time. This is because the disposition effect increases over time. Naturally, this may falsely give the impression that day traders are learning with experience.

Finally, in Chapter 4, we evaluate which signal (the polluted or the real one) day traders value more when finally deciding whether to keep day trading or quit the activity. As expected, we find that bad traders are more likely to quit than good traders and that the informative signal is a better predictor of market survival than the polluted one. However, this learning process is sluggish. Indeed, it takes several weeks of activity for the probability of the worst traders quitting the market to cross the 50% threshold.

2 Data, setting and model

In this chapter, we present the data and discuss its main characteristics. Then, we discuss why we need a model in this setting, compare the present work with previous papers, and present the model's trading structure. Finally, we explain how the distribution of profits derived from the model can be characterized as a mixture to facilitate the computation of moments (the details are left to Appendix A).

2.1 Data

The data comes from CVM, the Brazilian regulatory body equivalent to the SEC, and spans from 2012 to 2018. It includes (individual-day) day trading data for stocks listed in the *Bovespa* stock exchange. It also includes two mini options contracts: the *mini-índice* (mini-index) and *mini-dólar* (mini-dollar). Day traders operating in the latter two markets are trying to predict intraday variation in the market index (*Ibovespa*) and the exchange rate between dollars and Brazilian *reais*, respectively. These smaller contracts with lower margins are ideal for retail day trading and attract many individuals with this objective in mind.

Throughout most of the analysis, we restrict the sample to individuals who traded for at least 150 days and had their first observation from 2013 onwards. The goal is to select individuals that traded a minimum number of days, allowing an evaluation of performance over time and a better identification of the first trading day of these individuals. With this, we hope to determine their trading experience more accurately. In order to clean wealth heterogeneities, we also divide each trader's daily profit by their standard deviation. This normalization will give a better sense of what a "large profit" or a "small loss" means and make interpersonal comparisons easier.

Figure 1 presents a histogram of the normalized profits net of transaction costs. Three main characteristics stand out. First, the proportion of days individuals earn positive profits is below 50%. Worse than that: the mean normalized profit is negative and economically large (around 1/5 of the individual standard deviation). This result shows that, on average, individuals lose money considerably, considering their typical range of gains and losses. Indeed, it becomes even clearer when we look at the lifetime results at the end of their day-trading activity in Figure 2, where only a fraction of the traders have a positive profit over the long run.

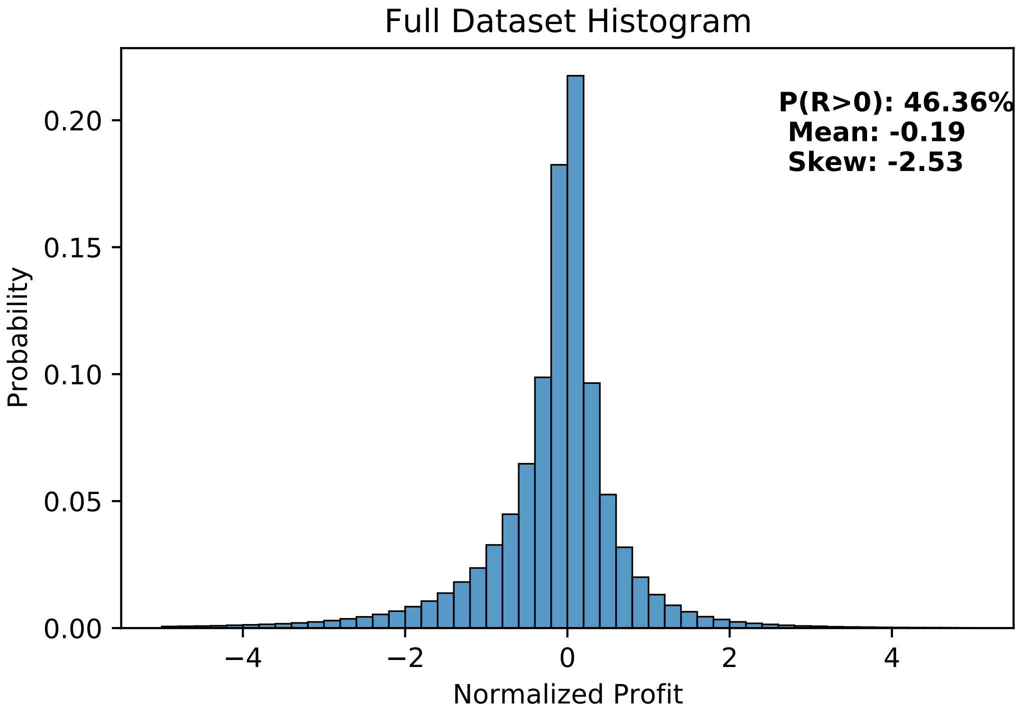


Figure 1 – Empirical distribution of normalized daily profits (net of transaction costs). On the top right corner, we also report: the proportion of positive profits; the empirical mean; and the empirical skewness. The limits in the x axis were cropped at $[-5, 5]$ to facilitate viewing.

Finally, there is a considerable asymmetry to the left of the data. This left skewness indicates that traders are expected to earn small amounts conditional on earning a positive profit. However, when they lose, they lose a lot. This asymmetry is compatible with the disposition effect: if individuals are prone to realizing gains and holding losses, one would expect a higher probability of earning smaller gains and a higher probability of incurring even more considerable losses (Odean, 1999). Notice that, by the nature of the data, this asymmetry does not come from asymmetries in the market or specific asset prices since these markets must clear out in the end. Throughout the paper, the focus of the analysis will be on these three moments, i.e., the proportion of positive results, average profits, and skewness.

Figure 3 presents the evolution of these moments over time. We divide the initial

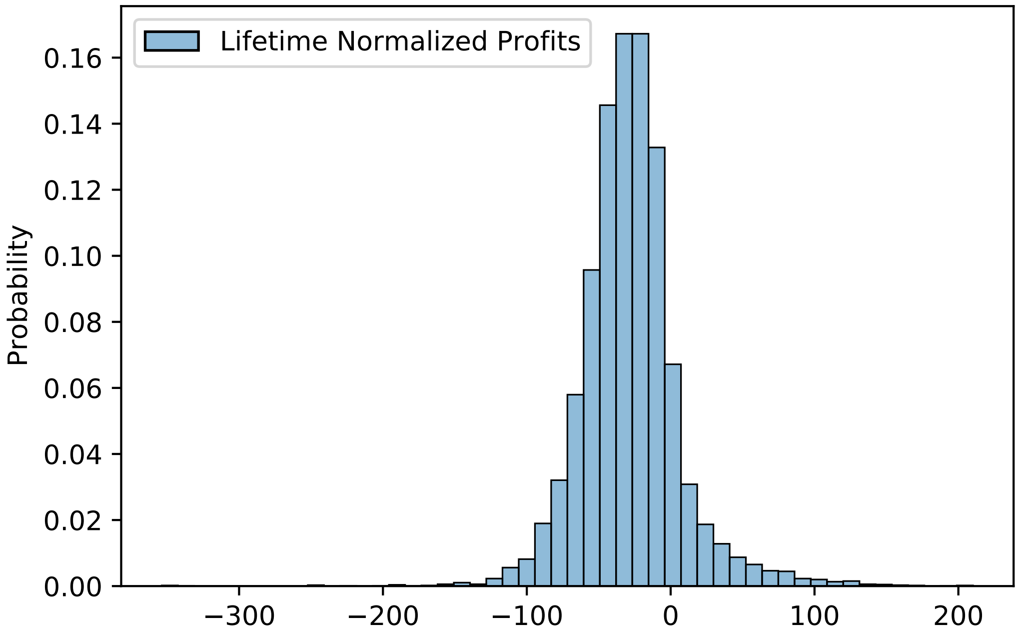


Figure 2 – Empirical distribution of total normalized profits (net of transaction costs) over the course of an individual day trader's career.

150 days of each trader over five sub-periods of 30 days each. Then, we compute the moments on all of these sub-panels.

The top graph shows that the proportion of positive days improves with time, from around 45% to 48%. At first glance, this indicates that day traders acquire ability with experience, which would be reflected in better results. However, they are still losing money on most days, even after more than 100 trading days. Furthermore, even though the middle graph shows an improvement in average daily profits, it is small. Indeed, the average profits remain in a considerably negative region, only increasing from around −0.21 to −0.18. Finally, skewness oscillates around a sizeable negative level, showing that the asymmetry between gains and losses persists even among day traders with greater experience.

Next, we perform some cross-section comparisons, which are shown in Table 1. The first two rows show the comparison between men and women. We see that women perform better than men, on average, on all dimensions considered: they receive positive profits in a larger proportion of days, have a larger average profit, and an asymmetry to the left

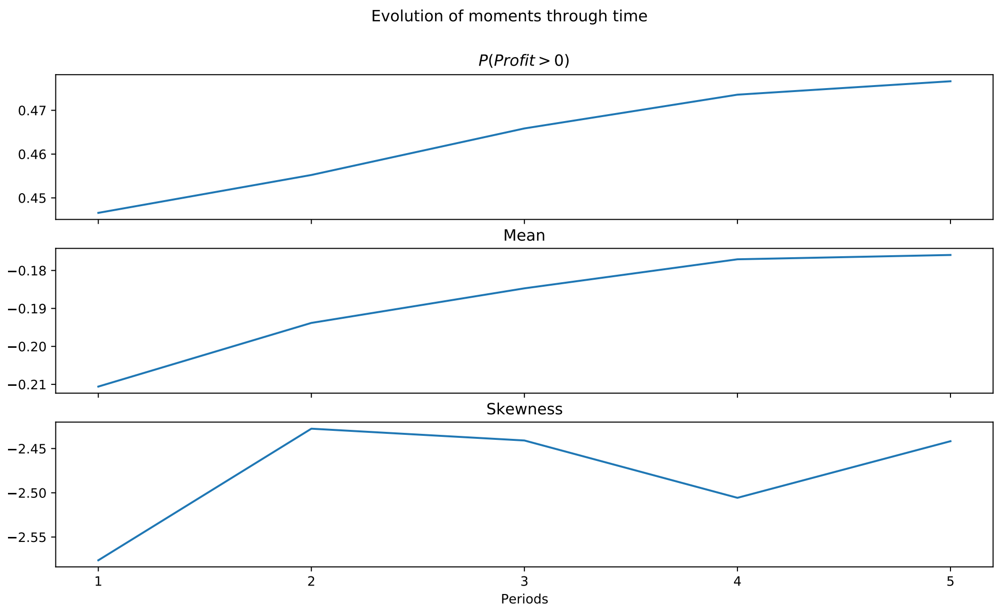


Figure 3 – Evolution of the three main moments (net transaction costs) over time. Each period comprises 30 trading days for all traders previously selected.

that is smaller in magnitude. However, they still have an average negative profit and a large negative skewness in absolute terms. This result is in line with the literature finding that men are more prone to overconfidence (Barber and Odean, 2001), which in turn is a possible theoretical explanation for the disposition effect. The more biased behavior of men, in turn, could explain, then, at least partially, their inferior performance.

In addition, we divide day traders into age terciles. The older traders (top tercile) have better average profits and a less pronounced asymmetry to the left than younger traders (bottom tercile). However, older traders still have a negative average profit and sizeable negative skewness.

Finally, we compare traders with different professions. Specifically, I compare traders identified with professions that deal with money more routinely *versus* professions that do not[1]. The data shows that day traders from the former group have a better average profit than those of the latter and earn positive profits more frequently. However, their skewness is very similar, and both of them have negative daily average profits. Thus, it seems that being in professions that deal with money captures abilities that are useful

[1] In the former category of occupations, we included the following: *manager*; *business director*; *actuary and mathematician*; *accountant technician*; *accountant*; *auctioneer*; *bank teller*; *realtor*; *capitalist*; *economist*; *diplomat*; *insurance agent*; and *businessperson*. The latter group includes the rest.

in day trading[2], but this is not enough to make the activity profitable for them. In the last line, we display the results from economists. Even though economists earn positive profits on most days, they also have negative average profits and a large asymmetry in the distribution of profits.

Sub-samples	Prob	Mean	Skew
Men	0.44	-0.22	-2.61
Women	0.49	-0.14	-2.12
Older	0.47	-0.16	-2.07
Younger	0.46	-0.21	-2.88
Money Prof.	0.48	-0.15	-2.55
Not Money Prof.	0.46	-0.20	-2.53
Econ	0.51	-0.10	-2.19

Table 1 – Comparison of moments from different sub-samples (net of transaction costs). *Older* and *Younger* refer to the top and bottom age terciles, respectively. *Money Prof.* refer to individuals in professions that deal routinely with money, and *Not Money Prof.* refers to the rest. *Econ* refers to economists.

Therefore, two main conclusions seem to stand out. First, the poor performance of day traders seems pervasive across different groups (even though it varies in magnitude among them). Second, and related to the first conclusion, the asymmetry between gains and losses is ubiquitous and high in all groups. If the disposition effect explanation is correct, this finding suggests that the biased behavior is present in all groups. Rather than a peculiarity of some specific subgroup driving the results, it seems to be the general behavior of individuals in day trading.

2.2 The Model

2.2.1 Why do we need a model?

The previous results have significant consequences for our analysis. Indeed, if most day traders exhibit the disposition effect, then the probability of earning positive profits in a single day will be a polluted measure of ability. As the market moves, a trader holding

[2] Of course, this does not mean that simply acquiring work experience in these professions causes the difference in ability observed. For instance, individuals in *Money* are generally more educated than the ones in *Not Money*, which could facilitate access to the trading platforms and better financial education.

a winning position will be more likely to liquidate it. The ex-ante effect of this behavior is reducing the probability of receiving larger profits as the trader closes his position prematurely. Another implication is that the distribution displays bunching to the right of the origin, which we also observe in Figure 1.

Likewise, if a trader is losing money on his position, he will be inclined to hold it, expecting the market to turn in his favor or even double down on his bets to try and recover his losses. This type of behavior will increase the probability of bigger losses, which will lead to fat tails on the left. This pattern is also compatible with the one seen in the histogram of Figure 1.

Figure 4 further reinforces this analysis by connecting profits to volume. First, we normalize the daily volume of each individual by its standard deviation, in a way similar to what was done previously to profit. Next, we sort observations on daily profit percentiles and compute the average profit in these percentile bins. Finally, we plot the average normalized volume in each of these correspondent bins and the ratio between the average profit and the average volume of each bin.

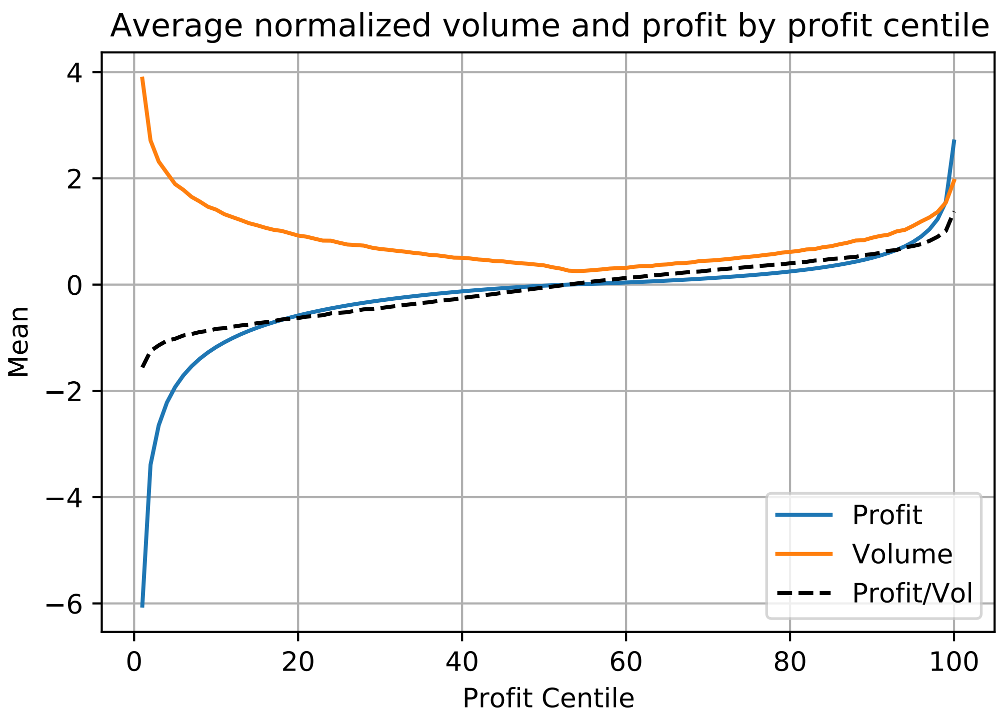


Figure 4 – Average of normalized profit percentiles (net of transaction costs) and average of volume percentiles.

We see that, in line with the disposition effect hypothesis, not only do day traders lose a lot when they lose; they also trade a lot when they lose a lot, comparatively. This pattern is compatible with day traders with losses in the day being more inclined to make more trades (or to increase the size of their position) to recover their losses. In comparison, day traders who earn positive profits in a day want to close their positions comparatively faster and refrain from making additional bets. Indeed, this generates an asymmetry in the volume dimension analogous to the one found in the profit dimension.

Therefore, observing the final daily results of day traders (which is what our data provides us) can lead to misleading conclusions about their performance. Indeed, the disposition effect will increase the probability of earning positive profits at the end of the day by the two channels described above. As a result, observing an increase in this statistic over time does not necessarily reflect an increase in ability. Thus, to construct a better measure of ability, we need a model that imposes more structure on the problem. With this, we hope to separate ability from the disposition effect by estimating the respective parameters that match the distribution of daily profits observed in the data.

2.2.2 Comparison to previous works

An earlier paper by Linnainmaa (2011) uses a structural model to argue that Finnish day traders trade to learn. In this model, traders use their performance to update beliefs about their abilities. Even though it acknowledges that investors in the sample may suffer from the disposition effect, the paper does not include any mechanisms capturing this behavior in its trading model.

In a more recent paper, Barber et al. (2019) use Taiwanese data to investigate if rational learning is compatible with day trading. However, instead of using a structural trading model, they present descriptive statistics, a survival analysis, and reduced-form regressions to conclude that day trading is compatible with overconfidence or biased learning but not rational learning.

Unlike these papers, we propose a simple structural trading model that tries to separate "earning positive profits because of superior ability" from "earning positive profits because of the disposition effect." With this, we seek to answer questions such as:

1. Do day traders improve with time? If so, by how much?

2. Does the disposition effect decrease as traders become more experienced?

3. How much is the contribution of the disposition effect, *versus* ability, to: men performing worse than women on average; economists receiving positive profits most of the time; younger traders performing worse than older traders?

2.2.3 Model description

The main idea of the model is to capture the asymmetry between gains and losses in a way consistent with the asymmetry in volume. This is done by including the disposition effect mechanism. In this model, traders are loss averse, more willing to realize gains than losses. Besides, they have a daily reference point, which seems reasonable given the nature of day trading.

Therefore, the setting is similar to that of the cab driver literature (Camerer et al., 1997; Farber, 2005, 2008), in which cab drivers establish pre-defined daily goals of revenue for a single working day and stop working as soon as that goal is reached. Here, day traders stop trading if their morning trades are successful; otherwise, they trade again in the afternoon, hoping to recover from the morning losses.

2.2.3.1 Trading structure

The data generating process of a trading session is as follows: at trading session t, agent i makes a bet of size b_t^i and earns the outcome:

$$y_t^i = b_t^i x_t^i, \tag{2.1}$$

where $x_t^i \sim t_{\mu,\nu}$, i.i.d. across i and $t_{\mu,\nu}$ denotes a t-Student distribution with with ν degrees of freedom and mean μ. The choice of the Student distribution is simply to clean out the kurtosis in the data through the parameter ν and has no economic significance.

There are at most two possible trading rounds in a single day. The size of the bet in the first trading round is normalized at one ($b_1^i = 1$). Therefore, $y_1^i = x_1^i$. The second round includes the disposition effect mechanism: if $y_1^i \geq 0$, i stops trading on that day. Otherwise, i bets $b_2^i = -\alpha y_1^i$ in the second session. Thus, the size of the second session bet is proportional to the losses in the first period, and α is this constant of proportionality.

Let R^i denote the total profits from these two sessions in a single day:

$$R^i = y_1^i + y_2^i \tag{2.2}$$

It is possible to derive an expression for the pdf of this final distribution; however, computing it will require a numeric integral. Therefore, directly calculating moments would be computationally burdensome, requiring the numeric computation of a double integral. Fortunately, we do not need the pdf if we are mainly interested in the three moments previously described. Indeed, we can obtain these moments quickly by characterizing the final distribution as a mixture.

To see this, let $B(p)$ denote a Bernoulli random variable with parameter p. Let $t^+_{\mu,\nu}$ denote $t_{\mu,\nu}$ truncated at values *above* zero (i.e., where negative values have been removed), with pdf:

$$f_{t+}(x;\mu,\nu) = \frac{f_t(x;\mu,\nu)}{1 - F_t(0;\mu,\nu)}\mathbb{I}(x \geq 0), \tag{2.3}$$

where f_t, F_t denote the pdf and cdf of the t distribution, respectively, and $\mathbb{I}$ denotes the indicator function. Analogously, let $t^-_{\mu,\nu}$ denote $t_{\mu,\nu}$ truncated at values *below* zero (i.e., where positive values have been removed), with pdf:

$$f_{t-}(x;\mu,\nu) = \frac{f_t(x;\mu,\nu)}{F_t(0;\mu,\nu)}\mathbb{I}(x < 0). \tag{2.4}$$

Moments for $t^+_{\mu,\nu}$ and $t^-_{\mu,\nu}$ are easily computed numerically.

We can define a mixture distribution:

$$R^i = B(p) \cdot t^+_{\mu,\nu} + (1 - B(p)) \cdot (t^-_{\mu,\nu}(1 - \alpha t_{\mu,\nu})), \tag{2.5}$$

where we specifically define

$$p := \mathbb{P}(t_{\mu,\nu} > 0) = 1 - F_t(0;\mu,\nu). \tag{2.6}$$

Since wish to normalize the variance of R^i to one, we are interested in

$$\tilde{R}^i := \frac{R^i}{\text{std}(R^i)} \tag{2.7}$$

where $\text{std}(R^i)$ is the standard deviation of R^i.

By characterizing the distribution as a mixture, all of the moments of interest are easily computed, requiring only moments from $t^-_{\mu,\nu}$, $t^+_{\mu,\nu}$ and $t_{\mu,\nu}$. The proof of this result is straightforward and left to Appendix A.

Therefore, $\mathbb{P}(\tilde{R} > 0)$ measures the probability of a trader winning in a single day[3]. As was discussed previously, this information is noisier because of the disposition effect.

[3] I omit the individual index for notation ease.

Likewise, $\mathbb{P}(\tilde{y}_t > 0)$ measures the probability that the trader obtains a positive profit in a single trading session t. This information captures the investor's ability better since it depends only on ability (captured by μ) and not on the disposition effect (captured by α). By recovering the model parameters (μ, ν, and α), we can estimate these quantities and better assess the evolution of day trading performance with time.

3 Learning to day trade

This chapter estimates the model's parameters, shows why the probability of a daily profit is a polluted signal for assessing skill and documents the prevalence of the disposition effect in day traders. We will also try to answer the first question posited in the introduction during the following sections: do day traders learn to trade? To do so, we estimate the evolution of the parameters, aiming to understand how ability evolves with time and the disposition effect's role in affecting individuals' trading decisions.

3.1 Estimating the model's parameters

As has already been stated, for the following estimations, we select the first 150 daily observations for individuals with at least 150 observations whose first trade was in 2013 or later. In the primary analysis, we use the gross profit data (that is, without considering transaction costs).

There are two main reasons to use gross profits when deriving the main results. First, transaction costs are not exact but estimated, and there may be errors in their computation. Second, it may be that day traders are not aware of the totality of costs they will face in the future or are not paying enough attention to how these costs impact their final performance through time. The changes caused by using net profits instead do not change the main conclusions and are reported in Appendix B.

We normalize individual profits by within-individual standard deviation and calculate the empirical equivalents of our three moments of interest. First, we estimate μ, ν and α for each subgroup. Next, for the dynamic estimations, we estimate $\mu_t, \alpha_t, t = 1, ..., 5$ and ν using the three moments for these five periods. The following section presents the static and dynamic estimations (using all pooled data).

The static results show that, on average, traders' ability parameter is negative as traders are expected to lose money in a single trading session. Therefore, the fact that traders earn positive profits most days comes from the magnitude of the disposition effect, not from ability. The dynamic results indicate that ability increases very little, while the disposition effect increases with experience instead of decreasing. Cross-section analyses are also presented, showing that the main results for the entire sample are prevalent in different groups.

3.2 Static and Dynamic estimations

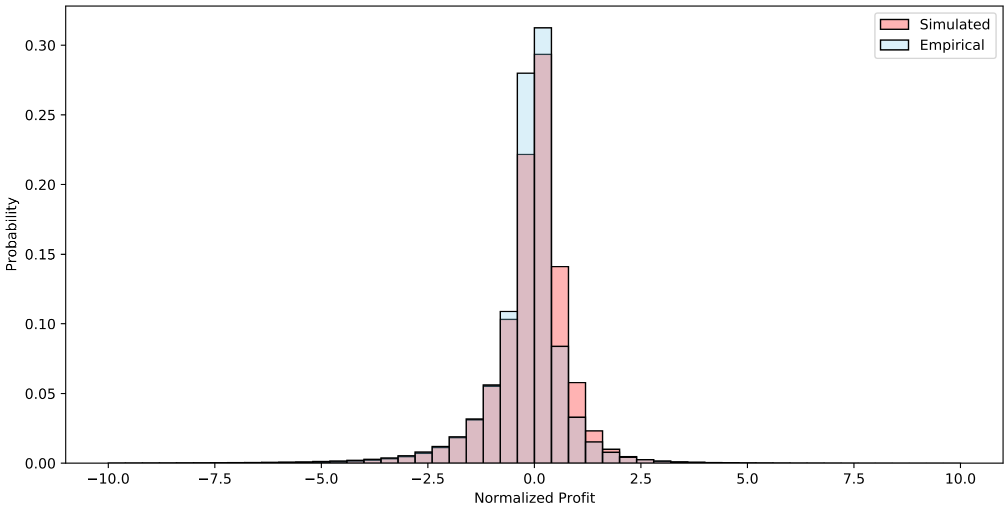
Table 2 presents the results of the pooled GMM estimation. It indicates that the average profit per trading session (our measure of ability) is significantly negative, both statistically and economically: on average, day traders lose about one-tenth of their daily profits standard deviation, even before discounting transaction costs. Besides, the estimate of α, which captures the magnitude of the disposition effect, is high. On average, the position size in the second session is around 4/5 of the loss from the first period. The low estimate of ν reflects the large kurtosis in the data.

The last two rows show the difference between the noisy ($\mathbb{P}(R > 0)$) and informative ($\mathbb{P}(y > 0)$) measures of ability. Whereas day traders win most of the days, they lose in most trading sessions. Thus, there are contradictory signals being sent to the trader. For example, if a trader looks at the proportion of days in which he is making positive profits, he may be inclined to conclude his performance is good. Nevertheless, by removing the disposition effect's role in this result, the conclusion would be that the average trader is likely to lose in a single trading session. If, in addition, traders are not sophisticated and have daily reference points, they may be more inclined to pay more attention to the first signal when they should value the second one more instead.

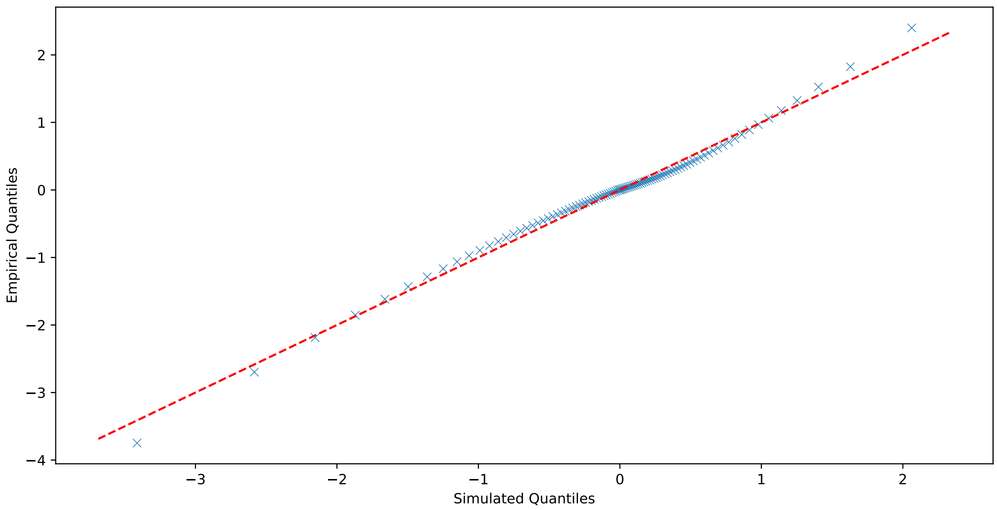
	Parameter	**95% CI**
μ	-0.0909	[-0.094, -0.088]
ν	4.1337	[4.054, 4.213]
α	0.8298	[0.822, 0.838]
$\mathbb{P}(R > 0)$	53.60%	[53.34%, 53.86%]
$\mathbb{P}(y > 0)$	46.59%	[46.47%, 46.71%]

Table 2 – GMM results for the static model using the full sample of traders. Observations (individual-day): 1,581,300.

Figure 5 compares the distribution generated by the fitted model with the empirical data. Panel 5a presents the histogram of data simulated from the model *versus* the empirical data. Panel 5b presents a Q-Q plot comparing the percentiles of the simulated and empirical distributions. We see that, even though we only matched three moments,



(a) Histogram comparing empirical and simulated data.



(b) Q-Q Plot of empirical and simulated data

Figure 5 – Simulated daily profits from the structural static model (using pooled data) *versus* the empirical profits.

the simulated distribution derived from the model seems to capture the whole distribution of the data fairly well (except, perhaps, in the extremes). Therefore, our model seems able to capture the main features of the empirical distribution despite its simplicity.

Figure 6 presents the dynamic GMM results. The graphs show an interesting pattern: even though there is a marginal increase in μ, there is also an increase in α. Even if day traders improve their trading skills with experience, they also become more biased. Nevertheless, μ is below 0 in every period. This result means that even for more experienced traders and before taking into account transaction costs, day traders are expected to lose money from a single trading session.

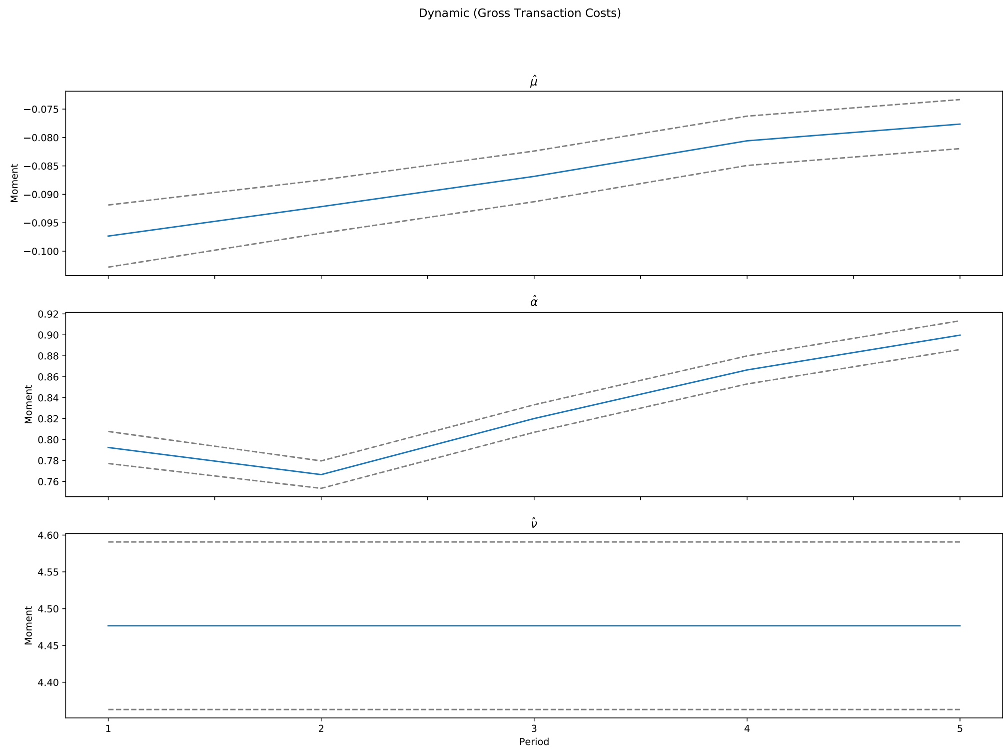


Figure 6 – Estimated parameters from the dynamic model (gross profits). Each period comprises 30 trading days. Dashed lines show 95% confidence intervals.

Figure 7 shows how well the moments implied by the estimated parameters fit the data. Overall, the fit of the GMM seems well adjusted for all moments, except perhaps for skewness. The iterative GMM causes this since the efficient weight matrix puts less weight on the quality of the skewness fit[1].

With these results, we can compute the implied $\mathbb{P}(y > 0)$ and compare it with $\mathbb{P}(R > 0)$. The results are shown in Figure 8. Confidence intervals are derived from the 95% intervals of the parameters.

[1] It is worth noting that the skewness depends on the cube of the value of the observations, which changes the relevant scale.

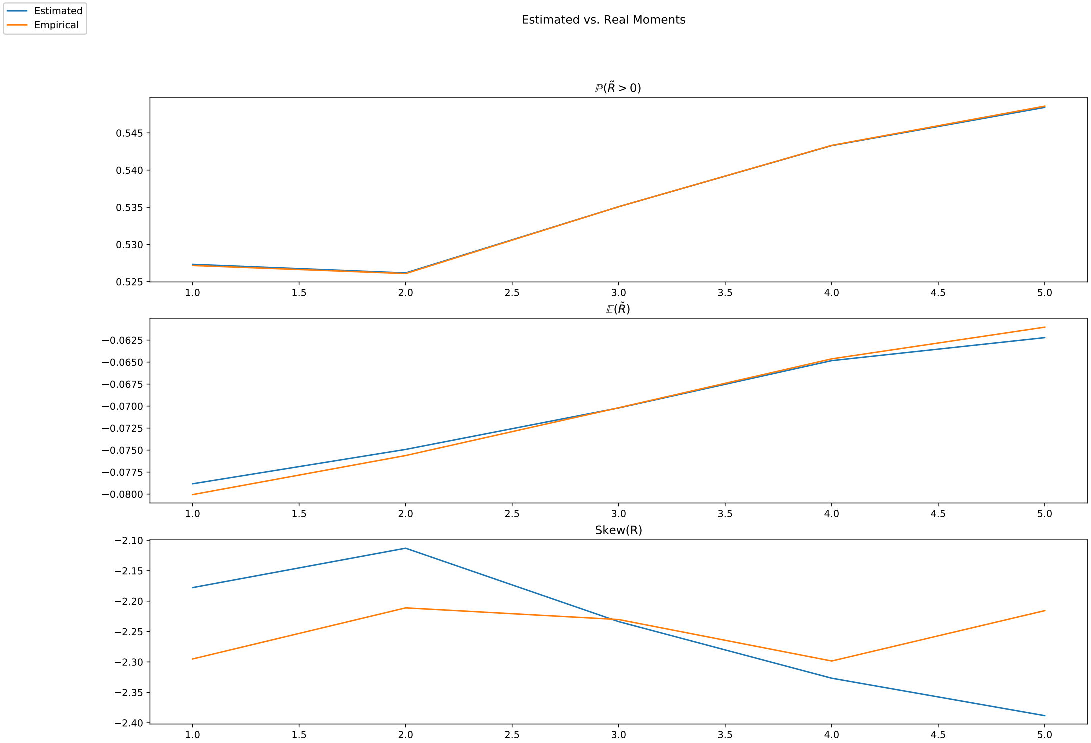


Figure 7 – Fit of the estimate dynamic moments *versus* the empirical dynamic moments. Each period comprises 30 trading days.

There are two critical differences between the evolution of each probability over time. The first one is a difference in level: the probability that a day trader will derive a positive profit at the end of the day is much higher than his probability of deriving a positive profit in a single trading session.

The second difference is a difference in slope: the improvement in the polluted measure of ability is more significant than in the clean measure. Thus, if day traders suffer from the disposition effect and use the polluted measure to evaluate their abilities, on average, they will conclude that their performance is improving. However, this does not come from an increase in ability: traders end most days with a positive trading balance because they tend to try again after starting the day with unsuccessful trades. As a result, the probability that they will earn positive profits in a single trading session increases only marginally.

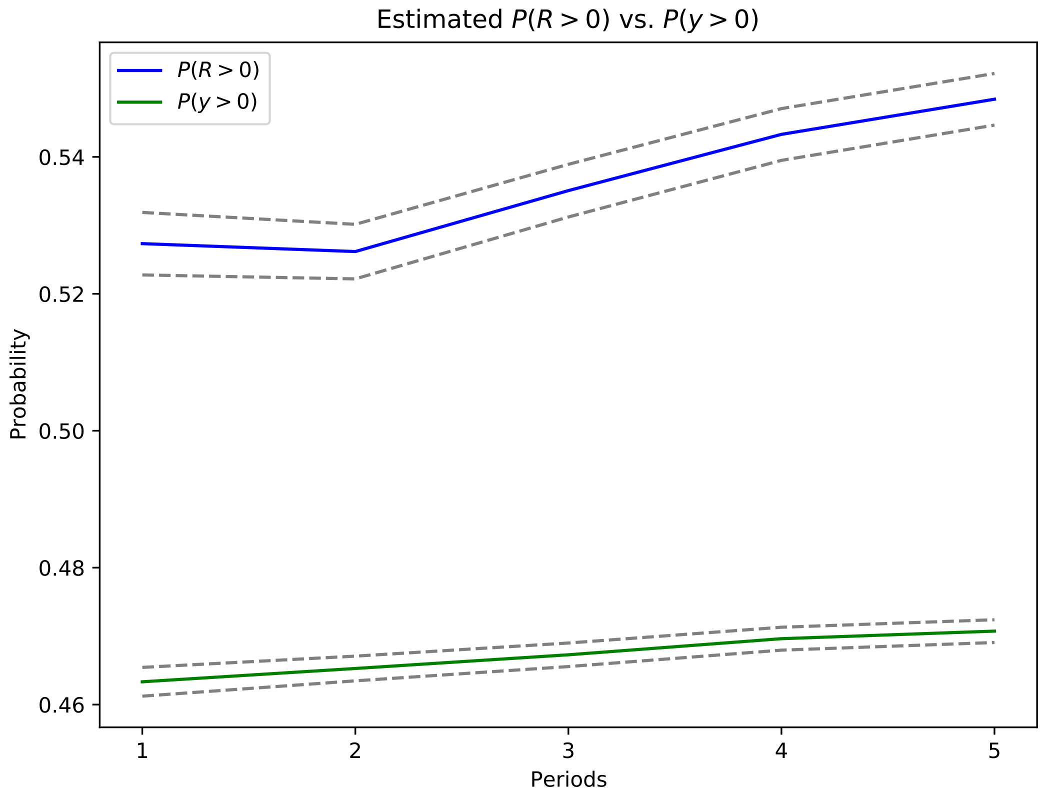


Figure 8 – Estimated "clean" measure of ability ($\mathbb{P}(y > 0)$) *versus* the estimated "polluted" measure of ability ($\mathbb{P}(R > 0)$). Each period comprises 30 trading days. Dashed lines are 95% confidence intervals.

These main findings come from the interaction of both μ and α increasing with experience. Indeed, they both would help account for part of the increase in the probability of a positive profit at the end of the day seen in Figure 8. In this case, they act in the same direction. However, they can act in opposite directions with respect to the expected daily profits. Effectively, the disposition effect helps day traders with a poor performance increase their probability of earning a positive daily gain at the expense of their average gains over time (which is what they should care about more).

This becomes clear by differentiating the expressions for $\mathbb{P}(R > 0)$ and $\mathbb{E}(\tilde{R})$ in Equations (A.3) and (A.4) with respect to α: the first derivative is positive, and the second is negative when μ is negative, which is empirically valid for the majority of day traders. If day traders assess their performance using their daily results, they may think they are getting better (when in fact, they are not).

3.3 Cross-section analysis

This section presents the results of the cross-section estimations, comparing ability and the disposition effect among different groups in our data according to three divisions: age, gender, and profession. With this, we hope to assess if the previous results regarding the prevalence of a high disposition effect and a low level of ability hold across different groups of day traders.

3.3.1 Men *versus* Women

In Table 3, we estimate the three parameters of interest by GMM separately in the subsample of men and the subsample of women. The level of the disposition effect captured by α is high for men and women and differs little between the two groups.

However, the biggest difference lies in ability: women have, on average, a higher level of ability than men. The ability parameter μ is still statistically below zero, even though this difference is small. Thus, women are still expected to lose money from a single trading session. The implications are that both $\mathbb{P}(R > 0)$ and $\mathbb{P}(y > 0)$ are higher for women. For both groups, we see again the pattern of $\mathbb{P}(R > 0) > 50\%$ and $\mathbb{P}(y > 0) < 50\%$. Therefore, our previous main remarks are still valid for women and men.

		Obs. (men):	1,406,250		
		Obs. (women):	140,400		
Men	Parameter	95% CI	**Women**	Parameter	95% CI
μ	-0.1014	$[-0.104, -0.098]$		-0.0178	$[-0.025, -0.011]$
ν	4.0871	$[4.010, 4.164]$		4.1722	$[4.094, 5.201]$
α	0.8332	$[0.826, 0.841]$		0.8319	$[0.813, 0.851]$
$\mathbb{P}(R > 0)$	53.23%	$[52.99\%, 53.49\%]$		56.44%	$[55.76\%, 57.14\%]$
$\mathbb{P}(y > 0)$	46.2%	$[46.08\%, 46.32\%]$		49.33%	$[49.06\%, 49.59\%]$

Table 3 – GMM results comparing men and women.

3.3.2 Older *versus* Younger traders

In Table 4, we estimate the model parameters analogously for the subsample of younger traders (bottom age tercile) and older traders (top age tercile) separately. The results show that younger traders have less ability and are more biased than older traders:

the expected loss in a single trade for younger traders is around 12% of a standard deviation, whereas it is around 5% for older traders. Moreover, second-period bets are close to 86% of first-period losses for younger traders, *versus* around 77% for older traders.

There are two implications for the implied probabilities. First, both the daily probability and the session probability of earning a positive profit are higher for older traders. Second, the gap between the two probabilities is wider for younger traders, resulting from their stronger disposition effect.

The results for the disposition effect are compatible with younger individuals having riskier behavior and being less willing to accept losses. However, the disposition effect is especially harmful to them since younger traders have less ability, making the impact of the disposition effect on average profits more significant.

Regarding the differences in ability, it is also possible that μ captures a higher previous trading experience for older traders or a higher level of financial education. Indeed, μ was defined as capturing both innate and acquired abilities that can be useful in trading. Regardless, the polluted measure of ability for both groups is significantly higher than the clean measure, and both groups have a poor performance on average. Thus, our previous conclusions do not substantially change with age.

		Obs. (younger): 538,680 **Obs. (older):** 467,676			
Younger	Parameter	95% CI	**Older**	Parameter	95% CI
μ	-0.1183	$[-0.123, -0.113]$		-0.0456	$[-0.05, -0.041]$
ν	4.1096	$[3.987, 4.232]$		4.2015	$[4.017, 4.386]$
α	0.86	$[0.847, 0.873]$		0.7688	$[0.758, 0.779]$
$\mathbb{P}(R > 0)$	52.85%	$[52.44\%, 53.27\%]$		54.65%	$[54.27\%, 55.05\%]$
$\mathbb{P}(y > 0)$	45.57%	$[45.37\%, 45.76\%]$		48.28%	$[48.11\%, 48.46\%]$

Table 4 – GMM results comparing younger traders (bottom age tercile) to older traders (top age tercile).

3.3.3 Comparing professions

Next, we divide our sample according to the occupation reported, comparing traders employed in professions that typically deal with monetary calculations and exchanges (which I label as *Money*) *versus* those that do not work in such professions (which I label as *Not Money*). The results are presented in Table 5.

		Obs. (*Money*): 343,200				
		Obs. (*Not Money*): 1,238,100				
		Obs. (*Econ*): 28,350				
	Money	95% CI	**Not Money**	95% CI	**Econ**	95% CI
μ	-0.046	$[-0.051, -0.041]$	-0.103	$[-0.107, -0.1]$	0.03	$[0.016, 0.044]$
ν	4.167	$[3.975, 4.36]$	4.125	$[4.039, 4.211]$	4.332	$[3.397, 5.267]$
α	0.845	$[0.832, 0.858]$	0.826	$[0.818, 0.834]$	0.817	$[0.774, 0.861]$
$\mathbb{P}(R > 0)$	55.6%	$[55.16\%, 56.04\%]$	53.05%	$[52.79\%, 53.32\%]$	58.32%	$[56.94\%, 59.75\%]$
$\mathbb{P}(y > 0)$	48.26%	$[48.07\%, 48.46\%]$	46.12%	$[46\%, 46.25\%]$	51.14%	$[50.61\%, 51.67\%]$

Table 5 – GMM results comparing traders who work in professions that deal with money (*Money*) *versus* those who aren't in those professions (*Not Money*). Additionally, the results for economists (*Econ*) are also reported separately.

As intuition would indicate, traders in *Not Money* have a lower level of trading ability, which is reflected in lower $\mathbb{P}(R > 0)$ and $\mathbb{P}(y > 0)$. However, this may also reflect the fact that individuals employed in professions dealing routinely with money also have distinct characteristics from the other group. For example, they are more likely to have better financial or mathematical education, making trading more accessible. In our analysis, the ability parameter μ captures all of these factors instead of measuring only innate aptitudes. However, the magnitude of the disposition effect for the two groups is quite similar (indeed, the confidence intervals overlap).

Looking specifically at the results for economists, we see that they have a higher ability level and display a lower disposition effect than the other groups of occupations. These estimations indicate that economists are less prone to biased behavior and better at day trading, on average. Indeed, their probability of receiving a positive profit in the session is slightly above 50%. However, even though the μ estimate is above zero, it is economically small (below 5% of a standard deviation). Therefore, even though economists earn money from most trading sessions, the expected profit per session is still low (even without considering transaction costs). In addition, the point estimates for α are still significant, indicating that economists also suffer from the disposition effect.

Thus, it does not seem to be the case that individuals unfamiliar with monetary calculations are driving the results. On the contrary, even traders from professions that usually deal with monetary calculations display low ability and a considerable disposition effect. Indeed, economists are not immune from this behavior either.

3.4 Conclusion

This chapter reinforces the already known result that day traders perform poorly on average. Indeed, this is true before we even take into account transaction costs. In addition, two main conclusions help us better understand how ability evolves with time and its role in determining performance.

First of all, day traders do not improve significantly, in an economic sense, with experience: they merely learn to lose slightly less money. This improvement is so tiny that it does not alter the sizable poor performance most traders exhibit through time. Thus, with this caveat, the answer to the question posed at the beginning of the chapter is negative.

Even though there is also an improvement in $\mathbb{P}(R > 0)$, this seems to come mainly from an increase in the disposition effect. Meanwhile, $\mathbb{P}(y > 0)$ stays pretty much the same throughout the whole period and is always below the 50% threshold. Similarly, even though μ increases marginally with experience, it stays below zero. As losses accumulate, the implication is that more experienced day traders tend to accumulate larger lifetime losses than less experienced day traders. This implication will be analyzed again later in the next chapter.

The second point refers to the disposition effect: it is economically significant and present in different groups, whether we separate the sample by age, gender, or profession. The significant asymmetry between gains and losses observed in the data does not seem to be led by some specific subgroup of traders but rather constitutes the standard behavior in day trading. What is worse is that, contrary to intuition and what Friedman's survival argument (Friedman, 1953) would lead us to believe, this bias does not fade out as experience increases. Rather than triggering a change in behavior over time as day traders suffer significant losses, the disposition effect actually increases over time.

This puzzling result is compatible with the idea that traders have difficulty in re-evaluating the financial implications of their decision to start day trading in the first place, which comes with a negative impact on their own self-image. Another possible explanation is that day traders are looking at the wrong signals: if traders are looking at the number of days they receive positive profits, they may falsely conclude they are improving, as Figures 6, 7 and the graph for $\mathbb{P}(R > 0)$ in Figure 8 indicate. However, as the first panel in Figure 6 and the graph for $\mathbb{P}(y > 0)$ in Figure 8 show, on average traders have sizable session losses which do not improve with time, and low probabilities of a positive session profit (which does not improve either).

A rational trader looking at the more informative signals (e.g., session probabilities, average gains over time, distribution of gains over time) should conclude that trying to recover losses by increasing the size of their position is damaging in the long run. On the other hand, a less sophisticated trader looking at the polluted signals (daily probabilities) may falsely conclude that they are becoming better day traders when in fact, they are not. What makes matters worse is that the same behavior detrimental to traders' average profits over time, i.e., the disposition effect, is the primary driver of this deceitful improvement in the polluted measure of ability.

4 Trading to learn

During the following sections, we will try to answer the second question posed in the introduction: how do day traders assess their performance and decide to keep trading or quit the market? We estimate the model parameters individually and construct a survival analysis model to understand how ability and the disposition effect influence the number of days a day trader operates in the market. Again, we will use normalized gross profits throughout the analysis.

4.1 Survival Statistics

To better understand what forces influence a day trader's decision to quit, we classify an individual as having "quit" if we do not observe any trading during the last two months of the sample (November and December 2018). Otherwise, we classify a trader as having "survived." After keeping the previous selection of 10,542 individuals, 2,579 day traders survived (24.46%).

Table 6 shows how performance during the first 150 trading days differs between the subsample that eventually quit *versus* the one that survived. In Panel 6a, we see that the former group received a positive profit in fewer days. This difference is especially apparent when we compare the top winners in each group: the top 25% winners in the former group won in more than 60% of the days, whereas the top 25% winners in the latter group won in at least 68% of days. Indeed, the group of quitters also performed worse in terms of average normalized profits during the same period, as Panel 6b shows. However, it is worth noting that even the median survivor had a negative average profit during the period. This finding illustrates again the extent to which how pervasive day trader's poor performance is.

Figure 9 plots the non-parametric Kaplan-Meier estimates of the survival function conditional on quintiles of performance during the initial 150 trading days. The Kaplan-Meier estimator of the survival function $S(T) = 1 - F(T)$ at time T is simply given by:

$$\hat{S}(T) = \prod_{t \leq T} \left(1 - \frac{q_t}{r_t}\right), \tag{4.1}$$

Win Rate	quits	survives
obs.	7963.0000	2579.0000
mean	0.5225	0.5780
std	0.1214	0.1608
min	0.0867	0.1000
25%	0.4400	0.4600
50%	0.5067	0.5533
75%	0.5933	0.6800
max	1.0000	1.0000

(a) Statistics for the individual win rate during the first 150 days of trading for individuals who quit *versus* those who survived at the end of the sample.

Avg. Profit	quits	survives
obs.	7963.0000	2579.0000
mean	-0.1029	0.0303
std	0.1721	0.2864
min	-0.9640	-0.7720
25%	-0.1990	-0.1438
50%	-0.1120	-0.0363
75%	-0.0317	0.1298
max	1.5544	1.7748

(b) Statistics for the average individual normalized gross profit during the first 150 days of trading for individuals who quit *versus* those who survived at the end of the sample.

Table 6 – Descriptive statistics for win rate (proportion of days in which a trader received a positive gross profit) and average normalized profit during the first 150 days, conditional on quitting in the future or not.

where, in our context, r_t is the number of traders who can quit at t and q_t the ones that quit at t. The graph shows that traders with better performance during this initial period tend to trade more days before quitting than those with worse performance, likely driven out of the market after successive losses.

Therefore, there is some early indication that day traders, at least to some extent, "trade to learn": that is, they observe their performance and use this assessment when pondering whether to keep day trading or quit in the future. However, this learning process is slow. Indeed, it takes a long time (even *after* the 150th day) for the survival curve of the bottom traders to cross the 50% threshold. Besides, around 20% of the bottom traders are expected to survive after 350 trading days (in total).

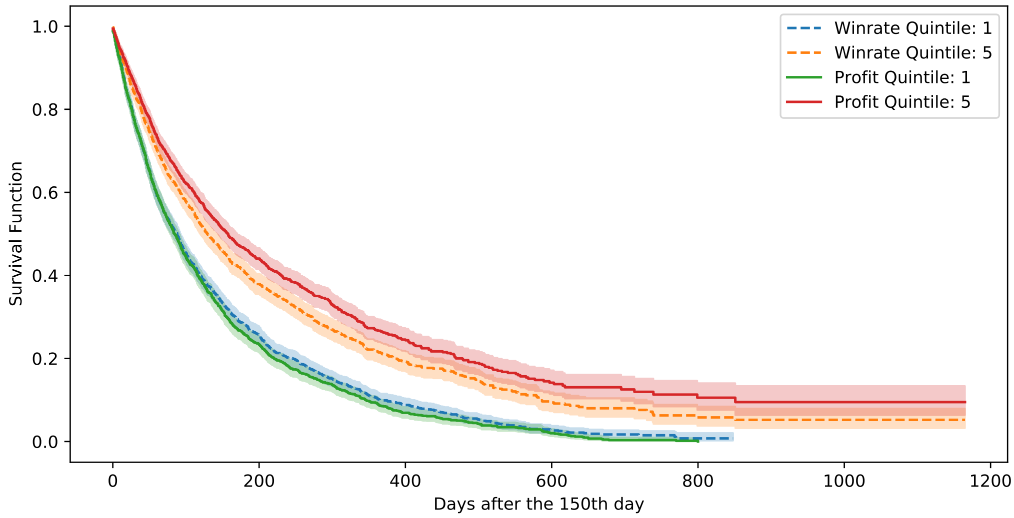


Figure 9 – Kaplan-Meier estimate of the survival function conditional on win rate (dashed) and average normalized profit quintiles during the initial 150 days. Shaded areas are 95% confidence intervals.

4.2 Ability, bias and the decision to quit

Given the previous selection of individuals, we estimate the model's parameters (μ, ν, and α) individually. As before, we use normalized gross profits from the initial 150 trading days. We restrict the observations to individuals to which the constrained optimization generated a fit sufficiently close to the empirical moments.

To understand how μ and α impact survival in the market, we add the estimated parameters as explanatory variables in a Cox proportional-hazards model. In our context, the relevant event is a day trader quitting. Therefore, the hazard rate $h(t)$ is defined by

$$h(t) = \lim_{\Delta t \to 0} \frac{\text{traders quitting in the interval } [t, t + \Delta t]}{\Delta t \times R(t)}, \tag{4.2}$$

where $R(t)$ are the individuals that didn't quit before t (i.e., at the risk of quitting at t). Therefore, the Cox model is:

$$\underbrace{h(t|x)}_{\text{hazard}} = \overbrace{b_0(t)}^{\text{baseline hazard}} \exp\left(\sum_{i=1}^{n} b_i x_i\right) \tag{4.3}$$

where t measures the number of days after the 150th and x_i the i-th vector of covariates. I add dummies for state, age tercile and sex as controls. The estimated hazard ratios are shown in Figure 10.

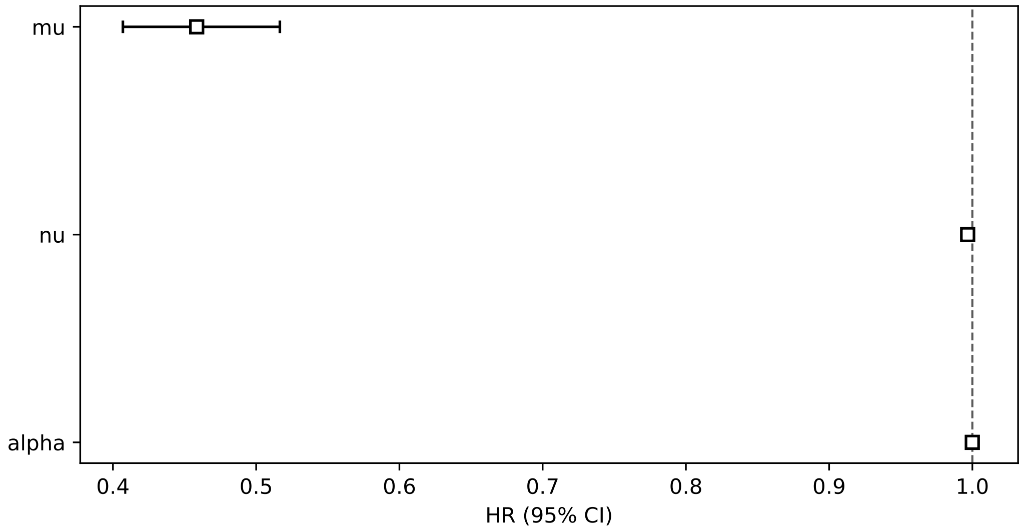


Figure 10 – Estimated hazard ratios from the Cox model, with the individually estimated μ, ν and α being the variables of interest. Controls include dummies for state of residence, sex and age tercile.

Keeping everything else constant, increasing α or ν does not affect the probability of quitting at t. On the other hand, an increase μ has a significant and sizable effect on reducing the probability of quitting. Indeed, a unit increase of μ makes a day trader less than half as likely to quit day trading at t. This finding reinforces the non-parametric results, indicating that day traders who have shown more remarkable ability in the past will be more likely to keep day trading in the future.

Figure 11 illustrates the impact of changing μ on the baseline survival function. Again, this graph supports the idea that day traders trade to learn, but this learning process is slow: the survival curves associated with negative μ parameters do not decline steeply but require many trading days (after day 150) to cross the 50% threshold.

Next we compute, individually, $\mathbb{P}(y_i > 0)$ and $\mathbb{P}(\tilde{R}_i > 0)$ as implied by the individually estimated parameters. Then, we estimate another Cox regression including these probabilities and the previous controls as explanatory variables to understand how the clean and polluted probabilities of winning affect the decision to quit. The estimated hazard ratios for them are presented in Figure 12.

An increase in the more informative, or "cleaner," probability $\mathbb{P}(y_i > 0)$ will decrease the probability of quitting, everything else constant. Indeed, the results imply

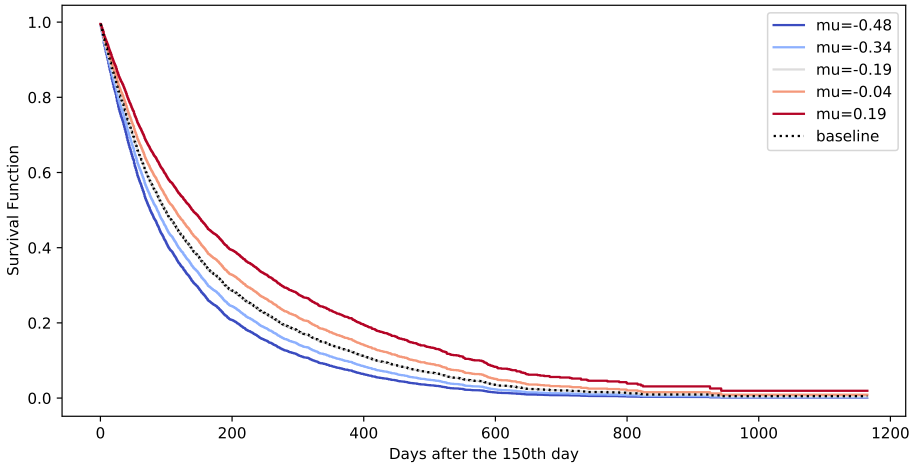


Figure 11 – Estimated survival functions varying μ and maintaining everything else constant. The selected values of μ are the 10%, 25%, 50%, 75% and 90% quantiles, respectively (rounded). The baseline curve is equal to the predicted curve at average values.

that a percentage point increase in this probability will make a day trader around 2% less likely to quit. A change in the noisier, or "polluted", probability $\mathbb{P}(\tilde{R}_i > 0)$, however, has a negligible and statistically insignificant effect on the probability of quitting, after controlling for $\mathbb{P}(y_i > 0)$. Figure 13 shows how the survival function changes when $\mathbb{P}(y_i > 0)$ changes, keeping everything else constant. The same pattern seen in Figure 11 is repeated.

Changing $\mathbb{P}(\tilde{R}_i > 0)$ while keeping $\mathbb{P}(y_i > 0)$ fixed means that ability is kept constant, while the disposition effect varies. What the previous results suggest is that the disposition effect has no direct impact on the decision to quit, acting only through its indirect effect on the performance of traders[1]. Indeed, replacing $\mathbb{P}(\tilde{R}_i > 0)$ by α as an explanatory variable produces very similar results (which are not reported here). Therefore, the daily (polluted) probability is not informative in explaining the decision to quit, but the session (clean) probability is.

A possible concern from the previous results is that the effect of ability and bias on the decision to quit is non-monotonic (Barber et al., 2019). Increasing μ, for instance,

[1] The effect of an increase of α on $\mathbb{P}(\tilde{R}_i > 0)$ and on extreme results is unambiguous: it increases both. However, the effect on $\mathbb{E}(R_i)$ is not: it will be positive if $\mu > 0$ and negative if $\mu < 0$. This can be easily seen by differentiating Equation (A.4) from the appendix and noting that $(1-p)\mathbb{E}(t^-_{\mu,\nu}) < 0, \forall p < 1$. Therefore, if day traders look at their average profits when assessing their ability and deciding to keep trading or not, this indirect impact of the disposition effect on the decision will also depend on the sign of μ.

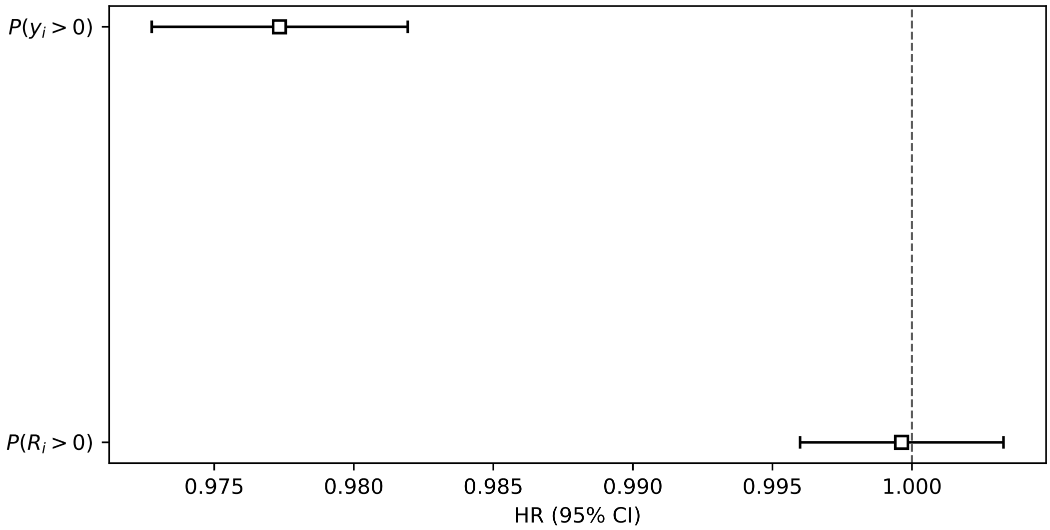


Figure 12 – Estimated hazard ratios from the Cox model, with the implied individual $\mathbb{P}(R_i > 0)$ and $\mathbb{P}(y_i > 0)$ being the variables of interest. Controls include dummies for state of residence, sex and age tercile.

could decrease the hazard ratio of quitting up to a point but increase it afterward. This pattern will emerge if the best traders are more likely to "cash out while winning" after many successful trades. Therefore, we repeat the Cox regressions but use dummies to sort the variables of interest into intervals.

Figure 14 presents the estimated hazard ratios using the estimated ν, μ and α as explanatory variables (the latter two discretized into intervals). Even though there is a non-monotonicity between the bottom interval of μ and the next interval, the estimated hazard ratios are mostly consistent with the previous results. Traders falling in the lower intervals of ability have a higher probability of quitting at any point in time than the baseline ($\mu \in (-0.07, 0]$).

This difference becomes smaller as the interval approaches the baseline and becomes negative for the top interval: traders with a high ability ($\mu > 0.2$) are significantly *less* likely to quit day trading than the baseline. However, the estimates for α show no clear pattern, even though the coefficient for $\alpha \in (2, 5]$ is statistically significant. Consequently, these findings are consistent with the previous results.

Likewise, I repeat the Cox regression with $\mathbb{P}(R_i > 0)$ and $\mathbb{P}(y_i > 0)$ as explanatory variables, this time discretized into intervals. The baseline interval for the former is

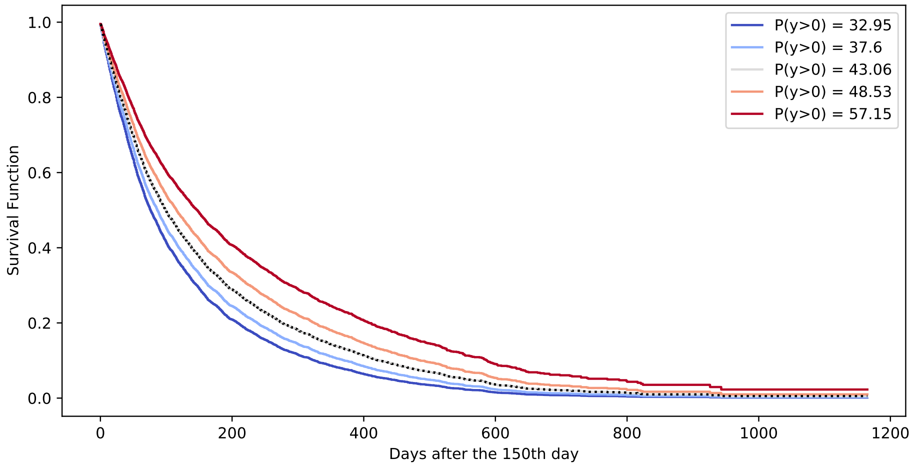


Figure 13 – Estimated survival functions varying $\mathbb{P}(y_i > 0)$ and maintaining everything else constant. The values of $\mathbb{P}(y_i > 0)$ are the 10%, 25%, 50%, 75% and 90% quantiles, respectively. The baseline curve is equal to the predicted curve at average values.

$\mathbb{P}(R_i > 0) \in (49\%, 50\%]$ and for the latter is $\mathbb{P}(y_i > 0) \in (48\%, 50\%]$.

Figure 15 presents the results and paints a very similar picture. The hazard ratios for the $\mathbb{P}(y > 0)$ intervals below the baseline are above 1 and tend to decrease when approaching the baseline. The two top probability intervals have associated hazard ratios below 1. In addition, except for $\mathbb{P}(y_i > 0) \in [0\%, 30\%]$ and $\mathbb{P}(y_i > 0) \in [34\%, 36\%]$, the point estimates are monotonic: the higher the probability intervals are, the smaller the hazard ratios. Like before, the estimated hazard ratios for the intervals associated with $\mathbb{P}(R_i > 0)$ do not show a clear pattern and are statistically insignificant. Thus, the polluted measure of ability is not important to explain whether to keep trading or quit after controlling for the cleaner measure. Again, this modification does not change our previous conclusions.

4.3 Conclusion

Do day traders trade to learn about their abilities when deciding to keep trading or quitting? The results in this chapter lead to two main conclusions related to the second question posed in the introduction.

The first conclusion is that the answer is positive (at least to some extent). Whether

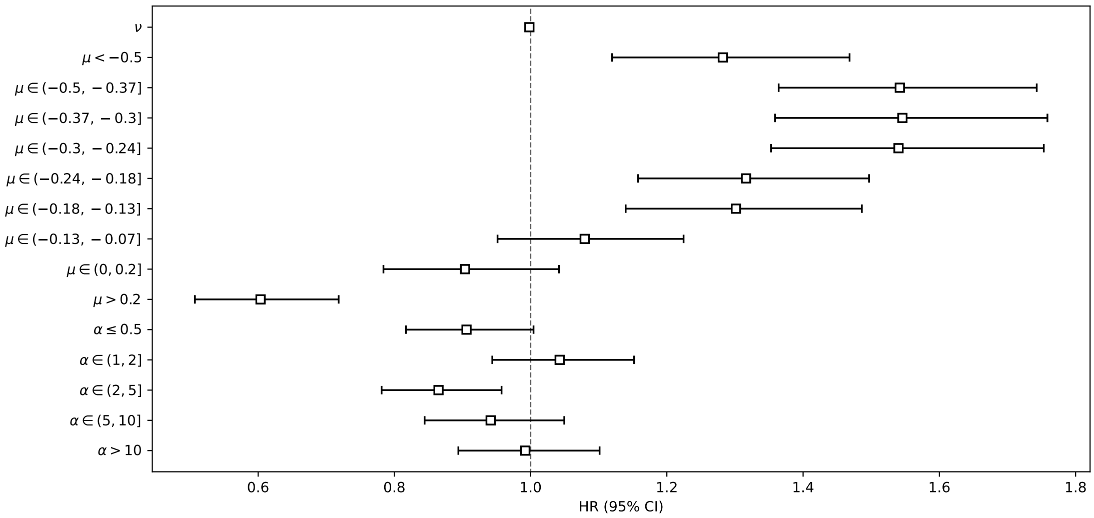


Figure 14 – Estimated hazard ratios from the Cox model, with the individually estimated μ, ν and α being the variables of interest. Parameters μ and α are sorted into intervals and included as dummy variables. The baseline dummy for μ is $\mu \in (-0.07, 0]$ and for α is $\alpha \in (0.5, 1]$. Controls include dummies for state of residence, sex and age tercile.

we use the past average profits, the estimated μ, or the implied probability of receiving a positive profit in a single trading session, the end results are the same: the more ability a trader displays during the initial 150 days, the less likely he will be to quit day trading in the future (and he will tend to trade more days before he does so).

After controlling for ability, the disposition effect seems to have no clear role in this decision. Likewise, after controlling for the implied $\mathbb{P}(y_i > 0)$, $\mathbb{P}(R_i > 0)$ seems to have a negligible impact as well. Thus, in line with the rational explanation, the market drives out the worst day traders by providing negative signals of their ability while keeping the best ones. Furthermore, it is the clean measure of ability, not the polluted one, the main driver behind the decision to quit.

The second conclusion is that this learning process is slow. Even after traders observe and analyze their performance through the first 150 trading days, it still takes a long time for many of the worst day traders to quit. For instance, Figure 9 shows that it still takes more than 100 days (after the initial 150) for the survival curve associated with the bottom quintile of average profits to cross the 50% threshold. The same is true with regards to the survival curves associated with negative μ in the Cox regression, as Figure

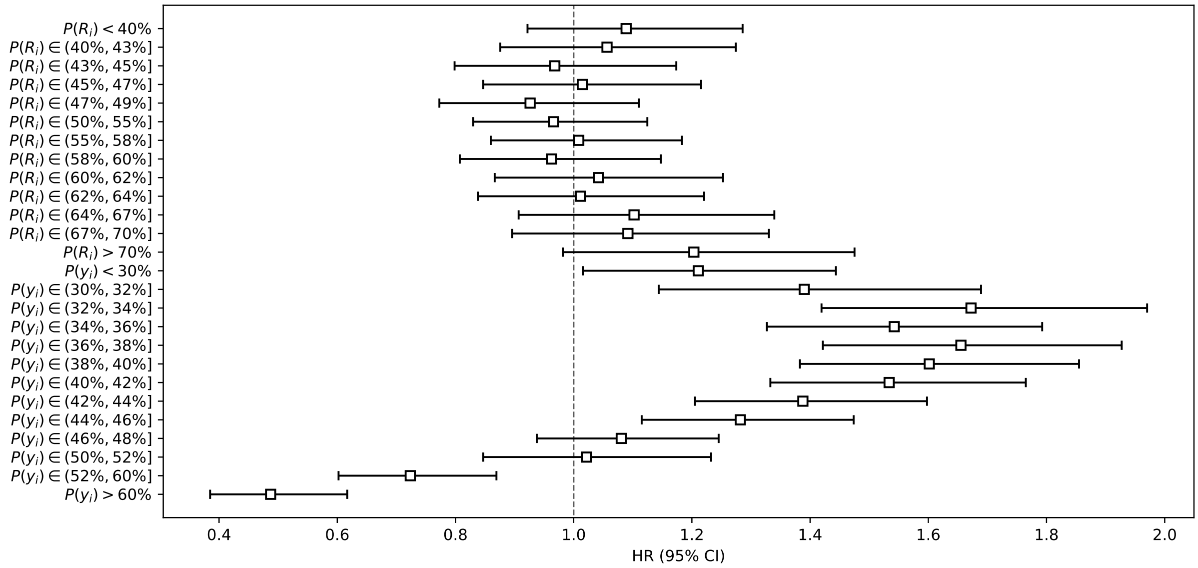


Figure 15 – Estimated hazard ratios from the Cox model, with the implied individual $\mathbb{P}(R_i > 0)$ and $\mathbb{P}(y_i > 0)$ being the variables of interest, sorted into intervals. The baseline dummy for $\mathbb{P}(R_i > 0)$ is $\mathbb{P}(R_i > 0) \in (49\%, 50\%]$ and for $\mathbb{P}(y_i > 0)$ is $\mathbb{P}(y_i > 0) \in (48\%, 50\%]$. Controls include dummies for state of residence, sex and age tercile.

11 indicates: traders with low levels of ability still take long to quit.

The observation that it takes so long for the great majority of day traders to realize that they are unskilled and quit is difficult to reconcile with rational models. It is easier the explain it in the light of behavioral explanations, such as overconfidence. The overconfidence story says that after taking sustained losses, the losing day traders still believe they can make the activity profitable for themselves. As a result, they underweight the signals the market is providing through the course of trading. After some time, however, they will likely quit. Moreover, they will be more likely to quit in inverse proportion to their ability. However, this learning process can be very costly for the vast majority of day traders, as Figure 2 indicates.

5 Conclusion

It is hard to reconcile day trading activity with rational behavior. The large majority of Brazilian day traders lose money in the long run; most of those who do not earn less than the minimum wage (Chague et al., 2019). Nevertheless, it is worth asking: do these day traders at least improve with time? Is there skill involved? Do unskilled traders eventually look back at their performance and quit?

The asymmetry between gains and losses observed seems consistent with the disposition effect. However, by the nature of our data, we cannot observe each trade's results but only the final daily results. In order to quantify this biased behavior, we propose a straightforward model that divides the day into two trading sessions and separates ability from the disposition effect. The results show that the disposition effect is economically meaningful: traders nearly try to recover all of their losses in the second period after a negative first period result. It is also prevalent among different groups when dividing the sample by age, gender and profession.

Chapter 3 shows that day traders become more skilled with experience (or, at least, learn to partially reduce losses). However, since the improvement is tiny, it does not seem to help their situation in any meaningful way: day traders would generally be better off quitting earlier instead of acquiring trading experience. Their bias, however, seems to increase instead of decrease. Indeed, even though the daily probability of receiving a positive profit increases with time, the implied probability of a positive session profit does not.

Traders using the daily probabilities to assess their ability may believe they are improving, but this happens mainly through the influence of the disposition effect. At the same time, the disposition effect reduces traders' gains over time (which is what they should care about). For this reason, the proportion of days on which day traders receive positive profits should not be used as a measure of ability.

The results in Chapter 4 support the conclusion that the less skilled traders in the past tend to quit faster in the future. Likewise, the more skilled traders tend to take longer to leave the market. However, this process is slow: it still takes a long time for the less skilled traders to quit even after the initial 150 trading days. These dynamics are challenging to reconcile with rational models and more easily comprehended if we think of day traders as overconfident about themselves, underweighting the negative signals that the market is providing them.

Combining these two answers, we can conclude that acquiring experience in this market is costly. It is not just that new traders are likely to lose money: experienced traders, on average, do not improve significantly from this baseline. There seems to be no point after which day trading starts paying off for the typical trader. Indeed, the performance improvement is so tiny that more experienced traders are expected to have larger career losses than novice traders. Moreover, for most traders, trading to learn does not pay off either. It takes a long time, and traders will most likely find out they are unskilled in the end. Consequently, in this case, the fruits of learning are mostly too little, too late.

Bibliography

Barber, B. M., Y.-T. Lee, Y.-J. Liu, T. Odean, and K. Zhang (2019, aug). Learning, fast or slow. *The Review of Asset Pricing Studies 10*(1), 61–93.

Barber, B. M. and T. Odean (2001, feb). Boys will be boys: Gender, overconfidence, and common stock investment. *The Quarterly Journal of Economics 116*(1), 261–292.

Camerer, C., L. Babcock, G. Loewenstein, and R. Thaler (1997, may). Labor supply of new york city cabdrivers: One day at a time. *The Quarterly Journal of Economics 112*(2), 407–441.

Chague, F., R. De-Losso, and B. Giovannetti (2019). Day trading for a living? *SSRN Electronic Journal*.

Farber, H. S. (2005, feb). Is tomorrow another day? the labor supply of new york city cabdrivers. *Journal of Political Economy 113*(1), 46–82.

Farber, H. S. (2008). Reference-dependent preferences and labor supply: The case of new york city taxi drivers. *The American Economic Review 98*(3), 1069–1082.

Friedman, M. (1953). The case for flexible exchange rates. In *Essays in positive economics*. Chicago: University of Chicago Press.

Heimer, R. Z. (2016, jul). Peer pressure: Social interaction and the disposition effect. *The Review of Financial Studies 29*(11), 3177–3209.

Jordan, D. and J. D. Diltz (2004, dec). Day traders and the disposition effect. *Journal of Behavioral Finance 5*(4), 192–200.

Linnainmaa, J. T. (2005). The individual day trader. Working paper, UCLA.

Linnainmaa, J. T. (2011, feb). Why do (some) households trade so much? *Review of Financial Studies 24*(5), 1630–1666.

Odean, T. (1999, dec). Do investors trade too much? *American Economic Review 89*(5), 1279–1298.

Appendix

APPENDIX A – Model Algebra

A.1 Model algebra

In this section, the goal is to show how the desired moments of $\tilde{R}^i$ can be computed using only moments from $t^+_{\mu,\nu}$, $t^-_{\mu,\nu}$ and $t_{\mu,\nu}$, and using the characterization from Equation (2.5). Throughout the following derivations, we omit the individual index i for the sake of notation.

A.2 First moment: $\mathbb{P}(\tilde{R} > 0)$

First, notice that:

$$\begin{aligned} \mathbb{P}(R > 0) = \quad & \mathbb{P}(B(p) = 1) + (1 - \mathbb{P}(B(p) = 1)) \cdot \mathbb{P}(t^-_{\mu,\nu} \cdot (1 - \alpha t_{\mu,\nu}) > 0) \\ & [1 - F_t(0; \mu, \nu)] + F_t(0; \mu, \nu) \cdot \mathbb{P}\left(t^-_{\mu,\nu}(1 - \alpha t_{\mu,\nu}) > 0\right). \end{aligned} \tag{A.1}$$

Simplifying the last term, we have that:

$$\begin{aligned} & \mathbb{P}\left(t^-_{\mu,\nu}(1 - \alpha t_{\mu,\nu}) > 0\right) = \mathbb{P}\left(1 - \alpha t_{\mu,\nu} < 0\right) = \\ & \mathbb{P}\left(t_{\mu,\nu} > \frac{1}{\alpha}\right) = 1 - F_t\left(1/\alpha; \mu, \nu\right) \end{aligned} \tag{A.2}$$

Plugging (A.2) into (A.1), we finally arrive that the expression:

$$\mathbb{P}(R > 0) = (1 - F_t(0; \mu, \nu)) + F_t(0; \mu, \nu)\left(1 - F_t\left(1/\alpha; \mu, \nu\right)\right) \tag{A.3}$$

The equation above only requires computations of the cdfs of the generalized and truncated t distributions. Indeed, notice that a crucial step for this result in (A.2) is that the probability of getting a positive profit in the second trading session does not depend on the size of the losses during the first session. Since $t^-_{\mu,\nu}$ is always negative, we only need $(1 - \alpha t_{\mu,\nu})$ to be positive to ensure that R is positive in this case.

A.3 Second Moment: $\mathbb{E}(\tilde{R})$

By simply applying the expectation in Equation (2.5) and using independence, we have:

$$\begin{aligned}\mathbb{E}(R) &= \mathbb{E}\left[B(p)t^{+}_{\mu,\nu} + (1-B(p))(t^{-}_{\mu,\nu}(1-\alpha t_{\mu,\nu}))\right] \\ &= p\mathbb{E}(t^{+}_{\mu,\nu}) + (1-p)\mathbb{E}(t^{-}_{\mu,\nu})(1-\alpha\mu) \qquad \text{(A.4)}\end{aligned}$$

A.4 Third Moment: Skew(R)

Let $\mathbb{E}(R) := \mu_R$. By expanding the term and applying the expectation, we have:

$$\begin{aligned}R^2 &= B(p)[t^{+}_{\mu,\nu}]^2 + (1-B(p))(1-\alpha t_{\mu,\nu})^2[t^{-}_{\mu,\nu}]^2 + 2B(p)(1-B(p))t^{+}_{\mu,\nu}t^{-}_{\mu,\nu}(1-\alpha t_{\mu,\nu}) \\ \implies \mathbb{E}(R^2) &= p\mathbb{E}\left(\left[t^{+}_{\mu,\nu}\right]^2\right) + (1-p)\mathbb{E}\left(\left[t^{-}_{\mu,\nu}\right]^2\right)\mathbb{E}\left([1-\alpha t_{\mu,\nu}]^2\right). \qquad \text{(A.5)}\end{aligned}$$

Notice that the cancellation of the third term comes from $B(p)(1-B(p)) \equiv 0$. Besides:

$$\mathbb{E}\left([1-\alpha t_{\mu,\nu}]^2\right) = \alpha^2\mathbb{E}(t^{2}_{\mu,\nu}) - 2\alpha\mu + 1 \qquad \text{(A.6)}$$

Plugging Equation (A.6) into (A.5), we have $\mathbb{E}(R^2)$ and, with that, we can compute

$$Var(R) := \sigma_R^2 = \mathbb{E}(R^2) - \mathbb{E}^2(R),$$

where $\mathbb{E}(R)$ comes from Equation (A.4). With this, we can compute $\mathbb{E}(R^3)$ analogously:

$$\begin{aligned}R^3 = B(p)\left[t^{+}_{\mu,\nu}\right]^3 + (1-B(p))\left[t^{-}_{\mu,\nu}\right]^3(1-\alpha t_{\mu,\nu})^3 \implies \\ \mathbb{E}(R^3) = p\mathbb{E}\left(\left[t^{+}_{\mu,\nu}\right]^3\right) + (1-p)\mathbb{E}\left(\left[t^{-}_{\mu,\nu}\right]^3\right)\mathbb{E}\left([1-\alpha t_{\mu,\nu}]^3\right), \qquad \text{(A.7)}\end{aligned}$$

where the cross terms involving $B(p)\cdot(1-B(p))$ cancel out as in Equation (A.5). By expanding the last term, we have:

$$\mathbb{E}\left([1-\alpha t_{\mu,\nu}]^3\right) = -\alpha^3\mathbb{E}(t^{3}_{\mu,\nu}) + 3\alpha^2\mathbb{E}(t^{2}_{\mu,\nu}) - 3\alpha\mu + 1, \qquad \text{(A.8)}$$

and by plugging Equation (A.8) into Equation (A.7) we can compute $\mathbb{E}(R^3)$. Finally, we are able to compute Skew(R):

$$\text{Skew}(R) := \mathbb{E}\left(\left[\frac{R-\mu_R}{\sigma_R}\right]^3\right) = \frac{\mathbb{E}(R^3) - 3\mu_R\sigma_R^2 - \mu_R^3}{\sigma_R^3}, \qquad \text{(A.9)}$$

which is invariant to scaling (Skew(cR) = Skew(R) for $c > 0$).

Thus, our final moments of $\tilde{R}$ will be given by:

1. $\mathbb{P}(R > 0)$, from Equation (A.3).

2. $\mathbb{E}(R/\sigma_R) = \mu_R/\sigma_R$, from Equation (A.4).

3. $\text{Skew}(R)$, from Equation (A.9).

APPENDIX B – Estimation of the model's parameters after taking into account transaction costs

Here we repeat the estimation of the model's parameters as in Chapter 3 but using profits net of transaction costs. Overall, the main results are similar, and the conclusion does not change. The most significant difference is that, while in Chapter 3 the proportion of days with positive profits was typically above 50% in most cases, here it is usually below 50%. Thus, there is one additional reason for skepticism towards using the proportion of days with positive profits to measure day traders' performance.

Table 7 presents the results of the pooled estimation. After comparing them with Table 2, there is not any economically meaningful qualitative difference. The estimated ability is more negative, which is expected: all investors would become more skilled in making money through investing if there were no additional costs per transaction.

	Parameter	**Standard Error**
μ	-0.2581	0.002
ν	3.8675	0.022
α	0.7719	0.005

Table 7 – GMM results for the static model using the full sample of traders, using profit data net of estimated transaction costs. Observations (individual-day): 1,581,300.

Figure 16 presents a Q-Q plot comparing the percentiles of the simulated distribution using the parameters from Table 7 against the empirical distribution. As in Figure 5b, the simulated and empirical distributions are very similar.

Figure 17 shows the dynamic results. As we would expect, the main difference from Figure 6 comes from estimated μ_t being more negative and estimated α_t being smaller than its counterparts from Chapter 3 . This difference is consistent with the model: a loss in a trading session translates into a bigger loss after subtracting transaction costs. Thus, we should expect the disposition effect to be more intense after a loss when transaction

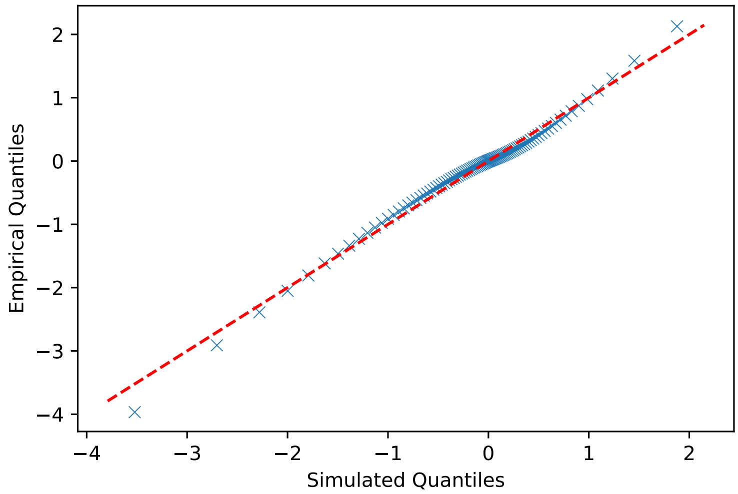


Figure 16 – Q-Q plot of the empirical percentiles against the simulated percentiles using the estimated parameters from Table 7.

costs are considered. Figure 18 shows how well the moments implied by the estimated parameters fit the data, and is similar to Figure 7. The results for the implied $\mathbb{P}(y > 0)$ and $\mathbb{P}(R > 0)$ are shown in Figure 19. As discussed before, the main difference from Figure 8 is that now $\mathbb{P}(R > 0)$ is below 50% for all periods.

Table 8 presents the comparison between men and women; Table 9 presents the comparison between older and younger traders; and, finally, Table 10 presents the comparisons between groups of professions. All of these results are qualitatively similar to what was shown in the main text.

	Obs. (men):		1,406,250		
	Obs. (women):		140,400		
Men	coef	std err	**Women**	coef	std err
μ	-0.2688	0.002		-0.1854	0.006
ν	3.8357	0.022		4.1722	0.130
α	0.7759	0.006		0.7599	0.014

Table 8 – GMM results comparing men and women (net of transaction costs).

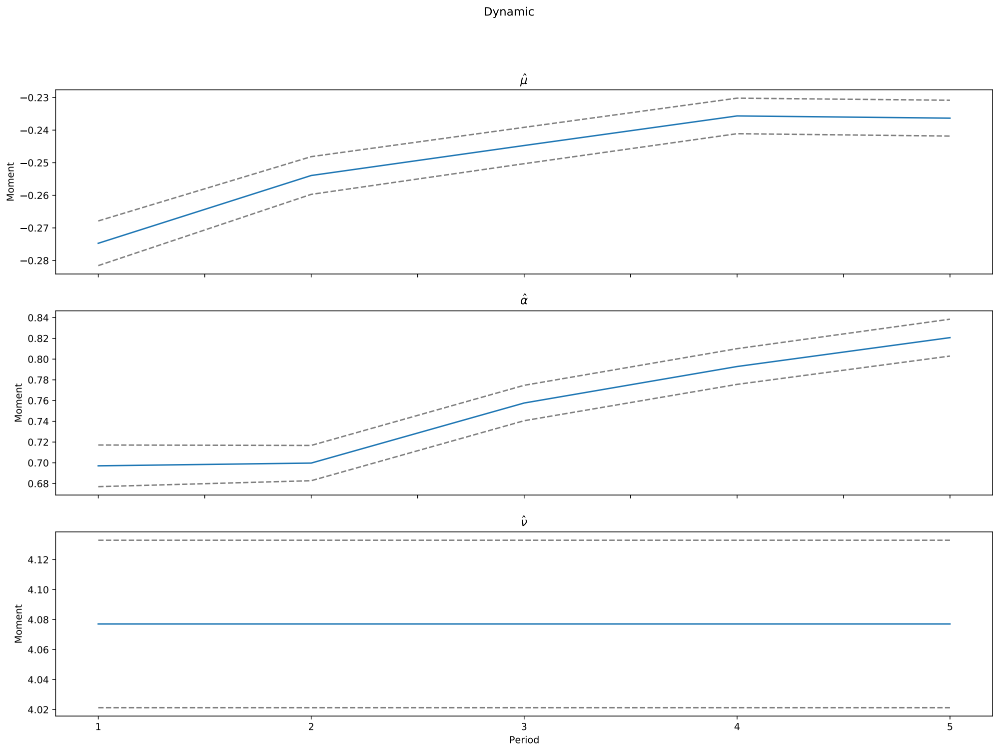


Figure 17 – Dynamic estimated parameters (net of transaction costs). Each period comprises 30 trading days. Dashed lines show 95% confidence interval.

	Obs. (younger):	538,680			
	Obs. (older):	467,676			
Younger	coef	std err	**Older**	coef	std err
μ	-0.2846	0.004		-0.2144	0.003
ν	3.8572	0.035		3.8877	0.047
α	0.8115	0.010		0.6996	0.008

Table 9 – GMM results (net of transaction costs) comparing younger traders (bottom age tercile) and older traders (top age tercile).

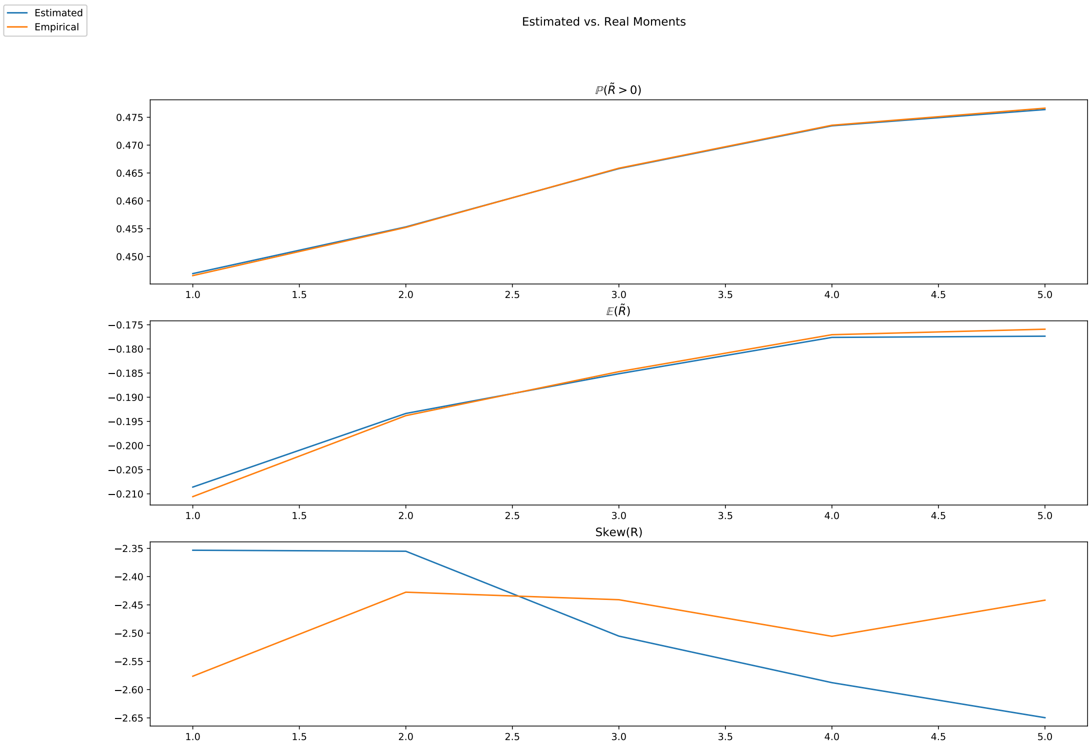


Figure 18 – Fit of the estimate dynamic moments *versus* the empirical dynamic moments (net of transaction costs). Each period comprises 30 trading days.

Obs. (money):	343,200
Obs. (not money):	1,238,100
Obs. (econ):	28,350

Money	coef	std err	**Not Money**	coef	std err	**Econ**	coef	std err
μ	-0.2018	0.004		-0.2739	0.002		-0.1295	0.011
ν	3.8765	0.053		3.8650	0.024		3.9703	0.230
α	0.7650	0.010		0.7743	0.006		0.7209	0.026

Table 10 – GMM results (net of transaction costs) comparing traders who work in professions that deal with money (*Money*) *versus* those who don't (*Not Money*). Additionally, the results for economists (*Econ*) are also reported.

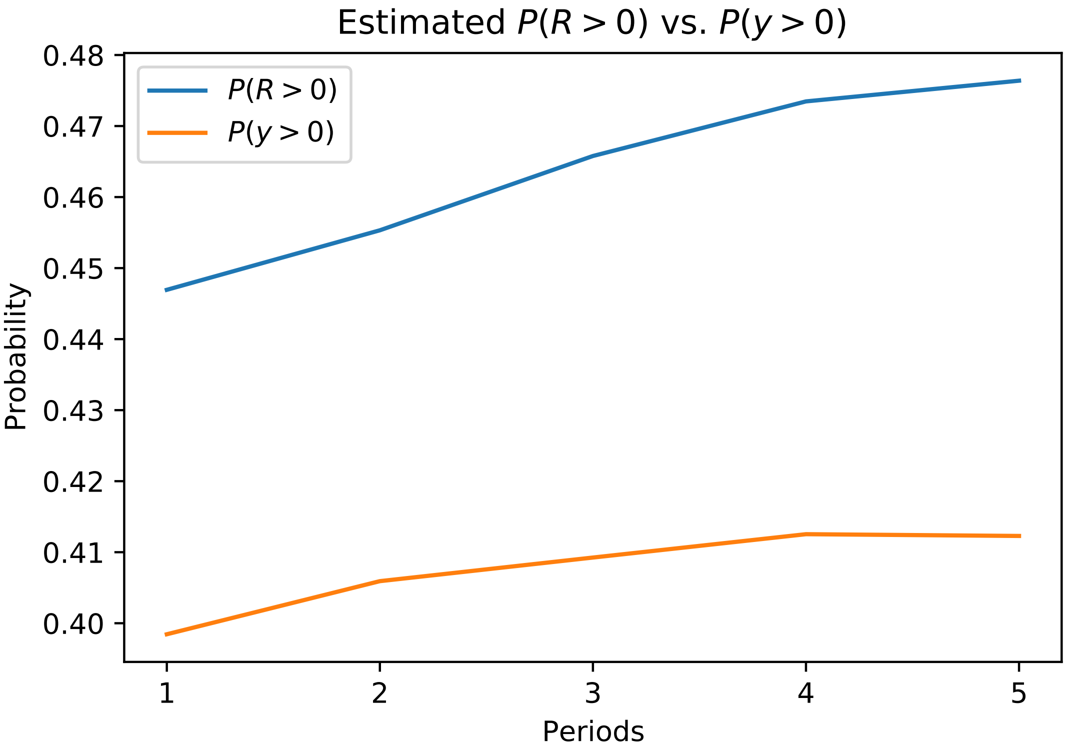


Figure 19 – Estimated "clean" measure of ability ($\mathbb{P}(y > 0)$) *versus* the estimated "polluted" measure of ability ($\mathbb{P}(R > 0)$) (net of transaction costs). Each period comprises 30 trading days.