Make hay while the sun shines: an empirical study of maximum price, regret and trading decisions*

Giovanni Burro† Julia Brettschneider‡ Vicky Henderson§

University of Warwick

November 6, 2019

Abstract

We carry on an empirical investigation of Regret Theory in dynamic financial decisions using data from actual stocks trading decisions (LDB dataset). We reject the prediction of Expected Utility theory that the optimal strategy for a decision maker implies stopping at a threshold. We investigate differences in adopting a threshold strategy in the loss and gain domain based on investors characteristics. Furthermore, we analyse the impact of price level, maximum price level and day of occurrence of the maximum on the propensity to sell a stock for a gain. We find that investors are more likely to sell a stock in a moment closer in time to maximum occurrence and at a price further from running maximum price of the investment episode.

Keywords: Regret Theory, Trading, Threshold, Maximum Price

JEL Codes: C55, D90, G40

*We would like to thank Terrance Odean for making the data available. We would like to thank David Hobson, Dario Spanò, Jo Kennedy, Jane Hutton, Daniel Read, James Griffin, David Firth, Robin Cubitt, Graham Loomes, Peter Wakker, Frédéric Schneider, Alessandro Bucciol (discussant), Gerrit Antonides (discussant) and seminar and conference participants at the University of Warwick (Bridges, YRM), University of Nottingham (CeDEx Brown Bag), Loughborough University (BEDS 2019), Roma (IISES 2018), University of Verona (Economics Department), Dublin (IAREP/SABE 2019), Dublin (2019 SABE Early Career Researchers Workshop), Bertinoro (Workshop for PhD students in Econometrics and Empirical Econonomics, Italian Econometric Society 2019), Amsterdam (SPUDM 2019), Trento (JDMx 2019), Milan (Advances in Decision Analysis 2019) for their comments and suggestions. Giovanni Burro gratefully acknowledges the support of the Leverhulme Trust, through the Bridges Doctoral Training Centre.

†University of Warwick Department of Statistics, g.burro@warwick.ac.uk

‡University of Warwick Department of Statistics, julia.brettschneider@warwick.ac.uk

§University of Warwick Department of Statistics, vicky.henderson@warwick.ac.uk

1 Introduction

Our work contributes to two broad fields: decision making and empirical analysis of the behaviour of financial investors. First, we are carrying on an empirical investigation of an extension of Regret Theory, RT (Loomes and Sugden, 1982; Strack and Viefers, 2013), one of the most successful theories of decision under risk (Bleichrodt and Wakker, 2015). Second, we add to the literature on the behavior of individual investors (see Barber and Odean, 2013 for a comprehensive review of the topic). Our work stems from the theoretical and experimental work of Strack and Viefers (2013), SV hereafter. In their work, they derive theoretical predictions for an Expected Utility (EU) agent and a RT agent in financial trading decisions. They show that a RT agent does not necessarily stops at a threshold when deciding to sell a stock and that her propensity to sell is influenced by the running maximum of the price of the stock she is trading. We test their predictions on the LDB dataset (Odean, 1998, 1999; Dhar and Zhu, 2006; Barber and Odean, 2013).

Regret theory has already been used to explain several phenomena in finance and economics. Gollier and Salanié (2006) and Muermann, Mitchell, and Volkman (2006) incorporate regret into models of asset pricing and portfolio choice, Braun and Muermann (2004) show that regret can explain the commonly observed preference for low deductibles in personal insurance markets. Barberis, Huang, and Thaler (2006) use regret to explain why people tend to invest too little in stocks. Michenaud and Solnik (2008) show that regret can explain the empirically observed behaviour that many investors do not adopt a full hedging policy with respect to currency risk, and Muermann and Volkman (2006) show how regret can explain the disposition effect.

Maximum price and its effect on individual decision has already been studied from other perspectives than ours. Grinblatt and Keloharju (2001) look at recent local maxima and the propensity to sell of investors. Heath, Huddart, and Lang (1999) show that exercise of option is higher when the price of the underlying stock is above its last year's peak. Barber and Odean (2008) and Huddart, Lang, and Yetman (2009) find that trading volume is high around both last year maximum and minimum. Finally, an experimental paper by Baucells, Weber, and Welfens (2011) investigates the impact of the highest level in a stream of payoffs on subjects formation of reference point.

We closely follow the work of Strack and Viefers (2013). They extend RT to dynamic trading decisions and test their predictions in a laboratory experiment. They compare a RT agent to an EU agent. An EU agent would find it optimal stopping at a threshold and her propensity to sell would not be influenced by the distance from past maximum. A RT agent does not necessarily stops at a threshold and her propensity to sell would be increasing in the price of the stock and decreasing in the price distance from past maximum.

Our results section consists of two parts. In the first part we investigate if investors follow a threshold strategy in the gain and in the loss domains. Consistently with SV we find that most of the traders do not follow it. We find that investors are more willing to adopt a threshold strategy in the gain with respect to the loss domain. That is a sort of a corollary of the higher propensity to realize gains with respect to losses, the disposition effect (Shefrin and Statman, 1985; Odean, 1998; Barber and Odean, 2013). On top of that we will go on to characterize those traders who are more likely to follow it and build some links with the literature on trading decisions and individual characteristics (Shapira and Venezia, 2001; Dhar and Zhu, 2006; Seasholes and Zhu, 2010; Korniotis and Kumar, 2011, 2013; Richards and Willows, 2018). We find that sophisticated investors and active traders are more likely to follow a threshold strategy and affluent and older investors are less likely to follow a threshold strategy. Males are more willing to realize losses at a threshold than females. In the second part, we investigate the impact of past maximum of a stock in a specific trading episode, on the propensity to sell that stock. We model the time to sell using a proportional hazard model (Cox, 1972). We take into account both the distance in price from past maximum, the distance in time from maximum realization and the ratio of price over running average price. We find that investors are more likely to realize a gain, the higher is the price over average price in the investment. The relationship of the propensity to realize a gain with the distance in price from past maximum follows an inverse U-shape. It peaks when distance in price is low but it decreases when it gets very low. We test also predictions which were not considered in the model by Strack and Viefers (2013). We take into account distance in time from past maximum. We find a strong and clear pattern: investors are less likely to sell a winning stock, the further in time the maximum price occurred. Moreover, we investigate the relation between distance in price and distance in time from past maximum and find that, when distance in time is low, investors are more likely to realize a gain, the higher is the distance in price from past maximum. We suggest that a panic effect is the main force leading to the realization of a stock.

Our paper consists of five sections, including this one. Section 2 reviews the literature. Section 3 describes our data and the methodology we use to analyze the data. In Section 4 we present our results and in Section 5 we conclude.

2 Literature Review

2.1 Too Proud to Stop: Regret in Dynamic Decisions

We start this section summarizing the findings in Strack and Viefers (2013). Their work is a pioneering work in the exploration of regret in dynamic decisions context. Regret has been widely studied in the context of static decisions while pretty neglected in the context of dynamic decisions (notable exceptions are the recent works by Descamps, Massoni, and Page, 2016; Fioretti, Vostroknutov, and Coricelli, 2018). The objective of their study is a stopping problem. In a stopping problem a decision maker observes a sequence of offers, which are realizations of a stochastic process X_t. After observing the n-th offer the decision maker is given the possibility to stop or keep seeing further offers. Once she chooses to go on she cannot accept forgone offers, while if she stops she cannot accept future offers. In such a decision framework, an EU maximizer would adopt a threshold strategy, meaning that she should stop the game as soon as the price reaches the gain threshold she fixed as her goal. The optimal threshold is ex-ante optimal and optimal at any point in time (there is no time inconsistency).

Definition 1. A threshold strategy $\tau(b)$ prescribes that agent stops at time t if the value of the process X_t exceeds the cut-off b and continues otherwise, where b is a given constant. If the agent uses the cut-off strategy $\tau(b)$ she will stop at the time $\tau(b, X) = min\{t \geq 0 : X_t \geq b\}$.

Building on Feng and Hobson (2016), SV show that a RT agent would not necessarily follow a threshold strategy. On top of that, she will adopt an optimal strategy which is different from the EU maximizer one. A RT agent will never find it optimal stopping under a past maximum while an EU maximizer who failed to stop at her threshold will not take into account how far above the threshold the stock went and will still find it optimal stopping at any value above her ex ante threshold. According to the hypothesis of the authors, the basic idea behind those concepts is that in expected utility the agent only evaluates current state of the process while in regret theory she takes into account the path of the process. It would be useful looking at figure (1). An EU agent would experience the same utility at the point denoted as bu, no matter when the stock hits it. A RT agent would experience a lower utility the second time the stock hits bu wrt to the first time. More formally, a regret agent faces a disutility due to regret because she compares her choice to the one that revealed to be ex-post optimal, namely

$$Regret = (\max_{s \in S} U_s) - U_t$$

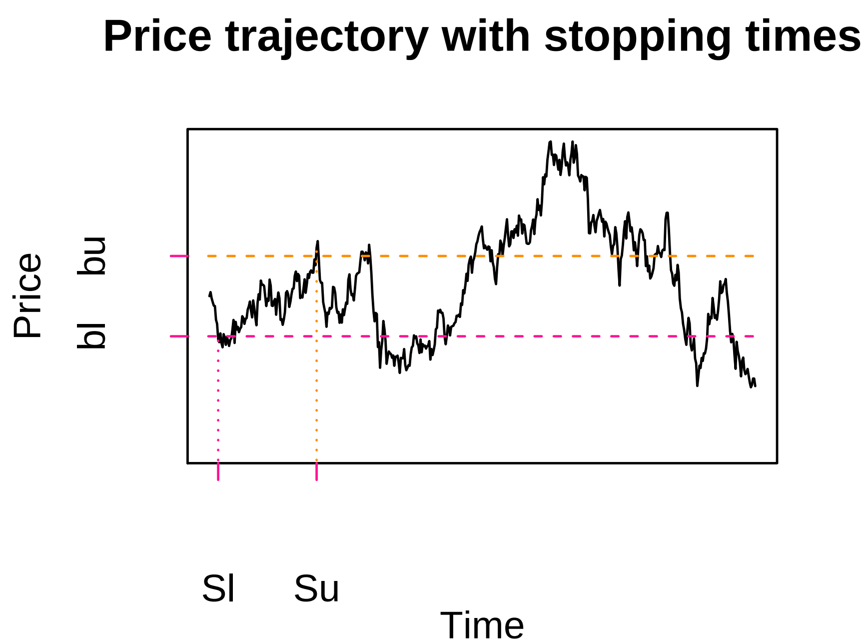

Figure 1: Example of a threshold investment strategy with upper and lower thresholds highlighted.

where S is the set of times relative to which the agent evaluates her regret. Thus the regret is driven by the difference between the maximum utility that could have been achieved and the utility actually achieved stopping at time t.

SV go on testing in the lab the following hypothesis, derived from their model:

• H1: Agents have a constant reservation level and they behave consistently in any repetition of the task, i.e. for all realized sequences (agents play 65 times the game) of offers $X = (X_1, X_2, ...) \neq \left(X_1', X_2', ...\right) = X'$ the level at which the agent stops the game is the same $\tau(X) = \tau\left(X'\right)$

• H2: Agents never stop under the running maximum, i.e. they never stop at a level where they decided to continue before;

• H3: Agents follow a threshold strategy, hence the stopping decision τ satisfies $\tau = \inf\{t : X_t = X_\tau\}$

• H4: Agents are more willing to stop the higher the level of the price and they are less willing to stop the higher the maximum level of the price up to the decision moment, i.e. the empirical frequency with which subjects stop at a given level x, given past maximum s is increasing in x and decreasing in s.

In their work they are able to reject the first three hypothesis and to confirm the fourth

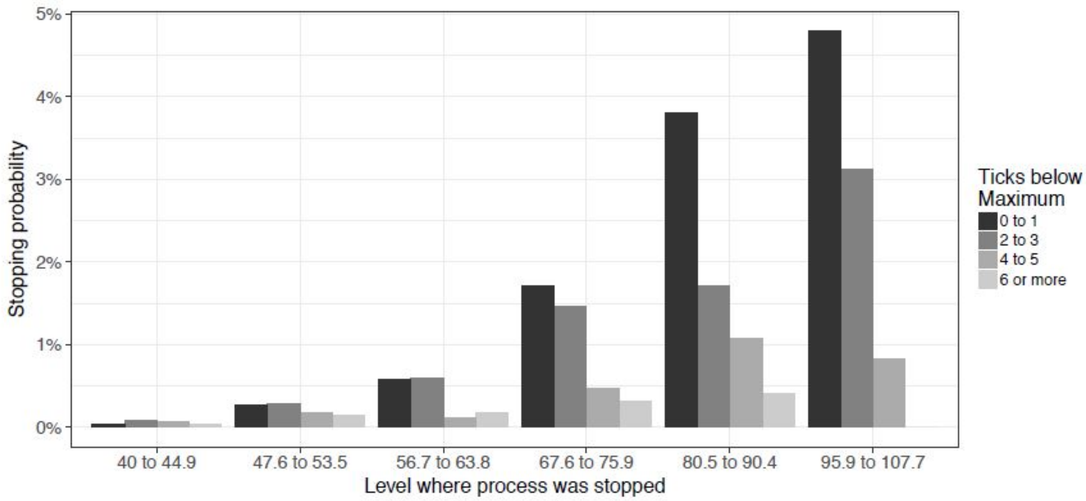

Figure 2: Experimental Results from Strack and Viefers (2013). Empirical stopping frequency per any level of the price and distance from past maximum (Figure 4 in Strack and Viefers, 2013).

one (see Figure 2. In our work we are interested in translating their hypothesis into testable hypothesis on a real world dataset. We will go back to this later in the discussion.

2.2 The behaviour of individual investors

Our work contributes to the understanding of the behaviour of individual investors. A very good review on the topic is Barber and Odean (2013). The most widely studied phenomenon in this area is the disposition effect (Shefrin and Statman, 1985), the tendency of investors to realize gains at a faster rate than losses. Odean (1998) examines trading records for 10,000 accounts at a large US discount brokerage for the period 1987–1993. He compares the rate at which investors sell winners (realized gains) and losers (realized losses) and compares the realization of gains and losses to the opportunities to sell winners and losers. He finds that, relative to opportunities, investors realize their gains at about a 50% higher rate than their losses and that this difference is not explained by informed trading, a rational belief in mean reversion, transactions costs, or rebalancing. The disposition effect is stronger among individuals and institutional investors, according to Brown, Chappel, Da Silva Rosa, and Walter (2006) and Barber, Lee, Liu, and Odean (2007). Dhar and Zhu (2006) show that wealthier individuals, individuals employed in professional occupations and frequent traders exhibit a lower disposition effect.

As far as maximum is concerned, there are several papers linked to the idea that the maximum point of a price process has an effect on agents behaviour. First of all, Grinblatt and Keloharju (2001) show that high recent past returns of stocks tend to increase propensity to sell of investors and propensity to sell is higher when a stock hits its last month maximum price. Heath et al. (1999) show that exercise of option is higher when the price of the underlying stock is above its last year's peak. Barber and Odean (2008) and Huddart et al. (2009) find that trading volume is high around both last year maximum and minimum. Finally, an experimental paper by Baucells et al. (2011) investigates the impact of the highest level in a stream of payoffs on the formation of the reference point.

The data we are investigating is a well established dataset in the economic literature. It is referred to as the *LDB Dataset*. Data contain information on trading activities of individual households in the period 1991-1996. Investors are from the USA. In particular, we can mention several studies that tried to explain a great variety of phenomena through those data. There are some papers which used the rich description in terms of investors characteristics to detect differences in investing ability among groups. For example, Barber and Odean (2001) conclude that women trade less than men and have better average performances. Korniotis and Kumar (2011, 2013) focus on demographic differences among people and differences due to age and financial knowledge. Ivković, Sialm, and Weisbenner (2008) document that investors with concentrated portfolios tend to outperform other investors. The authors attribute concentration to an informational advantage of those investors. A similar finding is due to Ivković and Weisbenner (2005). They argue that investors in the LDB dataset tend to rely upon local stocks and that leads to higher returns. However, Seasholes and Zhu (2010) dispute their finding blaming some methodological issues in the estimation of cross-sectional differences in investors performance. Investors' performance is investigated by Odean (1999), Barber and Odean (2000), Coval, Hirshleifer, and Shumway (2005) and Huang (2019).

3 Motivation, Data and methodology

Our data span the period from January 1991 to November 1996. We use two different samples of the data for the two sections of our analysis. We are going to explain the details of all our choices. Some of the data preparation steps are common between the two specifications and we will describe them here. We obtain price data from the CRSP (Centre for Research in Security Prices) of WRDS (Wharton Research Data Services). We exclude those stocks for which we were not able to recover price information. We remove investor-stocks records if at least one of the entries has negative commissions (which may indicate that the transaction was reversed by the broker). We remove from the sample investor-stocks that include short-sale transactions or that have positions that were opened before the starting point of our dataset. We remove those trades where the buy and sell dates coincide (Ivković, Poterba, and Weisbenner, 2005; Ben-David and Hirshleifer, 2012).

The starting point of an investment the first time an investor buys a stock or any time she buys it without the stock being present in the portfolio at that time. The end point of an investment the first sale date after that buy date as the end point of the investment[1] (Shapira and Venezia, 2001; Brettschneider and Burgess, 2017). We define an episode as all the day-stock information between a buy and a sell date. An episode is classified as a gain if the selling price is higher or equal than the buy price. It is classified as a loss otherwise.

We now introduce a variable: the distance from the extreme at time t, d_t, which we will refer to as "distance". Time t is defined in terms of trading days.

Definition 2. The distance from the extreme is defined as

$$d_t = \begin{cases} \frac{t-t_{max}}{t}, & \text{if episode ends up as a gain} \\ \frac{t-t_{min}}{t}, & \text{if episode ends up as a loss} \end{cases} \tag{1}$$

where t is the number of days since the episode started, and t_{max} and t_{min} are the days when the current maximum and minimum prices of the episodes realized, respectively. Hence, t is always bigger or equal than t_{max} and t_{min}. These are calculated taking the starting point of an episode equal to $t = 0$.

3.1 Threshold Investigation

We introduce the research question we will address in the corresponding subsection in Section 4.

- *Do investors follow a threshold strategy, as defined in definition (1)*?
- *Which categories of investors are more likely to follow a threshold strategy?*

Threshold strategy in the loss domain is not discussed from a theoretical point of view in SV, hence our analysis is more agnostic than the one we make for the threshold strategy in the gain domain. Optimality of the threshold strategy is defined theoretically only in the gain domain by SV. We will assume that a threshold strategy is rational also for the loss domain. It is linked to the concept of "stop-losses".

[1]We carried on part of the analyses using a different definition of the selling date and results were basically the same. On top of that, in the vast majority of cases investors only bought a stock once and then sold the entire share they held in their portfolio at once.

Table 1: Summary Statistics for the sample used in the Threshold analysis. Gain refers to the sample where only investments which resulted in a gain are considered (return higher or equal than 0). Loss refers to the sample where only investments which resulted in a loss are considered. All considers those portfolio where at least a gain and a loss trading episode were completed.

	Gain	Loss	All
Percentage of Threshold Episodes	0.316	0.258	0.293
Number Bank Account	15,624	11,390	8,674
Mean Rate of Threshold Consistency per Bank Account	0.275	0.216	0.257
Median Rate of Threshold Consistency per Bank Account	0.043	0	0.250
Mean Number of Episodes per Bank Account	4.640	3.954	11.493
Median Number of Episodes per Bank Account	2	2	6
Number of Cash Bank Accounts	2,591	1,812	1,218
Number of IRA Bank Accounts	2,798	1,674	1,227
Number of Keogh Bank Accounts	91	76	51
Number of Margin Bank Accounts	2,436	1,884	1,469
Number of Schwab Bank Accounts	7,708	5,944	4,709
Number of General Traders	10,368	7,080	5,085
Number of Affluent Traders	2,134	1,549	1,045
Active Trader (num.)	3,122	2,761	2,544
Mean Age	4.970	5.058	5.025
Median Age	4.800	4.800	4.800
Mean Income per Bank Account	6.219	6.224	6.213
Median Income per Bank Account	6	6	6
Number of Females	1,329	949	689
Number of Males	11,947	8,717	6,627
Number of Not Professional Traders	10,362	7,672	5,916
Number of Professional Traders	880	632	462
Number of Traders with Other Occupation (even NA)	4,382	3,086	2,296

We take into account episodes whose length is shorter or equal to 300 days from buy to sell date (209 trading days). We want to be sure we capture active trading decisions and that we are not looking at decisions of buy-and-hold long term investors (Benartzi and Thaler, 1995; Heath et al., 1999; Brettschneider and Burgess, 2017). We focus on the sample of portfolio for which demographics are available. The sample is summarized in Table 1. It is obviously the case that an episode where the investor follows a threshold strategy is characterized by the condition $d_T = 0$ where T is the selling date for that episode. This is the proxy we are going to use to define a threshold strategy. Namely,

Definition 3. A trading episode is said to be a threshold strategy episode if $d_T = 0$.

It is obviously the case that Definition 3 gives a necessary but not sufficient condition to define an investment as a threshold episode. However, since our aim is rejecting a threshold strategy, we are only adding more obstacles to our goal by taking into account a less stringent hypothesis. Being able to reject it would imply, *a fortiori*, that a threshold strategy does not hold. A perfect threshold strategy would imply that a stock is sold as soon as the maximum is realized.

In Table 1 we see the distribution of some statistics for our sample. We notice that they are in line with the idea that investors suffer from the disposition effect. In particular, d_t is lower for gains than for losses. This is a signal that investors tend to realize gains quicker and have a strong aversion to realize losses at a minimum. Overall, the vast majority of trades does not adhere to a threshold strategy. That is shown also in Figure 3. Our definition of threshold strategy applies only to 31.6% of episodes in the gain domain, and to 25.8% of trades in the loss domain.

These are descriptive statistics which clearly hint that a threshold strategy is not consistently followed by our population of investors. In Section 4 we are going to investigate how the propensity to follow a threshold strategy changes at individual level.

3.2 Maximum Investigation

Our second research question is:

- *How does the propensity to sell a stock vary with respect to the three following variables?*
    - *Level of the price;*
    - *Distance in time from the day of running maximum realization;*
    - *Distance in price from the running maximum.*

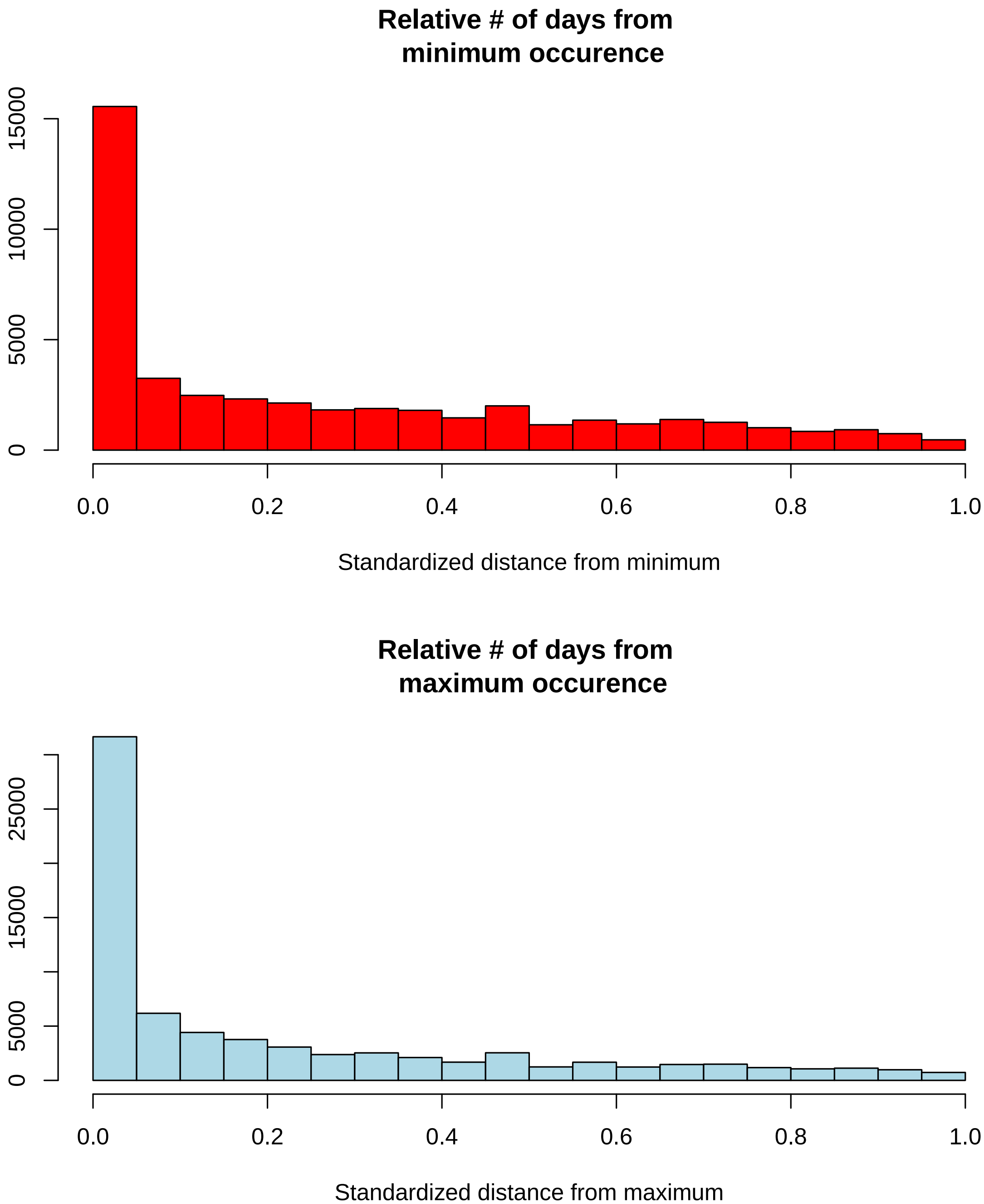

Figure 3: On the left, frequency of d_T, the rescaled distance in time from the minimum day calculated in the day when a sell for a loss took place. One observation per episode. On the right, frequency of d_T, the rescaled distance in time from the maximum day calculated in the day when a sell for a gain took place. One observation per episode.

In this section we restrict our attention to a random sample of 13000 episodes whose length is smaller or equal to 300 days (209 trading days). We made this choice to increase the probability we capture active decisions from investors and to get more stable estimation in our survival models. We only take into account trading episodes which ended up with a gain, in order to be consistent with SV. Their predictions only take into account maximum price and an upper threshold. On top of that, in their experiment the simulated price process is biased towards the gain domain (there is an higher probability of the price to increase than to decrease). We excluded the 10% of most volatile episodes[2]. Since we are interested in how the propensity to sell a stock changes when the distance of the price from past maximum increases, we did not want to focus on those trades where the likelihood of big intraday drops in price is high. The characteristics of our sample are summarized in Table 2.

Table 2: Summary Statistics for the sample used in the Maximum analysis

Number Bank Accounts	8704
Number Trading Episodes	13000
Mean Return per Episode	1.19
Median Return per Episode	1.12
Mean length per Episode	68.07
Median Return per Episode	52

We analyse data using the proportional hazard model (PH), developed by Cox (1972). Survival analysis models are widely used in medical research and they are relatively popular in demography and labour economics. They have recently been used in a series of financial applications (Ivković and Weisbenner, 2005; Deville and Riva, 2007; Panayi and Peters, 2014; Gepp and Kumar, 2015; Jiao, 2015; Brettschneider and Burgess, 2017; Richards, Rutterford, Kodwani, and Fenton-O'Creevy, 2017).

PH model is a semi-parametric model, aimed at describing the "time-to-event" of individuals. In our case the time to event is the time from the start to the end of an investment episode. It has the advantage of assessing the impact of covariates over the entire time axis, while for example a logistic regression only evaluates the odds of the event/non event with respect to a fixed time. Informally, a PH model incorporates all the information accumulated in time for a given episode. A logistic regression would evaluate the information in a given day independently from the information on other days of the same episode. We now add more details to our discussion. In particular, we need to define some objects

Definition 4. Let T be a non-negative continuous random variable, representing the time until the event of interest.

[2]We get the average daily ratio of minimum to maximum price per each investment episode and we exclude those trades where the ratio is lower or equal than 0.93

$F(t) = P(T \leq t)$ denotes the distribution function and $f(t)$ the probability density function of random variable T.

$S(t) = P(T > t) = 1 - F(t)$ is the survival function. It is the probability that a randomly selected individual will survive beyond time t. It is a decreasing function, taking values in $[0, 1]$ and it equals 1 at $t = 0$ and 0 at $t = \infty$.

$H(t) = \log S(t)$ is the cumulative hazard function.

The hazard function (or hazard rate) is defined as:

$$\begin{aligned} h(t) &= \lim_{\triangle t \to 0} \frac{P(t \leq T < t + \triangle t \mid T \geq t)}{\triangle t} \\ &= \frac{1}{P(T \geq t)} \lim_{\triangle t \to 0} \frac{P(t \leq T < t + \triangle t)}{\triangle t} \\ &= \frac{f(t)}{S(t)} = \frac{-d}{dt} \log S(t) = \frac{d}{dt} H(t) \end{aligned}$$

$h(t)$ measures the instantaneous risk of dying right after time t given the individual is alive at time t.

In particular, we will fit a regression model where we evaluate the change in the hazard rate with respect to a set of covariates. That is to say, given a set of covariates $\mathbf{x}_i = (x_{i1}, x_{i2}, ..., x_{ip})$ measured for subject i, the following model is fit to data

$$h_i(t) = h_0(t) \exp\left(\beta^t x_i\right);$$

with β, a $p \times 1$ vector of parameters and $h_0(t)$ which is the baseline hazard function (i.e. hazard for a subject i with $x_j = 0,\ j = (1,\ 2, ...,\ p)$).
The proportional hazards assumption states that the ratio of the hazards of two subjects with covariates x_i and $x_{i'}$ is constant over time:

$$\frac{h_i(t)}{h_{i'}(t)} = \frac{\exp(\beta^t x_i)}{\exp(\beta^t x_{i'})}$$

The Cox PH model is a semi-parametric model. It means that it leaves the form of $h_0(t)$ completely unspecified and it estimates the model in a semi-parametric way. Then, to estimate the model we maximize a partial likelihood. Finally, we should point out that we are estimating a model with time changing covariates so it is better to define it as

$$h_i(t) = h_0(t) \exp\left(\beta^t x_{it}\right)$$

To deal with unobserved heterogeneity we stratify the model based on the investor who holds the position. This means each investor has a different baseline hazard function, which can absorb any heterogeneity not captured by the model covariates Hence, the hazard function for the i_{th} position of the j_{th} investor is

$$h_{ij}(t) = h_{0j}(t) \exp\left(\beta^t x_{ijt}\right);$$

where $\mathbf{x}_{ijt}$ is the covariate vector for the position. A final word should be spent on why we need to stratify at investor level[3]. Possible reasons for there being a difference between investors include their preference for risk, their beliefs about the market (e.g. whether there is price momentum or not), their investment objectives and the particular strategy they are following. For example, some investors may trade very frequently and follow a strategy based on short term changes in stock prices. The holding periods of these investors will therefore be shorter than other investors in the sample. Differentiating investors baseline hazards can separate out this kind of difference from the effects of covariates included in the model, that are in theory common across all investors.

4 Results

4.1 Threshold strategy investigation

In this section an investment episode will be claimed to respect a threshold strategy if it satisfies Definition 3. As we said, this is only a necessary condition for a threshold strategy (it is not sufficient). However, since we want to show that a threshold strategy is not able to define the behaviour of traders, we are actually making harder to disprove this claim. We already saw in Section 3 that the vast majority of trading episodes is not consistent with a threshold strategy. The unit of analysis in this section will be the portfolio, since we are interested in the rate of consistency with threshold for each portfolio[4] In table 3 we analyze the rate of threshold consistency per bank account. We perform a negative binomial regression[5]. The dependent variable in our negative binomial regression is defined as:

[3]We will refer to investor level to simplify the discussion. We actually stratify at bank account level but in most cases investors only have one bank account.

[4]Each individual might hold more than 1 bank account but we consider each bank account separately.

[5]We performed a Likelihood Ratio Test to check that a negative binomial model is more approriate than a Poisson model. Negative binomial models assume the conditional means are not equal to the conditional variances. This inequality is captured by estimating a dispersion parameter (not shown in the output) that is held constant in a Poisson model. Thus, the Poisson model is actually nested in the negative binomial model. We can then use a likelihood ratio test to compare these two and test this model assumption.

- N_g, the number of investments in a portfolio which respect Definition 3 and resulted in a gain, for columns 1 and 2 of Table 3.

- N_l, the number of investments in a portfolio which respect Definition 3 and resulted in a loss, for columns 3 and 4 of Table 3.

- $N = N_g + N_l$, the total number of investments in a portfolio which respect Definition 3 for columns 5 and 6 of Table 3.

Hence, in columns 1 and 2 we only look at those bank accounts where at least an investment which resulted in a gain was recorded. In columns 3 and 4, we look at those bank accounts where at least an investment which resulted in a loss was recorded. In columns 5 and 6, we look at those bank accounts where at least an investment which resulted in a gain and at least one which resulted in a loss were recorded. To control for the fact that different investors completed different numbers of trading episodes, we take into account the logarithm of the number of completed episodes (completed gains, losses or all depending on the regression) as an offset. Hence, we measure how the rate of threshold consistency varies from one portfolio to the other. We now introduce the covariates we are taking into account. They are all defined at bank account level.

- Dummy for the account type: Cash Account, IRA and Keogh, which are two different types of retirement accounts, Margin accounts and Schwab One accounts;

- Client Segment: Affluent if at any point in time she has more than $100,000 in equity, active if she makes more than 48 trades in any year and General for the residual individuals. If traders could be classified as both affluent and active they were classified as active traders;

- Age in decades;

- Income is classified as a numeric variable which takes values from 1 to 9 and increases with income of the individual[6].

- Gender;

- Occupation, we follow Dhar and Zhu (2006): non-professional if the trader has a "white collar/clerical","blue collar/craftsman" or "service/sales" job; professional occupation

[6] 1 corresponds to less than $15,000 per year; 2 to $15,000 to $19,999; 3 to $20,000 to $29,999; 4 to $30,000 to $39,999; 5 to $40,000 to $49,999; 6 to $50,000 to $74,999; 7 to $75,000 to $99,999; 8 to $100,000 to $124,999; 9 to $125,000 or more

if the trader has a "professional/technical" or "administrative/managerial" occupation; the residual category is everyone else[7].

We see that investors with margin accounts and Schwab accounts are more likely to follow a threshold strategy. In particular, investors with margins accounts show a rate of threshold consistency which is, on average, 20% higher than cash accounts both for gains and for losses, separately (almost 30% when we consider overall rate). Schwab account holders have a consistency rate which is around 10% higher than cash accounts for gains and around 5% for losses. If we consider a threshold strategy as the rational choice for an investors, we see that more sophisticated investors (those who have margin accounts and Schwab accounts) are more likely to adopt it. This is in line with the idea that sophistication lowers investment biases (Grinblatt and Keloharju, 2001; Dhar and Zhu, 2006). Retirement accounts show a higher rate of threshold consistency than cash accounts but only when the overall rate is considered. Active traders have a higher rate of threshold consistency. That is especially true for losses (6 to 8% higher rate than general traders). That is to be expected since frequency of trading increases the chances that investors are constantly monitoring their investments. Consistency with threshold depends also on attention since investors might lose the possibility to stop at a threshold because of inattention. Affluent traders seem to be less consistent than general traders with a threshold strategy but the effect boils down when we take into account age. That is probably due to the fact that affluent traders are much older, on average, than general traders. Older investors are less likely to be consistent with threshold strategy. Every ten years, the rate of threshold consistency decreases by around 8%, 4% and 7% in the gain, loss and overall sample. This finding can be linked to the idea that older individuals have lower decision making abilities and make worse financial decisions (Korniotis and Kumar, 2011; Bruine De Bruin, 2017). The higher is the income of the traders, the less likely they are to follow a threshold strategy. Males are more likely to follow a threshold strategy than females for losses (their rate of threshold consistency is 14.1% higher). We do not see any differences based on the occupation of the traders.

The take home from this section is that investors do not seem to follow consistently a threshold strategy. That is in line with the experimental findings from SV. We tried to investigate differences at investor level. The main differences are due to age, client segment and account type. Sophisticated investors (margin and Schwab accounts) and active traders are more consistent with a threshold strategy. Affluent and older investors are less consistent than general investors with a threshold strategy. Males are more willing to realize losses at a threshold than females.

[7] To avoid having too many missing observations, also missing values were classified in the residual category.

Table 3: Odds ratios with 95% c.i. of a Negative Binomial regression where each observation corresponds to a bank account. The dependent variable is the number of times the investor stopped at a threshold in the gain, loss or overall. There is an offset equal to the number of episodes in the bank account (in the gain, loss, overall).

	Gain		Loss		All	
	(1)	(2)	(3)	(4)	(5)	(6)
Account Type (ref. Cash)						
Account Type IRA	1.068*	1.044	1.097*	1.082	1.105***	1.096**
	(0.997,1.143)	(0.965,1.129)	(0.998,1.207)	(0.971,1.206)	(1.031,1.185)	(1.012,1.187)
Account Type Keogh	1.111	1.180	1.267*	1.144	1.238**	1.262**
	(0.900,1.367)	(0.917,1.516)	(0.980,1.628)	(0.840,1.545)	(1.025,1.494)	(1.014,1.570)
Account Type Margin	1.194***	1.186***	1.237***	1.195***	1.289***	1.275***
	(1.119,1.274)	(1.099,1.280)	(1.135,1.350)	(1.080,1.324)	(1.210,1.373)	(1.184,1.374)
Account Type Schwab	1.129***	1.092***	1.152***	1.130***	1.202***	1.169***
	(1.068,1.194)	(1.023,1.167)	(1.069,1.244)	(1.033,1.236)	(1.137,1.270)	(1.094,1.248)
Client Segment (ref. General)						
Client Segment Affluent	0.905***	0.951	0.871***	0.952	0.863***	0.911**
	(0.851,0.962)	(0.884,1.023)	(0.798,0.950)	(0.860,1.053)	(0.810,0.920)	(0.844,0.981)
Client Segment Active	1.036*	1.092***	1.059**	1.078**	1.076***	1.117***
	(0.998,1.076)	(1.045,1.142)	(1.009,1.111)	(1.017,1.142)	(1.040,1.113)	(1.073,1.163)
Age (decades)		0.921***		0.958***		0.932***
		(0.906,0.936)		(0.938,0.978)		(0.919,0.946)
Income		0.990*		0.977***		0.985***
		(0.979,1.000)		(0.964,0.991)		(0.976,0.995)
Male		1.008		1.141**		1.058
		(0.931,1.093)		(1.021,1.277)		(0.979,1.144)
Occupation (ref. Other (also NA)						
Non Professional Occupation		1.060		1.004		1.031
		(0.977,1.150)		(0.898,1.121)		(0.954,1.113)
Professional Occupation		1.015		0.956		0.994
		(0.971,1.061)		(0.901,1.014)		(0.954,1.036)
McFadden Adj. R^2	0.24	0.46	0.25	0.46	0.25	0.46
Observations	15,624	11,477	11,390	8,315	8,674	6,280

Note: *p<0.1; **p<0.05; ***p<0.01

4.2 Maximum price investigation

In this section we propose a survival analysis estimation of the propensity to sell a stock. We look at a sample of investments which lasted no longer than 300 days (209 trading days) and resulted in a gain. The choice of restricting the sample to such a period comes from the fact that we want to guarantee the PH assumption holds and we base our estimation on the idea that investors attention does not span a very long period (Benartzi and Thaler, 1995; Brettschneider and Burgess, 2017). The choice of restricting attention to the gain domain is due to the fact that it would be difficult estimating the impact of maximum price on the propensity to sell a stock for a loss, since maximum price would often coincide with buy price. On top of that, SV test their prediction in a laboratory setting where stocks are on average in the gain domain. In our analysis we take into account several variables which are updated at daily level. The buy date of any investment episode is date 0. Every date is registered as the difference in trading days between that date and the buy date and denoted by t. Only propensity to sell on days when the stock is trading above the buy price is estimated. That means that information on those days in which the stock is trading at a loss is not incorporated in the estimate. We confine ourselves to estimate the propensity to sell for a gain. Hence, we thought it was not appropriate estimating the propensity to sell the stock on those days when it was trading at loss, since we constrained it to zero[8]. We define now our covariates of interest, measured at any given day t

- Distance: it is the rescaled distance from the occurrence of running maximum date, as we defined it in definition 2, $\frac{t-t_{max}}{t}$. t_{max} is the day when maximum price between day 0 and day t realized. We split it into tertiles based on the stock-portfolio-day distribtuion. Distance is defined as low in the interval [0; 0.07]; medium in the interval [0.07; 0.34) and high in the interval [0.34; 1];

- Ratio to Max Price (Ratiomax) is the ratio of daily closing price to maximum price up to that time in the investment episode. On the selling date it is equal to the ratio of selling price to maximum price in the episode. We split it into quartiles based on the stock-portfolio-day distribution. Defined as low in the interval [0.349; 0.918]; medium-low in the interval (0.918; 0.957]; medium-high in the interval (0.957; 0.981] and high in the interval (0.981; 1].

- Ratio to Average Price (Ratioavg): Ratio to Avg Price is the ratio of daily closing price to average daily closing price up to that time in the investment episode. On the selling

[8]To make a parallel with the medical literature, from which we borrow our estimation strategy, think about an allergy which we know can only occur during the spring. It would not make sense estimating the probability of occurrence based on the covariates measured during the winter.

date it is equal to the ratio of selling price to average closing price in the episode. We split it into quintiles based on the stock-portfolio-day distribtuion. Defined as low in the interval [0.58; 1.01]; medium-low in the interval (1.01; 1.03]; medium in the interval (1.03; 1.07]; medium-high in the interval (1.07; 1.14]; high in the interval (1.14, 3.20].

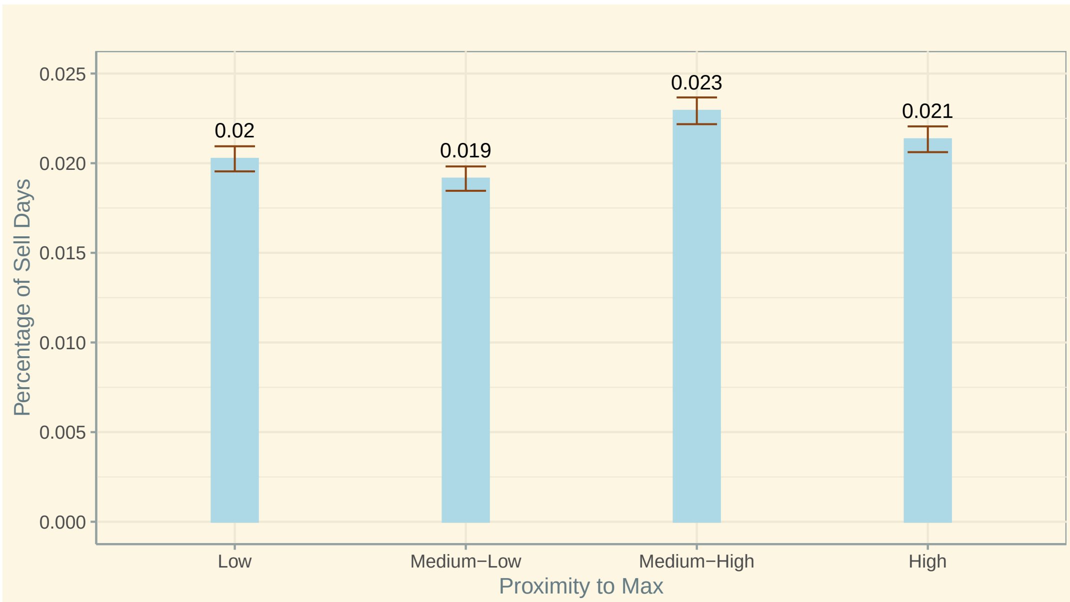

Figure 4: Percentage of selling days out of all stock-portfolio-days per each category (0.95 c.i.). Ratio to Max Price is the ratio of daily closing price to maximum price up to that time in the investment episode. On the selling date it is equal to the ratio of selling price to maximum price in the episode. Low [0.349; 0.918]; Medium-Low (0.918; 0.957]; Medium-High (0.957; 0.981]; High (0.981; 1].

We standardize all variables since we need consistency from one trading episode to another. Prices are really different from stock to stock. We need to normalize them in a way to make them comparable. On top of that, we decided to stratify variables instead of using their continuous version for two reason. First, to take into account non linearity. Second, to have a model which could be analyzed through Proportional Hazard technique[9]. The most important implication of the PH assumption is that the effect of a covariate is constant in time. Whilst violation of the assumption does not invalidate the model, it does significantly alter the interpretation, particularly when only hazard ratios are reported. If an effect does change over time then the hazard ratio is only an average of this process, and if it changes a lot then this average can be misleading.

[9]A careful fine tuning was made to identify the best number of bins to split each variable. We wanted to be parsimonious and at the same time accurate. The specification we chose was the one which led to the Proportional Hazard assumption not being violated for any of the models. On top of that, we were able to

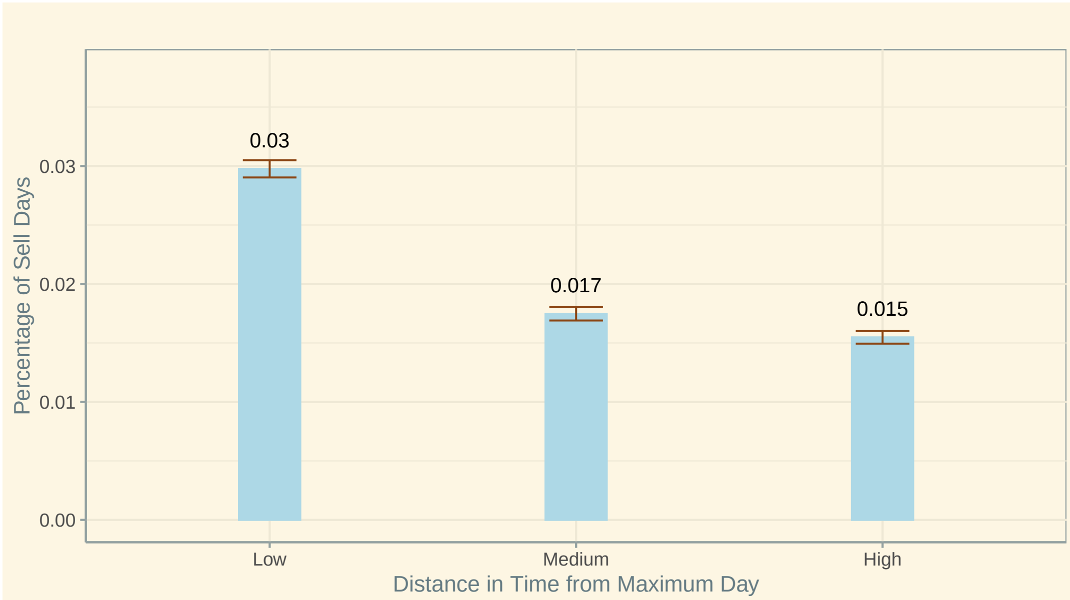

Figure 5: Percentage of selling days out of all stock-portfolio-days per each category (0.95 c.i.). Distance in Time from Max Day is the standardized distance as defined in Definition 2. Low [0; 0.07]; Medium [0.07; 0.34); High [0.34; 1].

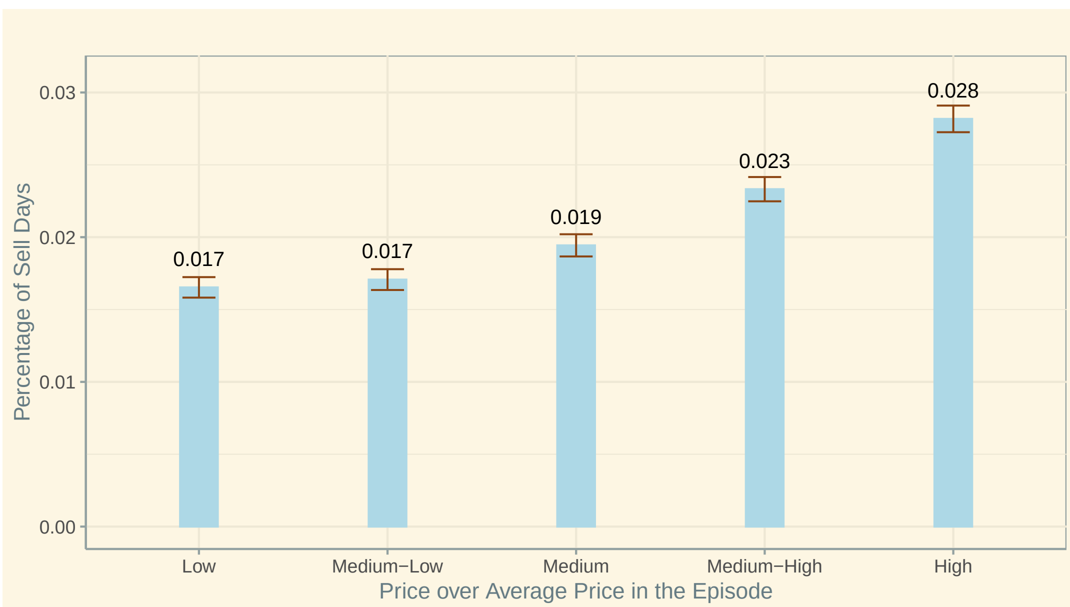

Figure 6: Percentage of selling days out of all stock-portfolio-days per each category (0.95 c.i.).Ratio to Avg Price is the ratio of daily closing price to average daily closing price up to that time in the investment episode. On the selling date it is equal to the ratio of selling price to average closing price in the episode. Low [0.58; 1.01]; Medium-Low (1.01; 1.03]; Medium (1.03; 1.07]; Medium-High (1.07; 1.14]; High (1.14, 3.20].

Table 4: Proportional Hazard model of the hazard of selling a stock. The event is the sale. Each day when the stock is not sold is a non-event. Each observation is at stock-portfolio-day level. Odds Ratios with c.i. Baseline hazard rate stratified at investor level. Clustered robust s.e. at bank account level. Ratio to Max Price is the ratio of daily closing price to maximum price up to that time in the investment episode. On the selling date it is equal to the ratio of selling price to maximum price in the episode. Low [0.349; 0.918]; Medium-Low (0.918; 0.957]; Medium-High (0.957; 0.981]; High (0.981; 1]. Distance in Time from Max Day is the standardized distance as defined in Definition 2. Low [0; 0.07]; Medium [0.07; 0.34); High [0.34; 1]. Ratio to Avg Price is the ratio of daily closing price to average daily closing price up to that time in the investment episode. On the selling date it is equal to the ratio of selling price to average closing price in the episode. Low [0.58; 1.01]; Medium-Low (1.01; 1.03]; Medium (1.03; 1.07]; Medium-High (1.07; 1.14]; High (1.14, 3.20].

	(1)	(2)	(3)
Ratio Price to Max Price (ref. Low)			
Medium-Low	0.909		
	(0.792,1.043)		
Medium-High	1.062		
	(0.932,1.210)		
High	0.720***		
	(0.619,0.837)		
Dist. in Time from Max Day (ref. Low)			
Medium		0.877**	
		(0.786,0.979)	
High		0.430***	
		(0.385,0.481)	
Ratio Price to Avg Price (ref. Low)			
Medium-Low			0.837***
			(0.741,0.946)
Medium			0.914
			(0.802,1.042)
Medium-High			0.914
			(0.801,1.044)
High			1.082
			(0.924,1.267)
Xu-O'Quigley R^2	0.020	0.061	0.013
Concordance	0.57	0.61	0.55
PH Assumption Valid (0.05)	YES	YES	YES
Time Controls	YES	YES	YES
Number of Trading Episodes	13,000	13,000	13,000
Number of Bank Accounts	8,704	8,704	8,704
Observations	621,849	621,849	621,849
Note:			*p<0.1; **p<0.05; ***p<0.01

Figures 4 to 6 show the average percentage of selling days for all the categories of Ratiomax, Distance and RatioAvg. We see that the propensity to sell is highest when Ratiomax is in the medium-high category. That means that investors are more likely to sell at a point when the price is relatively close to running maximum but not when it is very close to it (Figure 4). This is nor confirming the predictions of SV. Distance in time from maximum seems to have a clear effect. The further in time is the maximum, the less likely is the investor to realize a stock. It is possible that investors hope the stock to rebound when the maximum is far in time (Figure 5). Distance in time is not taken into account in SV. Probably, time effects would be difficult to capture in a laboratory setting. We believe that the distance in time from the maximum point has not the same psychological relevance if measured in trading periods in the lab or in days in the field. However, we feel it is worth investigating the effect of time and we could open the way for the development of new theories. We think that regret might arise not only because of price distance but also because of time distance from the past maximum. Figure 6 seems to confirm the finding by SV that investors are more likely to stop, the higher the level of the price process.

In Table 4 we estimate three Proportional Hazard model where we take into account the effect of Ratiomax, Distance and Ratioavg. On top of that we stratify the baseline hazard at bank account level, in order to take into account differences in the propensity to realize a stock due to fixed investors' characteristics. We also control for time effects (month and year). For example, it is well known that the disposition effect is not present in December (Odean, 1998) and taking into account the year we take into account market and economic conditions in that period. The time-span considered was a period during which the stock market experienced good returns in general. After taking into account all these effects, we see that the propensity to sell is lowest when the stock is trading close to the maximum price. The probability of selling at a high ratiomax is 28% lower than the probability of selling at other points. This contradicts the main conclusion of SV. Propensity to sell peaks at a medium-high level of the Ratiomax but differences among low, medium-low and medium-high categories are not significant. We conclude that regret does not bite as we expected from the theoretical discussion in SV. Propensity to sell is highest when price is close but not extremely close to the maximum. The pattern for Distance is quite strong and clear. Propensity to sell is 12.3% lower at a medium with respect to a low Distance and 57% lower at high distance. The effect for Ratioavg is different to what we saw in Figure 6. Propensity to sell is highest when Ratioavg is high but all effects are not significanlty different from zero, apart from the fact that propensity to sell is lowest at medium-low level of Ratioavg.

capture the main non-linear changes in the effect of the variables. In the appendix we repeat the analysis referring to the absolute distance in trading days from the maximum day.

Most of the effect are not singnificant but the propensity to sell is higest for very low or very high prices. This contradicts predictions in SV and hints to the idea that investors are somewhow attracted by extreme good or bad events (Barber and Odean, 2008).

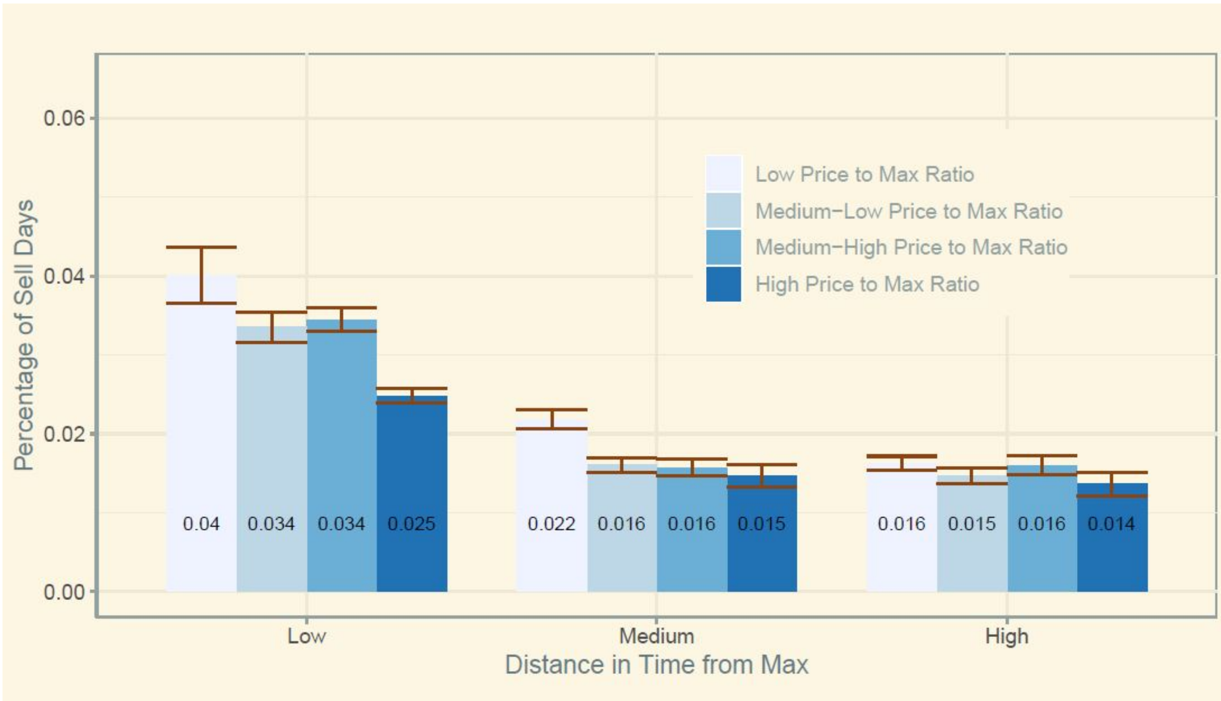

Figure 7: Percentage of selling days out of all stock-portfolio-days per each category (0.95 c.i.). Ratio to Max Price is the ratio of daily closing price to maximum price up to that time in the investment episode. On the selling date it is equal to the ratio of selling price to maximum price in the episode. Low [0.349; 0.918]; Medium-Low (0.918; 0.957]; Medium-High (0.957; 0.981]; High (0.981; 1]. Distance in Time from Max Day is the standardized distance as defined in Definition 2. Low [0; 0.07]; Medium [0.07; 0.34); High [0.34; 1]. Ratio to Avg Price is the ratio of daily closing price to average daily closing price up to that time in the investment episode. On the selling date it is equal to the ratio of selling price to average closing price in the episode. Low [0.58; 1.01]; Medium-Low (1.01; 1.03]; Medium (1.03; 1.07]; Medium-High (1.07; 1.14]; High (1.14, 3.20].

Since the effect of Distance seems to be the most relevant in terms of both strength and explanatory power, we are interested in the interaction of it with the effect of Ratiomax. Here, we leave the test of the hypotheses outlined in SV and we move to a closer investigation of the combined effect of time and price. In the appendix we report the distribution distribution of the combination of the two categories. At least 2% of stock-portfolio-day observations fall in each category. From Figure 7 we can see that the interaction between Ratiomax and Distance leads to an interesting scenario. When the Distance is low or medium, the average percentage of selling days is highest when Ratiomax is low. There is probably a panic effect, which can be framed as regret but is much deeper than what we described before, in the discussion of SV theory. Investors are more willing to realize a gain when it is closest in time to maximum but furthest in price. If regret is having a role, it is through this channel. Investors regret selling at a time far from maximum. However, they sell as soon as the price decreases significantly. We can see this by looking at the fact that the percentage of sale days is 2.5% at high Ratiomax and low Distance and it is 4% at low price and low distance. Hence, it looks like anticipated regret of incurring higher losses is higher than experienced regret given by the distance in price from past maximum. This is in line with the evidence of anticipated regret present in Fioretti et al. (2018).

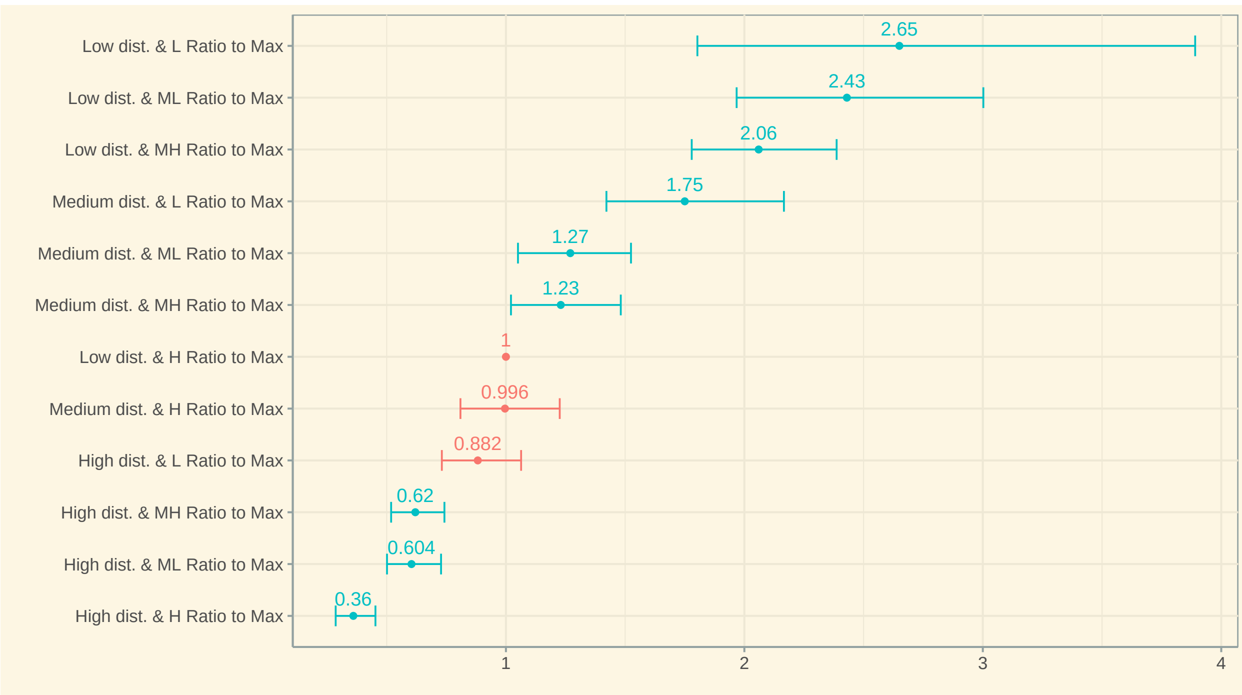

Figure 8: Odds Ratio from a Proportional Hazard model of the hazard of selling a stock. Each Odds Ratio corresponds to the effect of being in a given category. The event is the sale. Each day when the stock is not sold is a non-event. Each observation is at stock-portfolio-day level. Clustered robust s.e. at bank account level. Blue odds ratios are significantly different from 1 ($p < 0.05$), red are not. Ratio to Max Price is the ratio of daily closing price to maximum price up to that time in the investment episode. On the selling date it is equal to the ratio of selling price to maximum price in the episode. Low [0.349; 0.918]; Medium-Low (0.918; 0.957]; Medium-High (0.957; 0.981]; High (0.981; 1]. Distance in Time from Max Day is the standardized distance as defined in Definition 2. Low [0; 0.07]; Medium [0.07; 0.34); High [0.34; 1]. The baseline category is "Low Distance and High Ratio".

We fit a Proportional Hazard model where we take into account all the 12 combinations of Ratiomax and Distance categories. We stratify the baseline hazard at bank account level, in order to take into account differences in the propensity to realize a stock due to fixed investors' characteristics. We also control for time effects (month and year). By having 12 categories where both Ratiomax and Distance can change, we address the problem of collinearity we would have by fitting a model where we include Ratiomax and Distance as two distinct covariates. All categories in our regression are disjoint. The regression table is reported in the appendix. In Figure 8 we report the odds ratio of the 11 coefficients. The baseline category corresponds to "Low Distance and High Ratiomax", the ideal point to sell a stock, from an accounting perspective. We can see that all cases where distance is low are clustered at the top, when we rank categories based on the propensity to sell for each of them. All cases where distance is high are clustered at the bottom. The big exception is the baseline category, low Distance and high Ratiomax. Figure 8 offers a more complete interpretation to the sample averages we reported in Figure 7. When we take into account investors' fixed effects and time, the hazard of selling at a high price at a low distance from maximum day is even lower than expected. Hazard of selling is 2.65 higher for low distance and low Ratiomax stock-days than at the baseline. In general, when the stock is close in time but not close in price (i.e. low distance but not high Ratiomax) the hazard of selling is always estimated to be at least double than that at the baseline. The complete picture seems to suggest that reality is more complicated than how the lab describes it. Distance in price from the maximum plays a role but the effect is not nice and linear as the one observed by SV. When distance in price is considered in isolation, it looks like propensity to sell peaks at a point which is close but not the closest possible to maximum. However, it is really intriguing looking at the combined effect of time distance from maximum day and price distance from maximum day. We fell we can safely claim that traders are more willing to sell at a low distance in time from maximum. They probably wait for a new maximum when a lot of time has passed since the last one. Hence, it is always a salient figure in their mind. However, when the distance in time is short, they are more willing to sell stocks which are further from past maximum. We believe panic might play a big role. It looks like investors are not extremely good at catching the best time to realize a stock and decide to realize it only when the price path shows a defined descending trend. Predictions of regret theory in dynamic context by SV are only partially confirmed then.

5 Conclusions

Investigating investors behaviour in the light of decision making models is a fascinating challenge (Barberis and Xiong, 2009; Kaustia, 2010; Hens and Vlcek, 2011; Henderson, 2012; Strack and Viefers, 2013). Fitting a comprehensive model able to describe all aspects of financial decisions would be the ultimate goal of this stream of research. However, our understanding of financial decisions proceeds more as a constructive process. We contributed to the explanations of trading decisions from a very specific angle.

We focused on selling activities, in particular in the gain domain. We moved from theoretical considerations of what could explain selling decisions of investors. The idea that investors would behave as perfectly rational and optimizing decision makers, with no time inconsistency in their preferences, is fascinating both for scholars studying rational decision making and scholars studying optimal stopping behaviour from a more mathematical point of view.

However, based on empirical and experimental evidence gathered in the last 20 years, the idea of describing individual investors as rational decision makers was disputed (Barber and Odean, 2013). In our work we tested the idea that investors stop at an optimal ex ante threshold and are prone to regret in their decision to sell stocks. We moved from theoretical and experimental evidence found by Strack and Viefers (2013) and tested their predictions on field data.

Loosely speaking, a threshold strategy implies that an investor sells a stock as soon as her optimal goal in terms of returns is reached. We rejected that hypothesis for our sample of investors. We further investigated investors differences in the propensity to adopt a threshold strategy. First, we saw that investors are more willing to adopt a threshold strategy in the gain with respect to the loss domain. That is a sort of a corollary of the higher propensity to realize gains with respect to losses, formalized as disposition effect (Shefrin and Statman, 1985; Odean, 1998; Barber and Odean, 2013). Then, we found that the main differences in the rate of threshold consistency are due to age, client segment and account type. Sophisticated investors and active traders are more consistent with a threshold strategy. Affluent and older investors are less consistent than general investors with a threshold strategy. Males are more willing to realize losses at a threshold than females. Sophisticated investors seem to be less prone to investment biases (Grinblatt and Keloharju, 2001; Dhar and Zhu, 2006). Males seem to be more willing to accept losses at a minimum, that is partially at odds with Barber and Odean (2001), who find that males under-perform females because of overconfidence. Affluent and older investors are more likely to depart from a rational threshold strategy, this is linked to the idea that decision making of older individuals is poorer (Korniotis and Kumar, 2011; Bruine De Bruin, 2017).

We investigated the impact of running maximum price in an investment episode on the propensity of investors to realize gains. We fitted a Proportional Hazard model to the propensity to sell for a gain. Predictions of regret theory in a dynamic context (Strack and Viefers, 2013) are that propensity to sell increases with the level of the price and decreases with the distance of the price from past maximum. The first prediction was confirmed, but the evidence is not strong. Investors tend to realize more stocks, the higher is the price over average price in the investment. However, not all effects were significant, when we take into account investors and time fixed effects. The effect of distance in price from past maximum was not the same found in Strack and Viefers (2013). We did find that the propensity to sell is actually lower when the stock is trading very close to maximum price, while it is highest when the price is close to past maximum but not in the closest region. The relation between propensity to sell and distance in price from past maximum follows an inverse U-shape.

We did not only restrict to test ideas exposed in Strack and Viefers (2013). We investigated the impact that distance in time from past maximum has on the propensity to sell a stock. We find a very strong effect, with the propensity to sell a stock falling as the distance in time from past maximum increases. On top of that, we investigated the joint effect of distance in time and distance in price from past maximum, finding that investors are more willing to realize stocks which are closer in time but further in price from past maximum. When time distance from past maximum is short, the predictions of regret reverse. Investors are more willing to realize stocks, the further is the price from maximum. It looks like two forces are at work. Investors are willing to wait for a new maximum to occur if a long time has passed since the last one and they panic when the stock price drops a short time after maximum.

These findings open up the way to further experimental and theoretical work in this area. First, we believe that it would interesting isolating the effects of time and price in a controlled laboratory setting and second, we think that it would be worth incorporating the time dimension in theoretical dynamic stopping models.

References

Barber, B. M., Lee, Y.-T., Liu, Y.-J., & Odean, T. (2007). Is the aggregate investor reluctant to realise losses? evidence from taiwan. *European Financial Management*, *13*(3), 423–447.

Barber, B. M., & Odean, T. (2000). Trading is hazardous to your wealth: The common stock investment performance of individual investors. *The Journal of Finance*, *55*(2), 773–806.

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

Barber, B. M., & Odean, T. (2008). All that glitters: The effect of attention on the buying behavior of individual and institutional investors. *The Review of Financial Studies*, *21*(2), 785–818.

Barber, B. M., & Odean, T. (2013). The behavior of individual investors. In G. M. Constantinides, M. Harris, & R. M. Stulz (Eds.), (Chap. 22, Vol. 2, pp. 1533–1570). Handbook of the Economics of Finance. Elsevier.

Barberis, N., Huang, M., & Thaler, R. H. (2006). Individual preferences, monetary gambles, and stock market participation: A case for narrow framing. *American Economic Review*, *96*(4), 1069–1090.

Barberis, N., & Xiong, W. (2009). What drives the disposition effect? an analysis of a long-standing preference-based explanation. *The Journal of Finance*, *64*(2), 751–784.

Baucells, M., Weber, M., & Welfens, F. (2011). Reference-point formation and updating. *Management Science*, *57*(3), 506–519.

Bell, D. E. (1985). Disappointment in decision making under uncertainty. *Operations Research*, *33*(1), 1–27.

Ben-David, I., & Hirshleifer, D. (2012). Are investors really reluctant to realize their losses? trading responses to past returns and the disposition effect. *The Review of Financial Studies*, *25*(8), 2485–2532.

Benartzi, S., & Thaler, R. H. (1995). Myopic loss aversion and the equity premium puzzle. *The Quarterly Journal of Economics*, *110*(1), 73–92.

Bleichrodt, H., & Wakker, P. P. (2015). Regret theory: A bold alternative to the alternatives. *The Economic Journal*, *125*(583), 493–532.

Bordalo, P., Gennaioli, N., & Shleifer, A. (2012). Salience theory of choice under risk. *The Quarterly Journal of Economics*, *127*(3), 1243–1285.

Braun, M., & Muermann, A. (2004). The impact of regret on the demand for insurance. *Journal of Risk & Insurance*, *71*(4), 737–767.

Brettschneider, J., & Burgess, M. (2017). *Using a frailty model to measure the effect of covariates on the disposition effect* (tech. rep. No. 5). CRISM Working Paper Series.

Brown, P., Chappel, N., Da Silva Rosa, R., & Walter, T. (2006). The reach of the disposition effect: Large sample evidence across investor classes. *International Review of Finance*, *6*(1-2), 43–78.

Bruine De Bruin, W. (2017). Ageing and economic decision-making. In B. Ranyard (Ed.), *Economic psychology* (Chap. 23, pp. 371–386). John Wiley & Sons, Ltd.

Coval, J. D., Hirshleifer, D. A., & Shumway, T. (2005). *Can individual investors beat the market?* HBS Finance Working Paper.

Cox, D. R. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society. Series B (Methodological)*, *34*(2), 187–220.

Delquié, P., & Cillo, A. (2006). Disappointment without prior expectation: A unifying perspective on decision under risk. *Journal of Risk and Uncertainty*, *33*(3), 197–215.

Descamps, A., Massoni, S., & Page, L. (2016). Available at SSRN: https://ssrn.com/abstract=2883847.

Deville, L., & Riva, F. (2007). Liquidity and arbitrage in options markets: A survival analysis approach. *Review of Finance*, *11*(3), 497–525.

Dhar, R., & Zhu, N. (2006). Up close and personal: Investor sophistication and the disposition effect. *Management Science*, *52*(5), 726–740.

Feng, H., & Hobson, D. (2016). Gambling in contests with regret. *Mathematical Finance*, *26*(3), 674–695.

Fioretti, M., Vostroknutov, A., & Coricelli, G. (2018). *Dynamic regret avoidance*. USC-INET Research Paper.

Firth, D. (2003). Overcoming the reference category problem in the presentation of statistical models. *Sociological Methodology*, *33*(1), 1–18.

Firth, D., & De Menezes, R. X. (2004). Quasi-variances. *Biometrika*, *91*(1), 65–80.

Gepp, A., & Kumar, K. (2015). Predicting financial distress: A comparison of survival analysis and decision tree techniques. *Procedia Computer Science*, *54*, 396–404.

Gollier, C., & Salanié, B. (2006). *Individual decisions under risk, risk sharing and asset prices with regret*. Working paper, University of Toulouse.

Grambsch, P. M., & Therneau, T. M. (1994). Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika*, *81*(3), 515–526.

Grinblatt, M., & Keloharju, M. (2001). What makes investors trade? *The Journal of Finance*, *56*(2), 589–616.

Heath, C., Huddart, S., & Lang, M. (1999). Psychological factors and stock option exercise. *The Quarterly Journal of Economics*, *114*(2), 601–627.

Henderson, V. (2012). Prospect theory, liquidation, and the disposition effect. *Management Science*, *58*(2), 445–460.

Hens, T., & Vlcek, M. (2011). Does prospect theory explain the disposition effect? *Journal of Behavioral Finance*, *12*(3), 141–157.

Hirshleifer, D. A., Myers, J. N., Myers, L. A., & Teoh, S. H. (2008). Do individual investors cause post-earnings announcement drift? direct evidence from personal trades. *The Accounting Review*, *83*(6), 1521–1550.

Huang, X. (2019). Mark twain's cat: Investment experience, categorical thinking, and stock selection. *Journal of Financial Economics*, *131*(2), 404–432.

Huddart, S., Lang, M., & Yetman, M. H. (2009). Volume and price patterns around a stock's 52-week highs and lows: Theory and evidence. *Management Science*, *55*(1), 16–31.

Ivković, Z., Poterba, J., & Weisbenner, S. (2005). Tax-motivated trading by individual investors. *The American Economic Review*, *95*(5), 1605–1630.

Ivković, Z., Sialm, C., & Weisbenner, S. (2008). Portfolio concentration and the performance of individual investors. *Journal of Financial and Quantitative Analysis*, *43*(3), 613–655.

Ivković, Z., & Weisbenner, S. (2005). Local does as local is: Information content of the geography of individual investors' common stock investments. *The Journal of Finance*, *60*(1), 267–306.

Jia, J., Dyer, J. S., & Butler, J. C. (2001). Generalized disappointment models. *Journal of Risk and Uncertainty*, *22*(1), 59–78.

Jiao, P. (2015). *Losing from naive reinforcement learning: A survival analysis of individual repurchase decisions*. Available at SSRN: https://ssrn.com/abstract=2574141.

Kaustia, M. (2010). Prospect theory and the disposition effect. *Journal of Financial and Quantitative Analysis*, *45*(3), 791–812.

Kent, J. T., & O'Quigley, J. (1988). Measures of dependence for censored survival data. *Biometrika*, *75*(3), 525–534.

Korniotis, G. M., & Kumar, A. (2011). Do older investors make better investment decisions? *The Review of Economics and Statistics*, *93*(1), 244–265.

Korniotis, G. M., & Kumar, A. (2013). Do portfolio distortions reflect superior information or psychological biases? *Journal of Financial and Quantitative Analysis*, *48*(1), 1–45.

Kőszegi, B., & Rabin, M. (2006). A model of reference-dependent preferences. *The Quarterly Journal of Economics*, *121*(4), 1133–1165.

Kőszegi, B., & Rabin, M. (2007). Reference-dependent risk attitudes. *American Economic Review*, *97*(4), 1047–1073.

Loomes, G., & Sugden, R. (1982). Regret theory: An alternative theory of rational choice under uncertainty. *The Economic Journal*, *92*(368), 805–824.

Loomes, G., & Sugden, R. (1986). Disappointment and dynamic consistency in choice under uncertainty. *The Review of Economic Studies*, *53*(2), 271–282.

Michenaud, S., & Solnik, B. (2008). Applying regret theory to investment choices: Currency hedging decisions. *Journal of International Money and Finance*, *27*(5), 677–694.

Muermann, A., Mitchell, O., & Volkman, J. (2006). Regret, portfolio choice, and guarantees in defined contribution schemes. *Insurance: Mathematics and Economics*, *39*(2), 219–229.

Muermann, A., & Volkman, J. M. (2006). *Regret, pride, and the disposition effect*. PARC Working Paper Series.

O'Quigley, J., Xu, R., & Stare, J. (2005). Explained randomness in proportional hazards models. *Statistics in Medicine*, *24*(3), 479–489.

Odean, T. (1998). Are investors reluctant to realize their losses? *The Journal of Finance*, *53*(5), 1775–1798.

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

Panayi, E., & Peters, G. (2014). Survival models for the duration of bid-ask spread deviations. In *2014 ieee conference on computational intelligence for financial engineering economics (cifer)* (pp. 9–16).

Richards, D. W., Rutterford, J., Kodwani, D., & Fenton-O'Creevy, M. (2017). Stock market investors' use of stop losses and the disposition effect. *The European Journal of Finance*, *23*(2), 130–152.

Richards, D. W., & Willows, G. D. (2018). Who trades profusely? the characteristics of individual investors who trade frequently. *Global Finance Journal*, *35*, 1–11.

Seasholes, M. S., & Zhu, N. (2010). Individual investors and local bias. *The Journal of Finance*, *65*(5), 1987–2010.

Shapira, Z., & Venezia, I. (2001). Patterns of behavior of professionally managed and independent investors. *Journal of Banking & Finance*, *25*(8), 1573–1587.

Shefrin, H., & Statman, M. (1985). The disposition to sell winners too early and ride losers too long: Theory and evidence. *The Journal of Finance*, *40*(3), 777–790.

Strack, P., & Viefers, P. (2013). *Too proud to stop: Regret in dynamic decisions*. Available at SSRN: https://ssrn.com/abstract=2465840.

Xu, R., & O'Quigley, J. (1999). A r2 type measure of dependence for proportional hazards models. *Journal of Nonparametric Statistics*, *12*(1), 83–107.

Appendix A. Methodological notes

Negative Binomial Regression

The Poisson distribution may be generalized by including a gamma noise variable which has a mean of 1 and a scale parameter of ν. The Poisson-gamma mixture (negative binomial) distribution that results is

$$P(Y = y_i|\mu_i, \alpha) = \frac{\Gamma(y_i + \alpha^{-1})}{\Gamma(y_i + 1)\Gamma(\alpha^{-1})} \left(\frac{\alpha^{-1}}{\alpha^{-1} + \mu_i}\right)^{\alpha^{-1}} \left(\frac{\alpha^{-1}}{\alpha^{-1} + \mu_i}\right)^{y_i} \tag{2}$$

where $\mu_i = t_i * \mu$ and $\alpha = \nu^{-1}$. The parameter μ is the mean incidence rate of y per unit of exposure. Exposure may be time, space, distance, area, volume, or population size. Because exposure is often a period of time, we use the symbol t_i to represent the exposure for a particular observation. When no exposure is given, it is assumed to be one. The parameter μ may be interpreted as the risk of a new occurrence of the event during a specified exposure period, t.

In our case the exposure time t_i is the number of episodes in a bank account and μ_i is the number of threshold episodes in the same bank account. In negative binomial regression, the mean of y is determined by the exposure time t and a set of k regressor variables (the *x's*). The expression relating these quantities is

$$\mu_i = \exp(\log(t_i) + \beta_1 x_{1i} + ... + \beta_k x_{ki}) \tag{3}$$

often $x_1 = \mathbf{1}$, in which case β_1 is called the intercept. We estimate the vector of β coefficients through maximum likelihood.

Pseudo R squared for Cox models

In this section, we would like to spell out some details of the pseudo R squared proposed by Xu and O'Quigley (1999), which we report for our proportional hazard models. They start from a coefficient of explained randomness derived by Kent and O'Quigley (1988). The coefficient aims at explaining the variability on the outcome looking at the distribution of time to events, given covariates. That coefficient has the following properties and others which are not really relevant to our discussion but help realizing how accountable that measure is:

- When a covariate is unrelated to survival, and the corresponding regression coefficient it is equal to zero, it is equal to zero;

- When the effect of at least a coefficient is different from 0, it is between 0 and 1;
- It is invariant under linear transformations of covariates and under monotone increasing transformations of time.

The coefficient we use, used the same basic ideas but looking at the distribution of covariates at each time. The construction could be carried out using routine quantities calculated during a standard proportional hazards analysis. Inference was also greatly simplified. Most importantly, the presence of time-dependent covariates presents no difficulties for Xu and O'Quigley (1999) coefficient estimation, while the one by Kent and O'Quigley (1988) is not defined in that case. In O'Quigley, Xu, and Stare (2005) there are further discussions on the robustness of the pseudo R-squared we used.

Appendix B. Supplementary Analysis

Here, we include some extra analysis to complement the analysis in the main text. In Table 5 we present the same model we presented in Table 3, taking into account knowledge and experience of investors. As you can see, those figures are available for a relatively small subset of individuals. The sole novelty is represented by the effect of knowledge. Investors with extensive knowledge of investments have a rate of threshold consistency which is 12.7% lower than others, in the gain domain. That might be an effect of overconfidence. On top of that, we present a logit regression where the dependent variable takes the value of 1 if the investor followed a threshold strategy for at least one investment in her portfolio (at least one gain, at least one loss or at least one). Results are presented in Tables 6 and 7 and they confirm the analysis we made with the negative binomial. Effects are obviously stronger, since we are considering in the same group all investors who followed a threshold strategy at least once.

Figures 4 to 6 show the survival curves for Ratiomax, Distance and Ratioavg respectively. In our application, survival probability corresponds to the probability of holding the stock. We see that the effect of Distance is much more neat than the effect of Ratiomax and Ratioavg, as we saw in our analysis in the main text. As we mentioned in the text, we report the joint distribution of Ratiomax and Distance (Table 8) and the regression table which corresponds to Figure 8 (Table 9). Moreover, we reproduce Figure 8 estimating confidence intervals using the method of Quasi-variance developed by Firth (2003), Firth and De Menezes (2004). In this way, we overcome the problem of having a reference category for which confidence intervals cannot be estimated and we enhance the possibility of reproducing our work. In Figures 13 and 14 and tables 10 to 12 we reproduce the analysis of the main text, using absolute distance in trading days from maximum instead of the standardized distance. We see that the Proportional Hazard assumption is not valid for models in Tables 11 and 12. That means that the effect we report is not constant in time. Hence, the hazard ratio we report is only an average effect. That is to be expected since we believe that being 1 or 2 days far from maximum is not the same when an investor is considering a 7 days or a 50 day investment, intuitively. This back up the idea that a standardized distance, like the one we defined in Definition 2 is approriate. We again see that, on average, the propensity to sell is highest for small distance in time from maximum day and high distance in price from the maximum.

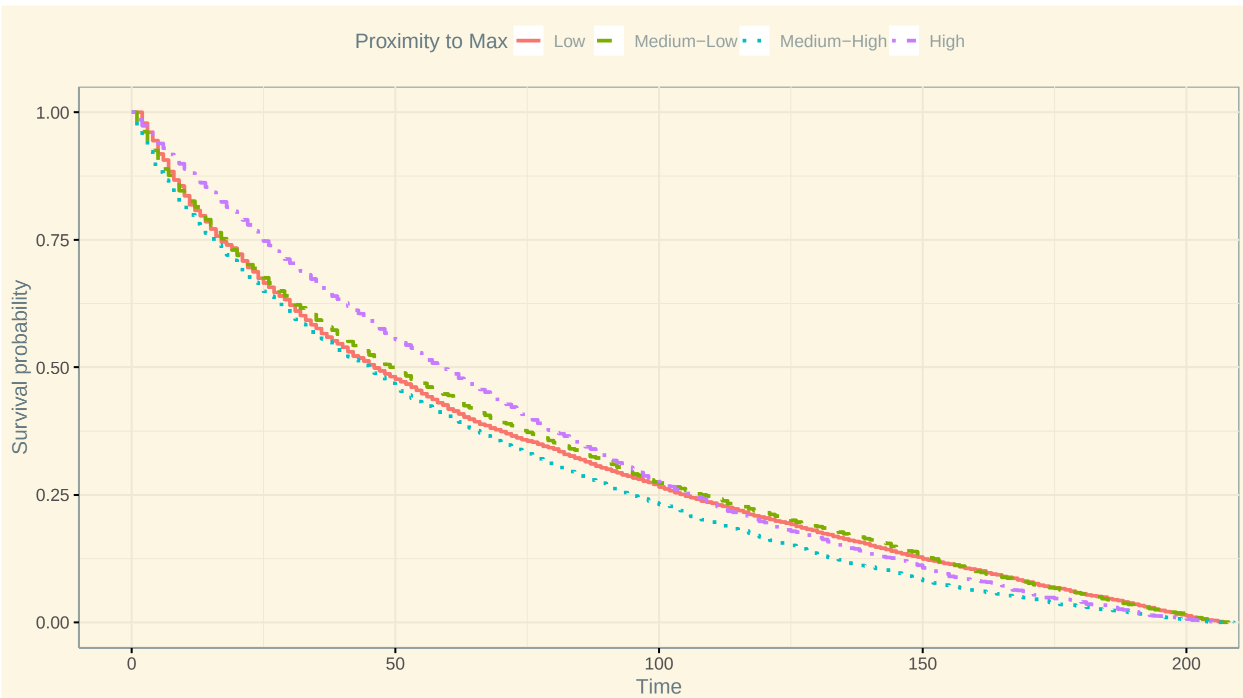

Figure 9: Survival curves for the four categories of Ratio to Max Price. Ratio to Max Price is the ratio of daily closing price to maximum price up to that time in the investment episode. On the selling date it is equal to the ratio of selling price to maximum price in the episode. Low [0.349; 0.918]; Medium-Low (0.918; 0.957]; Medium-High (0.957; 0.981]; High (0.981; 1]. A lower probability of survival implies an higher propensity to sell the stock.

Table 5: Odds ratios with 95% c.i. of a Negative Binomial regression where each observation corresponds to a bank account. The dependent variable is the number of times the investor stopped at a threshold in the gain, loss or overall. There is an offset equal to the number of episodes in the bank account (in the gain, loss, overall).

	Gain		Loss		All	
	(1)	(2)	(3)	(4)	(5)	(6)
Account Type (ref. Cash)						
Account Type IRA	1.032	1.029	1.081	1.081	1.076	1.068
	(0.904,1.178)	(0.905,1.172)	(0.910,1.286)	(0.914,1.281)	(0.949,1.220)	(0.946,1.207)
Account Type Keogh	1.182	1.212	1.261	1.219	1.255	1.266*
	(0.841,1.654)	(0.872,1.679)	(0.850,1.849)	(0.822,1.783)	(0.946,1.663)	(0.961,1.664)
Account Type Margin	1.219***	1.244***	1.140	1.182*	1.272***	1.292***
	(1.067,1.394)	(1.092,1.418)	(0.956,1.362)	(0.996,1.405)	(1.123,1.442)	(1.146,1.459)
Account Type Schwab	1.162**	1.154**	1.143*	1.160**	1.221***	1.216***
	(1.038,1.304)	(1.032,1.292)	(0.987,1.328)	(1.005,1.343)	(1.097,1.360)	(1.097,1.351)
Client Segment (ref. General)						
Client Segment Affluent	1.004	0.984	0.964	0.970	0.944	0.929
	(0.906,1.111)	(0.889,1.087)	(0.834,1.111)	(0.841,1.115)	(0.850,1.047)	(0.838,1.028)
Client Segment Active	1.081**	1.082**	1.080*	1.097**	1.115***	1.110***
	(1.008,1.158)	(1.012,1.158)	(0.987,1.182)	(1.005,1.197)	(1.049,1.186)	(1.045,1.178)
Age (decades)	0.907***	0.909***	0.931***	0.943***	0.910***	0.917***
	(0.883,0.931)	(0.886,0.932)	(0.900,0.962)	(0.914,0.973)	(0.889,0.931)	(0.897,0.938)
Income	0.994	0.992	0.976**	0.975**	0.989	0.987*
	(0.977,1.011)	(0.976,1.008)	(0.955,0.998)	(0.955,0.996)	(0.975,1.004)	(0.973,1.002)
Male	0.975	0.968	1.083	1.129	0.975	0.989
	(0.858,1.109)	(0.854,1.099)	(0.912,1.292)	(0.954,1.344)	(0.866,1.098)	(0.882,1.111)
Occupation (ref. Other (also NA)						
Non Professional Occupation	1.044	1.051	0.995	1.042	1.010	1.035
	(0.909,1.196)	(0.919,1.198)	(0.822,1.196)	(0.870,1.241)	(0.887,1.147)	(0.915,1.169)
Professional Occupation	1.010	1.002	0.916*	0.927*	0.975	0.973
	(0.945,1.079)	(0.939,1.070)	(0.839,0.999)	(0.852,1.010)	(0.919,1.034)	(0.918,1.031)
Experience (ref. Good)						
Experience Extensive	0.926*		1.041		0.979	
	(0.853,1.005)		(0.940,1.151)		(0.913,1.049)	
Experience Low	1.008		1.044		1.033	
	(0.935,1.086)		(0.944,1.154)		(0.964,1.107)	
Experience None	1.066		0.835		0.973	
	(0.881,1.285)		(0.628,1.095)		(0.815,1.157)	
Knowledge (ref. Good)						
Knowledge Extensive		0.873***		1.055		0.964
		(0.798,0.955)		(0.947,1.175)		(0.895,1.039)
Knowledge Low		1.006		1.012		1.028
		(0.934,1.084)		(0.916,1.118)		(0.961,1.100)
Knowledge None		1.039		1.114		1.079
		(0.932,1.158)		(0.967,1.281)		(0.980,1.188)
McFadden Adj. R^2	0.77	0.76	0.76	0.75	0.76	0.77
Observations	4,713	4,911	3,531	3,663	2,701	2,809

Note: *p<0.1; **p<0.05; ***p<0.01

Table 6: Odds ratios with 95% c.i. of a logit regression where each observation corresponds to a bank account. The dependent variable is 1 if the investor stopped at a threshold at least once in the gain, loss or overall.

	Gain		Loss		All	
	(1)	(2)	(3)	(4)	(5)	(6)
Account type (ref. Cash)						
Account Type IRA	1.109*	1.130*	1.048	1.041	1.161*	1.238**
	(0.994,1.237)	(0.997,1.282)	(0.909,1.208)	(0.884,1.225)	(0.987,1.366)	(1.027,1.494)
Account Type Keogh	1.174	1.321	1.731**	1.444	1.413	1.380
	(0.763,1.807)	(0.749,2.345)	(1.079,2.775)	(0.812,2.551)	(0.788,2.585)	(0.657,3.010)
Account Type Margin	1.423***	1.374***	1.475***	1.467***	1.531***	1.450***
	(1.271,1.594)	(1.204,1.569)	(1.287,1.690)	(1.251,1.721)	(1.310,1.790)	(1.205,1.746)
Account Type Schwab	1.453***	1.417***	1.369***	1.329***	1.540***	1.526***
	(1.325,1.593)	(1.272,1.579)	(1.223,1.535)	(1.164,1.519)	(1.352,1.755)	(1.307,1.782)
Client Segment (ref. General)						
Client Segment Affluent	0.790***	0.843***	0.758***	0.772***	0.708***	0.743***
	(0.718,0.869)	(0.753,0.944)	(0.673,0.854)	(0.670,0.888)	(0.618,0.810)	(0.631,0.873)
Client Segment Active	2.455***	2.737***	2.166***	2.122***	2.405***	2.690***
	(2.253,2.676)	(2.468,3.037)	(1.978,2.372)	(1.904,2.366)	(2.168,2.671)	(2.371,3.055)
Age (decades)		0.903***		0.996		0.881***
		(0.876,0.931)		(0.962,1.032)		(0.846,0.918)
Income		0.973***		0.965***		0.970**
		(0.954,0.993)		(0.942,0.988)		(0.944,0.997)
Male		1.122		1.115		1.135
		(0.976,1.290)		(0.942,1.324)		(0.930,1.384)
Occupation (ref. Other (also NA)						
Non Professional Occupation		0.982		0.909		0.964
		(0.846,1.140)		(0.759,1.087)		(0.784,1.187)
Professional Occupation		0.944		0.943		0.941
		(0.869,1.025)		(0.854,1.040)		(0.839,1.054)
McFadden Adj. R^2	0.03	0.29	0.029	0.29	0.038	0.31
Observations	15,624	11,477	11,390	8,315	8,674	6,280

Note: *p<0.1; **p<0.05; ***p<0.01

Table 7: Odds ratios with 95% c.i. of a logit regression where each observation corresponds to a bank account. The dependent variable is 1 if the investor stopped at a threshold at least once in the gain, loss or overall.

	Gain		Loss		All	
	(1)	(2)	(3)	(4)	(5)	(6)
Account Type (ref. Cash)						
Account Type IRA	1.173	1.170	1.030	1.054	0.999	0.981
	(0.949,1.451)	(0.951,1.440)	(0.790,1.344)	(0.814,1.368)	(0.735,1.359)	(0.728,1.322)
Account Type Keogh	1.710	1.652	1.684	1.566	1.669	1.818
	(0.814,3.692)	(0.804,3.475)	(0.791,3.599)	(0.745,3.295)	(0.619,4.999)	(0.686,5.387)
Account Type Margin	1.323**	1.401***	1.189	1.214	1.161	1.260
	(1.047,1.673)	(1.115,1.763)	(0.897,1.578)	(0.922,1.602)	(0.840,1.607)	(0.919,1.729)
Account Type Schwab	1.576***	1.580***	1.230*	1.267**	1.461***	1.494***
	(1.307,1.902)	(1.316,1.899)	(0.981,1.546)	(1.016,1.585)	(1.121,1.904)	(1.156,1.934)
Client Segment (ref. General)						
Client Segment Affluent	0.894	0.902	0.731***	0.745***	0.784**	0.799*
	(0.762,1.048)	(0.771,1.054)	(0.598,0.892)	(0.612,0.905)	(0.622,0.986)	(0.638,1.000)
Client Segment Active	2.496***	2.573***	1.991***	2.032***	2.564***	2.604***
	(2.154,2.894)	(2.228,2.974)	(1.700,2.334)	(1.741,2.374)	(2.135,3.086)	(2.176,3.122)
Age (decades)	0.872***	0.875***	0.984	0.991	0.826***	0.839***
	(0.829,0.916)	(0.833,0.918)	(0.930,1.041)	(0.937,1.047)	(0.773,0.882)	(0.787,0.895)
Income	0.977	0.975	0.965*	0.971	0.978	0.979
	(0.945,1.009)	(0.944,1.006)	(0.928,1.002)	(0.936,1.008)	(0.935,1.023)	(0.937,1.023)
Male	1.021	1.017	1.153	1.157	0.886	0.880
	(0.807,1.291)	(0.807,1.282)	(0.876,1.526)	(0.882,1.525)	(0.634,1.233)	(0.631,1.221)
Occupation (ref. Other (also NA)						
Non Professional Occupation	0.831	0.866	0.926	0.987	0.878	0.957
	(0.651,1.060)	(0.682,1.100)	(0.687,1.244)	(0.737,1.317)	(0.621,1.243)	(0.682,1.348)
Professional Occupation	0.850**	0.849***	0.919	0.924	0.839**	0.836**
	(0.749,0.965)	(0.749,0.961)	(0.792,1.066)	(0.798,1.069)	(0.706,0.996)	(0.706,0.990)
Experience (ref. Good)						
Experience Extensive	0.911		1.189*		0.945	
	(0.771,1.076)		(0.987,1.431)		(0.765,1.168)	
Experience Low	0.916		0.995		0.900	
	(0.798,1.052)		(0.843,1.173)		(0.742,1.092)	
Experience None	1.137		0.674*		1.310	
	(0.801,1.616)		(0.432,1.032)		(0.802,2.176)	
Knowledge (ref. Good)						
Knowledge Extensive		0.837*		1.124		0.887
		(0.700,1.001)		(0.921,1.371)		(0.706,1.116)
Knowledge Low		0.895		0.923		0.885
		(0.780,1.027)		(0.783,1.087)		(0.732,1.071)
Knowledge None		0.950		0.901		0.918
		(0.774,1.167)		(0.706,1.147)		(0.697,1.212)
McFadden Adj. R^2	0.03	0.29	0.029	0.29	0.038	0.31
Observations	4,713	4,911	3,531	3,663	2,701	2,809

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

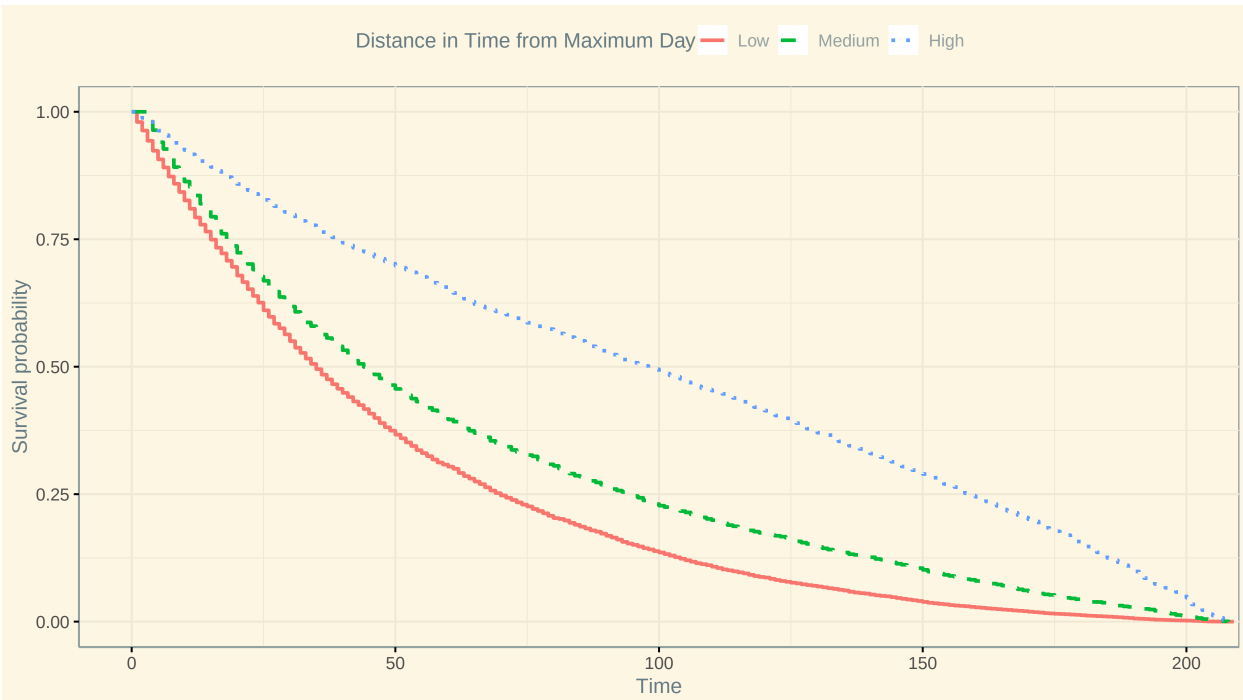

Figure 10: Survival curves for the three categories of Distance in Time from Max Day. Distance in Time from Max Day is the standardized distance as defined in Definition 2. Low [0; 0.07]; Medium [0.07; 0.34); High [0.34; 1]. A lower probability of survival implies an higher propensity to sell the stock.

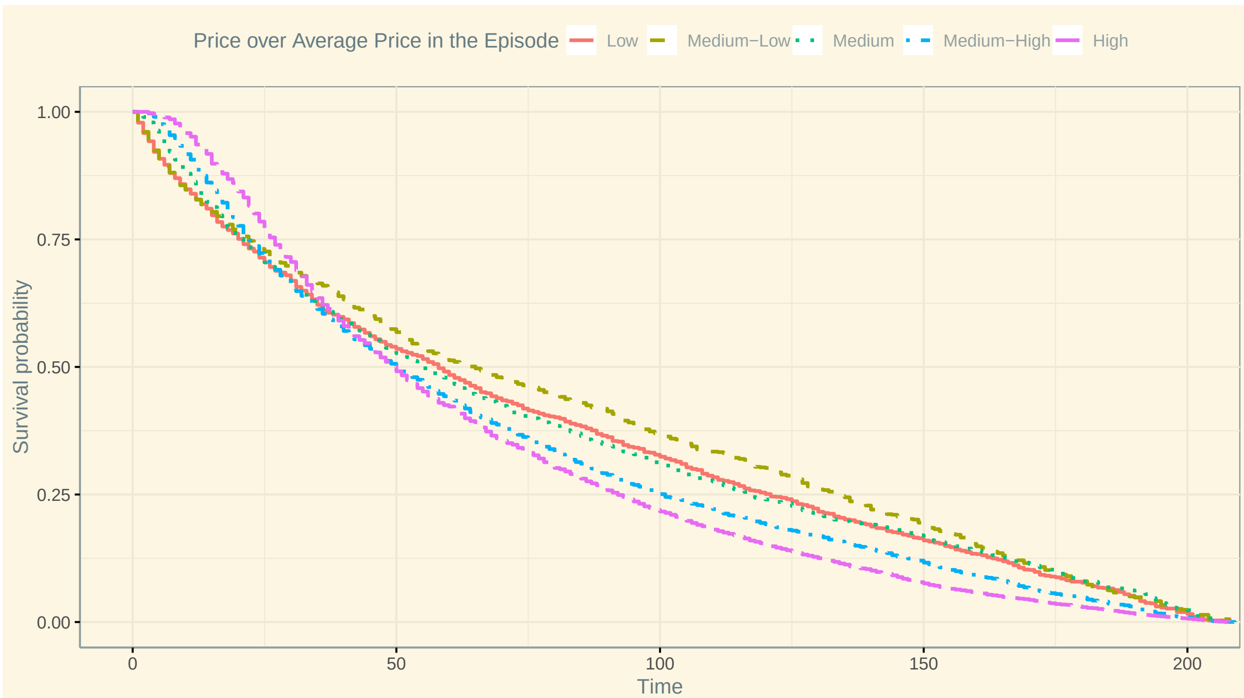

Figure 11: Survival curves for the five categories of Ratio to Avg Price. Ratio to Avg Price is the ratio of daily closing price to average daily closing price up to that time in the investment episode. On the selling date it is equal to the ratio of selling price to average closing price in the episode. Low [0.58; 1.01]; Medium-Low (1.01; 1.03]; Medium (1.03; 1.07]; Medium-High (1.07; 1.14]; High (1.14, 3.20]. A lower probability of survival implies an higher propensity to sell the stock.

Table 8: Percentage of stock-portfolio-days falling in each category. Ratio to Max Price is the ratio of daily closing price to maximum price up to that time in the investment episode. On the selling date it is equal to the ratio of selling price to maximum price in the episode. Low [0.349; 0.918]; Medium-Low (0.918; 0.957]; Medium-High (0.957; 0.981]; High (0.981; 1]. Distance from Max Day is the standardized time distance as defined in Definition 2. Low [0; 0.07]; Medium [0.07; 0.34); High [0.34; 1].

	Low Distance from Max	Medium Distance from Max	High Distance from Max
Low Price Ratio to Max	0.02	0.10	0.13
Medium-Low Price Ratio to Max	0.05	0.10	0.10
Medium-High Price Ratio to Max	0.09	0.09	0.07
High Price Ratio to Max	0.17	0.04	0.04

Table 9: Proportional Hazard model of the hazard of selling a stock. The event is the sale. Each day when the stock is not sold is a non-event. Each observation is at stock-portfolio-day level. Odds Ratios with c.i. Baseline hazard rate stratified at investor level. Clustered robust s.e. at bank account level. Ratio to Max Price is the ratio of daily closing price to maximum price up to that time in the investment episode. On the selling date it is equal to the ratio of selling price to maximum price in the episode. Low [0.349; 0.918]; Medium-Low (0.918; 0.957]; Medium-High (0.957; 0.981]; High (0.981; 1]. Distance in Time from Max Day is the standardized distance as defined in Definition 2. Low [0; 0.07]; Medium [0.07; 0.34); High [0.34; 1].

Dist. in time from Max and Ratio to Max (ref. Low and High)	
Low dist. and Low Ratio to Max	2.649***
	(1.803,3.891)
Medium dist. and Low Ratio to Max	1.755***
	(1.421,2.166)
High dist. and Low Ratio to Max	0.882
	(0.731,1.064)
Low dist. and Medium-Low Ratio to Max	2.430***
	(1.968,3.002)
Medium dist. and Medium-Low Ratio to Max	1.266**
	(1.051,1.524)
High dist. and Medium-Low Ratio to Max	0.604***
	(0.501,0.728)
Low dist. and Medium-High Ratio to Max	2.061***
	(1.779,2.387)
Medium dist. and Medium-High Ratio to Max	1.230**
	(1.021,1.482)
High dist. and Medium-High Ratio to Max	0.620***
	(0.519,0.742)
Medium dist. and High Ratio to Max	0.996
	(0.809,1.226)
High dist. and High Ratio to Max	0.360***
	(0.286,0.453)
Xu-O'Quigley R^2	0.095
Concordance	0.65
PH Assumption Valid (0.05)	YES
Time Controls	YES
Number of Trading Episodes	13,000
Number of Bank Accounts	8,704
Observations	621,849
Note:	*p<0.1; **p<0.05; ***p<0.01

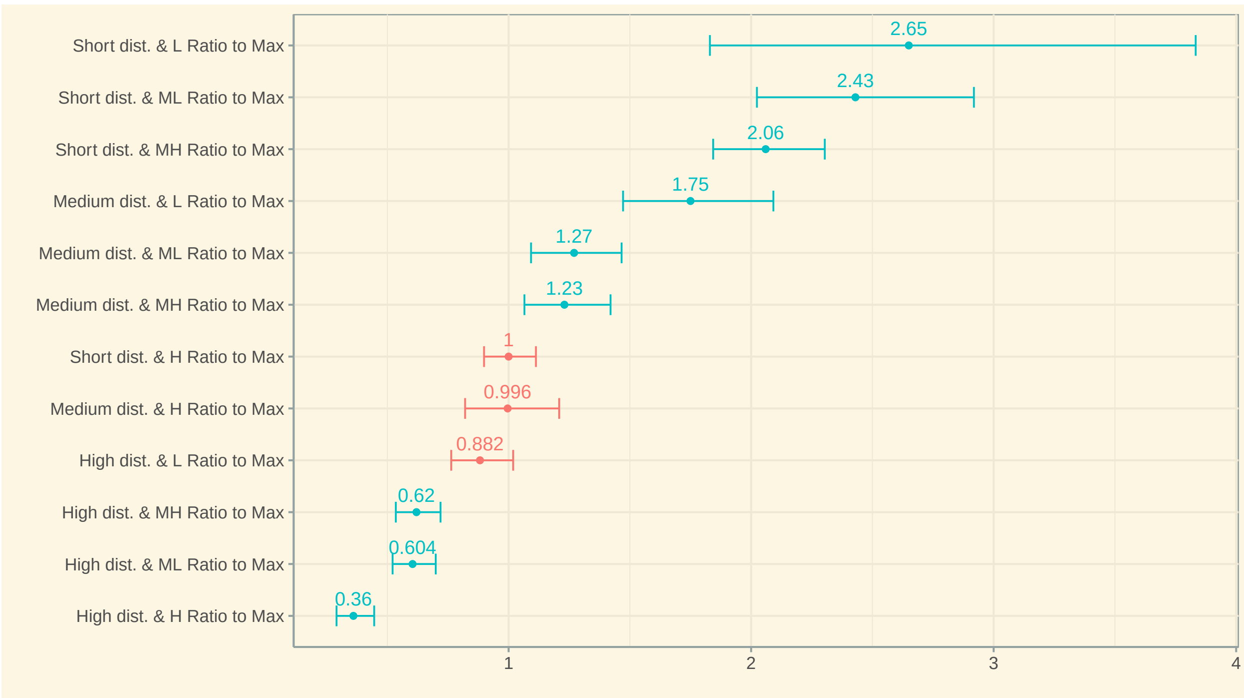

Figure 12: Odds Ratio from a Proportional Hazard model of the hazard of selling a stock. Each Odds Ratio corresponds to the effect of being in a given category. The event is the sale. Each day when the stock is not sold is a non-event. Each observation is at stock-portfolio-day level. Clustered robust s.e. at bank account level. Blue odds ratios are significantly different from 1 ($p < 0.05$) when relying upon exact standard errors, red are not. Confidence intervals were obtained with the Quasi-Variance method of Firth (2003), Firth and De Menezes (2004). Ratio to Max Price is the ratio of daily closing price to maximum price up to that time in the investment episode. On the selling date it is equal to the ratio of selling price to maximum price in the episode. Low [0.349; 0.918]; Medium-Low (0.918; 0.957]; Medium-High (0.957; 0.981]; High (0.981; 1]. Distance in Time from Max Day is the standardized distance as defined in Definition 2. Low [0; 0.07]; Medium [0.07; 0.34); High [0.34; 1]. The baseline category is "Low Distance and High Ratio".

Table 10: Percentage of stock-portfolio-days falling in each category. Ratio to Max Price is the ratio of daily closing price to maximum price up to that time in the investment episode. On the selling date it is equal to the ratio of selling price to maximum price in the episode. Low [0.349; 0.957]; Medium (0.957; 0.981]; High (0.981; 1]. Distance in Time from Max Day is measured in trading days.

	Max Day	1 Day After Max	2 Days After Max	3 to 5 Days After Max	5+ Days After Max
Low Price Ratio to Max	0.01	0.02	0.02	0.06	0.39
Medium Price Ratio to Max	0.04	0.03	0.02	0.05	0.11
High Price Ratio to Max	0.13	0.03	0.02	0.03	0.05

Table 11: Proportional Hazard model of the hazard of selling a stock. The event is the sale. Each day when the stock is not sold is a non-event. Each observation is at stock-portfolio-day level. Odds Ratios with c.i. Baseline hazard rate stratified at investor level. Clustered robust s.e. at bank account level. Distance in Time from Max Day is measured in trading days.

Dist. from Maximum Day (ref. Max Day)	
1 Day	1.093
	(0.943,1.266)
2 Days	0.929
	(0.783,1.101)
3 to 5 Days	0.742***
	(0.640,0.860)
More than 5 Days	0.413***
	(0.363,0.470)
Xu-O'Quigley R^2	0.061
Concordance	0.61
PH Assumption Valid (0.01)	NO
Time Controls	YES
Number of Trading Episodes	13,000
Number of Bank Accounts	8,704
Observations	621,849
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 12: Proportional Hazard model of the hazard of selling a stock. The event is the sale. Each day when the stock is not sold is a non-event. Each observation is at stock-portfolio-day level. Odds Ratios with c.i. Baseline hazard rate stratified at investor level. Clustered robust s.e. at bank account level. Ratio to Max Price is the ratio of daily closing price to maximum price up to that time in the investment episode. On the selling date it is equal to the ratio of selling price to maximum price in the episode. Low [0.349; 0.957]; Medium (0.957; 0.981]; High (0.981; 1]. Distance in Time from Max Day is measured in trading days.

Dist. in from Max and Ratio to Max (ref. Max Day and High)	
Max Day and Low Ratio to Max	3.051***
	(2.283,4.077)
1 Day After Max and Low Ratio to Max	1.702***
	(1.238,2.339)
2 Days After Max and Low Ratio to Max	1.791***
	(1.381,2.322)
3 to 5 Days After Max Low Ratio to Max	1.437***
	(1.166,1.771)
5+ Days After Max and Low Ratio to Max	0.706***
	(0.602,0.829)
Max Day and Medium Ratio to Max	2.529***
	(2.141,2.989)
1 Day After Max and Medium Ratio to Max	1.709***
	(1.362,2.143)
2 Days After Max and Medium Ratio to Max	1.339**
	(1.018,1.762)
3 to 5 Days After Max Medium Ratio to Max	0.908
	(0.720,1.146)
5+ Days After Max and Medium Ratio to Max	0.535***
	(0.442,0.647)
1 Day After Max and High Ratio to Max	1.348***
	(1.084,1.675)
2 Days After Max and High Ratio to Max	0.762
	(0.544,1.067)
3 to 5 Days After Max High Ratio to Max	0.741**
	(0.555,0.989)
5+ Days After Max and High Ratio to Max	0.366***
	(0.287,0.467)
Xu-O'Quigley R^2	0.1
Concordance	0.65
PH Assumption Valid (0.05)	YES
Time Controls	YES
Number of Trading Episodes	13,000
Number of Bank Accounts	8,704
Observations	621,849
Note:	*p<0.1; **p<0.05; ***p<0.01

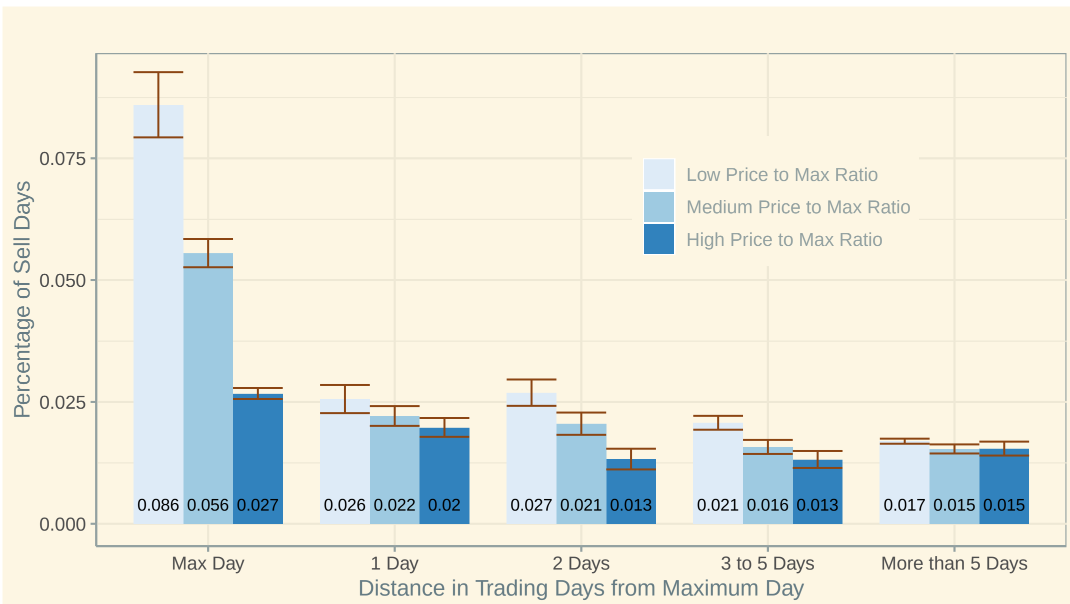

Figure 13: Percentage of selling days out of all stock-portfolio-days per each category (0.95 c.i.). Ratio to Max Price is the ratio of daily closing price to maximum price up to that time in the investment episode. On the selling date it is equal to the ratio of selling price to maximum price in the episode. Low [0.349; 0.957]; Medium (0.957; 0.981]; High (0.981; 1]. Distance in Time from Max Day is measured in trading days.

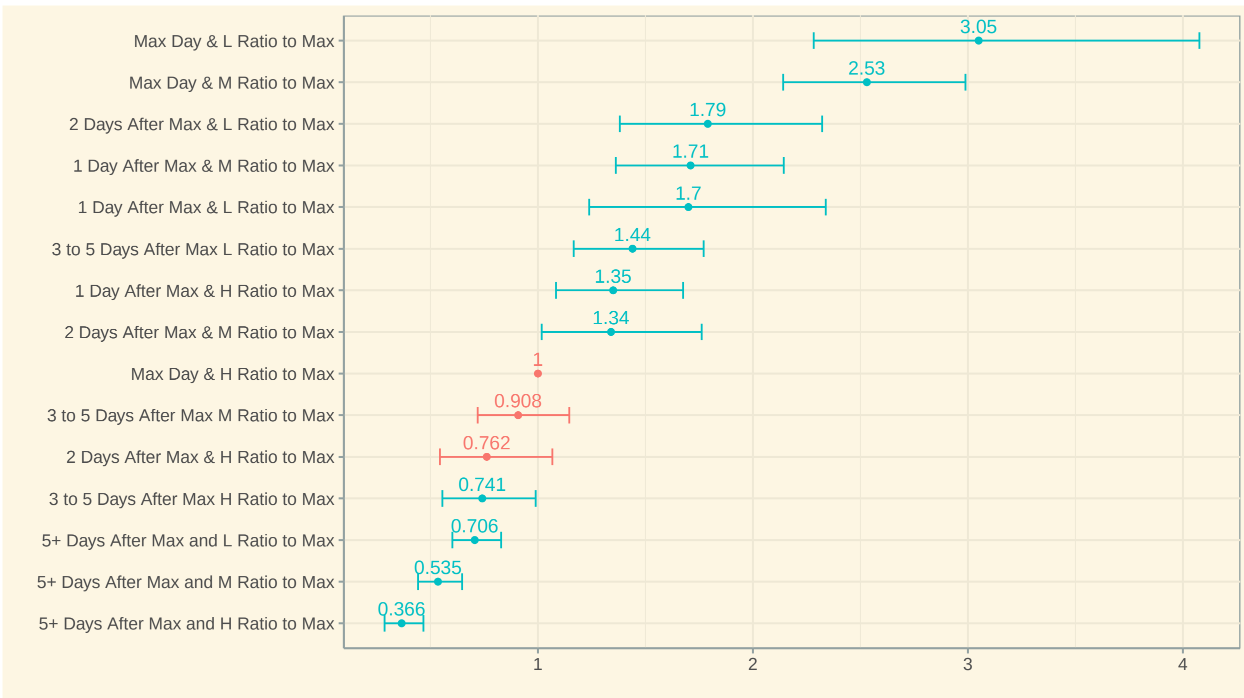

Figure 14: Odds Ratio from a Proportional Hazard model of the hazard of selling a stock. Each Odds Ratio corresponds to the effect of being in a given category. The event is the sale. Each day when the stock is not sold is a non-event. Each observation is at stock-portfolio-day level. Clustered robust s.e. at bank account level. Blue odds ratios are significantly different from 1 ($p < 0.05$), red are not. Ratio to Max Price is the ratio of daily closing price to maximum price up to that time in the investment episode. On the selling date it is equal to the ratio of selling price to maximum price in the episode. Low [0.349; 0.957]; Medium (0.957; 0.981]; High (0.981; 1]. Distance in Time from Max Day is measured in trading days. The baseline category is "Max Day and High Ratio".