



DEPARTMENT OF ECONOMICS
DISCUSSION PAPER SERIES

When is Market the Benchmark?
Reinforcement Evidence from Repurchase Decisions

Peiran Jiao and Heinrich H. Nax

Number 781
February 2016

Manor Road Building, Oxford OX1 3UQ

When is Market the Benchmark?

Reinforcement Evidence from Repurchase Decisions *

Peiran Jiao¹ and Heinrich H. Nax²

¹Department of Economics, University of Oxford

²Computational Social Sciences, ETH Zürich

Abstract

Reinforcement relative to an adaptive benchmark is a well-established model of behavior outside finance. Recently, reinforcement has been identified as an important driver of decisions to repurchase a stock. In this paper, we enrich the existing reinforcement model of repurchasing by an aspiration-based market benchmark. When choosing which stock to repurchase, investors' sources of reinforcement are weighted averages of absolute returns from previous sales and relative returns with respect to a market benchmark. The weights change according to market environments. We empirically identify the following crucial asymmetry that cannot be reconciled by simple reinforcement strategy, but is consistent with the model we propose: investors place more weight on relative returns when the market is performing well, and place more weight on absolute returns when the market is performing badly.

Keywords: Reinforcement, Stock Repurchasing, Aspiration Adjustment

JEL Classification Numbers: G02, G11, D01

*Peiran Jiao, Department of Economics, University of Oxford, Manor Road, OX1 3UQ Oxford, UK; Email: peiran.jiao@economics.ox.ac.uk. Heinrich H. Nax, Computational Social Sciences, ETH Zürich, Clausiusstrasse 37, 8092 Zurich, Switzerland; Email: hnax@ethz.ch.

1 Motivation

Financial economists have long been studying the factors that make individual investors buy and sell stocks. Several studies have addressed this question empirically. A difficult for empirical investigations of buying decisions stems from the fact that investors' choice sets, when buying, potentially include the universe of all securities. This paper contributes to the understanding of buying decisions by proposing a reinforcement learning model in the context of a special type of buying decision, repurchasing, for which the choice set can be unambiguously determined.¹ The novelty of our model, compared to prior reinforcement learning studies in finance, is that decision-makers do not solely care about the absolute return they received from prior investments, but also take the market into account (e.g. the S&P500 index) as a benchmark, allowing previously sold stocks to be absolute and/or relative winners or losers. Consistent with extant studies of investor psychology, we propose that investors pay more attention to the market performance, and thus the relative return to their investment, during an upward-trending market, but not during a downward-trending market. The main contribution of our study is twofold. First, we provide behavioral foundations based on a formal learning model for these reinforcement patterns. Second, we identify the varying source of reinforcement, absolute and relative stock performances, depending on market conditions, and find empirical evidence that cannot be generated by individual investors' simple reinforcement strategy, but is consistent with our adaptive-benchmark reinforcement model.

Reinforcement learning is a concept intensively studied in psychology. 'Classical reinforcement theory' (Bush and Mosteller, 1953; Suppes and Atkinson, 1959; Harley, 1981; Cross, 1983) assumes that, in accordance with the 'law of effect' (Thorndike, 1898), the probability of choosing an action is higher (lower) if that action, absolutely or compared with other actions or an adaptive aspiration benchmark, led to higher (lower) payoffs in the decision maker's past experience. Economists use reinforcement learning to explain behavior

¹We assume that short-selling is constrained, which is reasonable for individual investors.

in strategic interactions (see e.g. Roth and Erev, 1995; Erev and Rapoport, 1998). Beyond classical reinforcement learning, a wide array of generalized reinforcement learning models (e.g. Nevo and Erev, 2012; Camerer and Ho, 1999; Nax et al., 2013b) exist with various additional features that were derived from a rich experimental literature.² An important avenue of investigation has been to determine the nature of aspiration benchmarks according to which actions are reinforced, and how they are adapted.³

Recently, several empirical studies have investigated reinforcement-type behavior in financial markets. They find that average individual investors are more likely to subscribe to IPO auctions subsequent to successful personal experience, leading to deteriorating performances (Kaustia and Knüpfer, 2008). Sophisticated investors do not behave like this (Chiang et al., 2011). Moreover, investors are shown to allocate more to their 401(k) accounts after profitable experience in the past month (Choi et al., 2009). Closely related to our study, it was found that common stock investors repurchase more previously sold winners than previously sold losers (Strahilevitz et al., 2011). In an experiment, Jiao (2015a) uncovers that higher past reinforcements lead to increased propensities to repurchase and more optimistic beliefs.

What is missing in the literature on reinforcement investing is a model that clearly specifies the sources of the reinforcement benchmarks. The above empirical studies test joint hypotheses regarding both the source of reinforcement and the subsequent reinforcement-driven behavior; and they implicitly assume the reinforcement source to be the absolute return on previous sales, which, we believe, is not necessarily all that individual investors care about. Here, we provide a model that enriches existing studies by the introduction of a dynamic market benchmark. Stock performances are assessed vis-a-vis this benchmark, and decisions to repurchase a previously sold stock are made based on a weighted combination of (stock-) absolute and (market-) relative returns. Whether the decision to purchase a given stock is reinforced, therefore, depends on the price dynamics of (i) the stock it self, (ii)

²See Erev and Haruvy (2013) for an excellent review.

³See Bendor et al. (2001) for a review, and Selten (1998) for a review of aspiration adaptation theory.

the market, and (iii) the stock relative to the market. The model closest to ours is due to Dixon (2000) who introduce the concept of a dynamic market benchmark in the context of a repeated Cournot oligopoly.⁴ Basically, in our model, investors attach more weight to the relative performance when the market is going up, and care more about the stocks' absolute performance and less about the market when the market is going down.

We think that this extension of reinforcement theory, when applied to investment decisions, relies on important insights from behavioral economics and investor psychology. Conventional economic theories assume that an agent derives utility from final levels of wealth or consumption, and beliefs only enter into this framework as the weights on utilities from different states when the agent contemplates subjective expected utilities. However, people may also derive utility from information and beliefs (see e.g. Yariv, 2001). Particularly relevant is the finding of the 'ostrich effect' (Karlsson et al., 2009), according to which people are more willing to actively collect and process information under good prior news, but tend to ignore information following bad prior news. This is in accordance with the proposal of Galai and Sade (2006) that people may avoid "apparently risky financial situations by pretending they do not exist".⁵ The reason for behaving in this way is that people derive some sort of utility from holding an optimistic expectation, but face a tradeoff between benefits from biased belief and bad decisions (see e.g. Akerlof and Dickens, 1982; Brunnermeier and Parker, 2005). In our case, when the market is performing badly, investors may be less likely to search and process information about the market performance, which makes their belief-updating less frequent and the use of relative return as reinforcement sources less likely. Conversely, when the market is performing well, investors pay more attention to the market, placing more weight on relative returns, and potentially update their reference points (Kőszegi and Rabin, 2006) or aspiration levels (Bendor et al., 2001) more frequently according to the market benchmark. This type of behavior is also consistent with 'confir-

⁴See also Tietz and Weber (1972); Weber (2003) for related experimental evidence, and Nax et al. (2013a); Nax (2015); Nax and Pradelski (2015) for related aspiration adjustment models.

⁵Similarly, Brock and Balloun (1967) find that smokers attend to pro-smoking information while non-smokers do the opposite.

mation bias theory’ (Rabin and Schrag, 1999), according to which people hold optimistic prior belief about the market and misinterpret disconfirming evidence (in the spirit of various other theories including cognitive dissonance (Akerlof and Dickens, 1982), positive self-image (Eil and Rao, 2011), and wishful thinking (Mayraz, 2011)).

In this present study, we find that this proposition is not only plausible, but also empirically valid. We use data of individual trading records from a large discount brokerage firm covering trades in the years 1991 to 1996 (the same data as used in Strahilevitz et al. (2011)). We conduct survival analysis on all stocks sold to model their subsequent repurchases. Strahilevitz et al. (2011) find that investors repurchase more previously sold winners than losers, but we find that this is not the whole story. We computed the market returns during the holding periods before previous sales. By separating stocks sold during up/down markets, we identify that stocks with positive absolute returns lower than the market return were less likely to be repurchased than stocks that had negative return in up markets, a result that cannot be generated by a model with only absolute returns as the sources of reinforcement. By contrast, introducing relative returns in down markets did not create much difference. This pattern lends support to our model.

Another strand of related literature is the empirical study of investor buying and selling decisions. There is an agreement over the kind of stocks investors sell, consistent with the ‘disposition effect’ (Shefrin and Statman, 1985), i.e. they sell stocks that had good recent performance and hold those that had bad recent performance (see e.g. Odean, 1998), and this bias persists as investors become more experienced/‘sophisticated’ (Feng and Seasholes, 2005). In terms of buying, Odean (1999) finds that investors purchase stocks that have abnormally large positive or negative returns, possibly because these stocks attract their attention. Barber and Odean (2008) also find individual investors prefer to buy attention-grabbing stocks. Additionally, most studies agree that individual investors are net buyers of stocks that had bad recent performance, using a contrarian strategy (e.g. Grinblatt and Keloharju, 2001; Kaniel et al., 2008; Goetzmann and Massa, 2002) as if they believe in

mean reversion.⁶ A couple of things are noteworthy. Firstly, the adoption of reinforcement strategies as we describe here can reduce the cognitive load of investors with limited attention, because previously sold stocks are a natural subset of stocks that attract attention; and possibly due to cognitive dissonance (or ego utility), it is cognitively more justifiable to continue monitoring previously sold winners. On the other hand, reinforcement-based strategies differ from studies of simple contrarian (or ‘momentum’) strategies, because they require prior ownership and experienced gain/loss. But in order to account for the tendency to use contrarian or momentum strategies, our analysis controls for the performance of stocks subsequent to prior sales. On average, there should be no reason for investors to believe in continuation or mean reversion on stocks from which they gained or lost.

The rest of this note is structured as follows. Next, we introduce the formal model and formulate our hypotheses. In section three, we describe our data. Section four contains the results. Section five concludes.

2 The model

Suppose some investor i , in some period t , decides which stock S to purchase from some set of stocks $\mathcal{S}$ that he monitors, including potential repurchase of stocks he previously owned. Write $s_i^t = 1$ for the decision to purchase, and $s_i^t = 0$ for the decision not to purchase stock S . Based on Erev and Roth (1998), we define q_{iS}^t to be i ’s *propensity* to purchase any given stock $S \in \mathcal{S}$ at time t . This propensity is updated such that $q_{iS}^{t+1} = q_{iS}^t + R_{iS}^{t+1}$, where R_{iS}^{t+1} is the current *reinforcement* i received on S .⁷ Propensities translates into a probabilistic choice rule via $p_{iS}^{t+1} = e^{\lambda q_{iS}^{t+1}} / \sum_{R \in \mathcal{S}} e^{\lambda q_{iR}^{t+1}}$.⁸

The case considered in Erev and Roth (1998) is one of repeatedly playing the same

⁶Strahilevitz et al. (2011) find that investors are more likely to repurchase stocks that declined in value after previous sale. This can be a manifestation of a contrarian strategy as previously found in the literature.

⁷To make the reinforcement pattern realistic, we can easily include parameters for forgetting and experimentation in this formula.

⁸The linear probability rule is problematic when payoffs can be negative. Therefore, we use logistic probability rule to deal with that. See Camerer and Ho (1999) for more explanations.

underlying game, in which case it is convenient to express a reinforcement received by agent i from choosing strategy S in period t by the simple equation $R_{iS}^{t+1} = \phi_S^t - \underline{\phi}_{iS}$, i.e. as the difference between the realized payoff of playing S minus $\underline{\phi}_{iS}$, the smallest possible payoff i considers for S . So R_{iS}^{t+1} represents the reinforcement on not choosing strategy S . Here, the smallest possible payoff occurs in some way when the stock loses all its value. However, in our setting, it is more reasonable to think of $\underline{\phi}_{iS}$ as a ‘reference point’. In a market setting, reasonable reference points are not complete loss of the value of a stock, but rather either the status quo, i.e. a return of 0 as in Strahilevitz et al. (2011), or the market return.⁹ Denote by ϕ_{iS}^t , in our setup, the return to stock S for investor i in period t , i.e. as $\phi_{iS}^t = (\pi_{iS}^t(sell) - \pi_{iS}^t(buy)) / (\pi_{iS}^t(buy))$ where π is the price, and by $\bar{\phi}^t$ the *market performance* over the same period. Denote by Δ_S^t the difference between stock and market performances, i.e. the relative performance of stock S , $\Delta_S^t = \phi_S^t - \bar{\phi}^t$. Since our investors operate in a market setting, there are two ways to adapt this formulation:

Isolated view. $R_{iS}^{t+1} = -R_{iS}^{t+1} = \phi_S^t$; the agent takes in no information about the market performance.

Market perspective. $R_{iS}^{t+1} = -R_{iS}^{t+1} = \Delta_S^t$; the agent compares the stock performance to the market performance.

2.1 Hypotheses

Qualitatively, there are six different market situations depending on the direction of performances of (i) the market, (ii) the stock, and (iii) the stock relative to the market. If the market went up, own stock performance can be either negative or positive. When positive, it can be under- or overperforming relative to the market. Similarly, if the market went down, stock performance can be negative or positive. When negative, it can be under- or overperforming relative to the market.

⁹These are two natural reference points. Other potential benchmarks may exist, such as the investor’s previous performance, or return to any reference stock the investor chooses, but we do not consider these here.

Our hypotheses are summarized by the following model. We assume that every investor uses an admixture of the two views so that, for some variable weight $\varrho \in (0, 1)$, $R_{iS}^{t+1} = \varrho(\bar{\phi}^t)\Delta_S^t + (1 - \varrho(\bar{\phi}^t))\phi_S^t$. The precise weight $\varrho(\bar{\phi}^t)$ depends on and increases with the market performance such that $\varrho' > 0$.

Experience from previous performances matters. Importantly, the role of the market in influencing individual's assessment of their past performances, especially in terms of how much attention is paid to the market benchmark, is asymmetric as regards increasing versus decreasing markets. Our model presumes that investors dynamically adjust their benchmark and source of reinforcement according to how they wishfully allocate attention during different market conditions. Put bluntly, in increasing markets, investors appear to believe in the market and in market momentum, therefore reinforcing only those stocks that outperformed the market. By contrast, in decreasing markets, investors appear to be skeptical of the market overall and choose to focus on their own portfolio. Loosely in the spirit of Erev et al. (2008), we therefore ascribe behaviors with a flavor of loss-aversion (Tversky and Kahneman, 1991) to dynamic reinforcement adjustments. Future laboratory experiments are in planning that would allow to elicit these issues precisely.

3 The data

We rely on the well-known dataset first used in Barber and Odean (2000), and readers should consult their paper for details.¹⁰ The data consists of detailed trading records and position statements of 78,000 actively trading households in a discount brokerage firm in the US, together with details of the relevant demographic information for each account holder. Trading data comprises activity from January 1991 through December 1996. Recorded transactions include trading of common stocks, mutual funds, ADRs and some other financial assets, whereas we focus on common stock trading. Among all households, 66,465 of them had common stock positions within the sample period for at least one month, while on average

¹⁰See also Strahilevitz et al. (2011) and Jiao (2015b) for the use of the data to test repurchasing decisions.

60 percent of the market value of their accounts were invested in common stocks; and common stock trading accounted for slightly more than 60 percent of approximately 3 million trades. The data contain the amount and prices bought and sold together with commissions for each transaction. The position statements contain the end-of-month statements of stocks in each portfolio and their equity values. Other relevant information, such as market price, return and volume, is taken from the Center for Research on Securities Prices (CRSP) database.

For tax and other reasons, the average household in the dataset held roughly two accounts. In our analysis, we treat each household as an individual, which means all accounts that belong to one household are put together when counting repurchases.

4 Results

Strahilevitz et al. (2011) have used the same data as we use here to calculate the aggregate proportions of previous winners repurchased (PWR), and of previous losers repurchased (PLR). They show that the difference between these proportions is statistically significant, for both taxable and tax-exempt accounts. Then they used survival analysis (the Cox proportional hazard model) to perform a detailed comparison between previous winners and losers. Their results can be conveniently summarized in Figure 1. In the figure, the horizontal axis represents the return on previous sales. The red and blue curves are for stocks whose price increased and decreased respectively after sale. The (roughly) positively sloped curves demonstrate the naïve reinforcement learning bias in repurchases. However, their test implicitly assumes that investors use absolute returns on previous sales as the source of reinforcement. We argue that they did not uncover the true effects of prior experience, in that existence of the bias can depend on market performances, and thus relative return matters sometimes depending on market environment.

[Insert Figure 1 approximately here.]

In order to derive our empirical results, we adopt a similar methodology as in Strahilevitz

et al. (2011) with several important modifications. For each sale made before January 1, 1996, we generate a binary variable that is equal to 1 if the stock was subsequently repurchased within a year, and 0 otherwise. We focus on repurchases within a year because we want to allow all sales to have equal-sized opportunity windows for subsequent repurchases. Survival analysis explicitly models time to an event. With each sale as an observation, the event of interest is the repurchase, and the return from previous sale is the main explanatory covariate. Instead of using the nonparametric survival analysis as in Strahilevitz et al. (2011), we use the hazard model with Weibull error distribution function to capture the monotonic relationship between the probability of repurchase and time. The binary dependent variable of repurchasing is regressed on the baseline hazard rate and the matrix of other covariates, X , whereas the Weibull hazard and survival functions can be conveniently expressed as follows,

$$h(t, p, X) = pt^{p-1}e^{X\beta}, \quad (1)$$

$$S(t, p, X) = e^{-e^{X\beta}t^p}, \quad (2)$$

where t is time, and p the parameter to be estimated that accounts for the shape of the baseline hazard function. The baseline hazard rate is $h_0(t) = pt^{p-1}$. The results below also contain an intercept β_0 , which serves as a scaling factor for the baseline hazard function $h_0(t)\exp(\beta_0)$. To justify the use of this parametric form, Jiao (2015b) graphically shows that the number of repurchases is decreasing in time, which corresponds to a consistently significant estimate of $0 < p < 1$.

Our survival analysis includes the following control variables that might also influence the probability of repurchase. The inverse holding period, which accounts for the number of days a stock was held before it was previously sold. The log of relative trading volume, which is the volume of all trades of a stock measured in dollars in the previous year divided by the total volume of all trades of all stocks. Also included was a set of return measures for different time horizons (5, 10, 21, 42, 63, 126 trading days) after the stock was sold.

Strahilevitz et al. (2011) find that investors in this dataset were more likely to repurchase a stock if it decreased in value after being sold. So these return measures were included to capture that. Their coefficients were all significant, but will not be reported.

We first introduce market performances into the survival analysis; results are reported in Table 1. The main covariates are those that account for previous sales returns. We generate a pair of binary variables, *Pos* and *Neg*, equal to 1 when the previous sales return was positive and negative respectively. In regressions (1) and (2) for all stocks in all market conditions, we introduce another binary variable *SameD* that is equal to 1 when the stock moved in the same direction as the market during the holding period before the stock's previous sale, where market performance is calculated using a buy and hold strategy on the S&P 500 index within that holding period.¹¹ Then we interacted the sales return dummies with *SameD*. The coefficients represent the marginal effects of the covariates: for survival analysis, *ceteris paribus*, that is the effect of a change of a covariate by 1 unit on the hazard rate. Regression (1) is for previous winners: when the market is moving down, a previously sold winner was about 2.581 times more likely to be repurchased than other stocks previously held; but when the market was also moving up, the absolute winner was only $2.581 \times 1.472 \times 0.448 = 1.702$ times more likely repurchased. Regression (2) reveals that a previously sold loser was $1 - 0.412 = 0.598$ times less likely repurchased when the market moving up; and when the market was also moving down, the absolute loser was $1 - 0.412 \times 0.697 \times 2.117 = 0.392$ times less likely repurchased. Therefore, an absolute winner was more attractive and an absolute loser was less unattractive in a downward-trending market than in an upward-trending market.

[Insert Table 1 approximately here.]

To test it in another way, in regressions (3) and (4) of Table 1, we separate up and down market conditions and introduce a binary variable *PosrelSP* that is equal to 1 when the absolute performance of a stock is higher than the S&P 500 return during the same period,

¹¹To check robustness, we also tried using the value-weighted market performance, which yielded similar results.

and 0 otherwise. The results can be conveniently summarized in Table 2. Firstly, when the market is trending upward, the absolute return on a stock can be positive and better than the market (Region 1 in Table 2), positive but worse than the market (Region 2), or negative and of course worse than the market (Region 3). Region 3 is the default in baseline hazard ratio to which we are comparing with. Correspondingly, in Region 1, $Pos = 1$ and $PosrelSP = 1$, so that stocks with previous sales return in this category were $0.878 \times 2.207 = 1.938$ times more likely to be repurchased; in Region 2, $Pos = 1$ and $PosrelSP = 0$, and these stocks were $1 - 0.878 = 0.122$ times less likely repurchased. This is already surprising, given the finding in Strahilevitz et al. (2011) that stocks with previous positive absolute returns should be more likely repurchased than those with negative absolute returns. By contrast, we find that positive absolute returns were not always reinforcing, especially when they are worse than the market performance, which is hard to reconcile with the simple naïve reinforcement learning model. On the other hand, the effects of positive absolute returns are not so surprising when the market was trending down: stocks with positive absolute returns were repurchased more than stocks with negative absolute returns; stocks that performed better than the market were repurchased more than stocks that performed worse.

[Insert Table 2 approximately here.]

To check the robustness of these results, in regressions (5) and (6) we use the same method as above, but with a different measure of market performance, i.e. the performance of matching CRSP cap-based portfolios within the holding period of the stock. The rationale for checking this is that individual investors may care less about the overall performance of the market than about monitoring performance of stocks that are of similar market capitalization as the stocks they held. Therefore, according to the stock’s year-end capitalization portfolio assignment of the previous year, we calculate the performance of a buy-and-hold strategy on the stock’s matching cap-based portfolio within the holding period of the its original sale. Variable $PosrelCap$ is a binary variable that is equal to 1 when the performance of a stock is better than its corresponding cap-based portfolio, and 0 otherwise. The results are

consistent with those using the S&P 500 return as the market benchmark.

Therefore, we propose that the above-documented empirical regulation is consistent with our model in which investors adopt different sources of reinforcement under different market conditions. Still using the S&P 500 return as the benchmark, we introduce a measure of relative return that is the difference between the stock's absolute return and the market return. In order to be comparable to Strahilevitz et al. (2011), the following survival analysis investigates the effect of different levels of return, instead of just its sign. To do this, we introduce a series of binary variables, that are equal to 1 when the (absolute or relative) return falls into some return brackets. Particularly, the lowest bracket includes returns below -46%; starting from -46% each bracket has a width of 4% to up 66%; and the last bracket includes returns above 66%. The regression results are reported visually in Figure 2. Panel A plots the hazard rates with the absolute return on original sale measured on the horizontal axis. The red curve includes all sales, replicating the results in Strahilevitz et al. (2011) that stocks with positive absolute previous sales returns were more likely repurchased than those sold at absolute loss. The blue curve in Panel A is only for sales that yielded negative relative return when the market went up, i.e. cases corresponding to Regions 2 and 3 in Table 2.¹² The result clearly indicates that when the market return was positive, a positive absolute return that is worse than the market can be less reinforcing, making it even less likely repurchased than stocks with negative absolute returns. However, under such market condition, investors biased towards positive relative returns. Panel B of Figure 2 measures relative returns on the horizontal axis, whereas the blue and red curves represent upward and downward-trending market conditions, respectively. Both curves suggest that stocks with positive relative return on original sale were more likely repurchased than stocks with negative relative return, whereas the same value of relative returns were more reinforcing in downward market conditions. This suggests that investors were probably not always using the naïve reinforcement learning strategy based on absolute returns, but they kept an eye

¹²The estimated coefficients were not statistically significant at 5% level for the last four brackets, so we refrained from plotting them on the blue curve.

on the market, at least sometimes, as if they are using weighted averages of absolute and relative returns as the source of reinforcement, adjusting the weights dynamically according to the market conditions. When the market performed well, they switched from using absolute return to using relative return as the source of reinforcement, though still pursuing a reinforcement-type strategy. The reason for doing this could be that individual investors' allocation of attention subject to the ostrich effect, or that they update expectation as reference point and expect better outcomes on own portfolio when the market performs well. Further evidence is required to verify these potential underlying mechanisms.

[Insert Figure 2 approximately here.]

5 Conclusion

Despite the existence of more and less sophisticated forecasting models, the stock market remains a changing environment with uncertain prospects. Outside finance and economics, for example in behavioral psychology, it would be natural to assume that agents in such an environment follow satisficing and reinforcement behaviors instead of optimization and forecasting strategies. Only recently has the importance of such behaviors been acknowledged in economics and finance, in particular in the context of investors' decisions to repurchase stocks. Here, we identified a crucial new component in such decisions: the market performance as a benchmark, i.e the relative return as a source of reinforcement. We found that investors compare the performance of stocks they previously held with the performance of the market as a benchmark. Crucially, investors process this information in a way that places more weight on the relative performance when the market is going up, and more weight on a stocks' absolute performance (and less on the market) when the market is going down. This asymmetry indicates an important bias in impulse selection. There is enormous scope for furthering our understanding of such biases, and of reinforcement behaviors in general, in financial markets. Our own future work will focus on the role of the market in more general

investment decisions.

References

References

- Akerlof, G. A. and W. T. Dickens**, “The economic consequences of cognitive dissonance,” *American Economic Review*, 1982, pp. 307–319.
- Barber, B. M. and T. Odean**, “Trading is hazardous to your wealth: The common stock investment performance of individual investors,” *Journal of Finance*, 2000, pp. 773–806.
- and —, “All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors,” *Review of Financial Studies*, 2008, *21* (2), 785–818.
- Bendor, Jonathan, Dilip Mookherjee, and Debraj Ray**, “Aspiration-based reinforcement learning in repeated interaction games: An overview,” *International Game Theory Review*, 2001, *03*, 159–174.
- Brock, T. C. and J. L. Balloun**, “Behavioral receptivity to dissonant information,” *Journal of Personality and Social Psychology*, 1967, *6* (4), 413–428.
- Brunnermeier, M. and J. Parker**, “Optimal Expectations,” *American Economic Review*, 2005, *95*, 4.
- Bush, R. R. and F. Mosteller**, “A Stochastic Model with Applications to Learning,” *Ann. Math. Statist.*, 12 1953, *24* (4), 559–585.
- Camerer, C. F. and T. H. Ho**, “Experience-weighted Attraction Learning in Normal Form Games,” *Econometrica*, 1999, *67*, 827–874.

- Chiang, Y.-M., D. Hirshleifer, Y. Qian, and A. E. Sherman**, “Do investors learn from experience? Evidence from frequent IPO investors,” *Review of Financial Studies*, 2011, p. hhq151.
- Choi, J., D. Laibson, B. Madrian, and A. Metrick**, “Reinforcement learning and savings behavior,” *Journal of Finance*, 2009, 64 (6), 2515–2534.
- Cross, J. G.**, *A Theory of Adaptive Economic Behavior*, Cambridge University Press, 1983.
- Dixon, H. D.**, “Keeping up with the Joneses: competition and the evolution of collusion,” *Journal of Economic Behavior and Organization*, 2000, 43 (2), 223 – 238.
- Eil, D. and J. M. Rao**, “The good news-bad news effect: asymmetric processing of objective information about yourself,” *American Economic Journal: Microeconomics*, 2011, 3 (2), 114–138.
- Erev, I. and A. E. Roth**, “Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria,” *American Economic Review*, 1998, pp. 848–881.
- **and A. Rapoport**, “Coordination, ‘magic,’ and reinforcement learning in a market entry game,” *Games and Economic Behavior*, 1998, 23 (2), 146–175.
- **and E. Haruvy**, “Learning and the economics of small decisions,” *The Handbook of Experimental Economics*, 2013, 47, 501–512.
- **, E. Ert, and E. Yechiam**, “Loss aversion, diminishing sensitivity, and the effect of experience on repeated decisions,” *Journal of Behavioral Decision Making*, 2008, 21 (5), 575–597.
- Feng, L. and M. S. Seasholes**, “Do investor sophistication and trading experience eliminate behavioral biases in financial markets?,” *Review of Finance*, 2005, 9 (3), 305–351.

- Galai, D. and O. Sade**, “The “Ostrich Effect” and the Relationship between the Liquidity and the Yields of Financial Assets,” *The Journal of Business*, 2006, 79 (5), 2741–2759.
- Goetzmann, W. N. and M. Massa**, “Daily momentum and contrarian behavior of index fund investors,” *Journal of Financial and Quantitative Analysis*, 2002, 37 (03), 375–389.
- Grinblatt, M. and M. Keloharju**, “What makes investors trade?,” *Journal of Finance*, 2001, pp. 589–616.
- Harley, C. B.**, “Learning the evolutionarily stable strategy,” *Journal of Theoretical Biology*, 1981, 89 (4), 611 – 633.
- Jiao, P.**, “The Double-Channled Effects of Experience on Individual Investment Decisions: Experimental Evidence,” 2015. Department of Economics Discussion Paper Series No. 766, University of Oxford.
- , “Losing from Naive Reinforcement Learning: A Survival Analysis of Individual Repurchase Decisions,” 2015. Department of Economics Discussion Paper Series No. 765, University of Oxford.
- Kaniel, R., G. Saar, and S. Titman**, “Individual investor trading and stock returns,” *The Journal of Finance*, 2008, 63 (1), 273–310.
- Karlsson, N., G. Loewenstein, and D. Seppi**, “The ostrich effect: Selective attention to information,” *Journal of Risk and Uncertainty*, 2009, 38 (2), 95–115.
- Kaustia, M. and S. Knüpfer**, “Do investors overweight personal experience? Evidence from IPO subscriptions,” *The Journal of Finance*, 2008, 63 (6), 2679–2702.
- Kőszegi, B. and M. Rabin**, “A model of reference-dependent preferences,” *The Quarterly Journal of Economics*, 2006, pp. 1133–1165.
- Mayraz, G.**, “Priors and Desires,” *Available at SSRN 1805564*, 2011.

- Nax, H. H.**, “Equity dynamics in bargaining without information exchange,” *Journal of Evolutionary Economics*, 2015, 25 (5), 1011–1026.
- , **B. S. R. Pradelski, and H. P. Young**, “Decentralized dynamics to optimal and stable states in the assignment game,” *Proc. CDC - IEEE*, 2013, 52, 2398–2405.
- , **M. N. Burton-Chellew, S. A. West, and H. P. Young**, “Learning in a Black Box,” *University of Oxford, Department of Economics Working Papers*, 2013.
- Nax, Heinrich H. and Bary S.R. Pradelski**, “Evolutionary dynamics and equitable core selection in assignment games,” *International Journal of Game Theory*, 2015, 44 (4), 903–932.
- Nevo, I. and I. Erev**, “On surprise, change, and the effect of recent outcomes,” *Frontiers in Psychology*, 2012, 3.
- Odean, T.**, “Are investors reluctant to realize their losses?,” *Journal of Finance*, 1998, 53 (5), 1775–1798.
- , “Do Investors Trade Too Much?,” *American Economic Review*, 1999, 89 (5), 1279–1298.
- Rabin, M. and J. L. Schrag**, “First impressions matter: A model of confirmatory bias,” *Quarterly Journal of Economics*, 1999, pp. 37–82.
- Roth, A. E. and I. Erev**, “Learning in Extensive-Form Games- Experimental Data and Simple Dynamic Models in the Intermediate Term,” *Games and Economic Behavior*, 1995, 8, 164–212.
- Selten, R.**, “Aspiration Adaptation Theory,” *Journal of Mathematical Psychology*, 1998, 42, 191 – 214.
- Shefrin, H. and M. Statman**, “The disposition to sell winners too early and ride losers too long: Theory and evidence,” *Journal of Finance*, 1985, pp. 777–790.

- Strahilevitz, M., T. Odean, and B. M. Barber**, “Once burned, twice shy: How naïve learning, counterfactuals, and regret affect the repurchase of stocks previously sold,” *Journal of Marketing Research*, 2011, 48 (SPL), S102–S120.
- Suppes, P. and A. R. Atkinson**, *Markov Learning Models for Multiperson Situations*, Stanford University Press, 1959.
- Thorndike, E. L.**, *Animal intelligence : an experimental study of the associative processes in animals*, New York: Macmillan, 1898.
- Tietz, A. and H. J. Weber**, “On the nature of the bargaining process in the Kresko-game,” *H. Sauermann (Ed.) Contribution to experimental economics*, 1972, 7, 305–334.
- Tversky, A. and D. Kahneman**, “Loss Aversion in Riskless Choice: A Reference-Dependent Model,” *The Quarterly Journal of Economics*, 1991, 106 (4), 1039–1061.
- Weber, R. A.**, “‘Learning’ with no feedback in a competitive guessing game,” *Games and Economic Behavior*, 2003, 44, 134–144.
- Yariv, L.**, “Believe and Let Believe: Axiomatic Foundations for Belief Dependent Utility Functionals,” 2001. Yale University, Cowles Foundation Discussion Paper No. 1344.

Table 1: REINFORCEMENT LEARNING BIAS UNDER DIFFERENT MARKET CONDITIONS

This table reports the survival analysis results regarding the effects of positive and negative previous sales returns under different market conditions. *Pos* (*Neg*) is a binary variable that is equal to 1 when the absolute return on original sale was positive; *SameD* is a binary variable that is equal to 1 when the absolute return had the same sign as S&P 500 return within the holding period, and 0 otherwise; *PosrelSP* (or *PosrelCap*) is a binary variable that is equal to 1 when the return relative to S&P 500 (or matching CRSP cap-based portfolio) was positive, and 0 otherwise; control variables include inverse holding period, log relative abnormal volume and return subsequent to original sale. Robust standard errors clustered by households are in parentheses. Regressions (1) and (2) include all market conditions, (3) and (4) only for up markets, (5) and (6) only for down markets. (***) $p < 0.01$, ** $p < 0.05$, * $p < 0.10$)

	(1)	(2)	(3) UP	(4) DOWN	(5) UP	(6) DOWN
<i>Pos</i>	2.581*** (0.068)		0.878*** (0.020)	1.618*** (0.075)	0.895*** (0.019)	1.628*** (0.076)
<i>Neg</i>		0.412*** (0.011)				
<i>SameD</i>	1.472*** (0.027)	0.697*** (0.010)				
<i>Pos</i> × <i>SameD</i>	0.448*** (0.012)					
<i>Neg</i> × <i>SameD</i>		2.117*** (0.562)				
<i>PosrelSP</i>			2.207*** (0.044)	1.137*** (0.051)		
<i>PosrelCap</i>					2.239*** (0.040)	1.136*** (0.052)
<i>Control Varriables</i>	Yes	Yes	Yes	Yes	Yes	Yes
<i>p</i>	0.440 (0.004)	0.440 (0.004)	0.469 (0.003)	0.433 (0.004)	0.468 (0.003)	0.429 (0.004)
<i>N</i>	490,158	490,158	389,265	88,872	387,002	80,605

Table 2: COMPARISON OF HAZARD RATES IN UP AND DOWN MARKETS

This table is a comparison of hazard rates in up and down market conditions. We distinguish a total of six cases: 1-absolute return>market return>0; 2-market return>absolute return>0; 3-market return>0>absolute return; 4-absolute return>0>market return; 5-0>absolute return>market return; 6-0>market return>absolute return. The hazard rates were calculated based on survival analysis results in Table 1. Regions 3 and 6 are embedded in the baseline hazard rates in the survival analysis. The 95% Confidence Intervals of the hazard rates are reported. The market performance benchmark used are S&P 500 and cap-based portfolios respectively in the upper and lower panels.

Up Market			Down Market		
Region	Hazard Ratio	95% CI	Region	Hazard Ratio	95% CI
Market performance measured using S&P 500					
1	1.938	(1.782, 2.107)	4	1.840	(1.537, 2.201)
2	0.878	(0.840, 0.918)	5	1.137	(1.040, 1.242)
3	Default=1	N/A	6	Default=1	N/A
Market performance measured using cap-based portfolios					
1	2.004	(1.854, 2.164)	4	1.849	(1.540, 2.221)
2	0.895	(0.858, 0.933)	5	1.136	(1.037, 1.244)
3	Default=1	N/A	6	Default=1	N/A

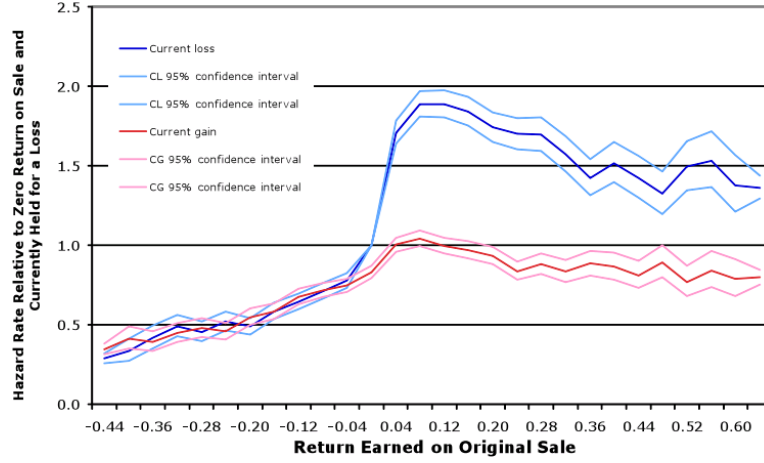
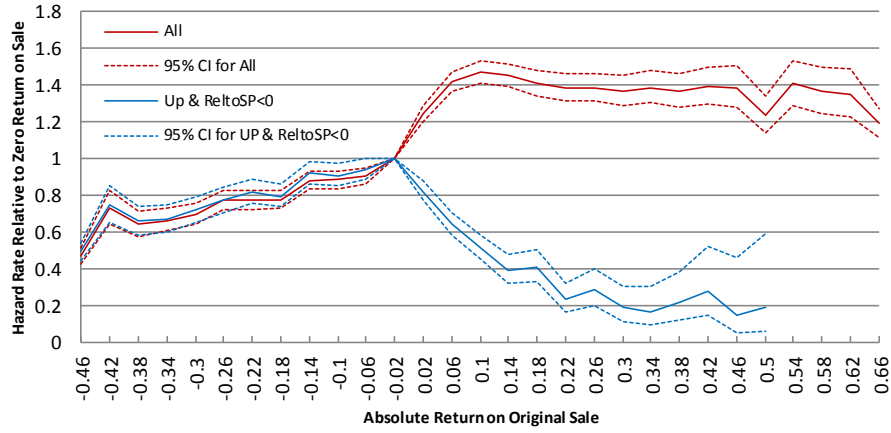
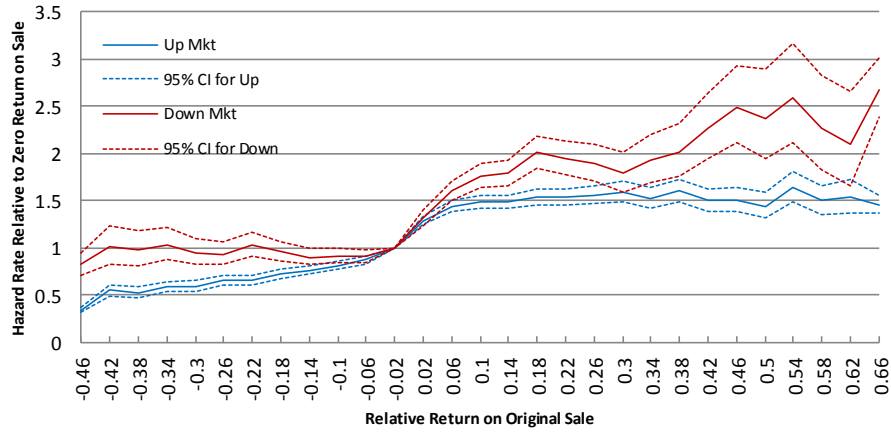


Figure 1: THE NAÏVE REINFORCEMENT LEARNING IN STRAHILEVITZ ET AL. (2011)
Note: This is Figure 1 in Strahilevitz et al. (2011) which demonstrates that stocks previously sold at gain were more likely repurchased than stocks previously sold at loss. The horizontal axis is the absolute return on original sale, and the vertical axis is the hazard rate relative zero return on original sale. The blue and red curves distinguish stocks that declined and those that increased in value subsequent to being sold. The 95% confidence intervals were added.



Panel A



Panel B

Figure 2: HAZARD RATES USING ABSOLUTE AND RELATIVE RETURNS

Note: This figure depicts the hazard rate of repurchases relative to zero return on sale against absolute return (in Panel A) and relative return (in Panel B) on original sale of the stock. The figures were plotted using the hazard rates from the survival analysis with the categorical return dummy variables as the covariates. In Panel A: the red line was plotted with all sales, and the blue line include only sales that yielded negative return relative to S&P 500 when the market return was positive. In Panel B: the red line was plotted for sales made when the market return was negative, and the blue line was plotted for sales made when the market return was positive. The blue line in PanThe dotted lines are 95% confidence intervals. In Panel A, the last four coefficient values were not plotted because of insignificance.