THE IMPACT OF SALIENCE ON INVESTOR BEHAVIOR:

EVIDENCE FROM A NATURAL EXPERIMENT

Cary Frydman and Baolian Wang*

September 2017

ABSTRACT: We test whether the salience of information causally affects investor behavior. Using investor level brokerage data from China, we estimate the impact of a shock that increased the salience of a stock's purchase price, but did not change the investor's information set. We employ a difference-in-differences approach and find that the shock causally increased the disposition effect by 20%. We document substantial heterogeneity across investors in the salience effect and we show that an investor level proxy for "salient thinking" can explain this heterogeneity. More generally, our results support a recently proposed class of models in which salience impacts choice.

*Frydman is at the Department of Finance and Business Economics, USC Marshall School of Business, cfrydman@marshall.usc.edu; Wang is at the Area of Finance and Business Economics, Gabelli School of Business, Fordham University, bwang46@fordham.edu. We are thankful for comments from Kenneth Ahern, Nicholas Barberis, An Li, Juhani Linnainmaa, Mark Seasholes and seminar participants from Ohio State University, University of Southern California, the Cognition and Decision Laboratory at Columbia University, the 2016 Caltech Junior Behavioral Finance conference, and the 2017 Interdisciplinary Symposium on Decision Neuroscience.

Over the past several years there has been a surge in interest on the impact of attention on economic choice. On the theoretical side, a new class of behavioral models has emerged which formalizes how attention to different attributes of financial assets, risky gambles, or consumer goods can have a systematic effect on the decision-making process and the resulting choice (Bordalo, Gennaioli, and Shleifer 2012; 2013a; 2013b; Koszegi and Szeidl 2013). For example, an investor may exhibit a preference for a financial asset with a positively skewed return distribution because his attention is drawn to the salient high payoff state of the world, which he then overvalues in the decision-making process.

This theoretical literature on “attribute-specific” attention has been accompanied by empirical work that is almost exclusively conducted in laboratory experiments. One reason for this is that the laboratory provides a controlled environment in which attention can both be measured and manipulated. For example, Hare, Malmaud, and Rangel (2011) manipulate attention towards the health attribute of a food item and find that subjects systematically choose healthier foods. In a financial decision-making context, recent experimental work demonstrates that manipulating the salience of a stock’s purchase price or the state space can have a causal effect on choice (Frydman and Rangel 2014; Frydman and Mormann 2017). However, it is unclear whether these attribute-specific salience effects extend outside the laboratory into higher stakes and more natural financial decision-making environments.

In this paper, we use a combination of account-level microdata from a Chinese brokerage house and a natural experiment to estimate the causal impact of salience on investor behavior. In our natural experiment, the brokerage house increases the *salience* of information about a stock’s capital gain; critically, the information provided to investors about the capital gain is held constant. Our dependent variable of interest is the disposition effect, which is the greater tendency to sell winning stocks compared to losing stocks (Shefrin and Statman 1985; Odean 1998)[1]. We hypothesize that the salience shock will increase the weight that investors attach to the capital gain during the decision-making process, and this in turn, will generate a stronger disposition effect.

There are three main challenges in testing for the effect of salience on investor behavior in the field. First, in many circumstances, a change in the salience of information is often correlated with a change in information itself. We exploit an exogenous shock to the display of a brokerage house's online trading platform, which did not affect the investor's information set. Specifically, in October 2004, the brokerage house began to display the capital gain on the online trading screen, for each stock held in the investor's portfolio. Critically, information about a stock's capital gain was available to the investor before this change in display (investors could access this information in three "mouse-clicks"), and thus this change represents a salience shock rather than an information shock.

Second, estimating the effect of salience on the average investor's behavior can mask important heterogeneity that makes it difficult to interpret the strength of the effect. For example, investors with large portfolios may exhibit greater salience effects because they are faced with more incoming information. This could lead to overestimating the average salience effect if larger investors contribute more observations to the analysis than smaller investors. Our data provide information on trades at the account level, and thus we can estimate the salience effect for each individual investor. Moreover, the microdata allow us to estimate the distribution of salience effects across our sample[2].

---

[1] The disposition effect has been documented both among individual and professional investors. It is an extremely robust effect that has been found among a wide variety of asset classes and international markets. For a comprehensive review of the disposition effect, see Kaustia (2010) and Barber and Odean (2013).

[2] Variation across investors in response to the salience shock can also have important welfare implications. See Taubinsky and Rees-Jones (2017) for a theoretical discussion on the importance of accounting for variation in tax salience effects, and see Finkelstein (2009) and Chetty, Looney, and Kroft (2009) for a more general discussion on welfare and tax salience effects.

The third challenge is that the salience shock may coincide with a distinct but unobserved shock that also affects trading behavior. Fortunately, the change in information display that we study generated a natural experiment that provides a control group, which can be used to control for common time series variation in the disposition effect. In particular, our data provide information on the method of trade: internet, phone or in-person (Barber and Odean 2002). Because the change in information display occurred online, the salience shock should not affect investors who traded by phone or in-person. Thus, we can employ a difference in differences approach to estimate the causal effect of salience on trading behavior.

There are two main results in this paper. First, after the salience shock, the disposition effect increases significantly among internet investors (treatment group), but there is no similar increase among non-internet investors (control group). We show that the two groups have parallel trends before the salience shock and that the disposition effect increases significantly more in the treatment group compared to the control group. We are thus able to interpret salience as causally affecting trading behavior. Before the change in information display, internet investors sold winning stocks with probability 7.4% and they sold losing stocks with probability 2.5%; this difference of 4.9% represents a significant disposition effect, and is similar in magnitude to previous work (Feng and Seasholes 2005). The salience shock caused the disposition effect to increase to 5.9% for the average internet investor, and we find substantial variation in this salience effect across investors.

Our second main result provides a source of this cross-sectional variation and presents an additional test of the salience mechanism. The logic of the test is as follows. We use a modified version of the model in Bordalo, Genniaoli, and Shleifer (2013) to generate a prediction that the degree to which an investor is prone to "salient thinking" (i.e., the degree to which he focuses attention disproportionately on salient attributes) is positively correlated with the change in the disposition effect. In other words, heterogeneity in the degree of salient thinking across investors should provide a source of variation in the salience effect we observe. In order to test this prediction, we need to estimate the degree to which each investor is prone to salient thinking.

Our proxy for an investor's degree of salient thinking is the so-called "rank effect," which refers to an investor's propensity to sell extreme ranked stocked in his portfolio (Hartzmark 2015). This effect has previously been attributed to a salience mechanism, in the sense that an investor allocates disproportionate attention to the extreme – and therefore salient – stocks in his portfolio. Thus, we assume that investors with high rank effects are more prone to salient thinking than investors with low rank effects. We estimate the rank effect at the investor level in two ways: one based on ranking by returns and one based on ranking by alphabetical company name. Regardless of how we measure the rank effect, we find that it does significantly correlate, across investors, with the change in the disposition effect after the salience shock. This result therefore provides additional cross-sectional evidence that is consistent with the salience mechanism.

We contribute to several lines of literature. First, we are able to identify a causal effect of salience on decision-making in the field. This jointly addresses the external validity concern from laboratory studies and the problem of non-identification in many field studies. Our field evidence also provides support for a key assumption in recent models of attention and economic choice. In particular, many of these models make two key assumptions: (i) attention is endogenously determined by the decision-maker's choice set and (ii) attributes of the choice set that receive greater attention will be overweighed in the choice process (Bordalo, Gennaioli, and Shleifer 2012; 2013a; 2013b; Koszegi and Szeidl 2013). Our data provide support for the second assumption, as we demonstrate that an increase in the salience of the capital gain increases the weight that investors attach to this attribute when making a trading decision. In addition, our cross-sectional results suggest that deriving the implications of heterogeneity in salient thinking may be an important direction for future theoretical work.

Second, we contribute to the literature on the disposition effect. Over the last decade there has been a resurgence in work on the disposition effect, with many researchers constructing formal models to understand its cause (Barberis and Xiong 2009; 2012; Ingersoll and Jin 2013; Meng and Weng 2017). There has also been a great deal of empirical and experimental work testing these newer models of the disposition effect (Ben-David and Hirshleifer 2012; Frydman et al. 2014; Birru 2015; Chang et al. 2016; Imas 2016; Heimer 2016; Fischbacher et al. 2017). We add to this literature by providing evidence that the salience of information can systematically increase the disposition effect. We also document a novel correlation between trading biases: the rank effect is correlated with the salience driven change in the disposition effect. This is important because a better understanding of the correlation structure between trading patterns has the potential to uncover a small set of psychological principles that can explain a growing set of behavioral effects (Barber and Odean 2013; Frydman and Camerer 2016) [3].

Finally this paper adds to the literature on information display and choice architecture (Thaler and Sunstein 2008). Choice architecture refers to the principle that there are a variety of ways to present a choice to a decision-maker, and importantly, the presentation mode has a systematic impact on choice (Johnson et al. 2012; Benartzi et al. 2017). While there has been an enormous amount of research on this topic in the marketing literature, there has been relatively less work in finance, and this work has concentrated mainly on studying saving and borrowing decisions (Madrian and Shea 2001; Bertrand and Morse 2011; Beshears et al. 2013; Choi et al. 2017). Our results provide a concrete example of how online information display can have a substantial effect on trading behavior at the investor level[4].

---

[3] While our data cannot distinguish between competing mechanisms of the disposition effect, we note that our results are consistent with a preference-based explanation of realization utility. As Barberis and Xiong (2012, p252) write, realization utility "is likely to play a larger role when the purchase price is more salient" and thus if realization utility is responsible for the disposition effect, we expect a larger disposition effect when the purchase price becomes more salient.

[4] Two other recent papers study the impact of a change in information display. Shaton (2017) studies the impact of a shock to the salience of past returns on aggregate outcomes such as fund-flow sensitivity and volume. Levi (2017) runs a randomized control trial experiment to investigate the effect of information

I. DATA AND NATURAL EXPERIMENT

A. Data Overview

We use an account level dataset to investigate trading behavior among investors at a security brokerage company in People's Republic of China.[5] Our data set is very similar to the Chinese dataset used by Feng and Seasholes (2004; 2005), and it is also similar to the US dataset used by Odean (1998). Our data come from a brokerage company that has multiple branches throughout China and serves approximately half a million investors in total. The main dataset comes from two of these branches and is comprised of three files: a trade file, a position file, and an investor demographic file, and our sample period is from January 2003 through December 2009. The trade file provides data at the account-date-stock level, and contains information about the trade time (hour: minute: second), stock ticker, buy/sell indicator, transaction price, number of shares purchased/sold, trading method (phone, internet, in-person), commissions and taxes. We restrict our analyses to trading of common stocks, and there are no short sales in our data since short selling was prohibited in China during our sample period.

A key variable we will use in our empirical methodology is the trading method. An investor can place an order either online, over the phone, or in person at their home branch of the brokerage company. Our empirical analysis focuses on one branch of this brokerage company because we only have the trading method data for this one branch. We have data on 16,809 accounts from this branch. Securities law in China during our sample period prohibits investors from holding more than one account, and it requires that each investor conduct all trades via the branch they opened the account with. This means that for every investor in our data set, we

display on individual consumption decisions. Our study is complementary to these two studies in that (i) we estimate the causal effect of a change in information display on individual investment decisions and (ii) we exploit cross-sectional variation in the response to the salience shock to better understand the mechanism through which behavior is affected.

[5] Investors need to open different security accounts to trade stocks listed in different exchanges. We identify investors based on their "fund account number" which is an internal code used by the brokerage company. The fund account number links a single individual to all the security accounts one may have.

observe all trades for that investor during our sample period. Chinese investors typically open an account close to where they live (Feng and Seasholes 2004), but if they are traveling to a different city and want to place an order, they can do so through the phone or the internet.

When a trade is executed, both sellers and buyers need to pay a set of fees[6]. The most important feature of the fee structure in our setting is that there is no capital gains tax. While the sale of a position will trigger a tax bill, the tax base is the amount of the entire position, not the capital gain or loss. In other words, selling both winners and losers in China during our sample period will generate a tax burden, and hence, we can rule out tax-loss selling as an explanation for our results.

B. Natural Experiment

In October 2004, one branch of the brokerage company for which we have data, altered the online display of their client's portfolio information. Before October 2004, when an online investor logged onto the online trading platform in this branch, the software displayed information about each of the investor's stock positions. The variables that were displayed for each stock included: stock name, stock ticker, number of shares held, current price, current value (current price times number of shares held), and two other variables that refer to frozen shares[7].

On October 1st 2004, the brokerage company added five new variables to the investor's portfolio page: 1) weighted average purchase price, 2) break-even price, 3) realized gain/loss, 4) paper gain/loss and 5) total gain/loss. The weighted average purchase price is calculated without

[6] There are three separate fees: 1) a stamp duty tax to the government, 2) a commission to the brokerage house, and 3) a fee to the stock exchange. The stamp duty tax rate from the beginning of our sample until January 23, 2005 was 0.2% of the value of the transaction, and this was reduced to 0.1% from January 24, 2005 until the end of our sample. The commission to the brokerage house ranges from 0.1% to 0.3% of the transaction value, with a minimum fee of ¥5. Finally, the stock exchange fee was 0.03% of the number of shares traded for the Shanghai Stock Exchange and 0.00255% of the transaction value for the Shenzhen Stock Exchange, respectively. The stock exchange fee was much smaller than the other two types of fees. Therefore, when summing all taxes and fees, the range of fees is approximately between 0.2% and 0.5%.

[7] The two variables are 1) number of shares that are frozen and 2) number of shares that are currently available for sale. Shares can be frozen because they are used for collateral or legal reasons. Shares that are not currently available for sales can occur because they are frozen, or because they were purchased on the same day (investors are not allowed to sell shares that are bought on the same day that they purchase them).

considering commissions, taxes, or cash dividends, but stock dividends and stock splits are taken into account. The break-even price provides the market price at which liquidating the entire (remaining) position will make the complete round-trip trade (since the initial purchase) break even, and accounts for commissions, taxes, cash dividends, stock dividends and stock splits. The realized gain/loss variable equals *#shares sold**(*sale price – weighted average purchase price*)+ *cash dividend*, and equals 0 if there has been no partial sale and no cash dividend[8]. The paper gain/loss variable equals *#shares held*(current price – weighted average purchase price).* Finally, the total gain/loss variable is the sum of the realized gain/loss and paper gain/loss. In addition to the variables that were added, the font color of some positions changed. Before October 2004, the font color of each position was blue; after the change, those positions trading at a gain were shown in red font while those positions trading at a loss were displayed in blue font. Figure 1 provides a screenshot of the trading screen before and after the change in information display.

Critically, the information provided by these newly added variables was already available to the investor before the change in information display. For example, an investor had access to both previous price history for all Chinese stocks as well as his own transaction history. Together, these two pieces of information are sufficient to compute all five variables that were added on the portfolio homepage. Hence, the change in information display reflects a shock to the salience of information rather than a shock to an investor's information set.

Moreover the five variables that were added are all closely related to one another. Once an investor is provided the weighted average purchase price, it becomes straightforward to compute realized and paper gains (and total gains are a sum of paper and realized gains). The weighted average purchase price and break-even variables are highly correlated and are very close to each other for most cases. These two variables typically exhibit a sizable difference only when there have been partial sales with significant realized gains/losses. Perhaps the more

[8] Dividends received are included in realized gain/loss.

important difference is that the term "break-even" may cause some investors to think that selling above the break-even price is the "right" thing to do. Later in the paper we address this concern by testing whether selling behavior is more sensitive to the weighted average purchase price compared to the break-even price. We also note that the break-even price information was available to investors prior to the information display change (though not on the account summary page), so any change in behavior due to this additional variable would still be driven by a salience shock rather than an information shock.

The change in information display was not driven by pressure from investors nor was it driven by a regulatory motive. In discussions with representatives from the brokerage company, we learned that the primary reason for the information display change was to give clients easier access to their account information, which was already provided by competing brokerage houses. The change was not considered to be an important event and no pilot programs were conducted. Investors who traded online prior to October 1, 2004 received a message the first time they logged onto the trading platform after the information display change alerting them to the easier access of their account information. Those traders who did not trade online before October 2004 but who subsequently moved online, did not receive any explicit message about the change in information display.

While our main hypothesis is that the change in information display affects behavior through a salience mechanism, it is possible that behavior can change through at least two other channels. First, an investor could (mistakenly) perceive the change in information display as advice from the brokerage company (Benartzi 2001). Second, the change in information display also affected the font color of each stock, and recent work shows that color per se can affect trading behavior (Bazley, Cronqvist, and Mormann (2017)). Later in the paper, we use cross-investor analyses to test whether these two mechanisms can fully explain our findings.

C. Methodology and summary statistics

Our main methodology follows Feng and Seasholes (2005), Ben-David and Hirshleifer (2012), and Barber and Odean (2013). Specifically, we use the transaction data set and the position data set to construct a holding sample containing an observation for each investor-stock-day. Investor-days when an investor does not trade are also included in the analysis. Buys and sales are aggregated on a daily basis for each account and each stock. For example, if an investor buys Sinopec stock on January 2, 2003 and held it until January 29, 2003, there will be 19 observations (19 business days).[9]

We flag the days when a position is opened and when shares are sold (including partial sales). We exclude positions for which we do not have information on the purchase price. This is mainly because investors bought their stocks before the start of our sample period. For an investor to be included in our analysis, we require that his position data are non-missing for both the pre and post periods. We also require that he traded at least once in the pre-period, so that we have information on his trading method.

We define an "internet investor" as an investor who placed at least one order via the internet before October 2004. A "non-internet investor" is defined as an investor who never traded online before October 2004. Table 1A provides summary statistics using this definition, and we see that there is roughly an even split between the two groups. Among the internet investor group, approximately two-thirds place orders exclusively through the internet, while one third uses a combination of internet and other methods. Among the non-internet investor group, most of the investors refrain from using the internet to place trades even after the salience shock. Specifically, among non-internet investors, only 9.2% of trades in the post period were placed through the internet. Table 1B provides summary statistics on investor characteristics for each of the two groups. The table shows there are clear differences between the two groups, as internet

[9] Odean (1998) adopts a different method and only includes an investor-day observation when the investor sells at least one stock. We conduct robustness tests using this method and report the results in Table 10.

investors hold portfolios with larger size and more stocks. Internet investors trade more frequently (as indicated by the higher unconditional selling propensity) and are also slightly older.[10] After presenting the main results, we provide a matching exercise to demonstrate that our results are not sensitive to these differences across groups.

II. EMPIRICAL FRAMEWORK

In this section we provide a basic empirical framework to guide our analysis of the effect of the salience shock on trading behavior. We assume that there is a single investor who must decide whether to sell or hold a stock. The investor's utility of selling the stock is a function of two attributes. The first attribute, and the one that is most important in our setting, is the sign of the stock's capital gain, *G*. The second attribute, *X*, can be any other stock-level attribute that carries nonzero weight in the investor's utility computation. For example, *X* could be the stock's expected return or it could summarize the tax consequences from selling the stock.

We assume the two attributes are additively separable and uncorrelated. The utility of selling relative to holding in the period before the salience shock is given by:

$$u_{pre} = \alpha G + X \tag{1}$$

where $G = 1$ if the stock is trading at a gain and $G = -1$ if it is trading at a loss. We also assume that the investor's selling propensity increases with the above utility[11]. Therefore, the first term, αG, captures the investor's motive for selling winning stocks and holding losing stocks (for $\alpha > 0$). The above assumptions can be microfounded with a variety of models of the disposition effect, including realization utility (Barberis and Xiong 2012; Ingersoll and Jin 2013; Frydman et

[10] The findings that internet investors have larger portfolios and trade more frequently are consistent with the US investors (Barber and Odean 2002), but the findings that internet investors are older contrasts with the US finding (Barber and Odean 2002). Barber and Odean (2002) also find that internet investors are more likely to be men, while sex is uncorrelated with internet status in our sample.
[11] E.g., because utility is also subject to a random shock that is i.i.d. across observations.

al. 2014) and cognitive dissonance (Chang, Solomon, and Westerfield 2016). The weight on G provides a measure of the strength of preference for selling winners and holding losing stocks; in other words, the weight on G is a measure of the investor's disposition effect. In the pre-period, α is therefore a measure of the investor's disposition effect.

In the period after the change in information display, the capital gain becomes more salient. We use a modified version of the theory in Bordalo, Gennaioli, and Shleifer (2013), and we model the change in information display as an increase in the weight attached to the first component relative to the second component. Specifically, the utility of selling, relative to holding, after the salience shock is given by:

$$u_{post} = \frac{1}{\delta}\alpha G + X \tag{2}$$

where $\delta \in (0,1]$ is a "salient thinking" parameter. As δ decreases, the investor attaches greater weight to the salient attribute (in this case, the sign of the capital gain). When $\delta = 1$, salience does not affect utility and equation (2) becomes equivalent to equation (1). It is important to emphasize that in this framework, the disposition effect (governed by α) and the salience effect (governed by δ) are driven by distinct mechanisms. The salient thinking parameter, δ, captures the general tendency for an investor to overweigh any attribute that is salient. If, for example, the change in information display had increased the salience of variable *X*, we would expect that the investor would increase the weight he attaches to *X* in the decision process.

In the post-period, the weight on G changes to $\frac{\alpha}{\delta}$, which is a measure of the investor's post-period disposition effect. Therefore, the disposition effect changes from α to $\frac{\alpha}{\delta}$ after the salience shock, which implies an increase in the disposition effect for any investor with $\delta < 1$. This leads to our first prediction about the change in the disposition effect after the salience shock:

Prediction 1: For any salient thinking investor with $\delta < 1$, the disposition effect will increase after the salience shock.

Note also that the disposition effect increases by a factor of, $\frac{1}{\delta}$, and therefore the change in the disposition effect is decreasing in δ. This implies that two investors with different levels of salient thinking will exhibit different trading responses to the salience shock – even if they exhibit the same disposition effect in the pre-period. An investor who is particularly prone to salient thinking, given by $\delta = \delta_{low}$, will exhibit a larger increase in the disposition effect compared to an investor who is less influenced by salient thinking, given by $\delta = \delta_{high}$. This leads to our second prediction:

Prediction 2: For any two investors with salience parameters given by δ_{high} and δ_{low} such that $0 < \delta_{low} < \delta_{high} < 1$, the investor parameterized by δ_{low} will exhibit a larger increase in the disposition effect compared to the investor parameterized by δ_{high}.

III. ESTIMATING THE CAUSAL EFFECT OF SALIENCE ON INVESTOR BEHAVIOR

A. Main Results

To test Prediction 1, which states that there should be an increase in the disposition effect after the salience shock for any salient thinking investor, we use a difference-in-differences methodology. Because the change in information display should only affect those investors who trade online, our treatment group is defined as the set of internet investors, and our control group is the set of non-internet investors (those who trade on the phone or in-person at their designated branch). The pre period starts at the beginning of our sample, January 1, 2003 and ends on

September 30, 2004; the change in information display occurs on October 1, 2004, and thus the post period is defined as October 1, 2004 through June 30, 2006. This definition of the pre and post period therefore uses only a subsample of our dataset, which is done for two reasons. First, it balances the lengths of the pre period and the post period at twenty-one months. Second, it allows us to run a placebo test using twenty-one months for both the placebo pre and post periods. Later in the paper, we check whether our main result is robust to including the entire sample from January 2003 through December 2009.

We estimate the effect of the salience shock in a regression framework using the following model:

$$Sell_{i,j,t} = \alpha + \beta_1 Gain_{i,j,t-1} + \beta_2 Post_t + \beta_3 Gain_{i,j,t-1} * Post_t + \beta_4 Internet_i + \beta_5 Gain_{i,j,t-1} * Post_t * Internet_i + \beta_6 Gain_{i,j,t-1} * Internet_i + \beta_7 Post_t * Internet_i + \varepsilon_{i,j,t} \quad (3)$$

where *i, j,* and *t* denote investor, stock, and day, respectively. $Sell_{i,j,t}$ is a dummy variable which equals 1 if investor *i* sells stock *j* (partially or fully) on day *t*. $Gain_{i,j,t-1}$ is a dummy variable which equals 1 if investor *i* has a gain on stock *j* on day *t*-1. $Post_t$ is a dummy variable which equals 1 if day *t* is after September 30, 2004 and 0 otherwise. The coefficient of the triple interaction term – β_5—captures the causal effect of the salience shock. Model (3) is estimated using linear regression and we cluster standard errors by day and by investor.

Columns (1) and (2) of Table 2A provide the estimation results for internet and non-internet investors, respectively. For presentation purposes, all coefficients are multiplied by 100. The coefficient on the *Gain* dummy for both groups is strongly positive, indicating that, in the pre period, both groups exhibit a significant disposition effect. In the pre period, the probability that an internet investor sells a winning stock is 7.4% compared to 2.5% for a losing stock. Non-internet investors exhibit a similar disposition effect as they sell winning stocks with probability 5.5% compared to 1.5% for losing stocks. Moreover, the coefficient on the *Gain*Post* variable is significantly positive for internet investors, but it is not statistically different from zero for non-internet investors. This indicates that internet investors exhibit a significant increase in the disposition effect after the salience shock, but non-internet investors do not. Column (3) shows the results from the full model estimation, and we see that the coefficient on the key triple interaction variable is significantly positive.

Table 2B adds several control variables that are known to affect selling propensities (Ben-David and Hirshleifer 2012). In columns (1) through (3), we add controls for the holding period, weighted average purchase price, and volatility (depending on sign of return). Specifically, the control variables are defined as following: $\sqrt{TimeOwned}$, square root of number of days since the position has been open; log(Buy price), the natural logarithm of the weighted average price; $Volatility^{-}$, the stock volatility calculated using the previous 250 days' daily returns if the return since purchase is negative, and zero otherwise; $Volatility^{+}$, the stock volatility calculated using the previous 250 days' daily returns if the return since purchase is positive, and zero otherwise.

Consistent with previous results using US data, we find that longer holding periods are associated with lower selling propensities and that volatility is associated with higher selling propensities (Ben-David and Hirshleifer 2012). After including these controls, the *t*-statistic on the key triple interaction becomes stronger. In columns (4) through (6) we add controls for returns and the interaction between returns and holding period. The two return variables are: $Return^{-}$, the return since purchase if the return since purchase is negative, and zero otherwise; $Return^{+}$, the return since purchase if the return since purchase is positive, and zero otherwise. We find that the coefficient on the triple interaction term in column (6) remains significantly positive at the 1% level.

The increase in the disposition effect among internet investors is economically large. In the analysis without any controls, the difference-in-differences estimate is 0.85%, which is 17.5% ($\frac{0.85\%}{4.86\%}$) of the baseline disposition effect among internet investors in the pre period. The difference-in-differences estimate increases to 1.54% with full controls, which represents an increase of more than a quarter of the pre period disposition effect level (6.09%). In Table 3, we re-estimate equation (3) and include both *Investor*Gain* and *Investor*Day* fixed effects. These control for individual level heterogeneity in the disposition effect and day-specific selling propensities. The *Investor*Day* fixed effect could be important, for example, if an investor allocates attention to his portfolio as a function of portfolio performance (Pagel 2017). Table 3 shows that the coefficient on the triple interaction remains significant after including these fixed effects.

B. Parallel trends assumption

In order for the difference in differences estimator to be valid, our key assumption is that, in the absence of the salience shock, the trends in the disposition effect are the same for the treatment and control groups. To investigate this parallel trends assumption, we provide both graphical and statistical tests. We begin by estimating the time series of the average disposition effect at quarterly frequency from January 1, 2003 to June 30, 2006. In particular, we run the following regression:

$$\begin{aligned} Sale_{i,j,t} = \alpha &+ \sum_{q=2003Q1}^{2006Q2} \beta_{int,q}\, Gain_{i,j,t-1} Internet_i D_q \\ &+ \sum_{q=2003Q1}^{2006Q2} \beta_{nonint,q}\, Gain_{i,j,t-1} NonInternet_i D_q + X\beta' + \varepsilon_{i,j,t} \end{aligned} \tag{4}$$

where D_q is a dummy equal to 1 if the observation is in quarter q, $Internet_i$ is a dummy equal to 1 if investor i is an internet investor and $NonInternet_i$ is a dummy equal to 1 if i is a non-internet investor. *Gain* is a dummy equal to 1 if the stock is trading at a positive capital gain and 0 otherwise. *Gain* is measured at the end of the previous trading day. *X* is a vector of control variables, including all the control variables in Ben-David and Hirshleifer (2012), the interaction terms between $Internet_i$ and D_q, and between $NonInternet_i$ and D_q. The coefficients of interest from this regression are $\{\beta_{int,q}\}_{q=2003Q1}^{2006Q2}$ and $\{\beta_{nonint,q}\}_{q=2003Q1}^{2006Q2}$ which provide the estimated time series of the disposition effect at quarterly frequency for internet and non-internet investors, respectively. Figure 2A plots these coefficients, along with the 95% confidence intervals, where standard errors are clustered by day and by investor. Upon visual inspection, the chart shows a divergence in the two time series immediately following the salience shock in the third quarter of 2004[12].

We can also formally test the requirement of the parallel trends assumption that the trends of the treatment and control groups are similar before the salience shock. To do so, we test whether there is a significant difference in $(\beta_{int} - \beta_{nonint})$ between consecutive quarters. Of the fourteen quarters plotted in Figure 2A, there is a change in $(\beta_{int} - \beta_{nonint})$ only between 2004Q3 and 2004Q4 (p=0.0001) and between 2004Q4 to 2005Q1 (p = 0.0973). This means that in the pre period, we cannot reject the assumption that there is a fixed difference in the disposition effect between the two groups. In addition, we find that there is a change in this difference in the two quarters immediately following the salience shock[13]. Taken together, these statistical tests

[12] One other pattern that is evident in Figure 2A is that there is substantial time series variation in the disposition effect. In US data, Odean (1998) documents that there is a decline in the disposition effect over the course of a calendar year, and it achieves a minimum in December as the annual deadline for tax-loss selling approaches. While the time series variation in the level of the disposition effect is not a major concern for our difference-in-differences methodology, its existence in a market without a capital gains tax is, to our knowledge, new to the literature.

[13] One potential concern with this "quarter by quarter" coefficient testing is that it is subject to a multiple comparisons problem. Because we run 13 separate hypothesis tests – one for each pair of consecutive quarters – a standard Bonferronni correction would render the 2004Q4 to 2005Q1 result insignificant, but the 2004Q3 to 2004Q4 result would remain highly significant. Therefore, by incorporating the multiple

combined with the divergence in the time series immediately following the treatment provide support for the parallel trends assumption.

One concern, however, from looking at Figure 2A is that the divergence in the disposition effects between the two groups is driven mainly by a decrease in the disposition effect of the control group. While this does not technically violate the validity of the difference in differences estimator, it heightens the concern that there may be a group specific shock to the disposition effect that coincides with the change in information display. An alternative to this group-specific shock explanation is that the drop in the disposition effect for the control group may simply be due to time series variation that is common to all groups.

Fortunately, we were able to obtain additional data to help distinguish which of these two explanations is more likely. In particular, we obtained data from a second branch, within the same brokerage firm, that is located in the same city as our main branch. We have data on 4,518 investors that satisfy the same data requirements we use for our main branch. We do not have information on whether these investors traded via phone or internet, but we do know that the purchase price and capital gain information was provided on the trading screen to investors in the second branch throughout our entire sample period. In other words, the trading platform for our main branch after the salience shock is identical to the trading platform for the second branch for the entire sample period.

If the drop in the disposition effect for the control group shown in Figure 2A was driven by a group specific shock to the control group, we would not expect to see a similar drop in the disposition effect among investors in the second branch. To test this, we estimate the time series of the disposition effect for the second branch using the model in (4), and plot the resulting coefficient estimates in Figure 2B, along with the previously estimated time series for the two groups in our main branch. The estimation shows that there is a clear decrease in the disposition

---

comparison correction, we find that the *only* quarter in which there is a significant change between the two groups is the one immediately following the salience shock.

effect among investors at the second branch at the time of the salience shock in October 2004. This suggests that the decrease in the disposition effect for the control group from the main branch is not due to a group specific shock, but is instead driven by time series variation that is common to multiple groups.

If there is indeed a time series component of the disposition effect that is common to all groups, as the data in Figure 2B suggests, we can purge this component from investors in our main branch by subtracting the disposition effect of investors from the second branch. Figure 2C provides the two time series from our main branch after this subtraction, where the vertical axis now provides the disposition effect for the treatment and control group, relative to the disposition effect in the second branch. As expected, we again see no significant difference between the two groups until immediately following the salience shock, and after controlling for the arguably common time series variation, the effect is driven mainly by an increase in the treatment group.

C. *Placebo test*

In order to check the robustness of our difference in differences methodology, we run a placebo test where we re-estimate equation (3), except that we re-define the treatment date using the "wrong" date. We choose the placebo treatment date such that the length of the placebo pre and post periods are identical to those used in our main difference in differences specification. Specifically, we define the placebo pre period from January 2006 through September 2007, and we define the placebo post period from October 2007 through June 2009. As in the main analysis, an investor must trade at least once to be included in the placebo analysis, and we define internet investors as those who submitted at least one trade via the internet during the placebo pre period, and we define non-internet investors as those who never submitted a trade via the internet during the placebo pre period.

Table 4A provides the results from the placebo regression without any controls. The first two columns show that there is still a significant disposition effect among internet and non-internet investors, respectively. Moreover, both groups exhibit a significant increase in the disposition effect after the placebo date. The key test is whether this increase in the disposition effect after the placebo date is similar between the treatment and control groups. The coefficient of interest is again on the triple interaction, *Gain*Post*Internet,* which the third column shows is not significantly different from zero. The economic magnitude is also small. Thus, while the disposition effect does increase in size after the "wrong" treatment date, the important result is that the size of this increase is not significantly different between the treatment and control groups. Table 4B adds a battery of controls that are identical to those used in Table 4B, and again we find that the coefficients of the triple interaction in Column (3) and (6) are not significantly different from zero.

In summary, the results in this section are consistent with Prediction 1. We find that the difference in differences estimate is significantly positive, indicating that the salience shock causes a significant increase in the disposition effect. We demonstrate that the parallel trends assumption holds, and thus our difference in differences estimator is valid. Finally, we show using a placebo test that the treatment effect is zero if we repeat our main analysis using the "wrong" treatment date.

IV. INDIVIDUAL DIFFERENCES AND A TEST OF THE MECHANISM

A. Estimating the Causal Effect at the Individual Level

Thus far we have focused on the change in the *average* investor's trading behavior after the salience shock. However, previous research shows that there is a large amount of heterogeneity in the disposition effect across investors (Dhar and Zhu 2006), and thus it may be that our results are driven by a small number of investors with large portfolios. For example, if investor A holds 10 stocks on day t and investor B holds 2 stocks on day t, investor A will have five times as many stock-day observations, and will be weighted more heavily when estimating the average disposition effect. Because our data enable us to observe trades at the account-day-stock level, we can investigate heterogeneity in the size of the disposition effect across investors and test whether our main result is robust to accounting for this heterogeneity.

To begin our individual level analyses, we estimate an individual measure of the disposition effect for each investor i and each period $p \in \{pre,\ post\}$. We do so by estimating the following model for each investor i, once for data in the pre period and once for data in the post period:

$$Sell_{i,j,t} = \alpha_{i,p} + \gamma_{i,p} Gain_{i,j,t-1} + \varepsilon_{i,j,t} \quad (5)$$

For each investor, this yields two pairs of coefficients, $(\alpha_{i,pre},\ \gamma_{i,pre})$ and $(\alpha_{i,post},\ \gamma_{i,post})$, and the individual level disposition effects are given by $DE_{i,pre} = \gamma_{i,pre}$ and $DE_{i,post} = \gamma_{i,post}$. To be included in the analysis, we require that each investor has, in both the pre and post periods, at least 50 stock-day observations in which he can sell a winning stock and at least 50 stock-day observations in which he can sell a losing stock. This sample restriction is non-trivial and only 39% of investors remain, but nonetheless it is necessary to provide meaningful individual level estimates.

In Figure 3, we plot DE_{pre} vs. DE_{post} for each of the two groups of investors. If an individual investor's disposition effect does not change over time, we expect the data to lie on the black forty-five degree line. There are two basic patterns worth noting from Figure 3. First, in both panels the data are concentrated in the first quadrant, indicating that most investors exhibit a disposition effect in both the pre and post periods; moreover, there is a large amount of heterogeneity in the size of the disposition effect across investors. Second, in both panels, there is a strong positive relationship between the two variables, indicating that the disposition effect is a persistent behavior. The main question we are interested in is whether $(DE_{i,post} - DE_{i,pre})$ is on average, larger for internet investors compared to non-internet investors. Graphically, this amounts to testing whether the degree to which the data lie above the forty-five degree line is larger for internet investors compared to non-internet investors.

To provide a formal test, we use our individual level estimates of the disposition effect as the dependent variable in the following regression:

$$\hat{\gamma}_{i,p} = \mu_i + \theta_1 Post_p + \theta_2 Internet_i * Post_p + \eta_{i,p} \quad (6)$$

$\hat{\gamma}_{i,p}$ is the estimated disposition effect from equation (5) for investor *i* in period $p \in \{pre,\ post\}$, and μ_i is an investor fixed effect. The coefficient of interest is θ_2, which provides an estimate of the difference in differences, taking into account investor fixed effects. Table 5 provides the estimation results for each of ten different sample criteria constraints. For the N=50 criteria that we use to construct Figure 3, the coefficient on the *internet*post* interaction is significantly positive, and therefore the main result is robust to controlling for individual differences in the pre-treatment level of the disposition effect. The table also provides additional specifications where we vary the sample criteria from N=10 to N=100. For sample sizes in the middle range we find significant effects, but the statistical significance gets smaller as we lower the threshold (allowing for noisier estimates) and as we increase the threshold (using only a small subset of the sample).

B. Explaining Variation in the Response to the Salience Shock

Figure 3 demonstrates that there is substantial variation in the disposition effect, but it also shows there is substantial heterogeneity across investors in the response to the salience shock. In other words, when the purchase price becomes more salient, some investors exhibit a strong increase in the disposition effect, while others exhibit little change in behavior. This can be seen by the variation in the distance between each point and the 45-degree line. Why are some investors more prone to the salience shock than others?

The framework we presented in section II provides some guidance to answer this question. The change in the disposition effect is driven by the salience parameter δ, and the size of the disposition effect change is given by $\frac{1}{\delta}$. As Prediction 2 states, variation in the salience effect size should be explained by variation in the salience parameter, δ.

Our empirical strategy to test Prediction 2 can be broken into two steps: first, we estimate a proxy for δ at the individual level, and then we test for a correlation between the proxy and the change in the disposition effect. Our proxy is a recently documented trading pattern, the rank effect (Hartzmark 2015), which is the empirical fact that investors have a higher propensity to sell stocks with highest or lowest returns in their portfolio, relative to stocks with returns in the middle. Hartzmark (2015) suggests that the mechanism behind this effect is driven by salience. If an investor holds multiple stocks in a portfolio, he allocates more attention to those stocks that are salient in the sense that the holding period return is farthest from a reference level return. By paying more attention to these extreme ranked positions, the investor is more likely to trade one of these positions when there is a need to sell. In other words, the salience mechanism is responsible for generating which stocks enter the "consideration set," and only stocks in this set are available to sell.

We estimate the rank effect by running the following regression at the investor level:

$$sell_{i,j,t} = \varphi_i + \delta_i \times extreme_{i,j,t} + \epsilon_{i,j,t} \tag{7}$$

where $sell_{i,j,t}$ is a dummy that takes on the value 1 if the investor *i* sold stock *j* on day *t,* and $extreme_{i,j,t}$ is a dummy that takes on the value 1 if stock *j* has the highest or lowest return in the investor *i*'s portfolio on day *t.* We restrict the sample to days where the investor sold at least one stock and there were at least five stocks in the investor's portfolio. An investor exhibits the rank effect if δ_i is significantly greater than zero.

We are able to estimate the rank effect for 1,376 investors in the pre period. The mean (median) rank effect is 4.687% (4.976%), which indicates that the probability of selling the highest or lowest ranked stock is 4.687% higher than the probability of selling a non-extreme ranked stock (a stock that does not exhibit the highest or lowest return in the portfolio.) The average rank effect is highly statistically significant with a t-statistic of 7.70 and Wilcoxon p-value lower than 0.001. We then perform a median-split on this sample of 1,376 investors, and classify those investors who have a rank effect below 4.976% (the median) as the low rank effect group; the remaining 688 investors are classified into the high rank effect group.

We interpret investors in the high rank effect group as being more prone towards salient thinking than investors in the low rank effect group. After all, investors with a high rank effect are particularly prone to sell extreme ranked stocks, presumably because their attention is heavily allocated towards these salient positions. Therefore, investors in the high rank effect group will have a lower value of δ (more severe salient thinking) compared to investors in the low rank effect group (less severe salient thinking).

Our key empirical test is whether investors in the high rank effect group exhibit a larger change in their disposition effect after the salience shock, compared to those investors in the low rank effect group. We use the individual level pre-period and post-period estimates of the disposition effect from equation (5), for each of the 1,376 investors for whom we were able to estimate the rank effect. Table 6A provides a formal test of the difference in differences[14]. For each specification (each column), the change in the disposition effect is significantly greater for investors in the high rank effect group compared to the low rank effect group. These results are therefore consistent with Prediction 2, as more salient thinking investors (high rank effect group)

[14] Each column in the table provides results from a regression where we require at least N gain observations and at least N loss observations for the investor to be included in the analysis, the same as Table 4.

exhibit a larger increase in the disposition effect after the salience shock compared to less salient thinking investors (low rank effect group).

One potential concern with this analysis is that the rank effect and disposition effect are each estimated as a function of the investor's stock-level holding period returns. It is therefore possible that the two effects have correlated estimation errors. This could lead to a mechanical correlation between the rank effect and the disposition effect, which could bias the results that we report in Table 6B, although the direction of the bias is unclear. To attenuate this concern, we follow Hartzmark (2015) and estimate a modified version of the rank effect that is based on the alphabetical ranking of company name in an investor's portfolio. We re-estimate equation (7), but the key variable $extreme_{i,j,t}$ is now defined to equal 1 if stock j has the highest or lowest ranking when sorted alphabetically by company name, in investor i's portfolio on day t, and 0 otherwise. The company name is prominently displayed in column 2 of the investor's trading platform (see Figure 1).[15]

Critically, we do not use the stock-level holding period return when estimating this version of the rank effect, and thus this should reduce any concern about a mechanical correlation between the disposition effect and the rank effect. We find that on average, investors are more likely to sell stocks at the top or bottom of their portfolio when ranked alphabetically by company name, which is consistent with Hartzmark (2015). Specifically, the mean (median) rank effect is 1.885% (1.586%), which indicates that the probability of selling the highest or lowest ranked stock is 1.885% higher than the probability of selling a non-extreme ranked stock (i.e., a stock that is not first or last when ranked alphabetically by company name.) Consistent with Hartzmark (2015), this version of the rank effect is smaller than the rank effect based on returns, but still

[15] The literal translation of the variable in column 2 of Figure 1 is "short name," which is typically composed of three or four Chinese characters. The "short name" is more comparable to the ticker symbol used in the U.S. market. The StockID in column 1 is a six-digit number and is more comparable to a CUSIP. In addition, the first digit of the StockID indicates a firm's stock exchange, and within each exchange, earlier listed firms also tend to have smaller StockID numbers. We do not find a significant rank effect if stocks are sorted by their StockID, perhaps because investors do not sort based on stock exchange or firm age.

highly statistically significant. The correlation between the two rank effect measures is 0.364 ($p<0.001$). This positive correlation provides extra support that our measures of the rank effect are indeed capturing a tendency for investors to allocate attention towards stocks with salient attributes, and thus, the rank effect is a reasonable proxy for the degree of salient thinking.

Table 5B shows that variation in this version of the rank effect does significantly explain variation in the salience effect. While the results are slightly weaker than in Table 5A, for nearly all columns, the change in the disposition effect is significantly greater for investors in the high rank effect group compared to the low rank effect group. This result therefore provides additional support that the change in the disposition effect is related to the proposed salience mechanism.

These cross-sectional results are also helpful in ruling out alternative explanations. One alternative explanation is that investors may perceive the change in information display as advice from the brokerage house (Benartzi 2001). A second potential explanation is that the effect is driven through the change in font color (Bazley, Cronqvist and Mormann 2017). However, these alternative mechanisms are distinct from salience, and thus they would not generate our second prediction that variation in the change in the disposition effect is correlated with the rank effect. Thus, our cross-sectional result is hard to explain under these alternative theories, but a salience explanation can explain both the average change and the cross-sectional variation result.

V. ROBUSTNESS CHECKS

In this section we provide robustness checks for our first main result of the treatment effect of the salience shock on the disposition effect.

A. Propensity-score matching

In Table 1 demonstrates that there are differences in observable characteristics between our treatment and control groups. Relative to non-internet investors, internet investors have larger portfolios, hold more stocks, trade more frequently, and are also slightly younger. Thus, to ensure that our results are not confounded by these systematic differences, we match investors on these characteristics. In addition, given our focus on the disposition effect, we also match on the disposition effect in the pre period.

Our matching exercise begins with a logit regression at the investor level of a binary variable (indicating whether an investor is an internet investor) on a host of investor characteristics. In particular, we include log portfolio size, average number of stocks, unconditional selling propensity, age, gender, and a measure for the disposition effect which is measured as $\gamma_{i,p}$ from equation (5). All of these variables are measured using data from the pre period. For each internet investor, we identify the non-internet investor that is closest to the internet investor in terms of the propensity score. The match is implemented using a nearest-neighbor propensity score match without replacement.

Of the 1754 internet investors, 576 cannot be matched to any eligible control investor within standard tolerances (specifically, a 0.005 caliper). The final matched sample therefore consists of 1178 internet investors and a corresponding sample of 1178 non-internet investors matched on investor characteristics measured in the pre period.

We estimate equation (3) based on this matched sample. Table 7 shows that the coefficient on the key triple interaction *Gain*Post*Internet,* without controls (Panel A) and with controls (Panel B) is significantly positive. The magnitudes are also similar to the estimates in Table 2. Figure 4 displays a graphical test of the parallel trends assumption among the matched sample. We see that in this matched sample, there is also no significant difference in the disposition effect between the two groups of investors. Taken together, these results suggest that the differences in characteristics between internet and non-internet investors are unlikely to explain our main results.

B. *Extending the post period*

Recall that our main results in Table 2 use a subset of our sample period in order to balance the length of the pre and post periods, and in order to allow sufficient data for a placebo test. We now examine whether the result is robust to including the whole sample, from January 2003 through December 2009. We re-estimate equation (3) with the full sample, and Table 8 shows that the coefficient on the key triple interaction *Gain*Post*Internet* is significantly positive, using three different subsets of control variables.

C. *Break-even price as the reference point*

We examine whether the main result is robust to a different specification of the reference point used to compute the disposition effect. In all previous analyses, we define a stock as trading at a gain if its current price exceeds the weighted average purchase price. However, the break-even price was also added to the display screen of the trading platform at the time of the salience shock. It is therefore important to check how trading behavior might be affected by the addition of this variable. Table 9A provides average selling propensities for gains and losses, defined for both the weighted average purchase price and the break-even price. We see that selling behavior is much more sensitive to the sign of the return when computed using the weighted average purchase price compared to the break-even price. Nonetheless, we still check whether our main difference in differences result is robust to using the break-even price as the reference point. Table 9B shows that the triple interaction is robust to using this alternative reference point. These results suggest that the weighted average purchase price is a more likely candidate for the reference point, but even under an alternative reference point specification, the salience shock still significantly impacts the disposition effect

D. Odean (1998) method

Finally, we check the robustness of our results using the Odean (1998) method. Odean (1998) only considers the investor*days where there is at least one sale and there are at least two stocks available for sale. We re-estimate equation (3) based on the sample satisfying the Odean (1998) criteria, and the results in Table 10 show that they are robust to the Odean (1998) sample selection method.

VI. CONCLUSION

In this paper we provide evidence of a causal link between salience and investor behavior in real world markets. We find that when a brokerage company increases the salience of a stock's capital gain, investors exhibit a significantly larger disposition effect. This change in trading behavior is economically meaningful as we observe a nearly 20% increase in the size of the disposition effect after the addition of the stock's capital gain to the investor's trading screen. While our sample period takes place in the early 2000s, our results are likely to be more applicable today as there has been a large migration towards making investment decisions on a digital platform (Benartzi and Lehrer 2015). In addition, our results complement a large body of empirical work that documents a correlation between proxies for attention and market level outcomes, as we are able to provide a causal interpretation for the impact of attention on investor behavior[16].

We also find substantial heterogeneity in the effect of the salience shock on investor behavior. Our second main result demonstrates that a portion of this variation can be explained with a separate trading behavior: the rank effect. Because the rank effect has previously been attributed to a salience mechanism (Hartzmark 2015), these cross-sectional results provide additional evidence that salience is driving the change in behavior. Moreover, documenting the

[16] See, for example, Barber and Odean (2008), Cohen and Frazzini (2008), DellaVigna and Pollet (2009), Hirshleifer, Lim, and Teoh (2009), Da, Engelberg, and Gao (2011), Lou (2014), and Wang (2017).

correlation structure between trading patterns may help uncover a core set of mechanisms – one of which could be salience – that explains a growing set of behavioral effects (Frydman and Camerer 2016).

Finally, our results provide some validation for an important assumption in several recent models of attention and economic choice (Bordalo, Genniaoli, Shleifer 2012; 2013a; 2013b; Koszegi and Szeidl 2013). In these models, the economic environment endogenously generates a decision-maker's attention allocation. In order for the economic environment to systematically influence choice through this channel, attention must have an impact on choice. Until now, evidence supporting this assumption has mainly come from controlled laboratory settings. Our results provide external validity for this assumption by using data from a more natural and higher stakes financial decision-making environment.

REFERENCES

Barber, Brad M., and Terrance Odean, "Online Investors: Do the Slow Die First?" *Review of Financial Studies,* 15 (2002), 455-487.

Barber, Brad M., and Terrance Odean, "All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors," *Review of Financial Studies,* 21 (2008), 785-818.

Barber, Brad M., and Terrance Odean. 2013. The behavior of individual investors. In *Handbook of the economics of finance*. Eds. G. Constantinides, M. Harris, and R. Stulz. Amsterdam: Elsevier.

Barberis, Nicholas, and Wei Xiong, "Realization utility," *Journal of Financial Economics,* (2012) 104, 251-271.

Bazley, William, Henrik Cronqvist and Milica Mormann, "In the Red: The Effects of Color on Investment Behavior," (2017), *Working paper.*

Benartzi, Shlomo, "Excessive Extrapolation and the Allocation of 401(k) Accounts to Company Stock," *Journal of Finance,* (2001) 56, 1747-1764.

Benartzi, Shlomo and Jonah Lehrer, *The Smarter Screen: Surprising Ways to Influence and Improve Online Behavior.* Penguin Press, (2015).

Benartzi, Shlomo, John Beshears, Katherine Milkman, Cass Sunstein, Richard Thaler, Maya Shankar, Will Tucker-Ray, William Congdon, and Steven Galin, "Should Governments Invest More in Nudging?", *Psychological Science,* (2017) 28, 1041-1055.

Ben-David, Itzhak and David Hirshleifer, "Are Investors Really Reluctant to Realize Their Losses? Trading Resposnes to Past Returns and the Disposition Effect," *Review of Financial Studies,* (2012) 25, 2485-2532.

Bertrand, Marianne and Adair Morse, "Information Disclosure, Cognitive Biases and Payday Borrowing," *Journal of Finance,* 66 (2011), 1865-1893.

Beshears, John, James Choi, David Laibson, and Brigitte Madrian, "Simplification and Saving," *Journal of Economic Behavior & Organization,* 95 (2013), 130-145.

Birru, Justin, "Confusion of Confusions: A Test of the Disposition Effect and Momentum," *Review of Financial Studies,* 28 (2015), 1849-1873.

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer, "Salience Theory of Choice Under Risk," *Quarterly Journal of Economics,* 127 (2012), 1243-1285.

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer, "Salience and Asset Prices," *American Economic Review,* 103 (2013a), 623-628.

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer, "Salience and Consumer Choice," *Journal of Political Economy,* 121 (2013b), 803-843.

Chetty, Raj, Adam Looney, and Kory Kroft, "Salience and Taxation: Theory and Evidence," *American Economic Review,* 99 (2009), 1145-1177.

Choi, James, Emily Haisley, Jennifer Kurkoski, and Cade Massey, "Small Cues Change Savings Choices," *Journal of Economic Behavior & Organization,* (2017), forthcoming.

Cohen, Lauren and Andrea Frazzini, "Economic Links and Predictable Returns," *Journal of Finance,* 63 (2008), 1977-2011.

Da, Zhi, Joseph Engelberg, and Pengjie Gao, "In Search of Attention," *Journal of Finance,* 66 (2011), 1461-1499.

DellaVigna, Stefano and Joshua Pollet, "Investor Inattention and Friday Earnings Announcments," *Journal of Finance,* 64 (2009), 709-749.

Feng, Lei and Mark Seasholes, "Correlated Trading and Location," *Journal of Finance,* 59 (2004), 2117-2144.

Feng, Lei and Mark Seasholes, "Do Investor Sophistication and Trading Experience Eliminate Behavioral Baises in Finance Markets?" *Review of Finance,* 9 (2005), 305-351.

Finkelstein, Amy, "*E-ZTAX*: Tax Salience and Tax Rates," *Quarterly Journal of Economics,* 124 (2009), 969-1010.

Fischbacher, Urs, Gerson Hoffmann, Simeon Schudy, "The Causal Effect of Stop-Loss and Take-Gain Orders on the Disposition Effect," *Review of Financial Studies,* 30 (2017), 2110-2129.

Frydman, Cary, Nicholas Barberis, Colin Camerer, Peter Bossaerts, and Antonio Rangel, "Using Neural Data to Test a Theory of Investor Behavior: An Application to Realization Utility," *Journal of Finance,* 69 (2014), 907-946.

Frydman, Cary and Antonio Rangel, "Debiasing the Disposition Effect by Reducing the Saliency of Information About a Stock's Purchase Price," *Journal of Economic Behavior & Organization,* 107 (2014), 541-552.

Frydman, Cary and Colin Camerer, "The Psychology and Neuroscience of Financial Decision Making," *Trends in Cognitive Science,* 20 (2016), 661-675.

Hartzmark, Samuel, "The Worst, the Best, Ignoring All the Rest: The Rank Effect and Trading Behavior," *Review of Financial Studies,* 28 (2015), 1024-1059.

Heimer, Rawley, "Peer Pressure: Social Interaction and the Disposition Effect," *Review of Financial Studies,* 29 (2016), 3177-3209.

Hirshleifer, David, Sonya Lim, and Siew Hong Teoh, "Driven to Distraction: Extraneous Events and Underreaction to Earnings News," *Journal of Finance,* 64 (2009), 2289-2325.

Imas, Alex, "The Realization Effect: Risk-Taking after Realized versus Paper Gains," *American Economic Review,* 106 (2016), 2086-2109.

Johnson, Eric J., Suzanne B. Shu, Benedict G. C. Dellaert, Craig Fox, Daniel G. Goldstein, Gerald Häubl, Richard P. Larrick, John W. Payne, Ellen Peters, David Schkade, Brian Wansink, and Elke U. Weber, "Beyond nudges: Tools of a choice architecture," *Marketing Letters,* 23 (2012), 487-504.

Kaustia, Markku. 2010. Disposition Effect. In Behavioral Finance, Robert W. Kolb Series in Finance. John Wiley Sons, Inc., Hoboken, NJ.

Koszegi, Botond and Adam Szeidl, "A Model of Focusing in Economic Choice," *Quarterly Journal of Economics,* 128 (2013), 53-104.

Levi, Yaron, "Information Architecture and Intertemporal Choice: A Randomized Field Experiment in the United States," (2017), *Working paper.*

Lou, Dong, "Attracting Investor Attention through Advertising," *Review of Financial Studies,* 27 (2014), 1797-1829.

Madrian, Brigitte and Dennis Shea, "The Power of Suggestion: Inertia in 401k Participation and Savings Behavior," *Quarterly Journal of Economics,* 116 (2001), 1149-1187.

Meng, Juanjuan and Xi Weng, "Can Prospect Theory Explain the Disposition Effect? A New Perspective on Reference Points?," *Management Science,* (2017), forthcoming.

Odean, Terrance, "Are Investors Reluctant to Realize Their Losses?" *Journal of Finance,* 53 (1998), 1775–1798.

Pagel, Michaela, "A News-Utility Theory for Inattention and Delegation in Portfolio Choice," (2017), *Working paper.*

Shaton, Maya, "The Display of Information and Household Investment Behavior," (2017), *Working paper.*

Taubinsky, Dmitry and Alex Rees-Jones, "Attention Variation and Welfare: Theory and Evidence from a Tax Salience Experiment," *Review of Economic Studies,* (2017), forthcoming.

Thaler, Richard and Cass Sunstein, *Nudge.* Yale University Press, (2008).

Wang, Baolian, "Ranking and Salience," (2017), *Working paper*.

Table 1. Internet vs. non-internet investors

This table reports the summary statistics of internet and non-internet investors. Internet investors are defined as those investors that submitted at least one order via computer links in the pre period. Non-internet investors are those who submitted all orders via phone service or cashier window in the pre period. Panel A reports the summary statistics on how investors submit their trades, and Panel B reports the summary statistics on investors' trading activities and demographics. For all the statistics, we first calculate the average of each investor, and then calculate the average across investors.

Panel A. Trade submission types

This panel reports the fraction of trades submitted via computer link for the internet and non-internet investors. We further classify internet investors into two groups: those who submitted *all* of his orders via computer links in the pre period and the ones who submitted part of his orders via computer links in the pre period. The pre period is from January 2003 to September 2004, and the post period is from October 2004 to June 2006. For the pre period, we only report the mean. For the post period, we report the mean, min, P10 (the 10th percentile), Q1 (the 25th percentile), median, Q3 (the 75th percentile), P90 (the 90th percentile), and max across different types of investors.

	Number of investors	Fraction of trades submitted via computer links in the pre period (Fraction Pre)	Fraction of trades submitted via computer links in the post period							
		Mean	Mean	Min	P10	Q1	Median	Q3	P90	Max
Internet Investors	1754	0.877	0.916	0	0.797	0.985	1	1	1	1
Fraction Pre=1	1153	1	0.932	0	0.8	1	1	1	1	1
0<Fraction Pre<1	601	0.640	0.889	0	0.759	0.926	1	1	1	1
Non-internet Investors	1715	0	0.092	0	0	0	0	0	0.375	1

Panel B. Other characteristics

This table presents the summary statistics of investors' trading activities and demographics. Portfolio size is the value (in RMB) of one's portfolio. Number of stocks is the number of stocks in one's portfolio. Unconditional selling propensity is total number of stock*days where there is a sell (partial or full) divided by the total number of stock*days where there is open position at the beginning of the day. Age is the age of an investor. Sex is equal to 1 for male and 0 otherwise. We conduct two tests on the significance of the difference between the internet investors and the non-internet investors: the *t*-test and the Wilcoxon sum rank test.

	Internet Investors		Non-internet Investors		Difference		
	Mean	Median	Mean	Median	Mean	t-stat	Wilcoxon p-value
Portfolio size (RMB)	63916	21853	38458	17553	25458	5.90	<.0001
Number of stocks	2.707	2.047	1.950	1.543	0.758	10.38	<.0001
Unconditional selling propensity (%)	4.225	2.101	2.403	1.163	1.822	9.86	<.0001
Age	51.412	51.000	49.056	48.000	2.356	5.71	<.0001
Sex	0.521	1.000	0.545	1.000	0.023	1.33	0.183

Table 2. The Change in the Disposition Effect

The table provides estimation results from an OLS regression where the dependent variable is a dummy which is equal to 1 if a stock is sold (fully or partially) on a given day, and 0 otherwise. Gain is a dummy variable equal to 1 if the return since purchase is positive and 0 otherwise. Post is a dummy variable equal to 1 for the period between October 2004 and June 2006, and 0 for the period between January 2003 and September 2004. Internet investors are defined as those investors that submitted at least one order via computer links in the pre period. Non-internet investors are those who submitted all orders via phone service or cashier window in the pre period. $\sqrt{TimeOwned}$: Square root of number of days since the position has been open. Log(Buy price): the natural logarithm of the weighted average price. $Return^{-}$: The return since purchase if the return since purchase is negative, zero otherwise. $Return^{+}$: The return since purchase if the return since purchase is positive, zero otherwise. $Volatility^{-}$: The stock volatility calculated using previous 250 days' daily returns if the return since purchase is negative, zero otherwise. $Volatility^{+}$: The stock volatility calculated using previous 250 days' daily returns if the return since purchase is positive, zero otherwise. All coefficients are multiplied by 100. Standard errors are clustered at two dimensions: one at the investor level and one at day level. ***, **, * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A. Difference in Differences Estimation without Controls

	(1) Internet	(2) Non-Internet	(3) All
Gain	4.861***	4.053***	4.053***
	(22.51)	(19.65)	(19.65)
Gain*Post	1.042***	0.192	0.192
	(4.59)	(0.71)	(0.71)
Post	-0.066	-0.289***	-0.289***
	(-0.68)	(-3.98)	(-3.98)
Gain*Post*Internet			0.850**
			(2.41)
Internet			1.036***
			(7.44)
Gain*Internet			0.807***
			(2.70)
Post*Internet			0.224*
			(1.84)
Constant	2.515***	1.479***	1.479***
	(22.49)	(17.87)	(17.87)
Adj-R^2	0.015	0.013	0.017
Obs.	3492530	2354476	5847006

Panel B. Difference in Differences Estimation with Controls

	(1) Internet	(2) Non-internet	(3) All	(4) Internet	(5) Non-internet	(6) All
Gain	6.456***	5.508***	5.642***	6.090***	5.050***	5.124***
	(24.88)	(19.21)	(25.31)	(21.99)	(17.20)	(22.60)
*Gain*Post*	1.612***	0.300	0.028	1.648***	0.285	0.181
	(8.17)	(1.31)	(0.12)	(9.00)	(1.28)	(0.80)
Post	1.751***	0.917***	1.535***	1.553***	0.858***	1.320***
	(13.66)	(8.61)	(14.33)	(13.35)	(8.91)	(14.02)
$\sqrt{TimeOwned}$	-0.275***	-0.151***	-0.225***	-0.377***	-0.223***	-0.322***
	(-21.88)	(-15.35)	(-26.15)	(-19.88)	(-12.57)	(-23.33)
Log (Buy price)	0.047	-0.011	0.071	0.129	0.070	0.141*
	(0.45)	(-0.12)	(0.93)	(1.27)	(0.71)	(1.89)
$Volatility^{-}$	1.539***	0.603	0.936**	2.332***	0.733	1.462***
	(2.61)	(1.22)	(2.25)	(4.51)	(1.62)	(3.89)
$Volatility^{+}$	8.044***	11.185***	8.268***	9.142***	11.080***	8.919***
	(3.31)	(3.89)	(4.14)	(4.22)	(3.97)	(4.89)
$Gain*\sqrt{TimeOwned}$	-0.301***	-0.278***	-0.307***	-0.320***	-0.303***	-0.320***
	(-18.28)	(-13.69)	(-22.41)	(-16.10)	(-13.23)	(-20.14)
$Return^{-}$				18.776***	11.114***	15.925***
				(20.76)	(14.42)	(25.13)
$Return^{+}$				-27.078***	-14.122***	-22.624***
				(-15.02)	(-8.05)	(-16.62)
$Return^{-} * \sqrt{TimeOwned}$				-0.797***	-0.463***	-0.677***
				(-18.18)	(-11.92)	(-21.81)
$Return^{+}*\sqrt{TimeOwned}$				1.344***	0.884***	1.169***
				(13.70)	(8.91)	(15.56)
Gain*Post*Internet			1.738***			1.537***
			(5.55)			(5.24)
Internet			0.866***			0.762***
			(6.89)			(6.86)
Gain*Internet			0.927***			1.038***
			(3.60)			(4.22)
Post*Internet			-0.110			-0.003
			(-0.93)			(-0.03)
Constant	5.026***	3.059***	3.651***	7.274***	4.537***	5.755***
	(18.98)	(11.37)	(18.69)	(21.25)	(11.68)	(22.75)
Adj-R^2	0.040	0.028	0.038	0.046	0.031	0.043
Obs.	3411949	2305713	5717662	3411949	2305713	5717662

Table 3. Main Difference in Differences Estimation with Investor Fixed Effects

The table provides estimation results from a regression where the dependent variable is a dummy which is equal to 1 if a stock is sold (fully or partially) on a day, and 0 otherwise, and in includes both *investor*day* and *investor**gain fixed effect. All coefficients are multiplied by 100. Standard errors are clustered at two dimensions: one at the investor level and one at day level. ***, **, * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A. Difference in Differences Estimation without Controls

	(1) Internet	(2) Non-Internet	(3) All
Gain*Post	0.734***	0.011	0.011
	(2.84)	(0.04)	(0.04)
Gain*Post *Internet			0.723**
			(2.05)
Adj-R^2	0.198	0.225	0.198
Obs.	3070498	1782234	4852732

Panel B. Difference in Differences Estimation with Controls

	(1) Internet	(2) Non-Internet	(3) All	(4) Internet	(5) Non-Internet	(6) All
Gain*Post	1.027***	0.147	0.109	1.011***	0.188	0.135
	(4.25)	(0.62)	(0.40)	(4.21)	(0.79)	(0.50)
$\sqrt{TimeOwned}$	-0.171***	-0.081***	-0.143***	-0.204***	-0.086***	-0.170***
	(-10.50)	(-6.16)	(-11.89)	(-10.33)	(-5.87)	(-11.43)
Log (Buy price)	-0.028	-0.058	-0.012	0.019	0.019	0.038
	(-0.27)	(-0.93)	(-0.17)	(0.17)	(0.31)	(0.50)
$Volatility^{-}$	0.301	0.791*	0.338	0.595	0.964**	0.585
	(0.57)	(1.71)	(0.85)	(1.19)	(2.15)	(1.55)
$Volatility^{+}$	0.513	5.871**	1.301	0.999	5.834**	1.558
	(0.26)	(2.06)	(0.76)	(0.50)	(2.04)	(0.91)
$Gain*\sqrt{TimeOwned}$	-0.183***	-0.125***	-0.179***	-0.189***	-0.117***	-0.178***
	(-9.91)	(-4.19)	(-11.23)	(-9.80)	(-4.21)	(-10.85)
$Return^{-}$				7.752***	2.967***	6.274***
				(9.51)	(5.78)	(10.20)
$Return^{+}$				-11.301***	3.480*	-7.342***
				(-5.03)	(1.85)	(-4.09)
$Return^{-} * \sqrt{TimeOwned}$				-0.316***	-0.098***	-0.253***
				(-8.27)	(-4.84)	(-8.92)
$Return^{+}*\sqrt{TimeOwned}$				0.519***	-0.124	0.345***
				(5.08)	(-1.27)	(4.19)
Gain*Post*Internet			0.955***			0.918***
			(2.83)			(2.75)
Adj-R^2	0.205	0.226	0.204	0.206	0.227	0.205
Obs.	2986018	1729927	4715945	2986018	1729927	4715945

Table 4. Placebo Test for Change in the Disposition Effect

The table provides estimation results from an OLS regression where the dependent variable is a dummy which is equal to 1 if a stock is sold (fully or partially) on a day, and 0 otherwise. The treatment is based on the "wrong date" of October 2007. We define the placebo pre period from January 2006 through September 2007, and we define the placebo post period from October 2007 through June 2009. Please see the caption of Table 2 for the definition of the control variables. All coefficients are multiplied by 100. Standard errors are clustered at two dimensions: one at the investor level and one at day level. ***, **, * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A. Placebo Difference in Differences Estimation without Controls

	(1) Internet	(2) Non-Internet	(3) All
Gain	4.454***	3.774***	3.774***
	(41.18)	(29.76)	(29.76)
Gain*Post	2.270***	2.132***	2.132***
	(17.55)	(13.49)	(13.49)
Post	-1.888***	-0.502***	-0.502***
	(-21.91)	(-4.61)	(-4.61)
Gain*Post*Internet			0.138
			(0.69)
Internet			2.314***
			(14.79)
Gain*Internet			0.680***
			(4.08)
Post*Internet			-1.386***
			(-9.99)
Constant	5.438***	3.124***	3.124***
	(54.03)	(26.09)	(26.10)
Adj-R^2	0.017	0.013	0.017
Obs.	13541625	4268413	17810038

Panel B. Placebo Difference in Differences Estimation with Controls

	(1) Internet	(2) Non-internet	(3) All	(4) Internet	(5) Non-internet	(6) All
Gain	6.421***	5.146***	5.112***	4.264***	4.243***	3.157***
	(49.12)	(25.76)	(39.39)	(32.67)	(18.49)	(23.36)
*Gain*Post*	2.905***	2.651***	2.749***	1.312***	1.184***	1.119***
	(26.81)	(18.00)	(17.68)	(13.15)	(8.65)	(7.50)
Post	-0.701***	-0.223**	-0.120	0.843***	1.013***	1.465***
	(-9.09)	(-2.06)	(-1.01)	(10.23)	(9.51)	(13.31)
$\sqrt{TimeOwned}$	-0.325***	-0.217***	-0.297***	-0.469***	-0.327***	-0.437***
	(-41.70)	(-23.69)	(-47.86)	(-35.39)	(-19.15)	(-40.13)
Log (Buy price)	-0.528***	-0.369***	-0.492***	-0.210***	-0.163**	-0.196***
	(-9.44)	(-4.98)	(-10.56)	(-3.81)	(-2.21)	(-4.24)
$Volatility^{-}$	1.508***	1.594**	1.584***	1.869***	1.882***	1.906***
	(4.65)	(2.08)	(5.31)	(6.11)	(2.60)	(6.77)
$Volatility^{+}$	3.364***	1.711***	2.983***	2.509***	0.929	2.219***
	(5.47)	(2.66)	(6.13)	(3.92)	(1.44)	(4.49)
$Gain*\sqrt{TimeOwned}$	-0.273***	-0.234***	-0.266***	-0.151***	-0.158***	-0.151***
	(-24.27)	(-11.05)	(-24.26)	(-14.21)	(-8.17)	(-14.36)
$Return^{-}$				25.481***	18.279***	23.905***
				(43.95)	(25.45)	(50.06)
$Return^{+}$				-0.182***	-4.480***	-0.194***
				(-6.19)	(-4.64)	(-6.46)
$Return^{-} * \sqrt{TimeOwned}$				-1.153***	-0.813***	-1.080***
				(-33.54)	(-19.98)	(-38.64)
$Return^{+}*\sqrt{TimeOwned}$				0.034***	0.177***	0.036***
				(6.67)	(5.37)	(7.11)
Gain*Post*Internet			0.189			0.312
			(1.00)			(1.06)
Internet			1.537***			1.613***
			(10.44)			(11.97)
Gain*Internet			1.299***			1.199***
			(9.15)			(8.48)
Post*Internet			-0.679***			-0.800***
			(-4.95)			(-6.43)
Constant	9.597***	6.460***	7.709***	11.203***	7.735***	9.188***
	(46.22)	(22.66)	(39.15)	(46.63)	(22.55)	(42.56)
Adj-R^2	0.039	0.030	0.038	0.045	0.035	0.043
Obs.	13461360	4251253	17712613	13461360	4251253	17712613

Table 5. Individual Level Analyses of Disposition Effect

For each investor, we estimate the regression $Sell_{i,j,t} = \alpha + \gamma_{i,p} Gain_{i,j,t-1} + \varepsilon_{i,j,t}$, once for the pre period and once for the post period. This generates two individual specific measures of the disposition effect: $\gamma_{i,Pre}$ and $\gamma_{i,Post}$. We require at least N gain observations and at least N loss observations for the investor to be included in the analysis. The individual level disposition effects are then entered into an OLS regression that includes an investor fixed effect. Each column of the table provides regression results for a different observation cutoff level, N.

	10	20	30	40	50	60	70	80	90	100
Post	0.001	0.127	0.161	0.119	0.096	0.195	0.266	0.375	0.463*	0.304
	(0.00)	(0.64)	(0.79)	(0.56)	(0.43)	(0.85)	(1.09)	(1.47)	(1.70)	(1.05)
Post*Internet	0.502*	0.560**	0.457	0.566**	0.645**	0.575*	0.504	0.392	0.376	0.532
	(1.73)	(2.05)	(1.65)	(1.97)	(2.19)	(1.90)	(1.58)	(1.18)	(1.07)	(1.45)
Adj-R^2	0.389	0.465	0.493	0.492	0.504	0.508	0.488	0.475	0.455	0.445
Obs.	5070	4660	4338	4045	3807	3563	3342	3133	2936	2775

Table 6. Individual Level Analyses of Rank Effect and Disposition Effect

For each investor, we estimate the regression $Sell_{i,j,t} = \alpha + \gamma_{i,p} Gain_{i,j,t-1} + \varepsilon_{i,j,t}$, once for the pre period and once for the post period. This generates two individual specific measures of the disposition effect: $\gamma_{i,Pre}$ and $\gamma_{i,Post}$. We require at least N gain observations and at least N loss observations for the investor to be included in the analysis. In Panel A the individual level disposition effects are then entered into an OLS regression, where "High Rank" is a dummy that takes on the value 1 if the investor exhibits a rank effect above the median (where rank is defined by the stock's holding period return). Panel B is identical to Panel A, except that rank is defined by the alphabetical ordering of the company's name. All regressions include an investor fixed effect. Each column of the table provides regression results for a different observation cutoff level, N. ***, **, * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A. Rank Effect by Returns

	10	20	30	40	50	60	70	80	90	100
Post	0.677	0.901*	1.061**	1.196**	1.193**	1.309**	1.353**	1.495***	1.552***	1.153**
	(1.28)	(1.80)	(2.08)	(2.28)	(2.29)	(2.50)	(2.46)	(2.68)	(2.74)	(2.01)
Post*Internet*	2.267**	2.065**	2.155**	2.306***	2.083**	2.115**	2.226**	2.245**	2.316**	1.910*
High Rank	(2.55)	(2.45)	(2.52)	(2.64)	(2.38)	(2.37)	(2.39)	(2.39)	(2.39)	(1.95)
Post*	-0.765	-0.546	-0.756	-0.869	-1.060	-1.080	-1.189	-1.245	-1.303*	-0.817
High Rank	(-1.07)	(-0.80)	(-1.09)	(-1.22)	(-1.49)	(-1.49)	(-1.56)	(-1.63)	(-1.66)	(-1.03)
Post*	-0.698	-0.785	-0.830	-0.834	-0.546	-0.572	-0.575	-0.737	-0.744	-0.401
Internet	(-1.09)	(-1.30)	(-1.36)	(-1.33)	(-0.88)	(-0.91)	(-0.88)	(-1.11)	(-1.11)	(-0.59)
adjR^2	0.417	0.472	0.478	0.469	0.481	0.487	0.479	0.477	0.459	0.464
Obs.	2501	2409	2338	2269	2208	2140	2067	2020	1956	1893

Panel B. Rank Effect by Alphabetical Company Name

	10	20	30	40	50	60	70	80	90	100
Post	0.342	0.888*	0.840	1.033*	1.038*	1.048*	1.089*	1.345**	1.406**	1.244**
	(0.66)	(1.78)	(1.64)	(1.95)	(1.96)	(1.96)	(1.95)	(2.40)	(2.45)	(2.12)
Post*Internet*	1.340	1.765**	1.450*	1.780**	1.573*	1.184	1.398	1.678*	1.611*	1.859*
High Rank	(1.51)	(2.10)	(1.69)	(2.04)	(1.79)	(1.32)	(1.50)	(1.79)	(1.67)	(1.91)
Post*	-0.156	-0.528	-0.341	-0.560	-0.746	-0.560	-0.666	-0.955	-1.005	-0.956
High Rank	(-0.22)	(-0.78)	(-0.49)	(-0.79)	(-1.04)	(-0.77)	(-0.87)	(-1.24)	(-1.27)	(-1.20)
Post*	-0.276	-0.720	-0.530	-0.632	-0.348	-0.172	-0.229	-0.522	-0.487	-0.514
Internet	(-0.43)	(-1.18)	(-0.85)	(-0.99)	(-0.54)	(-0.26)	(-0.34)	(-0.77)	(-0.70)	(-0.73)
adjR^2	0.415	0.471	0.476	0.467	0.479	0.484	0.476	0.475	0.456	0.463
Obs.	2501	2409	2338	2269	2208	2140	2067	2020	1956	1893

Table 7. The Change in the Disposition Effect-Propensity Score Matched Sample

The table provides the estimation results on the change of the disposition effect of the propensity score matched sample. Internet investors and non-internet investors are matched based on log portfolio size, number of stocks, unconditional selling propensity, age, gender and the disposition effect as estimated from equation (5). The dependent variable is a dummy which is equal to 1 if a stock is sold (fully or partially) on a day, and 0 otherwise. For variable definition, please see the caption of Table 2. All coefficients are multiplied by 100. Standard errors are clustered at two dimensions: one at the investor level and one at day level. ***, **, * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A. Difference in Differences Estimation without Controls

	(1) Internet	(2) Non-Internet	(3) All
Gain	4.437***	4.352***	4.352***
	(19.70)	(17.66)	(17.67)
Gain*Post	1.228***	0.194	0.194
	(4.80)	(0.57)	(0.57)
Post	0.063	-0.347***	-0.347***
	(0.62)	(-3.85)	(-3.85)
Gain*Post*Internet			1.033**
			(2.43)
Internet			0.372**
			(2.49)
Gain*Internet			0.085
			(0.26)
Post*Internet			0.410***
			(3.01)
Constant	1.970***	1.598***	1.598***
	(18.03)	(15.70)	(15.70)
Adj-R^2	0.015	0.014	0.015
Obs.	1866469	1777832	3644301

Panel B. Difference in Differences Estimation with Controls

	(1) Internet	(2) Non-Internet	(3) All	(4) Internet	(5) Non-Internet	(6) All
Gain	5.819***	5.862***	5.703***	5.664***	5.376***	5.326***
	(20.38)	(16.79)	(20.74)	(18.25)	(15.51)	(19.41)
Gain*Post	1.618***	0.348	0.172	1.595***	0.322	0.226
	(7.35)	(1.20)	(0.58)	(7.91)	(1.14)	(0.80)
Post	1.692***	0.968***	1.289***	1.567***	0.908***	1.168***
	(11.65)	(7.15)	(10.51)	(11.79)	(7.40)	(10.46)
$\sqrt{TimeOwned}$	-0.231***	-0.162***	-0.198***	-0.317***	-0.240***	-0.283***
	(-18.13)	(-13.32)	(-22.22)	(-16.30)	(-10.94)	(-19.31)
Log (Buy price)	0.206*	0.000	0.124	0.296***	0.092	0.211**
	(1.76)	(0.00)	(1.45)	(2.63)	(0.74)	(2.48)
$Volatility^{-}$	0.101	0.781	0.371	0.897*	0.727	0.739*
	(0.15)	(1.27)	(0.80)	(1.69)	(1.33)	(1.88)
$Volatility^{+}$	12.705***	10.360***	11.650***	12.571***	10.051***	11.441***
	(4.18)	(2.79)	(4.92)	(4.37)	(2.81)	(5.07)
$Gain{*}\sqrt{TimeOwned}$	-0.279***	-0.295***	-0.291***	-0.323***	-0.318***	-0.322***
	(-13.68)	(-11.46)	(-17.98)	(-13.61)	(-11.18)	(-17.54)
$Return^{-}$				16.478***	11.963***	14.337***
				(17.36)	(12.58)	(21.29)
$Return^{+}$				-24.119***	-15.473***	-20.294***
				(-11.84)	(-7.40)	(-13.77)
$Return^{-} * \sqrt{TimeOwned}$				-0.677***	-0.503***	-0.598***
				(-14.98)	(-10.43)	(-18.11)
$Return^{+}{*}\sqrt{TimeOwned}$				1.273***	0.954***	1.122***
				(11.78)	(8.01)	(13.95)
Gain*Post*Internet			1.584***			1.446***
			(4.27)			(4.14)
Internet			0.289**			0.235*
			(2.10)			(1.92)
Gain*Internet			0.309			0.407
			(1.08)			(1.50)
Post*Internet			0.156			0.202
			(1.16)			(1.62)
Constant	3.829***	3.248***	3.375***	5.792***	4.817***	5.208***
	(14.10)	(9.54)	(14.58)	(16.31)	(9.90)	(17.12)
Adj-R^2	0.037	0.030	0.035	0.042	0.034	0.039
Obs.	1828350	1739037	3567387	1828350	1739037	3567387

Table 8. Main Difference in Differences Estimation Using Full Sample

The table provides estimation results from an OLS regression where the dependent variable is a dummy which is equal to 1 if a stock is sold (fully or partially) on a day, and 0 otherwise. We use the entire sample period from January 2003 through December 2009, where the pre-treatment period starts in January 2003 and ends in September 2004. The post-treatment period starts in October 2004 and ends in December 2009. All coefficients are multiplied by 100. Standard errors are clustered at two dimensions: one at the investor level and one at day level. ***, **, * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A. Difference in Differences Estimation without Controls

	(1) Internet	(2) Non-Internet	(3) All
Gain	4.862***	4.075***	4.075***
	(22.59)	(19.94)	(19.95)
Gain*Post	1.289***	0.308	0.308
	(5.41)	(1.30)	(1.30)
Post	0.416***	1.090***	1.090***
	(4.08)	(11.24)	(11.24)
Gain*Post*Internet			0.981***
			(2.91)
Internet			1.029***
			(7.44)
Gain*Internet			0.807***
			0.787***
Post*Internet			(2.65)
			-0.674***
Constant	2.514***	1.485***	(-4.79)
	(22.57)	(18.09)	1.485***
Adj-R^2	0.017	0.011	0.016
Obs.	5425048	4009224	9434272

Panel B. Placebo Difference in Differences Estimation with Controls

	(1) Internet	(2) Non-internet	(3) All	(4) Internet	(5) Non-internet	(6) All
$Gain$	5.993***	4.306***	4.801***	3.762***	2.192***	2.462***
	(25.26)	(16.00)	(21.98)	(15.06)	(7.88)	(10.10)
$Gain*Post$	2.376***	0.558***	0.823***	2.463***	0.682***	0.860***
	(11.35)	(2.89)	(4.13)	(11.62)	(3.46)	(4.17)
$Post$	1.857***	2.422***	2.449***	1.827***	2.464***	2.489***
	(15.54)	(17.73)	(21.73)	(16.60)	(19.89)	(24.47)
$\sqrt{TimeOwned}$	-0.273***	-0.266***	-0.270***	-0.376***	-0.392***	-0.382***
	(-22.35)	(-20.75)	(-30.54)	(-20.39)	(-17.90)	(-27.08)
Log (Buy price)	0.168*	0.213**	0.177***	0.469***	0.572***	0.510***
	(1.81)	(2.17)	(2.64)	(5.04)	(5.96)	(7.53)
$Volatility^{-}$	-0.540*	-0.155	-0.342	0.282	0.371	0.342
	(-1.74)	(-0.52)	(-1.59)	(1.00)	(1.15)	(1.59)
$Volatility^{+}$	4.646***	2.184**	3.553***	4.078***	1.097	2.842***
	(3.26)	(2.13)	(3.66)	(3.00)	(1.15)	(3.09)
$Gain*\sqrt{TimeOwned}$	-0.236***	-0.128***	-0.200***	-0.158***	-0.042**	-0.115***
	(-17.61)	(-6.71)	(-16.14)	(-11.90)	(-2.24)	(-9.24)
$Return^{-}$				20.954***	19.637***	20.301***
				(23.10)	(20.39)	(30.80)
$Return^{+}$				-2.802*	-5.741***	-3.258**
				(-1.86)	(-8.69)	(-2.16)
$Return^{-} * \sqrt{TimeOwned}$				-0.849***	-0.836***	-0.841***
				(-18.98)	(-17.06)	(-25.50)
$Return^{+}*\sqrt{TimeOwned}$				0.104**	0.202***	0.120***
				(2.32)	(8.18)	(2.70)
Gain*Post*Internet			1.381***			1.425***
			(4.73)			(5.01)
Internet			0.836***			0.713***
			(6.77)			(6.65)
Gain*Internet			0.963***			1.118***
			(3.66)			(4.31)
Post*Internet			-0.610***			-0.664***
			(-4.77)			(-5.83)
Constant	4.831***	3.812***	3.938***	6.861***	5.907***	6.041***
	(17.61)	(13.15)	(20.01)	(20.41)	(15.25)	(24.41)
Adj-R^2	0.046	0.035	0.043	0.051	0.043	0.049
Obs.	5309350	3934767	9244117	5309350	3934767	9244117

Table 9. Robustness Checks Using Break-Even Price as Reference Point

Panel A. Probability of Sale by Sign of Gain for Break-Even and Weighted Average Purchase Price

		Weighted Average Purchase price	
		Gain	Loss
Break-even price	Gain	0.0727	0.0256
	Loss	0.0680	0.0196

Panel B. Difference in Differences Estimation Using Break-Even Price to Compute Gains and Losses without controls

The table provides estimation results from an OLS regression where the dependent variable is a dummy which is equal to 1 if a stock is sold (fully or partially) on a day, and 0 otherwise. *Gain* is a dummy that equals 1 if the current price is above the break-even price. All coefficients are multiplied by 100. Standard errors are clustered at two dimensions: one at the investor level and one at day level. ***, **, * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1) Internet	(2) Non-Internet	(3) All
Gain	4.137***	3.654***	3.654***
	(21.60)	(17.22)	(17.22)
Gain*Post	1.432***	0.182	0.182
	(6.43)	(0.74)	(0.74)
Post	-0.124	-0.371***	-0.371***
	(-1.12)	(-4.50)	(-4.50)
Gain*Post*Internet			1.250***
			(3.78)
Internet			1.206***
			(7.59)
Gain*Internet			0.482*
			(1.69)
Post*Internet			0.246*
			(1.79)
Constant	2.869***	1.663***	1.663***
	(22.14)	(18.05)	(18.05)
Adj-R^2	0.011	0.009	0.013
Obs.	3492530	2354476	5847006

Panel C. Difference in Differences Estimation Using Break-Even Price to Compute Gains and Losses with controls

	(1) Internet	(2) Non-Internet	(3) All	(4) Internet	(5) Non-Internet	(6) All
Gain	6.057***	5.247***	5.479***	5.292***	4.277***	4.498***
	(24.98)	(18.18)	(23.34)	(19.70)	(14.82)	(18.08)
Gain*Post	1.596***	0.189	-0.107	1.549***	0.189	0.020
	(8.24)	(0.88)	(-0.47)	(8.27)	(0.89)	(0.09)
Post	1.862***	1.022***	1.646***	1.723***	0.931***	1.497***
	(13.68)	(8.26)	(13.85)	(13.55)	(8.43)	(13.89)
$\sqrt{TimeOwned}$	-0.303***	-0.169***	-0.249***	-0.435***	-0.257***	-0.366***
	(-21.75)	(-15.47)	(-26.17)	(-19.95)	(-13.15)	(-23.43)
Log (Buy price)	-0.018	0.123*	0.006	0.175**	0.182**	0.160***
	(-0.35)	(1.92)	(0.12)	(2.51)	(2.30)	(2.88)
$Volatility^{-}$	1.500**	0.391	0.942**	1.817***	0.500	1.180***
	(2.37)	(0.78)	(2.12)	(3.06)	(1.03)	(2.77)
$Volatility^{+}$	9.264***	10.767***	8.878***	7.659**	11.023***	7.797***
	(3.33)	(3.45)	(3.99)	(2.41)	(3.54)	(3.11)
$Gain*\sqrt{TimeOwned}$	-0.308***	-0.262***	-0.303***	-0.297***	-0.237***	-0.276***
	(-17.52)	(-13.72)	(-22.27)	(-14.30)	(-11.89)	(-17.75)
$Return^{-}$				14.236***	10.600***	13.187***
				(17.53)	(13.60)	(22.01)
$Return^{+}$				-9.088***	-5.701***	-7.921***
				(-14.49)	(-8.04)	(-16.04)
$Return^{-} * \sqrt{TimeOwned}$				-0.671***	-0.471***	-0.608***
				(-15.64)	(-11.68)	(-19.26)
$Return^{+}*\sqrt{TimeOwned}$				0.520***	0.348***	0.443***
				(15.67)	(10.09)	(17.23)
Gain*Post*Internet			1.862***			1.642***
			(6.10)			(5.49)
Internet			1.019***			1.015***
			(7.10)			(7.44)
Gain*Internet			0.599**			0.671***
			(2.48)			(2.83)
Post*Internet			-0.139			-0.092
			(-1.03)			(-0.72)
Constant	5.716***	3.115***	4.158***	7.673***	4.720***	6.015***
	(23.15)	(13.52)	(24.89)	(22.73)	(12.77)	(25.86)
Adj-R^2	0.038	0.026	0.036	0.043	0.030	0.041
Obs.	3380169	2301054	5681223	3380169	2301054	5681223

Table 10. Robustness checks using Odean (1998) specification

The table provides estimation results from an OLS regression where the dependent variable is a dummy which is equal to 1 if a stock is sold (fully or partially) on a day, and 0 otherwise. The pre-treatment period starts in January 2003 and ends in September 2004. The post-treatment period starts in October 2004 and ends in June 2006. We only include the investor*days where there is at least one sale and there are at least two stocks available for sale. All coefficients are multiplied by 100. Standard errors are clustered at two dimensions: one at the investor level and one at day level. ***, **, * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1) Internet	(2) Non-internet	(3) All
Gain	13.951***	18.049***	18.049***
	(13.45)	(17.26)	(17.26)
Gain*Post	3.172***	-0.567	-0.567
	(3.32)	(-0.46)	(-0.46)
Post	-0.649	3.135**	3.135**
	(-0.88)	(1.98)	(1.98)
Gain*Post*Internet			3.739**
			(2.39)
Internet			-2.740
			(-1.20)
Gain*Internet			-4.098***
			(-2.78)
Post*Internet			-3.784**
			(-2.16)
Constant	20.860***	23.600***	23.600***
	(24.61)	(11.13)	(11.14)
Adj-R^2	0.031	0.032	0.033
Obs.	420689	114305	534994

Figure 1. Change in Information Display. The top panel shows the trading platform screen prior to the salience shock in October 2004. The bottom panel shows the trading platform screen after the salience shock, when five variables were added: 1) break-even price (column #5), 2) weighted average purchase price (column #6), 3) paper gain/loss (column #9), 4) realized gain/loss (column #10) and 5) total gain/loss (column #11). After the salience shock, the font color for each stock was a function of whether the stock was trading at a gain (red) or at a loss (blue). The English translations in both panels and the red rectangles highlighting the added variables in the bottom panel are for display purposes only and were not presented to investors.

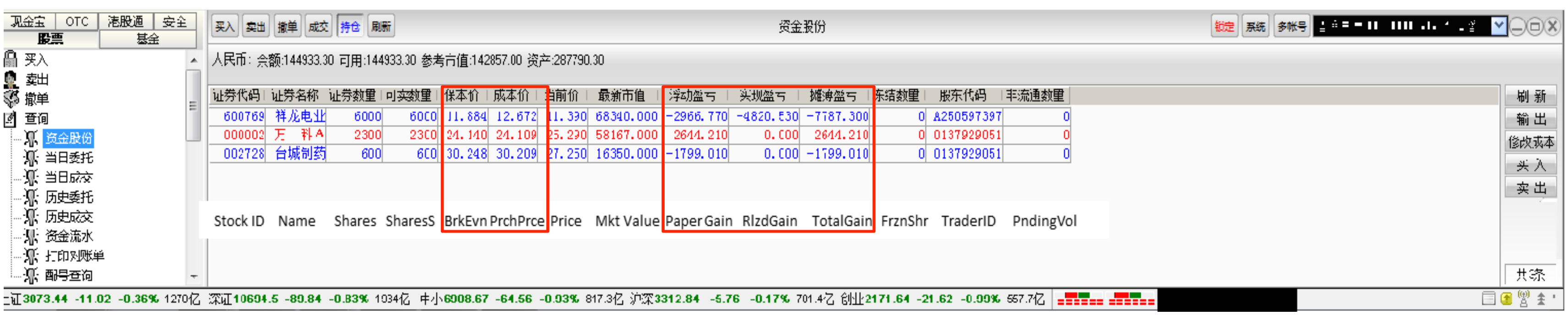
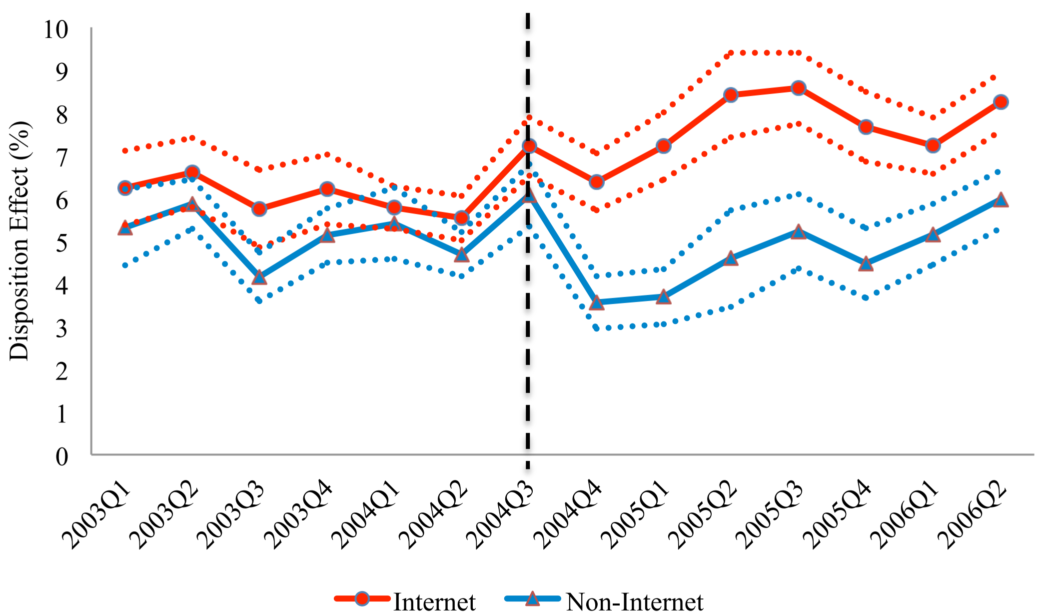


Figure 2. Time Series of the Average Disposition Effect.

Panel A: The figure displays estimation results from the following equation:

$$Sale_{i,j,t} = \alpha + \sum_{q=2003Q1}^{2006Q2} \beta_{int,q}\, Gain_{i,j,t-1} Internet_i D_q + \sum_{q=2003Q1}^{2006Q2} \beta_{nonint,q}\, Gain_{i,j,t-1} NonInternet_i D_q + X\beta' + \varepsilon_{i,j,t}$$

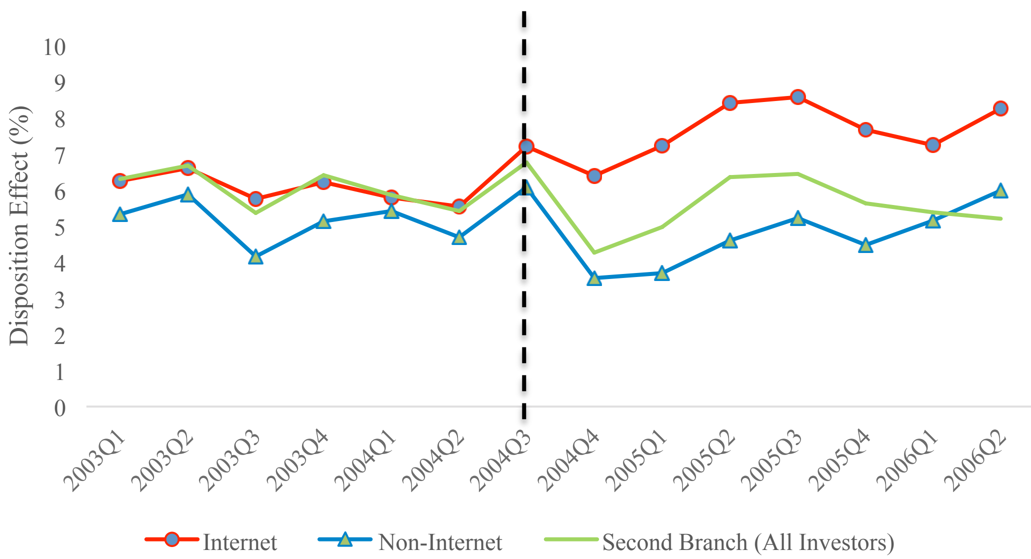
where D_q is a dummy equal to 1 if the observation is in quarter q, $Internet_i$ is a dummy equal to 1 if i is an internet investor and $NonInternet_i$ is a dummy equal to 1 if i is a non-internet investor. X is a vector of control variables, including all the control variables as Ben-David and Hirshleifer (2012), the interaction terms between $Internet_i$ and D_q, and between $NonInternet_i$ and D_q. $\beta_{int,q}$ represents the average disposition effect for internet investors in quarter q, and is plotted below in the red time series. $\beta_{nonint,q}$ represents the average disposition effect for non-internet investors in quarter q, and is plotted below in the blue time series. Colored dotted lines denote 95% confidence intervals around the estimated average disposition effect. The black vertical line denotes the onset of the salience shock.



Panel B. The figure displays estimation results from the following equation:

$$Sale_{i,j,t} = \alpha + \sum_{q=2003Q1}^{2006Q2} \beta_{int,q}\, Gain_{i,j,t-1} D_q + X\beta' + \varepsilon_{i,j,t}$$

X includes time fixed effects D_q and the full controls in Ben-David and Hirshleifer (2012).



Panel C: Time series of the disposition effect relative to investors in the second branch. Same as Panel A except that, for both the internet and non-internet groups, we subtract the time series of the disposition effect from the second branch.

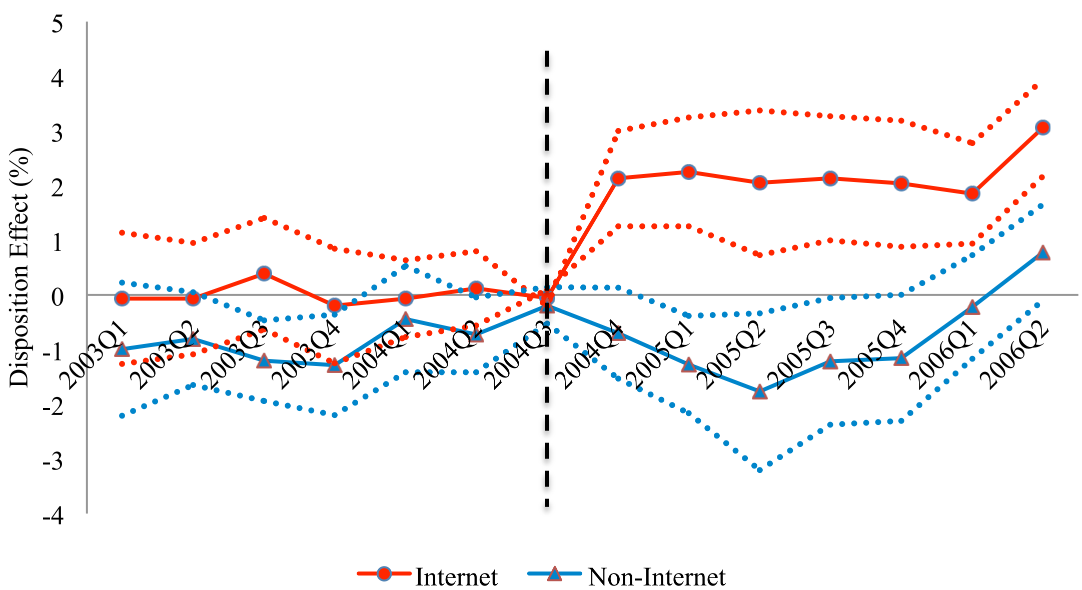
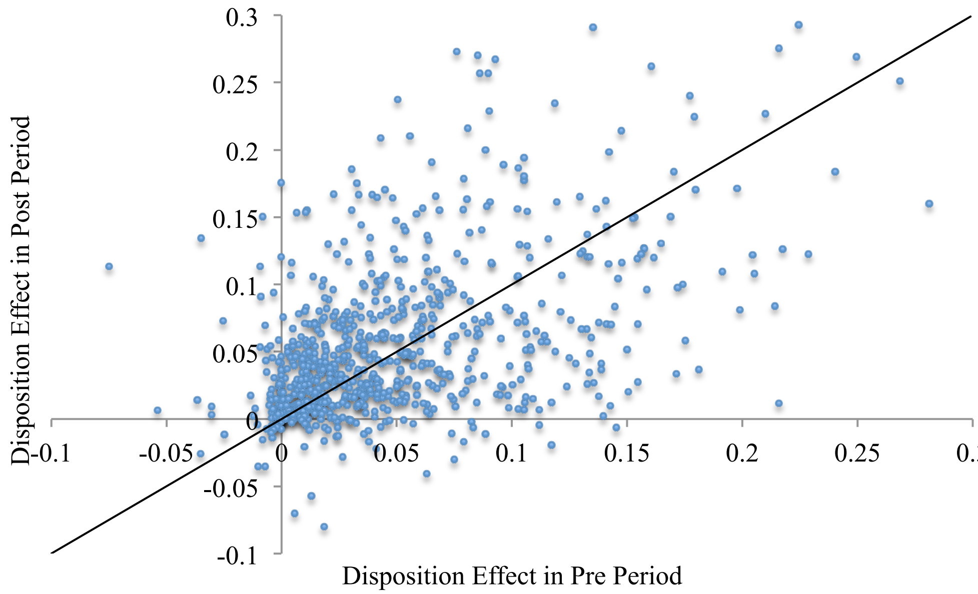


Figure 3. Individual Level Estimates of the Disposition Effect. We estimate the disposition effect for each investor, separately for the pre and post periods using equation (5). Each point represents an investor, and the *x*-axis measures the disposition effect in the pre period while the *y*-axis measures the disposition effect in the post period. The black diagonal line is the 45-degree line.

Panel A. Internet investors



Panel B. Non-Internet Investors

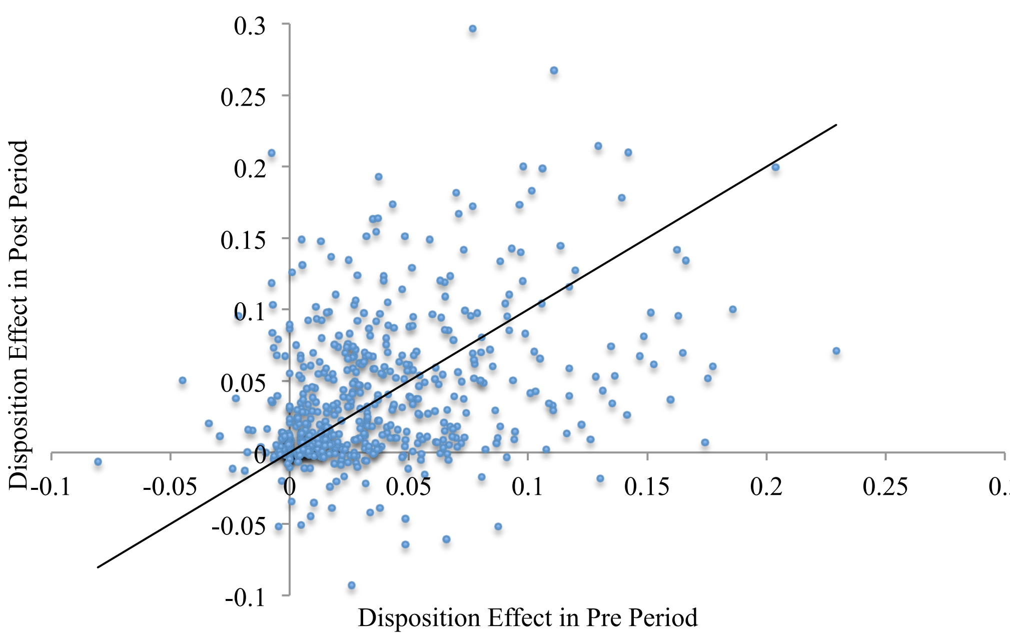


Figure 4. Time Series of the Average Disposition Effect-the Propensity Score Matched Sample

The figure displays estimation results from the following equation based on the propensity score matched sample:

$$Sale_{i,j,t} = \alpha + \sum_{q=2003Q1}^{2006Q2} \beta_{int,q}\, Gain_{i,j,t-1} Internet_i D_q + \sum_{q=2003Q1}^{2006Q2} \beta_{nonint,q}\, Gain_{i,j,t-1} NonInternet_i D_q + X\beta' + \varepsilon_{i,j,t}$$

where D_q is a dummy equal to 1 if the observation is in quarter q, $Internet_i$ is a dummy equal to 1 if i is an internet investor and $NonInternet_i$ is a dummy equal to 1 if i is a non-internet investor. X is a vector of control variables, including all the control variables as Ben-David and Hirshleifer (2012), the interaction terms between $Internet_i$ and D_q, and between $NonInternet_i$ and D_q. $\beta_{int,q}$ represents the average disposition effect for internet investors in quarter q, and is plotted below in the red time series. $\beta_{nonint,q}$ represents the average disposition effect for non-internet investors in quarter q, and is plotted below in the blue time series. Colored dotted lines denote 95% confidence intervals around the estimated average disposition effect. The black vertical line denotes the onset of the salience shock.