The Disposition Effect and The Character of Reference Dependence: Theory and Evidence *

Youssef Benzarti
UC Santa Barbara

Alexander Gelber
UC San Diego

Sreeraahul Kancherla
UC Berkeley

Patrick Langetieg
Internal Revenue Service

Daniel Reck
London School of Economics

Alisa Tazhitdinova
UC Santa Barbara

January 29, 2020

Preliminary and incomplete. Please do not cite or circulate.

Abstract

A central tenet of behavioral finance is that investors are more reluctant to realize losses than gains – the *disposition effect*. Using US population tax data on capital gains realizations, we find that the disposition effect is only exhibited by day traders. The disposition effect disappears for assets held for more than 30 days, and is virtually entirely driven by investors who primarily trade on very short horizons. This finding questions the previously held belief that the disposition effect is a wide-ranging phenomenon affecting most investors. Empirically, the disposition effect results from a large discontinuity in the probability of realizing a gain or loss exactly a zero return. We show theoretically that this discontinuity implies a notch in traders' valuation of the prospect of realizing a gain or loss at zero – conventional prospect theory predicts a kink, not a notch. We conclude that the disposition effect is driven by a small minority of overactive traders, who are subject to a peculiar form of reference dependence.

*Corresponding Author: Daniel Reck: d.h.reck@lse.ac.uk. The authors wish to thank Thummim Cho, Erik Eyster, Francois Gerard, Xavier Jaravel, Camille Landais, Joana Naritomi, and Cameron Peng. Jeanne Bomare provided excellent research assistance. Any remaining errors are our own.

1 Introduction

The *disposition effect* is the propensity to sell assets with accrued gains and hold assets with accrued losses. Beginning with Shefrin and Statman (1985) and Odean (1998), a large literature in behavioral finance reports empirical evidence of such an effect.[1] Much research in the literature also attempts to use theory to explain the disposition effect; theoretical work on this topic has influenced prospect theory and behavioral finance more generally.[2] The disposition effect is by now a core tenet of behavioral finance. It is often exhibited as an important challenge to the efficient markets hypothesis and even as a potential cause of asset price bubbles.

In this paper, we study the disposition effect using population tax data from the United States. Our main empirical result is that the disposition effect is driven by a small subset of highly active, short-horizon investors, the type of investors commonly called day traders. At short holding periods, i.e. assets held less than 10 days before being sold, the probability of realizing a tiny gain is almost twice as high as the probability of realizing a tiny loss, resulting in a discontinuity in the probability of realizing a gain or loss conditional on the paper return at zero. However, the disposition effect, measured as either the overall relative probability of realizing a paper loss versus a paper gain or as a discontinuity in the probability of realizing at zero, declines rapidly at longer holding periods and it disappears beyond 30 days. Moreover, we find that the disposition effect is almost entirely driven by a small subset of very frequent traders who tend to trade on very short horizons, and consequently that the disappearance of the effect at longer horizons is primarily an investor composition effect.

Our main theoretical result explains the discontinuity we observe in the probability of realization at a zero return for short holding periods. Theoretically, the disposition effect is often attributed to loss aversion and "realization utility", whereby investors' incur a separate payoff from each realized rate of return rather than from their entire portfolio value (Barberis and Xiong, 2012; Imas, 2016). We show that the apparent discontinuity in the probability of realization at the threshold between a gain or loss must be driven by a *notch* in agents' valuation of realizing a gain versus a loss.[3] Conventional prospect theory would include a kink in the valuation function at the reference point - zero in our model - but not a notch (Kahneman and Tversky, 1979). Most theoretical literature on the disposition effect, and on reference dependence more generally (see O'Donoghue and Sprenger (2018)), does not consider the implications of such a discontinuity.

We make our main empirical contribution using data reported by brokers to the Internal Revenue Service to administer capital gains taxes. These data cover the population of US taxable investors and have been used by only one other paper (Hoopes et al., 2016). For these investors, we observe realized capital gains

[1] See Barber and Odean (2013) and Hirshleifer (2015) for more comprehensive reviews of the empirical literature.
[2] See Barberis (2018) for a more comprehensive review of the theoretical literature.
[3] Ben-David and Hirshleifer (2012) allude to the possibility of a "sign realization preference" but do not examine it formally.

and losses reported by brokerages to the IRS, separately for each taxpayer, asset, and date of sale. These data are subject to some limitations, but they are far more comprehensive than any financial data used to study the disposition effect. Most of the literature uses data from a single discount broker, as in Odean (1998).[4] Discount brokers have low transaction costs by definition, so they tend to cater to a selected group of disproportionately active traders, especially day traders. Our results suggest that these highly active traders are important for understanding overall stock market activity because they trade so frequently, but also that *most* individual investors behave quite differently - and arguably more rationally.

Methodologically, we primarily analyze the empirical distribution of realized rates of return. We observe realizations directly, as each realization triggers a taxable event. Analyzing realized gains and losses directly allows us to very finely distinguish between gains and losses, which helps to detect a discontinuity. We do not directly observe unrealized capital gains by day. In some analysis, however, we do reconstruct individuals' portfolios based on all observed realized capital gains. To obtain the value of paper asset holdings that are not realized on a given day, we use price data from CRSP. Doing so allows us to estimate the distribution of paper returns and the probability of realizing a paper gain or loss. Our main empirical result is apparent from the observed distribution of realized rates of return alone; reconstructing paper portfolios adds depth to the results and helps us relate them to prior literature and underlying theory.

We document the existence of a large discontinuity in realized rates of return at zero. This discontinuity is apparent in overall data pooling realizations from many different holding periods, and it explains most of the difference in the overall probability of realizing a gain versus a loss. With some exceptions discussed below, much of the literature focuses on the overall probability of realizing a loss versus a gain to study the disposition effect. Rather than focusing on the sign of the return, in contrast, we analyze the rate of return continuously using non-parametric methods. Doing so reveals that the main driver of variation in the overall differential probability of realizing a gain versus a loss is the discontinuity at a zero return.

When we condition on the holding period, we find that the discontinuity at zero declines rapidly as the holding period increases, and it vanishes beyond holding periods of 30 days. Mostly as an apparent result of the vanishing discontinuity, the overall probability of realizing a loss is the same or slightly larger than the probability of realizing a gain at holding periods longer than 30 days. Prior literature has shown in various ways that the disposition effect declines or vanishes at longer holding periods, though we find a more rapid decline than most other papers (e.g. Ivković et al. (2005); Barber et al. (2007); Ben-David and Hirshleifer (2012); Brown et al. (2006)). We are the first to connect explicitly this vanishing of the disposition effect to an active minority of investors. We do so by calculating, for each individual investor, the fraction of the individual's trades done at short horizons - 5 days or less. We then decompose our main result by these

[4]See Table 2 of Barber and Odean (2013) for a review listing some exceptions, which use data from other countries.

investor-level characteristics. We find that the disposition effect is primarily driven by individuals with more than 75% of their trades done on short horizons. Moreover, day traders with more than 75% short-horizon trades display a large disposition effect discontinuity *even when they trade at much longer horizons than 30 days*. Mechanically, short-horizon traders comprise a large fraction of trades at short holding periods and a small fraction of trades at longer holding periods. The absence of these traders at long holding periods thus accounts for the disappearance of the disposition effect. We do find a smaller disposition effect for longer-horizon investors on the infrequent occasions when those investors trade on short horizons, suggesting that such individuals are subject to a disposition effect tendency when they dabble infrequently in day trading.

Evidence for a discontinuity in the probability of realization at zero from the literature is mixed. The only two papers that, to our knowledge, directly test for a discontinuity, find sharply contrasting results[5]. The contrast is especially surprising given that both papers apparently use data from the same discount brokerage - that of Odean (1999) and Barber and Odean (2000)[6]. Ben-David and Hirshleifer (2012) do not find evidence of a discontinuity in the probability of realization at zero using their preferred methods. However, an alternate test for a discontinuity in the Appendix Figure 4 of their paper shows evidence in favor of at least a small discontinuity.[7]Meng (2010) uses a similar test to that of Ben-David and Hirshleifer (2012) main method, and a discontinuity in the probability of realization at a zero return is strongly suggested by the results. The author does not discuss this discontinuity or decompose it, as the paper is focused on other questions. Nevertheless, the reasons for the major difference in the results of these two prior papers is unclear to us. We note, however, that the overall rate of realizing gains and losses is much higher in Meng (2010) and our analogous figures than in the Ben-David and Hirshleifer (2012) figures. We speculate that the divergent results are attributable to a methodological difference between the papers that reduces the overall probability of realization in the latter, which mechanically makes a discontinuity at zero much harder to detect using these methods.

Having shown the importance of a discontinuity in the rate of realization at zero for the overall disposition effect, we next ask theoretically what might be causing such a discontinuity. The early literature informally attributes the disposition effect to prospect theory (Odean, 1998; Shefrin and Statman, 1985): traders are averse to realizing losses and risk-taking in the loss domain, and therefore hold on to losers. This expla-

[5]Two other papers report suggestive evidence for a discontinuity at zero, both using Finnish data. Grinblatt and Keloharju (2001) analyze the distribution of realized rates of return. Kaustia (2010) analyzes the conditional probability of realization given the rate of return. As noted by Ben-David and Hirshleifer (2012), the use of broad bin widths of 1 and 2 percent, respectively, make it difficult to distinguish between a discontinuity and a smooth change in these papers. We use a much smaller bin width of 0.2% below, which minimizes this concern. Our result is apparent with even narrower bin widths, but rounding of purchase and sale prices in the data to the nearest dollar starts to introduce significant noise at narrower bin widths.

[6]Based on the descriptions in these two papers, they apparently use slightly different data files from the same discount broker, which may help explain the divergent results.

[7]We discuss the relationship between the different tests and the contrasting results on this point formally at the beginning of Section 3.3.

nation was challenged by Barberis and Xiong (2009; 2012), who showed that standard prospect theory can only consistently explain the disposition effect with realization utility, where the agent receives a payoff from realization itself (see also Imas (2016)). The fact that the probability of realization depends so sharply on the rate of return in our data is strongly suggestive of realization utility. We show that a model with realization utility and reference dependence can explain the discontinuous disposition effect we observe, but not using the standard formulation of reference dependence, as defined both in its static form in Kahneman and Tversky (1979) or in its stochastic form (Köszegi and Rabin, 2006, 2007, 2009).

The main departure from standard prospect theory required to rationalize what we observe is a *notch* in realization utility. Standard prospect theory supposes that agents evaluate gains and losses relative to a reference point, with a *kinked* utility function, such that the marginal utility of a small loss is discretely larger than the marginal utility of a small gain.[8] This kink is a central feature of prospect theory, and is used to rationalize important departures from canonical expected-utility-maximizing behavior such as risk aversion over small stake lotteries (see O'Donoghue and Sprenger (2018) for a comprehensive discussion), and even some tax sheltering behavior (Rees-Jones, 2017). However, realizing a minuscule gain and realizing a minuscule loss yield an arbitrarily similar payoff with a kinked realization utility function, so the kinked utility function cannot explain why individuals are empirically much more likely to realize the small gain. A utility function with a notch at the reference point, however, makes a small loss realization a discretely worse prospect than a small gain realization, and thus it can explain the discontinuity we observe.

All prior theoretical literature on the disposition effect that we know of, and much other prior work on reference dependence in general, uses a utility function that is continuous, if kinked, at the reference point. There are some exceptions in the broader literature on reference dependence, papers that study the case where reference dependent valuations exhibit a discontinuity at the reference point (Diecidue and Van De Ven, 2008; Payne, 2005; Fishburn, 1977). Our finding that explaining the disposition effect requires notched and reference-dependent utility suggests that further work should analyze notched reference dependence for investor behavior, and possibly for other behaviors. Relatedly, Allen et al. (2016) show that the bunching-type evidence that is attributed to reference dependence in many contexts could be driven by kinked or notched utility. However, our finding that only a minority of highly active traders exhibit this form of reference dependence suggests that further work should account for heterogeneity as well. We discuss some possibilities for future work along these lines in the conclusion.

In spite of two decades of research documenting and explaining the disposition effect, we find that this empirical pattern only holds for a small subset of highly active traders, i.e. day traders. A large majority of

[8] In the standard formulation, utility is usually also concave in the gain domain and convex in the loss domain, a property called *diminishing sensitivity*. We do not focus on this differential curvature, as it cannot explain our observed discontinuity.

investors display little to no disposition effect, and insofar as we believe that the reluctance to realize losses is sub-optimal, most individual investors are more rational than suggested by prior literature on the disposition effect. However, this finding does not imply that the disposition effect is irrelevant for our understanding of financial markets, because a highly disproportionate fraction of all trades are transacted by day traders. For example, we find that 30 percent of all asset sales (in dollars) in our data were sold after being held for one day or less, and nearly half were held for 10 days or less. In back-of-the-envelope counterfactual calculations, we find that these individuals' reluctance to realize very-short-run losses may have reduced sales of losing assets by $200 billion or more over the five-year period of our data, and due to the tax value of loss realizations, it could cause these individuals to pay tens of millions of dollars more in taxes on realized capital gains. Day traders trade so frequently that despite comprising less than 10% of all individual investors, their trades cause us to observe a strong, visible disposition effect in transaction-level distributional analysis that does not condition on holding period or trader characteristics. The fact that day traders dominate transaction-level data so strongly helps explain why prior studies did not uncover the importance of day traders for the overall disposition effect, and the fact that such a large share of transactions are conducted by individuals with strong and peculiar behavioral tendencies may matter a great deal for our understanding of the stock market.

The rest of the paper proceeds as follows. Section 2 discusses the institutional background and our dataset. Section 3 presents the main empirical results. Section 4 derives the theoretical implications of our results. Section 5 concludes.

2 Institutions and Data

2.1 Institutional Background

In the U.S., sales of financial assets generate realized capital gains and losses that are part of taxable income. The amount of taxable income from a realized capital gain or loss is the difference between the sale price and the acquisition price, the latter of which is often called the *cost basis*. The tax treatment of a particular realized gain or loss depends on the taxpayer's earnings, filing status, the amount of other gains and losses realized in a year, and the length of time between acquisition and sale - the *holding period*.

Profits from sales in which the asset was held for a year or less are called short-term gains and are treated as ordinary income subject to the individual's ordinary marginal tax rate (MTR). Assets that were held for over a year qualify for reduced rates, with the top rate ranging between 15% and 20% in 2011-2016. Starting January 2013, high-income individuals are liable for an additional 3.8% Net Investment Income Tax on the

lesser of their net investment income and the amount by which their modified adjusted gross income exceeds $250,000.[9] State taxation of capital gains varies greatly. Many states treat capital gains as ordinary income, while others exempt them altogether. Finally, a number of states offer a preferential treatment by excluding a portion of capital gains from taxation or by offering a credit.[10]

Capital gains taxes are verified through third-party information reporting by brokers on the sales of financial capital assets, which occurs on the Form 1099-B. On the modern Form 1099-B, brokers report the asset sold, including a Committee on Uniform Securities Identification Procedures (CUSIP) number identifying the asset, the sale price, the cost basis, the date sold, and, if all assets sold were purchased on the same day, the date of acquisition, along with a few other details. Cost basis reporting on the form is a relatively recent addition, which went into effect in 2011. Many brokers voluntarily reported cost bases to their customers - but not the IRS - prior to 2011, in order to help customers report these transactions on their tax returns. After the 2011 reform, however, reporting cost basis requires that brokers keep records of the basis or face penalties, so brokers were not required to report the cost basis for assets acquired before January 1, 2011.

As the 1099-B is our main source of data, we should take note of situations where taxpayers might realize capital gains or losses that would not be reportable on this form. First, the 1099-B only reports on trades in taxable financial accounts. We do not observe realizations in tax-exempt pension plans, such as 401(k) pensions, where taxation on the principal or accumulated returns does not occur until funds are distributed from these plans, i.e. usually near retirement. Often, pension funds are not actively managed by taxpayers, with little discretion on the part of individuals as to the exact timing of sales of specific assets in the fund. For further discussion of this issue, see Hoopes et al. (2016). Second, sales of non-financial assets, such as real estate, can trigger capital gains tax obligations, but these do not appear on 1099-B. Our results therefore only describe individuals' behavior with regard to *financial* assets. Finally, mutual fund distributions of capital gains or losses realized within the funds are also taxed as capital gains and losses, but these are reported, together with dividend distributions, on the Form 1099-DIV. As the timing of these distributions are outside the taxpayer's control and they do not appear on Form 1099-B, these are justifiably outside the scope of our analysis.

Taxpayers can use several well-known strategies to minimize their capital gains tax liabilities. In general, tax minimization strategies seek to realize losses early on and defer taxation of gains (Stiglitz, 1983; Constantinides, 1984). Notably, this is exactly the opposite of the disposition effect tendency. First, capital losses can be used to offset both realized gains and some ordinary income. Since each form of losses (short- or

[9]The threshold is set at $250,000 for married individuals, $125,000 for married filing separately and at $200,000 for single individuals or heads of household.

[10]Full exemption: Alaska, Florida, Nevada, New Hampshire, South Dakota, Tennessee, Texas, Washington, and Wyoming; preferential treatment: Arizona, Arkansas, Hawaii, Montana, New Mexico, North Dakota, South Carolina, Vermont, Wisconsin.

long-term) must first be used to offset gains of its own type, short-term losses are most valuable. Any excess losses can then be applied to capital gains of the other type, and the remaining losses can be used to offset ordinary income by up to $3,000 ($1,500 if married filing separately). Any further excess losses – if any – are carried over to future years. Tax-loss harvesting is one of the most popular strategies to minimizing tax liability, and results in a large number of realized losses in December (Ivković et al., 2005).[11] Second, holding on to winning investments for at least one year can reduce taxes directly, because long-term capital gains are subject to lower tax rates. Third, individuals who have purchased multiple sets of the same security can use one of several methods to calculate their cost basis, including the average cost of all the shares, the first-in-first-out method, and others. Selling stock with the highest purchase price first, sometimes called the high-cost method, is often the most advantageous for loss harvesting as it allows taxpayers to maximize the amount of realized losses. Fourth, there are a number of tax considerations surrounding the timing of mutual fund purchases and sales. A purchase of a mutual fund just prior to its annual distribution can result in a unavoidable tax liability on the fund's distribution, which may contain highly-taxed short-term capital gains. On the other hand, a sale just before the distribution results in a larger capital gains tax from the sale itself but avoids taxes on fund distributions. Overall, mutual funds that trade frequently result in larger tax liabilities that are completely outside of an investor's control.

2.2 Data

We use administrative records on the contents of all Form 1099-B's filed by brokers with the IRS, from 2011 to 2015. Brokers are required to file a separate Form 1099-B for each seller, asset, and date of sale. The underlying dataset is therefore large, containing approximately 2.4 billion transactions that record capital gain or loss realizations per tax year. Due to the large sample size, all results in this paper are based on a random sample of transactions by 1% of taxpayers. We make several restrictions to this random sample to arrive at our *main sample* described below.[12]

First, we restrict our sample to sales occurring during 2011-2015 because 2011 is the first year in which cost basis is observed, and because 2015 is the most recent year for which complete data were available at the time of sampling. Second, in order to calculate a rate of return and a holding period, we require that both cost basis and the date of acquisition are present. Since these are generally not available for pre-2011 acquisitions, this requirement eliminates about 48.5% of transactions and 57% of dollars transacted. However, this share decreases significantly over time, as summarized in Appendix Table A.2. In addition to pre-2011 transactions,

[11] An important exception to the loss offset provision are capital losses that result from a sale of stocks that are repurchased back within 30 days. Such "wash sales" losses cannot be used to offset capital gains.

[12] Appendix Table A.1 describes how each of these restrictions affects the amount of trading in the sample: by the number of transactions, taxpayers and brokers, as well as by dollar volume. The underlying dataset is the same as that used in Hoopes et al. (2016), which used 2008-2009 data to study stock sales during the financial crisis.

the acquisition date may be left empty if the securities sold were acquired on a variety of dates. Third, we restrict the sample to transactions where the traders filed an individual tax return - Form 1040 - in the given tax year. [13] Overall, these restrictions result in the main sample of 12,757,981 transactions corresponding to 208,226 individual taxpayers. Finally, we are able to match approximately 92% of these transactions to price data using CUSIP numbers.

Summary statistics for the main sample are available in Table 1. These statistics highlight large heterogeneity: capital gains and losses vary greatly in value, cost basis, and holding periods, while taxpayers exhibit large differences in income, marginal tax rates and trading frequencies. For the purposes of our analysis, two observations are worth highlighting. First, approximately 50% of realizations represent losses rather than gains. However, as it only pertains to realized rather than paper returns, this share is not immediately informative about the disposition effect. Second, within our sample, 99% of realized gains and losses are realized within 3.6 years, and 25% are realized within 30 days of purchase.

Our data is far more comprehensive than most data previously used to study the disposition effect and other questions in behavioral finance, but its construction leads to a few limitations that merit discussion. First, we restrict our attention to U.S. individuals trading in taxable accounts that they own directly. As such, our analysis will not describe trading by institutional traders like pension funds, or individuals trading via private investment companies or offshore accounts that they own. These investments are in general more sophisticated than individual investors trading in taxable accounts. Second, we do not observe cost basis for pre-2011 transactions. In many parts of our analysis we will restrict our attention to relatively short sets of holding periods, such as those less than 30 days, which limits the potential concerns arising from this restriction substantially. Consistent with this idea, in Appendix Figure A.1, we replicate our main result by year, noting that in the year 2015, the number of pre-2011 acquisitions is much more limited than in the overall main sample. Third, we require that the date of acquisition is present, which further eliminates about 14% of observations. We need this restriction because conditioning on holding period is a key part of our analysis. However, in Appendix Figure A.2, we show that, not conditioning on holding period, the main result holds if we include these observations in the estimated distribution of realized rates of return.

[13] Brokers report capital gains or losses realized on assets owned by pass-through entities on Form 1099-B, but we exclude all such sales from our data. Realized capital gains or losses in pass-through entities are taxed as individual capital gains or losses to the ultimate owners. We exclude these sales, however, because we cannot reliably link them to their ultimate individual owners - see Cooper et al. (2016) for a discussion of the difficulties here. Based on what we know about the ownership of these entities, this necessary restriction implies that we do not observe some trades by typically wealthy and highly sophisticated investors.

Transaction-Level details: based on 12,757,981 Form 1099B observations									
Variable	Mean	SD	p10	p25	p50	p75	p90	p95	p99
Value of Stocks/Bonds Sold ($)	5,779	44,471	13	82	522	2,932	10,748	22,527	83,788
Cost Basis ($)	5,660	41,954	14	80	506	2,884	10,539	22,200	82,443
Capital Gain ($)	119	12,960	-233	-16	1	65	404	1,012	4,902
Capital Gain Return	.2	36	-.17	-.045	.0026	.11	.34	.59	1.7
Holding Period (Days)	297	1,181	2	29	154	395	695	904	1,318
Taxpayer-Level details: based on 208,226 Form 1099B-1040 matches									
Age	53	18	28	39	54	67	78	84	92
Male	.63	.48	0	0	1	1	1	1	1
AGI	187,267	1,738,855	14,761	41,816	89,874	167,462	312,992	503,498	1,570,199
Marginal Tax Rate	.21	.081	.1	.15	.23	.26	.33	.35	.4
Long-term Capital Gains Rate	.082	.074	0	0	.11	.15	.15	.16	.2
$\mathbb{1}$(Used Tax Preparer)	.65	.47	0	0	1	1	1	1	1
Net Capital Gain Profit/Loss ($)	8,136	116,232	0	8.4	446	3,137	12,366	25,966	102,313
N of Person-Level Transactions	61	423	1	2	7	30	109	225	946
Person-Level Dollar Sale Volume ($)	354,085	6,765,762	69	1,340	13,887	84,469	378,494	908,498	4,831,547

Table 1: Summary Statistics for the Main Sample.

Notes: The main sample includes observations with nonmissing cost basis and acquisition date from 2011-2015 that were matched to Form 1040 returns. Taxpayers' demographic characteristics were obtained from the matched Form 1040 data. All dollar values are not adjusted for inflation.

Another limitation of the main sample is that we do not observe unrealized capital gains or losses. While we can conduct much of our analysis on only observations of realizations of realized gains or losses, we also construct a *paper returns sample* that does contain data on unrealized gains or losses. As the process is computationally intensive, we do this for 50,000 randomly selected individuals from the main sample. For these individuals, we construct their paper portfolio using data on all their observed purchases and acquisitions. We then match these data to daily asset prices from the Center for Research in Security Prices (CRSP) to assign these assets a daily value based on their CUSIP. The resulting dataset contains the paper holdings at the asset-day-individual level, along with an indicator for whether the asset was sold on that day. The main limitation of this dataset is that we cannot observe assets that were acquired and held during the sample period that were never sold. This concern is most prominent for acquisitions toward the end of the sample period, which are less likely to be sold in the sample period. Like the issue with missing cost basis for pre-2011 acquisitions, this concern is mechanically limited in situations where the investor holds

the asset for very short periods of time. For example, our data should cover essentially all asset holdings that are sold with a holding period of less than 30 days, which is material for our result.

Paper return sample summary statistics are available in Table 2.

Variable	Mean	SD	p10	p25	p50	p75	p90	p95	p99
Transaction-Level details: based on 2,937,331 Form 1099B observations									
Value of Stocks/Bonds Sold ($)	6,168	43,068	13	81	510	2,899	11,064	23,711	94,408
Cost Basis ($)	6,030	41,698	13	79	493	2,844	10,826	23,354	92,810
Capital Gain ($)	138	8,402	-228	-15	1	65	413	1,057	5,316
Capital Gain Return	.19	34	-.17	-.045	.0025	.11	.33	.56	1.4
Holding Period (Days)	308	1,188	4	37	167	407	708	916	1,321
Taxpayer-Level details: based on 50,000 Form 1099B-1040 matches									
Age	53	18	28	39	54	67	78	84	92
Male	.63	.48	0	0	1	1	1	1	1
AGI	181,773	1,113,695	14,794	41,601	89,417	167,441	312,446	505,558	1,550,138
Marginal Tax Rate	.21	.082	.1	.15	.23	.26	.33	.35	.4
Long-term Capital Gains Rate	.082	.074	0	0	.11	.15	.15	.16	.2
$\mathbb{1}$(Used Tax Preparer)	.65	.47	0	0	1	1	1	1	1
Net Capital Gain Profit/Loss ($)	7,977	92,322	0	8.5	445	3,120	12,243	25,755	103,693
N of Person-Level Transactions	59	260	1	2	7	30	106	225	929
Person-Level Dollar Sale Volume ($)	362,338	8,461,625	67	1,277	13,591	82,291	369,958	906,914	4,930,988

Table 2: Summary Statistics for the Paper Returns Sample.

Notes: Transaction-level details are based on a random sample of 50,000 taxpayers from the main sample. All dollar values are not adjusted for inflation.

Altogether, we conclude that despite some caveats and limitations, our data are much more comprehensive than many datasets used to examine the disposition effect, such as the transaction-level data from a single discount broker (see Barber and Odean (2013) for a review).

3 Empirical Results

This section presents our main empirical results. We begin by showing a discontinuity in realized rates of return at zero and an overall disposition effect. We then decompose the distribution of realized rates of return several ways to show how it vanishes at longer holding periods, and how it is driven primarily by day traders. Finally, we analyze the probability of realizing a gain or loss conditional on the rate of return, and formally test for a discontinuity in regression analysis.

3.1 Main Results

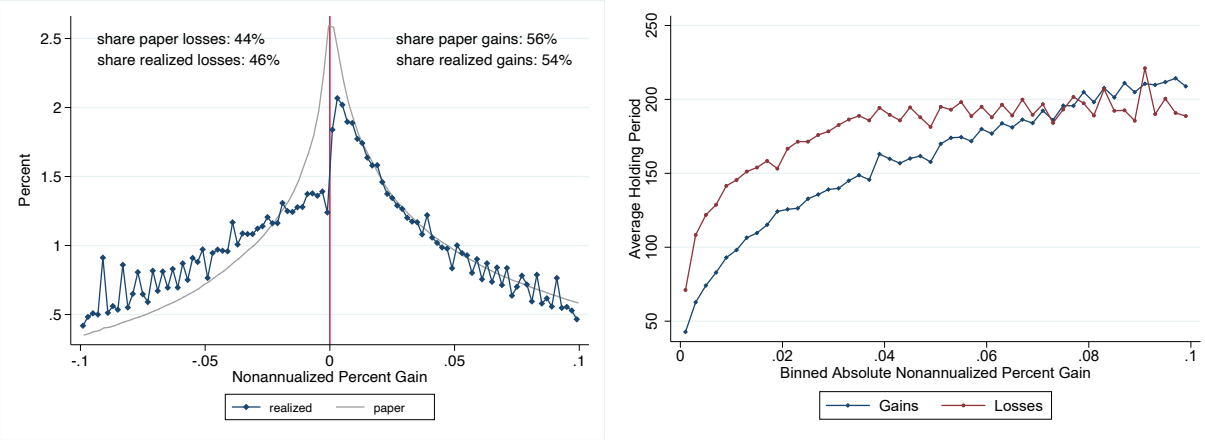
Figure 1 presents our key evidence documenting the disposition effect. The blue curve in Figure 1(a) shows the distribution of realized nonannualized returns for all transactions from our main sample with holding periods between 1 and 800 days.[14] Since we are interested in the relative likelihoods of selling winning and losing securities, we restrict our attention to transactions with nonannualized returns between -10% and 10%, and show the relative probabilities of realizing securities within this range. Figure 1(a) shows a sharp discontinuity at zero: the fraction of realizations in the bins just above zero is approximately 1 percentage point, or about 40 percent, higher than the fraction of realizations in the bins just below zero. Otherwise, the distribution appears to have a slightly heavier right tail, but no obvious missing or excess mass around zero.

The distribution of realized returns, however, depends on the decision to realize and the distribution of available paper returns. A discontinuity at zero could reflect a discontinuity in the realization decision or a discontinuity in the distribution of available returns. To address this possibility, Figure 1(a) overlaps the distribution of realized returns with a comparable distribution of paper returns.[15] This comparison provides clear evidence that the discontinuity at zero is driven by the realization decision rather than the availability of returns: the paper distribution shows no visible discontinuity at zero.

Finally, Figure 1(b) offers a more direct test of the disposition effect by comparing the holding periods of transactions with gains and losses. Here we plot the average holding period by bins of nonannualized returns, separately for transactions with losses and gains. Figure 1(b) confirms that for transactions with nonannualized returns less than about 8 percent, winning securities indeed have shorter holding periods than losing ones. For example, securities that were eventually realized with a positive return of less than 4% were held for approximately 50 days longer than securities that were eventually realized with a negative

[14] Throughout the paper, we omit transactions that were bought and sold on the same day, in order to provide a valid comparison of realized and paper returns. However, the distributions of realized returns show similar patterns when holding periods of zero days are included. See Appendix Figures B.4 and B.5. We also exclude observations of an exactly zero rate of return. Due to rounding sales amounts and cost basis to the nearest dollar on Form 1099-B's, we cannot split the zeros into gains and losses.

[15] While the distribution of realized returns is based on the full main sample, the distribution of paper returns is based on a random sample of paper returns described in Section 2.2. We confirm the comparability of two samples in Appendix Figure A.3. The distributions of realized returns based on the main and on the paper returns samples are very similar.



(a) Distributions of Paper and Realized Returns

(b) Average Holding Period

Figure 1: Evidence of the Disposition Effect

Notes: Figure (a) shows the distributions of nonannualized realized returns and paper returns for transactions with holding periods between 1 and 800 days. Figure (b) shows the average holding period of transactions with holding periods between 1 and 800 days. Realizations with zero returns are excluded. Nonannualized percent gain is calculated as the difference between the sale price and the cost basis divided by the cost basis. The y-axis shows the percent of sales for nonannualized returns between -10% and 10% in 0.2% bins.

return of a similar size. However, for larger returns, the holding periods for losses appear to be similar to or slightly shorter than that of gains. That the overall disposition effect is driven by small returns in absolute value supports our claim, based on Figure 1(a), that the disposition effect is primarily driven by the discontinuity at zero.

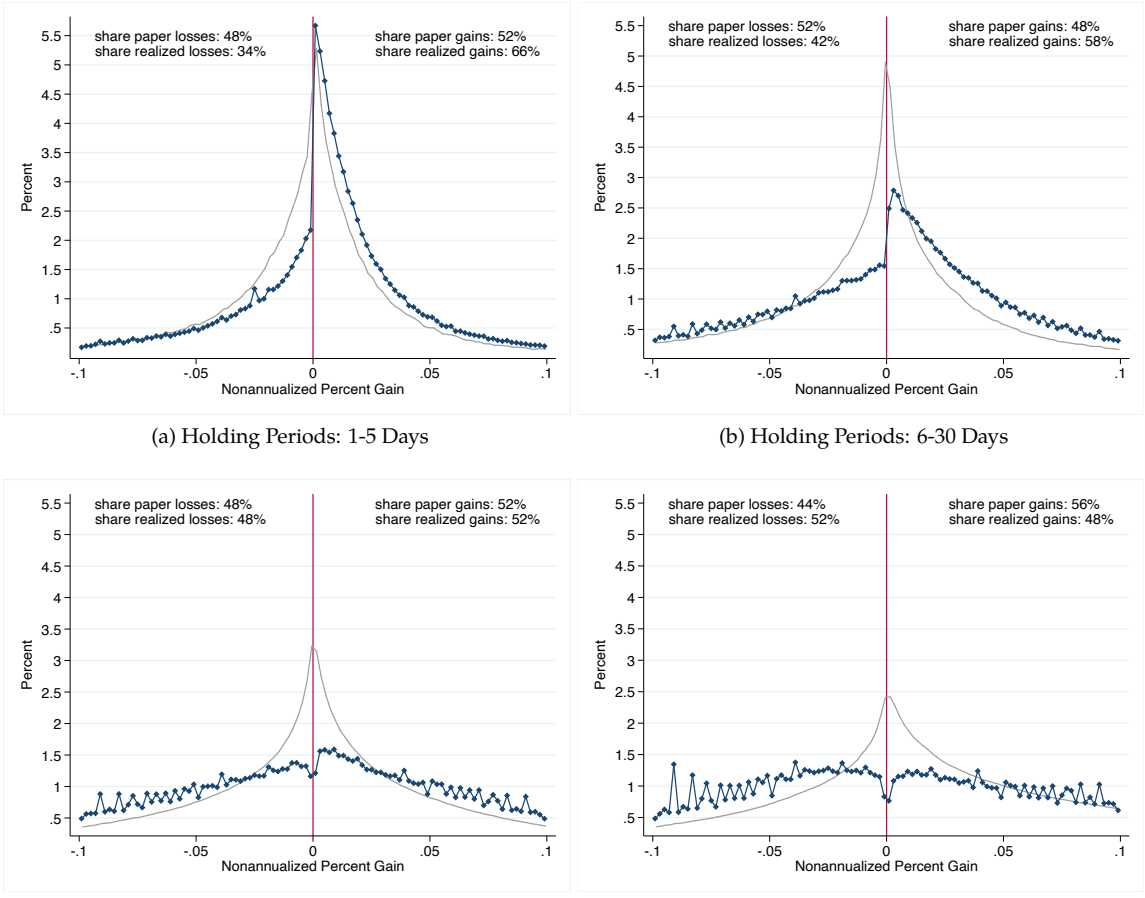
3.2 Heterogeneity and Robustness of Main Results

To better understand what drives the disposition effect and which individuals are prone to it, we now explore how the distributions of nonannualized returns depend on the holding period, cost basis and security type, investors' income and their frequency of trading. Our most important finding is that the disposition effect is limited to transactions with short holding periods and short-horizon traders.

Previous literature documented that the disposition effect appears to be the strongest for shorter holding periods (e.g. Ivković et al. (2005)). Figure 2 breaks down the distribution in Figure 1 by holding period duration. The results show that the entire discontinuity at zero is driven by transactions with holding periods of 30 days or less. We observe a large discontinuity for transactions with holding periods of less than 5 days – a winning security is 79% more likely to be sold than a losing security.[16] This likelihood decreases to 50% for holding periods between 6 and 30 days. At longer holding periods, the disposition effect disappears completely, or even reverses. The likelihood of selling winning and losing securities are the same for holding

[16] To see this, note that the ratio of realized gains to realized losses divided by the ratio of paper gains to paper losses is 1.79.

periods of 31 to 180 days, and at very long holding periods individuals are 27% less likely to sell a winner than a loser.



(a) Holding Periods: 1-5 Days

(b) Holding Periods: 6-30 Days

(c) Holding Periods: 31-180 Days

(d) Holding Periods: 181-800 Days

Figure 2: Distributions of Nonannualized Returns by Holding Period

Notes: This figure shows the distribution of nonannualized realized returns (blue) and the distribution on nonannualized paper returns (grey line) for transactions with holding periods between 1-5, 6-30 days, 31-180, and 181-800 days. Realizations with zero returns are excluded. Nonannualized percent gain is calculated as the difference between the sale price and the cost basis divided by the cost basis. The y axis shows the percent of sales for nonannualized returns between -10% and 10% in 0.2% bins.

The prevalence of the disposition effect at shorter durations cannot be attributed to tax incentives. In fact, the opposite is true: loss harvesting is more attractive at shorter holding periods than longer ones because long-term capital gains are subject to lower marginal tax rates than short-term capital gains. Short-term capital gains – whether held for less than 30 days or more – are subject to the same marginal tax rates. The differences in Figure 2 are also not driven by the relative availability of losses over time. If individuals delay loss realizations, then in the long run, taxpayers' portfolio may accumulate more paper losses than paper gains. This could then mechanically result in more realizations of losses and thus lead to the disappearance of the disposition effect. As we can see in Figure 2, the opposite is true – the distribution of paper returns

shows more winners than losers at higher holding periods, likely because of overall positive drift in stock prices.

Figure 2 thus implies that the disposition effect is not as widespread as has been previously believed. Table 3 allows us to gauge the economic significance of the disposition effect. Among individual investors, 25% of transactions have a holding period of 30 days or less, and 18% have a holding period of 10 days or less. Economically, these transactions represent 55% and 46% respectively of all dollars transacted. These shares are even higher for frequent traders, suggesting that individuals who "day-trade" are most likely to exhibit the disposition effect. The fact that the short-horizon trades are such a large share of overall volume helps explain why we saw an overall disposition effect when pooling across holding periods in Figure 1.

	HP≤1 day	HP≤10 days	HP≤30 days
All traders			
Share of total sale volume in dollars	28%	46%	55%
Share of transactions	9%	18%	25%
Frequent traders (Number of Realizations ≥ 1,000 in 2011-2015)			
Share of total sale volume in dollars	38%	64%	72%
Share of transactions	15%	25%	32%

Notes: This table shows what share of transactions have holding periods of 1 day or less, 10 days or less, or 30 days or less.

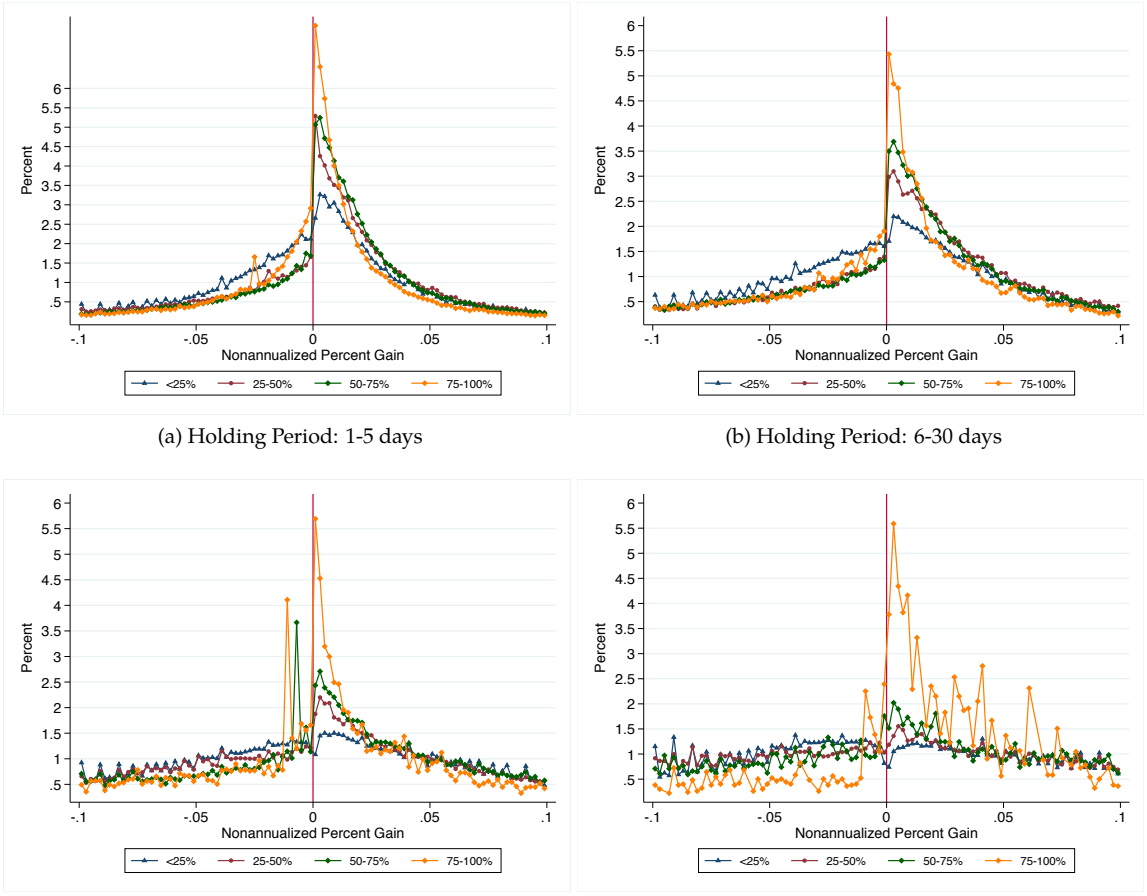
Table 3: Distribution of sales and realizations for the main sample.

Next we explore whether only certain types of traders or transactions are prone to the disposition effect. Since the disposition effect is only observed at short holding periods, it could be driven by investor heterogeneity, i.e. the behavior of day traders. Alternatively, the disposition effect could fade over time for virtually all investors because of the combination of diminishing memory recollection and loss aversion.

Figure 3 provides evidence of both effects, but the dominant force behind fade-out is clearly the investor heterogeneity channel. The figure breaks down individuals into four groups based on their frequency of "day-trading." Specifically, we compare individuals for whom transactions with holding periods of 5 days or less constitute less than 25% of all realizations, 25-50%, 50-75%, or more than 75%.

A clear pattern emerges. For day traders, we see a sharp discontinuity at all holding periods, including for transactions with holding periods of 181 days or more. Infrequent day traders, on the other hand, exhibit no disposition effect at longer holding periods, but are subject to the disposition effect at shorter holding periods (30 days or less), albeit at a lesser degree. Finally, there appears to be a strong increasing relationship between the magnitude of the disposition effect and individuals' frequency of day trading. For example, individuals

who day-trade only 25% of the time exhibit substantially larger disposition effect than those who day trade less.



(a) Holding Period: 1-5 days

(b) Holding Period: 6-30 days

(c) Holding Period: 31-180 days

(d) Holding Period: 181-800 days

Figure 3: Distributions of Nonannualized Returns by Trading Frequency

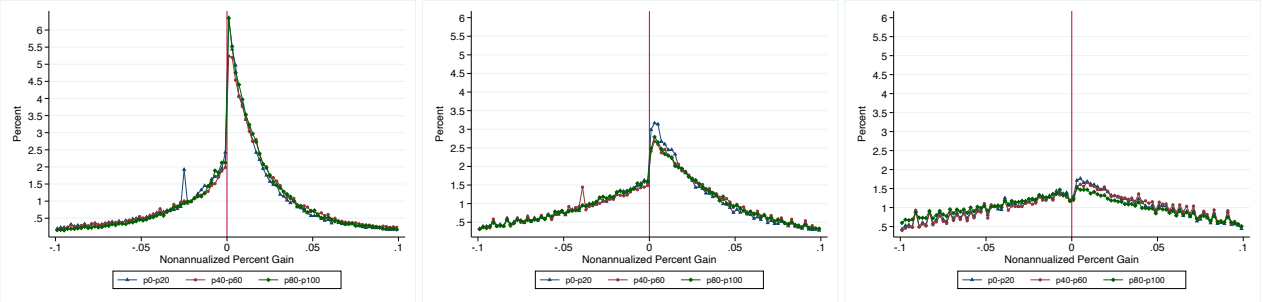
Notes: This figure shows the distributions of nonannualized realized returns for transactions with holding periods between 1-5, 6-30, 31-180, 180-800 days by the share of day-trading transactions. The latter is defined as any transactions with holding period of 5 days or less. Realizations with zero returns are excluded. Nonannualized percent gain is calculated as the difference between the sale price and the cost basis divided by the cost basis. The y axis shows the percent of sales for nonannualized returns between -10% and 10% in 0.2% bins.

As our last piece of evidence, Figure 4 explores how income, cost basis and security type affect the magnitude of the disposition effect.[17] Panel A breaks down individuals into three groups by their adjusted gross income: bottom 20th percentile, middle 40-60th percentiles, and top 20th percentile. We observe a clear discontinuity at zero for transactions with holding periods of 30 days or less for all income groups, and no discontinuity for longer holding periods. While we observe some heterogeneity across income, the disposition effect appears to be driven by holding period duration rather than individuals' income levels.

Panel B explores the importance of cost basis. Since sales amounts and cost bases are rounded to the

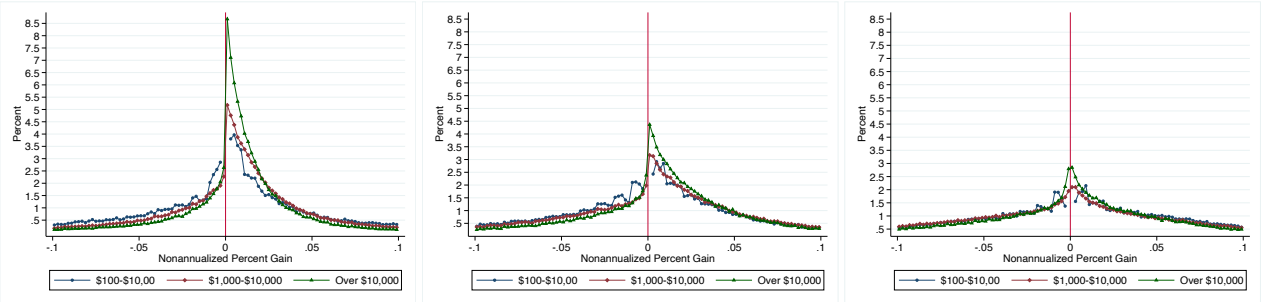
[17] Additional figures are available in Appendix B.

Panel A: By Adjusted Gross Income (AGI)



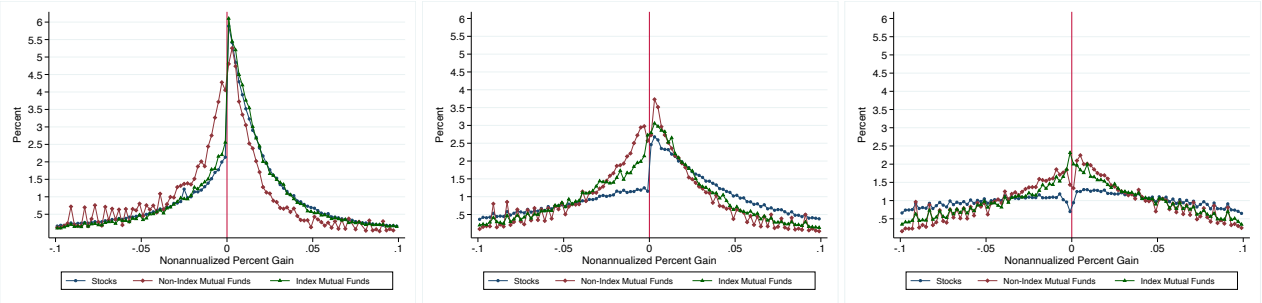
(a) Holding Periods: 1-5 Days (b) Holding Periods: 6-30 Days (c) Holding Periods: 31-180 Days

Panel B: By Cost Basis



(d) Holding Periods: 1-5 Days (e) Holding Periods: 6-30 Days (f) Holding Periods: 31-180 Days

Panel C: By Security Type



(g) Holding Periods: 1-5 Days (h) Holding Periods: 6-30 Days (i) Holding Periods: 31-180 Days

Figure 4: Distributions of Nonannualized Returns by AGI, Cost Basis, and Security Type

Notes: This figure shows the distributions of nonannualized realized returns for transactions with holding periods between 1-5, 6-30, and 31-180 days, broken down by adjusted gross income (AGI) percentiles, cost basis, and asset type. Realizations with zero returns are excluded. Nonannualized percent gain is calculated as the difference between the sale price and the cost basis divided by the cost basis. The y axis shows the percent of sales for nonannualized returns between -10% and 10% in 0.2% bins.

nearest dollar in our data, we cannot test for a discontinuity for transactions with small cost basis and small nonannualized returns.[18] For this reason, the distributions shown in Panel B omit some observations close to zero. Panel B documents that the disposition effect is strongest for transactions with large cost basis, but is otherwise driven by holding periods.

Finally, following Chang et al. (2016), Panel C investigates differences in realizations for four types of securities: stocks, mutual funds, and index mutual funds. Our results are consistent with the previous work on this point: stocks and index mutual funds exhibit large disposition effects, while mutual fund transactions – even with short holding periods – exhibit little to no disposition effect. Chang et al. (2016) attribute this difference to the fact that individuals can blame others for mutual fund losses, while they internalize losses for assets like individual stocks.

3.3 Discontinuity in the Conditional Probability of Realization

So far, we have focused on the distribution of realized rates of return around zero. In this section, we instead estimate the discontinuity in the probability of realizing a capital gain or loss, conditional on the rate of return.

Discontinuities in realized rates of return and in the probability of realizing conditional on the rate of return are closely related, but the latter is more commonly studied in prior literature, and we show in Section 4 that it has a more straightforward relationship to theoretical primitives. To see the relationship between the two discontinuities, let x be the rate of return and let $a \in \{0, 1\}$ denote whether the individual realizes a rate of return in a given period.[19] In the previous section, we analyzed the distribution $p(x|a = 1)$. In this section, we analyze $p(a = 1|x)$. These two are mechanically related by Bayes' rule, $p(a = 1|x) = p(x|a = 1)p(a = 1)/p(x)$, where $p(x)$ is the distribution of paper rates of return, without conditioning on a realization. We assume $p(x)$ is continuous, consistent with the evidence in Figures 1 and 2. Obviously, $p(a = 1)$ does not depend on x. Therefore, a discontinuity at $x = 0$ in $p(x|a = 1)$ implies that there *must* be a discontinuity at $x = 0$ in $p(a = 1|x)$.

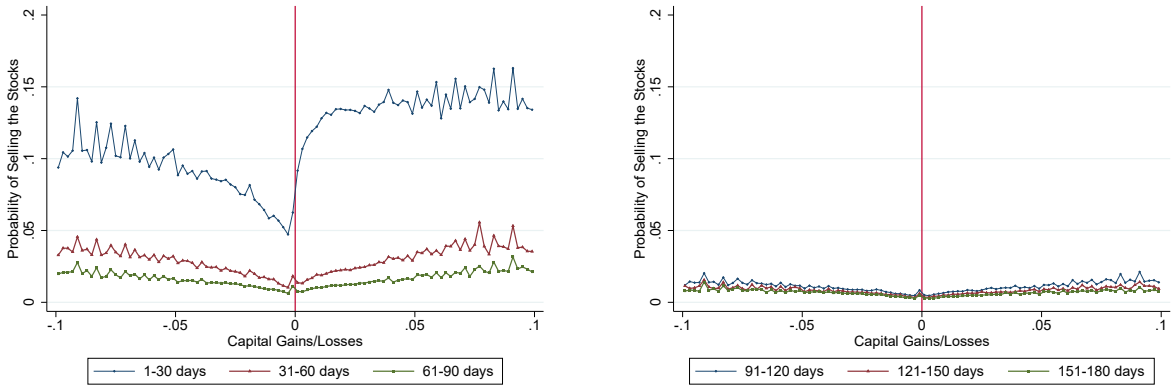
As we alluded to in the introduction, in their work on the data from Barber and Odean (2000), Ben-David and Hirshleifer (2012) do not find evidence for a discontinuity in $p(a = 1|x)$. They do find evidence for a discontinuity in $p(x|a = 1)$, however. They report the latter in Appendix Figure 4 of their paper, obtaining a result that is very similar to what we observe in Figure 1. The above Bayes rule logic suggests that they should also observe a discontinuity at $x = 0$ in their analysis of $p(a = 1|x)$ in the main part of their paper, but they do not detect one. We do not know why Ben-David and Hirshleifer (2012) do not detect a discontinuity

[18] We also show in Appendix Figure B.6 that some of the apparent noise in our figures is driven by rounding issues and small cost bases.

[19] We are implicitly conditioning on the holding period being in a specified range here, which we make explicit in the empirical results.

using similar specifications to those we employ in this section. Meng (2010) estimates $p(a = 1|x)$ on the same data, and the results suggest both a discontinuity and an overall higher probability of realization. One possibility is that the difference in results is driven by the overall lower probability of realization, $p(a = 1)$, in Ben-David and Hirshleifer (2012), which would mechanically make a discontinuity more difficult to detect, even if we know from estimating $p(x|a = 1)$ that such a discontinuity exists.[20]

Because we do not observe unrealized rates of return in the main sample, we estimate $p(a = 1|x)$ on the paper returns sample. [21] Figure 5 plots this distribution. At holding periods between 1 and 30 days, we observe a large discontinuity in the probability of selling stocks where the rate of return just exceeds 0. This discontinuity vanishes at holding periods longer than 30 days, exactly as in Figure 2.



(a) Holding Periods 1-90 Days

(b) Holding Periods 91-180 Days

Figure 5: The Probability of Realization Conditional on the Rate of Return

Notes: This figure plots the probability of selling stock and realizing a gain or loss conditional on the rate of return - $p(a = 1|x)$. The Capital Gain or Loss x is calculated as the difference between the paper value of the stock and the cost basis divided by the cost basis. We observe a large discontinuity at 0 for holding periods less than 30 days but no discontinuity for longer holding periods, consistent with the results in Figure 2.

We next estimate the discontinuity in $p(a = 1|x)$ in a regression context. Doing so allows us to formally test for the presence of a discontinuity at various holding periods, to ensure that the vanishing discontinuity at longer holding periods is not driven by a reduction in the overall baseline probability of realization, and to construct some simple back-of-the-envelope counterfactuals to assess the overall importance of the discontinuity for total sales.

[20] Ben-David and Hirshleifer (2012) suggest that the different result from Meng (2010) is driven by the decision to pool data from many holding periods. We show in Appendix B.7 that the decision to pool across holding periods or analyze them separately does not affect our result of a discontinuity that vanishes at longer holding periods.

[21] Appendix Figure A.3 shows that the distribution of realized rates of returns $p(x|a = 1)$ is virtually identical in the paper returns sample and in the main sample, which is unsurprising given that the paper returns sample is constructed from a random sample of individuals in the main sample.

We estimate regressions of the following form:

$$p(a = 1|x) = \alpha + \beta\{x > 0\} + f(x)\mathbb{1}\{x < 0\} + g(x)\mathbb{1}\{x > 0\}. \quad (1)$$

The parameter of interest is the discontinuity at $x = 0$, denoted by β, and $f(\cdot)$ and $g(\cdot)$ are specified functional forms, such as a quadratic function of x on either side of $x = 0$, with $f(0) = g(0) = 0$.

The methods for estimating β are analogous to those used for regression discontinuity designs (e.g. Gelman and Imbens (2019)). Table 4 presents the baseline results for several alternate specifications. We estimate a discontinuity of between 5 and 7 percentage points for all but the most conservative specification, whereby the discontinuity is 3 percentage points. Dividing by a baseline probability of roughly 8% - using the midpoint method and the specified functional form - we find that this is a 65 to 85 percent increase in the number of realizations just to the right of the zero, compared to just to the left. These point estimates are consistent with what we found using the distribution of realized rates of return at short holding periods in Figure 2.

The most important source of ambiguity across specifications concerns how we treat the bins immediately to the left and right of zero. Due to intra-day price volatility, it is possible that, according to the CRSP data we use to compute returns in the paper loss sample, an asset had a positive return from the date of acquisition to date of observation when the individual was actually in a losing position, or vice versa. This ambiguity causes the estimated conditional probability in the bins just to the left and right of zero to move toward one another, potentially shrinking the estimated discontinuity.[22] Panel A of Table 4 therefore presents results including the bins just to the left and right of zero - i.e. returns within 0.001 of zero - while Panel B excludes these bins. Column (1) and (2) of the Table use simple global linear or quadratic specifications (with different polynomials to the left and right of zero). Column (3) and (4) use a linear and quadratic specification with triangular kernel weighting around zero.[23] We estimate a similar magnitude of the discontinuity across polynomial specifications and a slightly larger magnitude when the bins to the right and left of zero are excluded.

Column (5) presents a much more conservative estimate: a simple difference between the bins immediately to the left and right of zero.[24] We believe this to be an underestimate given the aforementioned measurement issues immediately around zero, but we report it here to show that even despite these issues

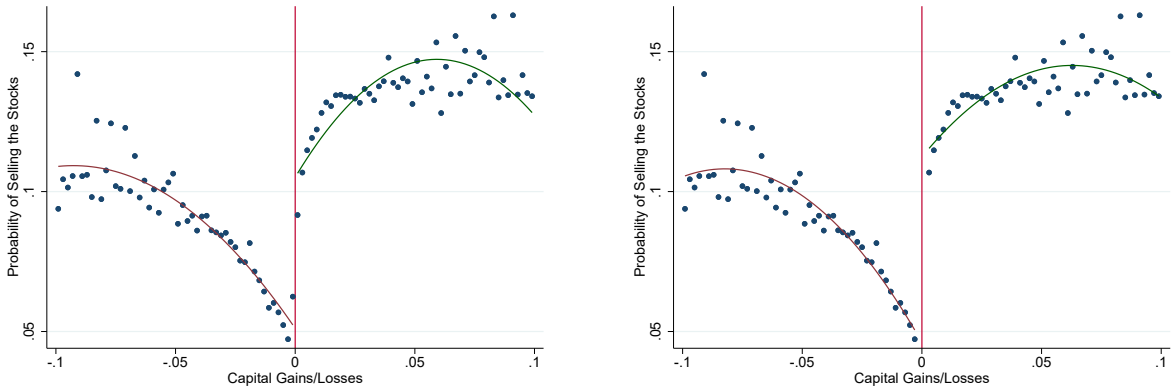
[22] That this pattern is in fact a bias seems plausible when we compare Figure 5 to Figure 2, which is not subject to these measurement concerns because we use actual sale and purchase amounts. Figure 2 does not exhibit a sudden shrinking of the discontinuity in the bins just to the right and left of zero. The closest bins to zero area also subject to rounding problems with small cost basis, see Figure B.6 Panel C in the Appendix.

[23] The bandwidth b is chosen to be 0.1, so that the triangular kernel assigns positive weight to all observations in the range depicted in our figures. For more details on bandwidth selection and kernel weighting, refer to Appendix C.

[24] The estimate from Column (5) is similar to what we would obtain using a kernel-weighted local linear specification with the optimal bandwidth selection algorithms favored in the Regression Discontinuity literature. The large size of our sample favors extremely narrow bandwidths according to the standard selection methods, which makes the aforementioned measurement issues right around zero a concern. We report results from local linear specifications with various bandwidths and discuss this issue further in Appendix C.

we still estimate a sizable discontinuity at zero.

Henceforth, we designate the global quadratic specification from Table 4 as our preferred specification, because it is very simple and we observe in Figure 6 that this specification fits well. Fitted polynomials for the other specifications in Table 4 are in Appendix C.

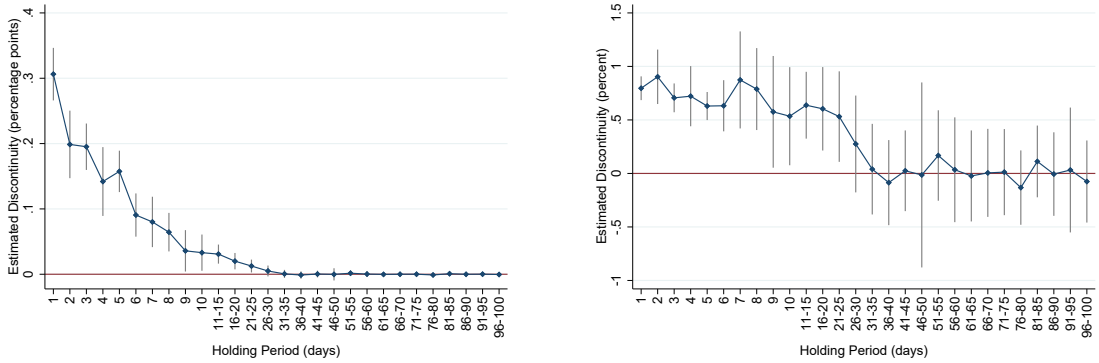


(a) Whole sample

(b) Excluding bins just left and right of 0

Figure 6: Global polynomial of order 2, HP 1- 30

Notes: This figure plots the results of the estimation of equation 1 with quadratic control functions, for transactions with holding periods between 1 and 30 days. Panel (a) includes the whole sample while Panel (b) excludes the bins just left and right of 0 (i.e. Capital gains/losses of -0.001 to 0.001)



(a) Discontinuity (percentage points)

(b) Discontinuity (percent)

Figure 7: Estimated discontinuities as a function of the holding period

Notes: Figure (a) plots the coefficient β of equation 1 with quadratic control functions, estimated over different holding periods. Figure (b) presents the discontinuity in percent estimated over the same holding periods, which corresponds to β divided by the estimated average of the limits on the left and right of 0.

	(1) Global linear	(2) Global pol. order 2	(3) Global linear triangular kernel	(4) Global pol. order 2 triangular kernel	(5) Naive coefficient
Panel A: Whole Sample					
Discontinuity (percentage points)	0.058^{***}	0.054^{***}	0.056^{***}	0.051^{***}	0.029^{***}
	$(1.5 \cdot 10^{-5})$	$(1.5 \cdot 10^{-5})$	$(4.9 \cdot 10^{-6})$	$(4.9 \cdot 10^{-6})$	$(1.0 \cdot 10^{-4})$
Baseline probability (midpoint)	0.086^{***}	0.078^{***}	0.083^{***}	0.077^{***}	-
	$(7.6 \cdot 10^{-06})$	$(7.6 \cdot 10^{-06})$	$(6.3 \cdot 10^{-05})$	$(6.7 \cdot 10^{-05})$	-
Discontinuity (percent)	0.672^{***}	0.691^{***}	0.680^{***}	0.661^{***}	-
	$(1.9 \cdot 10^{-4})$	$(2.0 \cdot 10^{-4})$	$(6.3 \cdot 10^{-5})$	$(6.6 \cdot 10^{-5})$	-
Observations	3,380,180	3,380,180	3,380,180	3,380,180	3,380,180
Panel B: Excluding bins left and right of 0					
Discontinuity (percentage points)	0.065^{***}	0.066^{***}	0.065^{***}	0.065^{***}	-
	$(1.4 \cdot 10^{-5})$	$(1.5 \cdot 10^{-5})$	$(4.3 \cdot 10^{-6})$	$(4.2 \cdot 10^{-6})$	-
Baseline probability (midpoint)	0.088^{***}	0.079^{***}	0.085^{***}	0.078^{***}	-
	$(7.1 \cdot 10^{-6})$	$(7.3 \cdot 10^{-6})$	$(5.4 \cdot 10^{-5})$	$(5.9 \cdot 10^{-5})$	-
Discontinuity (percent)	0.735^{***}	0.836^{***}	0.771^{***}	0.836^{***}	-
	$(1.7 \cdot 10^{-4})$	$(2.0 \cdot 10^{-4})$	$(5.4 \cdot 10^{-5})$	$(5.9 \cdot 10^{-5})$	-
Observations	3,061,524	3,061,524	3,061,524	3,061,524	3,061,524

Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Table 4: Main regression results, HP 1-30

Notes: The analysis sample is restricted to holding periods between 1 and 30 days. The "Naive coefficient" corresponds to a simple comparison of coefficients β_{-1} and β_1, obtained when running the regression: $P_{sale_i} = cst + \sum_{i=-50}^{50} \beta_i X_i + u_i$, X_i a dummy taking one if the observation lies in the i-th bin around 0. The baseline probability corresponds to the mean of the left-limit and right-limit of capital gains/losses around 0. The discontinuity in percent is obtained by dividing the discontinuity (percentage points) by the baseline probability.

In Figure 7, we fully interact our preferred specification with a fine binning of the holding period and then plot the estimated discontinuity as a function of the holding period. Panel 7a reports the percentage point effect and Panel 7b reports the percent change, using the midpoint baseline probability as in the third row of Table 4. We observe that the estimated discontinuity in percentage point terms falls rapidly as the holding period lengthens, from 30 p.p. at 1 day to less than 5 p.p. at 10 days. By 30 days, the discontinuity is a precisely estimated zero. The estimated discontinuity in percentage point terms implies that the probability of selling nearly doubles at the threshold between a gain and a loss at holding periods less than 10 days. We see a similar but less rapid decline in percent terms - the baseline probability of sale is much higher at 1

day than at 10 days. Still, by 30 days, the point estimate in percent terms also falls to zero. The fact that the discontinuity measured as a *percent* falls to zero at longer holding periods ensures that the vanishing of the discontinuity at longer holding periods apparent in Figures 2 and 5 is not attributable to the decline in the baseline probability of selling.

4 Theoretical Implications

In this section we derive the theoretical implications of our main empirical result. Our goal is to understand the implications of the large estimated discontinuity at a zero rate of return for traders' motives to sell. We begin by explaining the implications of this discontinuity with a simple reduced-form model. We then consider how the interpretation of the result changes with the addition of other concerns, such as dynamics and tax planning considerations. Finally, we combine the theoretical and empirical results to construct some simple counterfactual aggregates.

The model considers the binary decision problem of whether to sell or hold an asset in a given time period. We use a static model and focus on a single time period conceptually, but one can also think of this model generating data for many different periods, assuming a stationary environment. A mass of individuals hold an asset with some rate of return, and the individual must decide whether or not to sell the asset and realize the resulting gain or loss, or hold the asset and realize whatever return at some point in the future. We primarily wish to understand how the decision to sell the asset depends on the rate of return that would be realized if the asset is sold in the current period.

Setup. Each individual i holds an asset with a holding period rate of return x. We let the binary action a denote whether the individual sells ($a = 1$) or holds ($a = 0$) the asset. Each individual i chooses a to maximize utility $u_i(a)$. The agent takes action based on the difference in utility between the two actions, which we express in a reduced-form fashion as follows:

$$u_i(1) - u_i(0) = v(x) + \varepsilon. \tag{2}$$

All motives for realizing that depend on the rate of return are therefore summarized by $v(x)$; we normalize $v(0)$ to 0. The variable ε captures *all other motives to sell*, with $x \perp \varepsilon$. We assume the distribution of ε is everywhere continuous with positive support over all $\varepsilon \in \mathbb{R}$; we denote its cdf by $F(\varepsilon)$ and its pdf by $f(\varepsilon)$.

Our payoff from selling nests the notions of "realization utility" in the literature on the disposition effect (Barberis and Xiong, 2012; Imas, 2016). Prior work in the literature suggests that rationalizing a tendency to hold losers and sell winners requires some form of realization utility, where the agent gets a payoff from

realizing that is related to the realized rate of return - and not, for instance, on total wealth. The fact that the probability of realization depends so strongly on the rate of return in our context, especially at zero, is strongly suggestive of realization utility. We remain agnostic about the functional form of $v(x)$, but we note that Barberis and Xiong make $v(x)$ linear (i.e. $v(x) = x$), while Imas (2016) uses a loss averse utility function with a kink at $x = 0$, which accords with canonical prospect theory (Kahneman and Tversky, 1979).

Results. In this model, the probability of selling conditional on holding period rate of return x is given by the following:

$$\Pr(a_i = 1|x) = 1 - F(-v(x)). \tag{3}$$

Direct inspection of equation (3) and our assumption that ε_i has everywhere strictly positive and continuous support leads to the following intuitive result:

Proposition 1. *If* $\Pr(a_i = 1|x)$ *is discontinuous in* x *at* x_0*, then* $v(x)$ *must be discontinuous at* x_0*.*

Proposition 1 states that the large observed discontinuity in $\Pr(a = 1|x)$ at $x = 0$ must be due to the presence of a *discontinuity* in realization utility at $x = 0$.[25] Our empirical result thus contradicts the conventional way of modelling reference dependence.

Dynamics and Tax Planning The discussion so far supposes that the jolt of realization utility that accompanies the realization of a gain or loss is the main concern dictating whether an individual realizes the gain or loss. Here we discuss two other potentially relevant factors: dynamic concerns and tax concerns. We elaborate on these arguments in Appendix D, which extends the model to incorporate concerns that have nothing to do with reference dependence and formally addresses the dynamics issue.

Fundamentally, we argue that no classically rational individual would evaluate the prospect of realizing an extremely small tiny gain as *discretely* better than realizing an extremely small loss. These factors can and almost surely do shape the distribution of $\Pr(a_i = 1|x)$ away from 0 but they cannot introduce a discontinuity at zero.

First we consider dynamic concerns, building on the idea that realizations that are not undertaken today must be undertaken at some point in the future, and that stock prices naturally evolve over time. In Appendix D.2, we consider the sell/hold decision in a model where assets not sold today are sold in the future, potentially triggering a jolt of realization utility in the future. The main insight is that, when calculating the forward-looking value of holding, asset price movements will smooth out any discontinuity in realization

[25] A discontinuity in realization utility could come directly from payoffs, or it could come from belief in regression to the mean in price movements. To generate our results, the individual's beliefs about the future performance of the asset would need to be *discretely* more optimistic conditional on a small loss versus a small gain, which is a strong condition. We do not test whether the form of reference dependence at zero that explains our data is ultimately driven by beliefs or more primitive preferences.

utility in future periods. Thus, a discontinuity in realization utility today implies a discontinuity in utility from realization today, but not in future periods. This implies that the main theoretical result is unaffected by incorporating payoffs from future realizations.

Above, we abstract away from substitution across assets, which is a key feature of portfolio choice models. Barberis and Xiong (2012) show that in a dynamic portfolio choice model, realization utility creates a motive to realize gains today to "lock in" a jolt of positive realization utility, and to hold losers to defer the jolt of potentially negative realization utility into the future. In the model described by Barberis and Xiong, an agent optimally realizes any positive gain immediately and then reinvests in a different stock, and only sells losing positions in response to an exogenous liquidity shock. The model therefore predicts that many agents should sell just as their position enters positive territory (net of transaction costs). Because the only concern governing trading behavior in their model is realization utility, this model cannot explain why some agents hold onto winning positions for a very long time; incorporating broader concerns, especially tax concerns, might make the predictions along these lines more realistic. It is unclear whether a model modified to account for such factors would predict a discontinuity in $p(a = 1|x)$ just above zero. In any case, we note that in their model, agents that sell as soon as they are in a winning position to lock in a positive jolt of realization utility will also re-invest in stock immediately afterward, while those forced to sell a losing position will not. We therefore test, using the paper returns sample, whether the probability that agents purchase stock just after realizing a small gain is higher than after realizing a small loss.

The results of this test are contained in Figure 8. We note that overall reinvestment rates hover around 33 percent, consistent with other evidence (see e.g. Hoopes et al. (2016)). More importantly, we find no discontinuity in the reinvestment rate after a short gain, relative to after a small loss. If anything, reinvestment rates after short-holding-period realizations are slightly higher after a loss than after a gain. This finding suggests that the discontinuity in the main result is unlikely to be driven by the potential consequences of realization utility for forward-looking dynamic portfolio choice highlighted by Barberis and Xiong, but rather is driven by differences in how the agent directly evaluates the prospect of realizing a small gain versus a small loss in the contemporaneous period.

Next, we address tax concerns. As discussed in Section 2.1, loss realizations usually receive a different tax treatment than gains realizations (Ivković et al., 2005). The exact effective tax rates on gains versus losses will differ by individual: those with lots of realized gains already in the tax year will have an especially favorable tax rate on loss realizations, while those with many realized losses will tend to prefer realizing gains. We could incorporate tax planning into our theoretical analysis by supposing there are differential effective tax rate on losses versus gains, which create a payoff from a given realization. However, a realization of zero dollars obviously has no tax consequences. Tax concerns can therefore generate a kink in the utility of selling

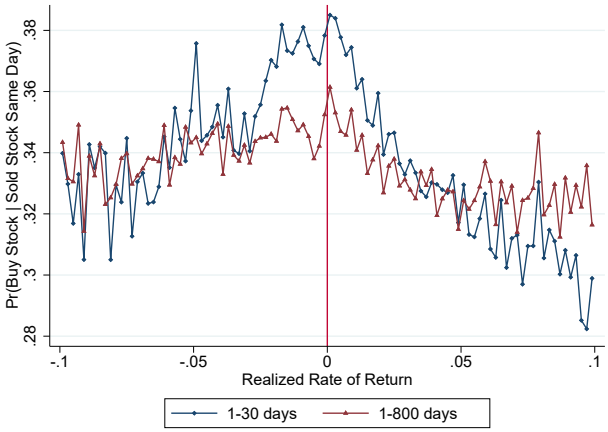


FIGURE 8: PROBABILITY OF BUYING STOCK IF STOCK SOLD THE SAME DAY

Notes: This figure plots the probability that a taxpayer reallocates sale proceeds to another security. Realizations with zero returns are excluded. Nonannualized percent gain is calculated as the difference between the sale price and the cost basis divided by the cost basis.

versus holding at $x = 0$, but they cannot make a tiny gain realization discretely more (or less) valuable than a tiny loss realization.

We therefore conclude that tax considerations may influence the shape of the distribution of realized gains and losses at various holding periods, but they cannot explain the presence of a discontinuity at zero or the fact that it fades out at longer holding periods. Additionally, tax considerations for most investors likely work in the opposite direction: as the vast majority of individuals in our data end the year in a gain position, they would benefit more from realizing losses than gains (see Table 1). To gain some suggestive evidence on the influence of tax concerns for the shape of distributions, we leverage the fact that a significant share of loss harvesting happens at the end of the tax year, when individuals' overall position for the year is known, in December (Poterba and Weisbenner, 2001). Many papers on the disposition effect document a reduced effect in December attributable to tax loss harvesting (e.g. Odean (1998); Barber and Odean (2004); Ivković et al. (2005)). Figure 9 splits the main result from Figure 1 by whether the realization occurred in December. Consistent with prior work, the overall number of realized losses is much larger in December, when significant loss harvesting occurs. There is still a discontinuity at zero in the December distribution of realized gains and losses, but the discontinuity is somewhat smaller because of the overall increase in loss realizations.

In summary, portfolio allocation strategies and tax concerns will shape the distributions we have been analyzing, but accounting for these factors does not alter the main implication of the discontinuity at zero. Altogether, we cannot come up with no reason why the empirical discontinuity we observe in the probability of realizing a gain or loss at zero does not imply that investors' loss aversion causes them to evaluate the

prospect of realizing a very small gain as discretely better than a very small loss. We acknowledge that further research with this data could help us to better understand the influence of the other factors discussed here on investor behavior.

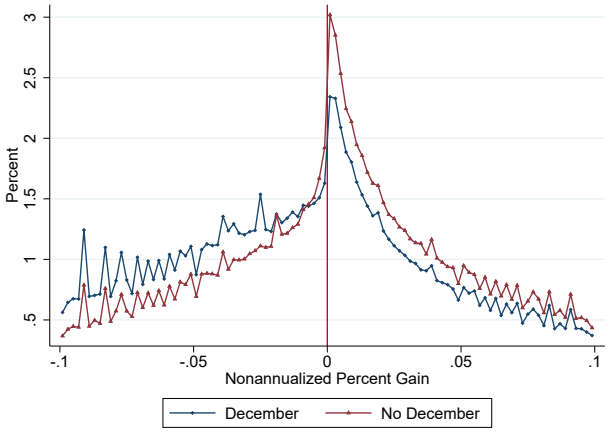


Figure 9: Distribution of Nonannualized Percent Gain of December vs No December transactions

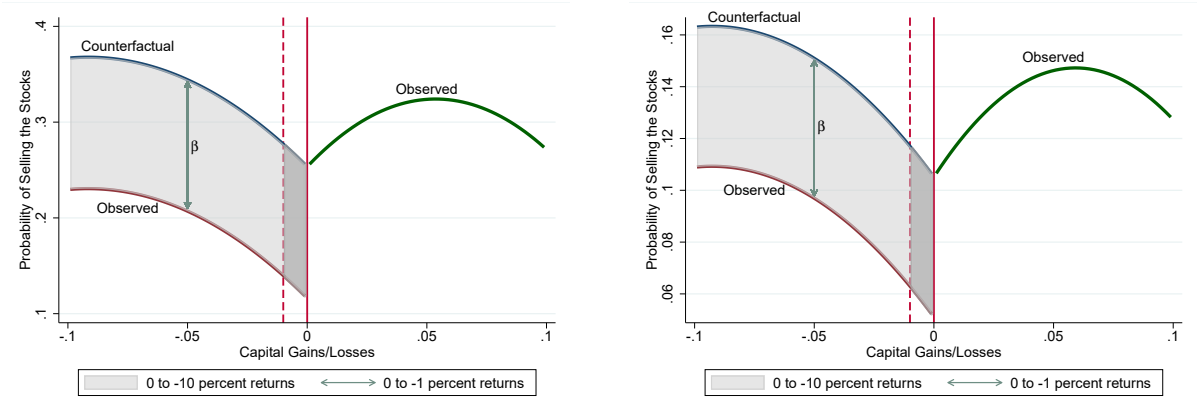
Notes: This figure shows the distribution of nonannualized realized returns for transactions with holding periods between 0 and 800 days, for December trades (blue line) and non-December trades (red line). Nonannualized percent gain is calculated as the difference between the sale price and the cost basis divided by the cost basis. The y-axis shows the percent of sales for nonannualized returns between -10% and 10% in 0.2% bins.

Counterfactual Analysis We next conduct some simple counterfactual analysis using our estimated discontinuity to gain a better sense of the overall consequences of this discontinuity for aggregate sales volume by the taxable individual traders that comprise our sample. Supposing that $v(x)$ contains a discontinuity at zero, our goal is to compare aggregates to a counterfactual in which that discontinuity were eliminated by increasing $v(x)$ in the loss domain.

Figure 10 illustrates the set of potential sales that we presume would have occurred in absence of the disposition effect for each of our counterfactual calculations. We construct a back-of-the-envelope counterfactual in five simple steps. First, we specify a holding period window, e.g. 1 to 10 days, and a range of negative rates of return x, e.g. 0 to -1 percent, over which to calculate counterfactual dollar amounts. Second, we estimate the regression from Equation (1) for our preferred (quadratic) specification within the specified holding period window. Third, for all bins of x in the specified range of negative returns, we calculate the total sales amounts and total capital losses in those bins during the specified holding period window for the full sample period.[26] Fourth, we multiply these totals by the estimated regression coefficient, β in Equation (1), to simulate the effect of increasing the probability of sale in the loss domain on total sales and total

[26]This counterfactual assumes a homogeneous discontinuity by cost basis and over time for simplicity. The finding in Figure 4 that our results are mostly driven by large cost bases suggests that the dollar amounts here are, if anything, a lower bound.

capital losses over the range of specified negative returns x. Fifth, we scale these up by the appropriate sampling weight - about 414 - to account for the fact that these data are coming from a random sample of the full population of individual taxable investors covered by Form 1099-B.



(a) Holding Period 1-10 days

(b) Holding Period 1-30 days

FIGURE 10: COUNTERFACTUAL FIGURE OF THE PROBABILITY OF SELLING STOCKS, FOR DIFFERENT HOLDING PERIODS

Notes: This figure plots a range of potential sales that could have happened absent the disposition effect. The light-grey and dark-grey areas represent the additional sales that could have been realized between 0 and -10% returns. The dark-grey area represents the additional sales that could have been realized between 0 and -1% only. Panel (a) plots this counterfactual for transactions with holding periods between 1 and 10 days, while Panel (b) uses transactions with holding periods between 1 and 30 days.

As shown in Figure 10, we perform the calculation for a narrow and a broad range of negative rates of return, either 0 to -1 percent or 0 to -10 percent, and for a short and long range of holding periods. We report the results of these calculations for all five years of our sample, 2011-2015, in Table 5. Even within the narrow range of rates of return, we estimate that between $150 to $200 billion in sales do not occur in the first 30 days because of individuals' aversion to realizing losses. This corresponds to between $650 and $770 million in additional capital losses, which corresponds to between $62 to $72 in realized losses per additional trade. Assuming that all taxpayers had short-run capital gains that could be offset by those realized losses - which seems sensible given the fact from Table 1 that over 90% majority of individuals ended the year in a net gain position - and using the tax rate on short-term gains/losses from Table 1, we find that taxpayers could have saved $140 to $160 million in taxes by realizing these losses.

When we calculate a counterfactual over a wider range of rates of return, from 0 to -10 percent, these aggregates are much larger, but more uncertain, as there is obviously more potential bias in the estimated counterfactual far away from zero. Bearing that uncertainty in mind, we estimate that in the counterfactual individuals would have sold roughly $500 billion more assets, realized roughly $10 billion more losses, or $300 per transaction, and saved roughly $2 billion in tax over the five-year period.

	Return between 0 and -10 percent		Return between 0 and -1 percent	
Holding Period	1-10	1-30	1-10	1-30
Additional sales amount (in billion $)	382	471	161	206
Additional capital losses (in million $)	8,600	10,600	650	770
Taxes saved from additional capital losses (in million $)	1,800	2,200	140	160
Additional loss realizations (in million)	27	39.5	8.9	12.6
Dollars of realized loss per additional transaction (in $)	317	269	72.4	61.5

Notes: This table presents estimations ofthe range of potential sales that could have happened absent the disposition effect. The results cover a five-year period, from 2011 to 2015.

Table 5: Counterfactual results

5 Conclusion

We find that in the population of US taxable investors, the disposition to sell winners and hold losers is driven by day trading. The disposition effect overall is mostly driven by a large discontinuity in the probability of realization at zero. This discontinuity fades rapidly with the holding period and is invisible beyond the horizon of a thirty-day holding period. Further analysis suggests that the discontinuity and its vanishing are both the result of a composition effect: investors that tend to trade at very short frequency are driving almost the entire effect, and the disposition effect reappears over longer horizons for this small group of investors, though they very seldom trade at such long horizons.

The traders primarily responsible for the disposition effect are a small minority of all individual investors, but they trade extremely frequently, so much so that the movement of aggregate distributions without conditioning on holding period is strongly influenced by their behavior. As a result, we conclude that most individual investors are not subject to the extreme loss aversion that apparently affects day traders, suggesting that most individuals are much more rational than one would be tempted to conclude by examining transaction-level data without accounting for investor heterogeneity.

Our results are broadly consistent with notions of realization utility, that individuals narrowly bracket their investments and evaluate the realization of each gain or loss in a separate mental account. However, unlike prior literature, explaining our results in terms of realization utility requires a notched form of reference dependence. We acknowledge, however, that our result regarding realization utility is a reduced-form one. Why exactly the notch exists for some investors is unclear, and we cannot entirely rule out explanations based on beliefs - though generating a notched valuation of realizations would require a peculiar form of beliefs - or, potentially, more exotic factors.

Why might traders with extreme loss aversion trade at short horizons? We conjecture that a portfolio choice model with realization utlity along the lines of Barberis and Xiong (2012), but with a notch in realization utility at zero for some investors, could explain essentially all of our results. Intuitively, such a model

should predict that investors with notched realization utility will tend to realize gains extremely quickly: once a paper gain is accrued, the investor is highly averse to the risk of it becoming a loss, and thus realizes the gain right away. The presence of the notch in realization utility could thereby cause some investors, endogenously, to behave as short-horizon traders. However, there are some nuances that need to be explored to flesh out this intuition, so we hope that future theoretical work will explore this question further.

While our results challenge the view of the disposition effect as something that is highly relevant for a typical investor, they do not necessarily diminish the importance of the disposition effect for our understanding of overall financial markets. In fact they introduce interesting questions for future research on this topic. Several papers have considered how the disposition effect, potentially in combination with other factors, can contribute to asset price bubbles or other price dynamics (Grinblatt and Han, 2005; Liao et al., 2020; Wang et al., 2017). Examining the role of day trading specifically, and the implications of the extreme form of loss aversion that explains our results, are interesting avenues for future research along these lines.

A Sampling and Selection Issues

	Number of Transactions By Year:					**Total**
Sample Restriction	2011	2012	2013	2014	2015	
F1099B Data, No Wash Sales	4,741,112	4,318,804	4,663,549	5,158,838	5,896,903	24,779,206
+ Non-missing Cost Basis	1,320,744	2,066,757	2,970,254	3,789,461	4,699,264	14,846,480
+ Non-missing Acq. Date	1,204,203	1,770,980	2,511,953	3,254,891	4,018,358	12,760,385
+ 1040 Matches Only	1,203,838	1,770,778	2,511,520	3,254,158	4,017,687	12,757,981
+ CRSP/OTC Matches Only	1,178,483	1,710,161	2,427,921	2,946,064	3,533,899	11,796,528

	Number of Observations by:			
Sample Restriction	Transactions	Taxpayers	Brokers	Dollar Volume
F1099B Data, No Wash Sales	24,779,206	360,556	6,280	171,897,554,095
+ Non-missing Cost Basis	14,846,480	246,134	4,211	93,884,554,908
+ Non-missing Acq. Date	12,760,385	208,308	3,307	73,750,226,394
+ 1040 Matches Only	12,757,981	208,226	3,307	73,729,610,420
+ CRSP/OTC Matches Only	11,796,528	199,442	2,948	69,629,944,446

Notes: This table shows how we arrive at our main sample, and how each individual restriction affects the number of observations in the sample.

Table A.1: Main Sample Construction: Importance of Restrictions

Data Issue	2011	2012	2013	2014	2015	All Years
Missing Cost Basis						
Percent	.721	.521	.363	.265	.203	.400
Count	3,420,368	2,252,047	1,693,295	1,369,377	1,197,639	9,932,726
Missing Acq. Date						
Percent	.736	.577	.455	.346	.288	.467
Count	3,492,063	2,492,873	2,123,436	1,788,065	1,698,869	11,595,306
Missing Both						
Percent	.711	.508	.357	.242	.172	.383
Count	3,375,522	2,197,096	1,665,135	1,253,495	1,017,963	9,509,211

Notes: This table summarizes the number of dropped observations due to missing cost basis or acquisition date, by year.

Table A.2: Main Sample Construction: Missing Cost Basis and Acquisition Date by Year

This section contains supplementary data on the construction of the sample and potential selection bias.

Graph of $p(a = 1|x)$ in the paper returns sample by year. Supplement with data on the dist of holding periods, implications for how much data should be missing.

Distribution of $p(x|a = 1)$ in paper returns sample vs in main sample.

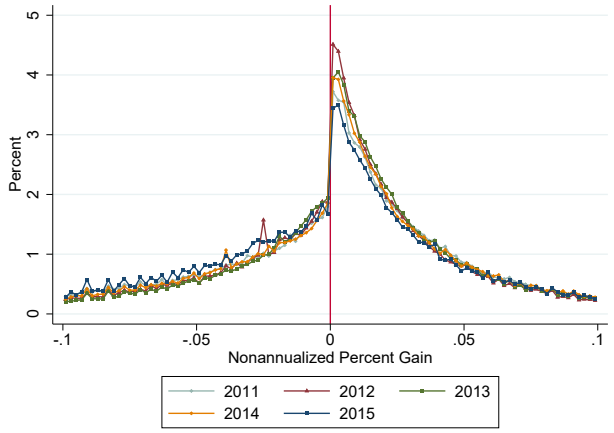


Figure A.1: Distribution of realized rates of return, by year

Notes: This figure shows the distribution of nonannualized realized returns for transactions with holding periods between 1 and 30 days, for different years. Nonannualized percent gain is calculated as the difference between the sale price and the cost basis divided by the cost basis. The y-axis shows the percent of sales for nonannualized returns between -10% and 10% in 0.2% bins

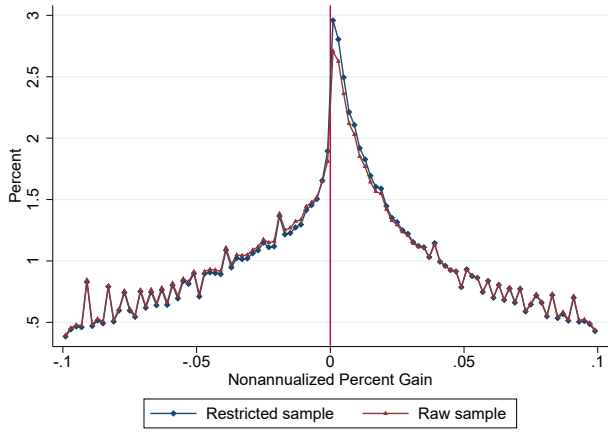


Figure A.2: Distribution of realized rates of return, With and Without Unknown Holding Period Observations

Notes: This figure shows the distribution of nonannualized realized returns for transactions with holding periods between 0 and 800 days (blue line) and also for all transactions, including the ones for which the holding period is not known (red line). Nonannualized percent gain is calculated as the difference between the sale price and the cost basis divided by the cost basis. The y-axis shows the percent of sales for nonannualized returns between -10% and 10% in 0.2% bins

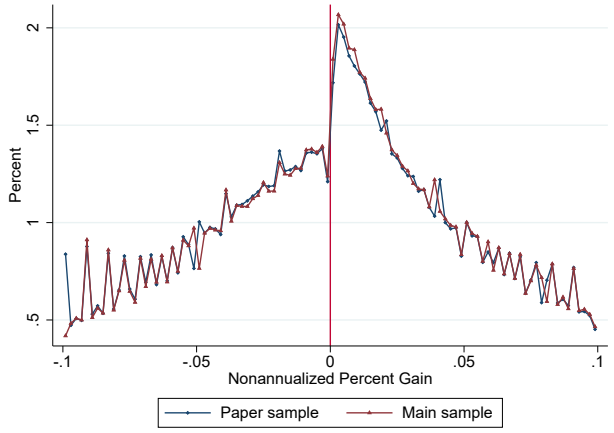
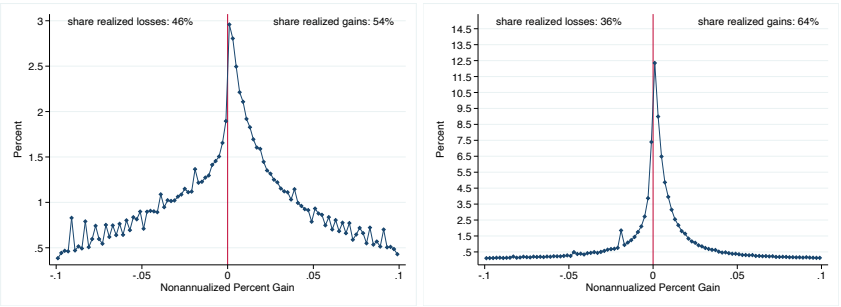


Figure A.3: Distribution of realized rates of return in the paper returns and in the main sample

Notes: This figure shows the distribution of nonannualized realized returns for transactions with holding periods between 1 and 800 days, in the main sample (blue line) and in the paper returns sample (red line). Nonannualized percent gain is calculated as the difference between the sale price and the cost basis divided by the cost basis. The y-axis shows the percent of sales for nonannualized returns between -10% and 10% in 0.2% bins

B Additional results



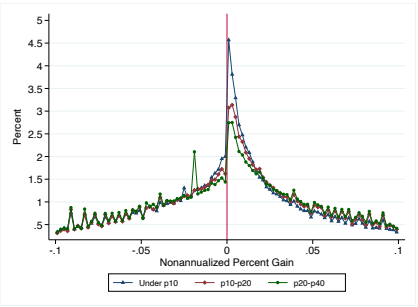
(a) Holding Periods: 0-800 Days (b) Holding Period: 0 Days

Figure B.4: Distributions of Nonannualized Returns: Including Zero-Days Holding Periods

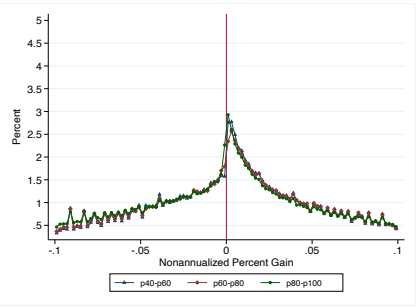
Notes: This figure shows the distribution of nonannualized realized returns for transactions with holding periods between 0 and 800 days (a) or 0 days (b). Realizations with zero returns are excluded. Nonannualized percent gain is calculated as the difference between the sale price and the cost basis divided by the cost basis. The y axis shows the percent of sales for nonannualized returns between -10% and 10% in 0.2% bins.

Graph of $p(a = 1|x$ in the paper returns sample, for additional intervals and single days

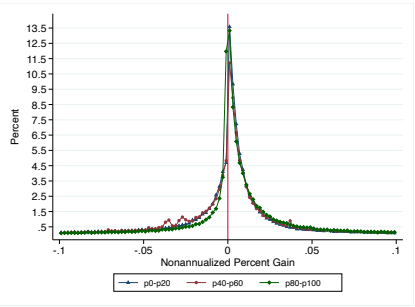
Panel A: By AGI



(a) Holding Periods: 0-800 Days

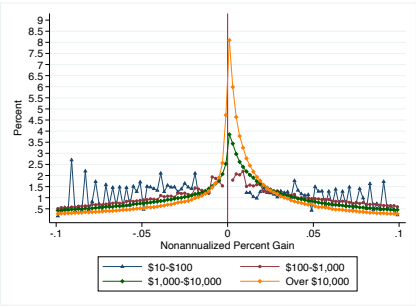


(b) Holding Periods: 0-800 Days

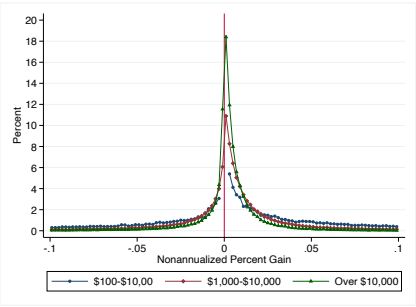


(c) Holding Period: 0 Days

Panel B: By Cost Basis

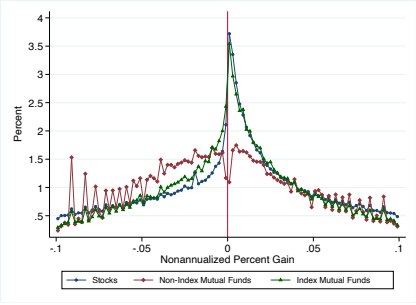


(d) Holding Periods: 0-800 Days

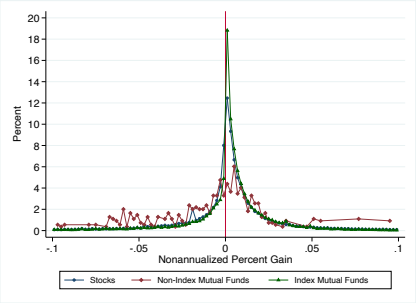


(e) Holding Period: 0 Days

Panel C: By Asset Type



(f) Holding Periods: 0-800 Days

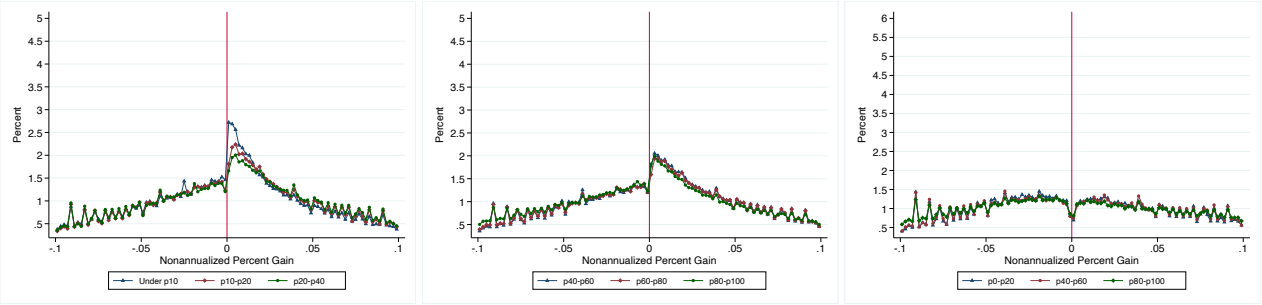


(g) Holding Period: 0 Days

Figure B.5: Distributions of Returns by AGI, Cost Basis, and Security Type: Including Zero-Days Holding Periods

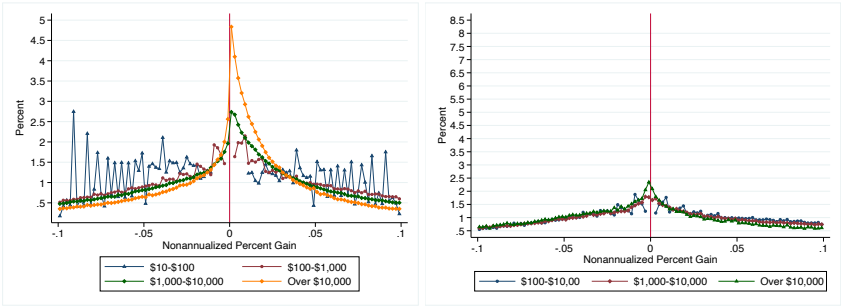
Notes: This figure shows the distributions of nonannualized realized returns for transactions with holding periods between 0-800, or 0 days, broken down by adjusted gross income (AGI) percentiles, cost basis, and asset type. Realizations with zero returns are excluded. Nonannualized percent gain is calculated as the difference between the sale price and the cost basis divided by the cost basis. The y axis shows the percent of sales for nonannualized returns between -10% and 10% in 0.2% bins.

Panel A: By AGI



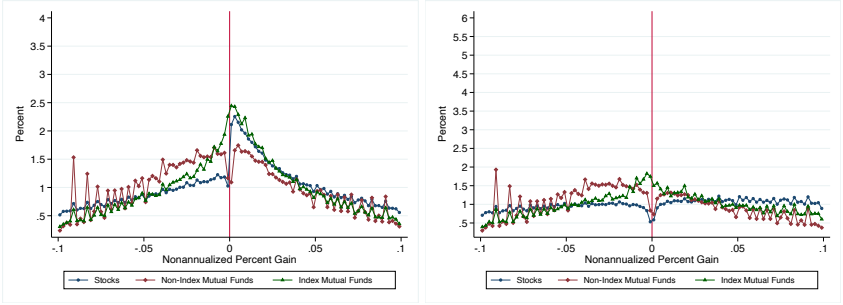
(a) Holding Periods: 1-800 Days (b) Holding Periods: 1-800 Days (c) Holding Periods: 181-800 Days

Panel B: By Cost Basis



(d) Holding Periods: 1-800 Days (e) Holding Periods: 181-800 Days

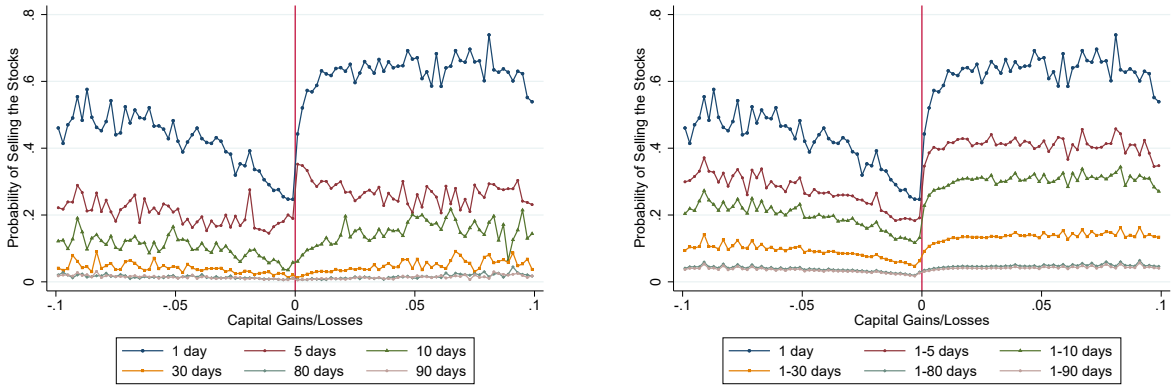
Panel C: By Asset Type



(f) Holding Periods: 1-800 Days (g) Holding Periods: 181-800 Days

Figure B.6: Distributions of Returns by AGI, Cost Basis, and Security Type: Excluding Zero-Days Holding Periods

Notes: This figure shows the distributions of nonannualized realized returns for transactions with holding periods between 1-800, or 181-800 days, broken down by adjusted gross income (AGI) percentiles, cost basis, and asset type. Realizations with zero returns are excluded. Nonannualized percent gain is calculated as the difference between the sale price and the cost basis divided by the cost basis. The y axis shows the percent of sales for nonannualized returns between -10% and 10% in 0.2% bins.



(a) Heterogeneity among the days

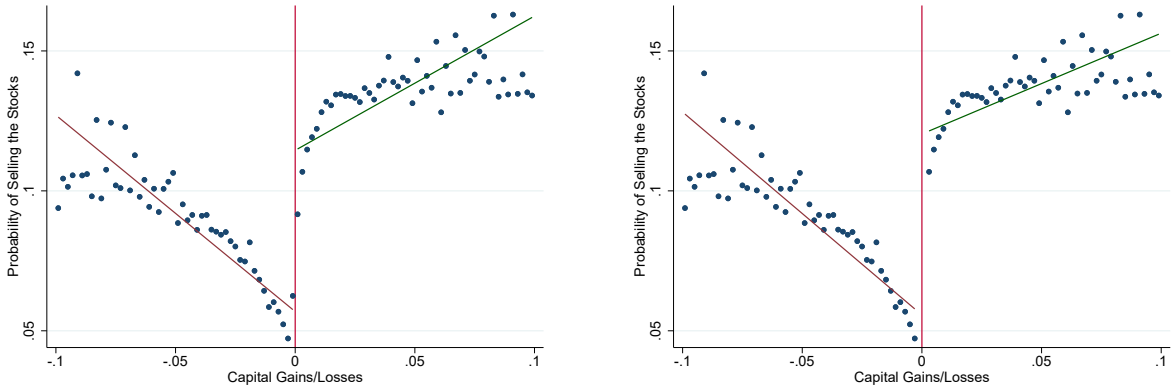
(b) Heterogeneity among the periods

Figure B.7: The probability of realization conditional on the rate of return

Notes: This figure plots the the probability of selling stock and realizing a gain or loss conditional on the rate of return - $P(a = 1|x)$ - for different holding periods. The Capital Gain or Loss x is calculated as the difference between the paper value of the stock and the cost basis divided by the cost basis. We observe a large discontinuity at 0 for holding periods less than 30 days but no discontinuity for longer holding periods, consistent with the results in Figure 2.

C Measuring Discontinuity in the Conditional Probability of Realization – Robustness Checks

Global polynomials

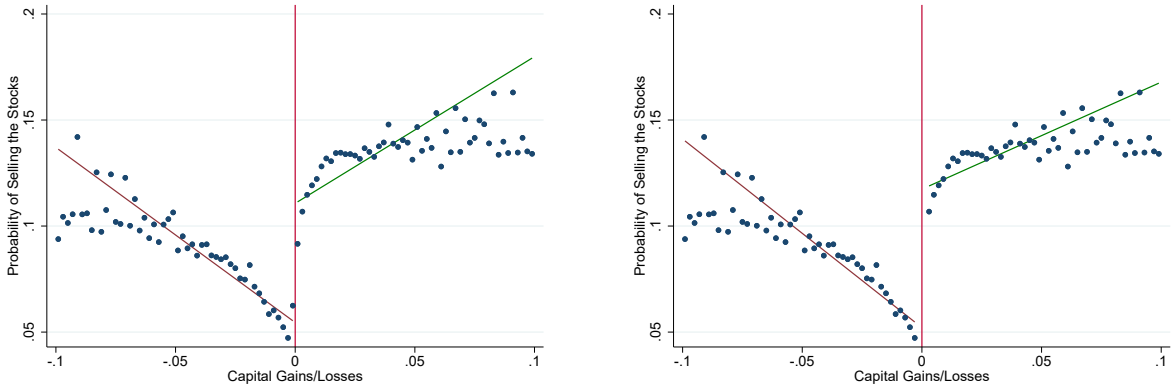


(a) Whole sample

(b) Excluding bins just left and right of 0

Figure C.8: Global polynomial of order 1, HP 1- 30

Notes: This figure plots the results of the estimation of equation 1 with linear control functions, for transactions with holding periods between 1 and 30 days. Panel (a) includes the whole sample while Panel (b) excludes the bins just left and right of 0 (i.e. Capital gains/losses of -0.001 to 0.001).



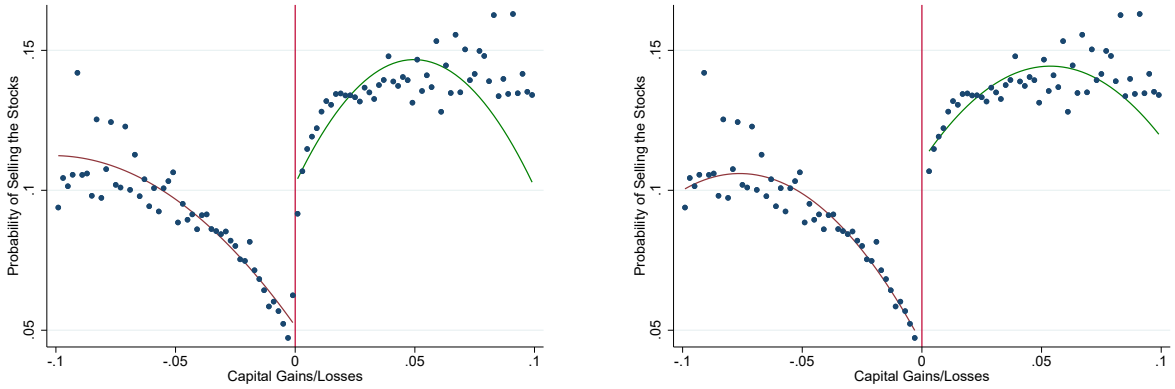
(a) Whole sample

(b) Excluding bins just left and right of 0

Figure C.9: Global polynomial of order 1, triangular Kernel weighting, HP 1- 30

Notes: This figure plots the results of the estimation of equation 1 with linear control functions, for transactions with holding periods between 1 and 30 days. The observations have been re-weighted with a kernel function of bandwidth 0.1. Panel (a) includes the whole sample while Panel (b) excludes the bins just left and right of 0 (i.e. Capital gains/losses of -0.001 to 0.001).

Local polynomials



(a) Whole sample

(b) Excluding bins just left and right of 0

Figure C.10: Global polynomial of order 2, triangular Kernel weighting, HP 1- 30

Notes: This figure plots the results of the estimation of equation 1 with quadratic control functions, for transactions with holding periods between 1 and 30 days. The observations have been re-weighted with a kernel function of bandwidth 0.1. Panel (a) includes the whole sample while Panel (b) excludes the bins just left and right of 0 (i.e. Capital gains/losses of -0.001 to 0.001).

	(1) -rd OB	(2) -rd OB/2	(3) -rd OB*2	(4) -rd OB*10
Discontinuity	0.014***	0.029***	0.021***	0.048***
	$(1.8 \cdot 10^{-15})$	$(6.1 \cdot 10^{-17})$	$(7.6 \cdot 10^{-06})$	$(1.9 \cdot 10^{-05})$
Bandwidth	0.0034	0.0017	0.0069	0.0343
Observations	3,380,180	3,380,180	3,380,180	3,380,180

t-statistics in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Notes: This table shows the results of a local linear polynomial regression estimated on transactions with holding periods between 1 and 30 days. The bandwidth has been chosen with -rd stata command, using Imbens and Kalyanaraman's algorithm.

Table C.3: Regression results with local polynomials, HP 1-30

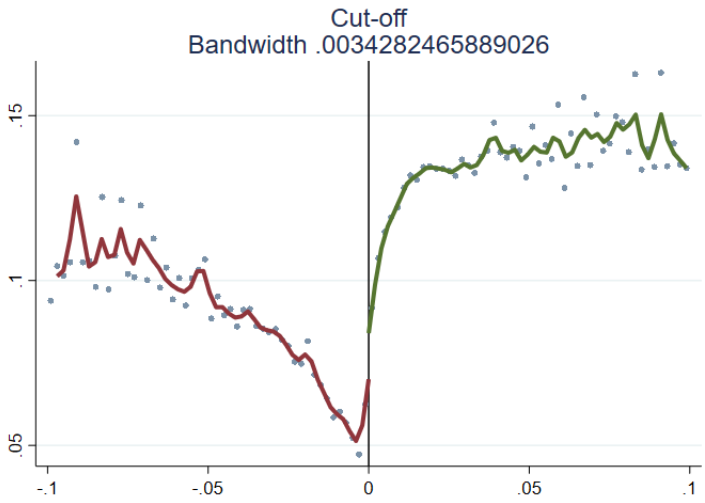
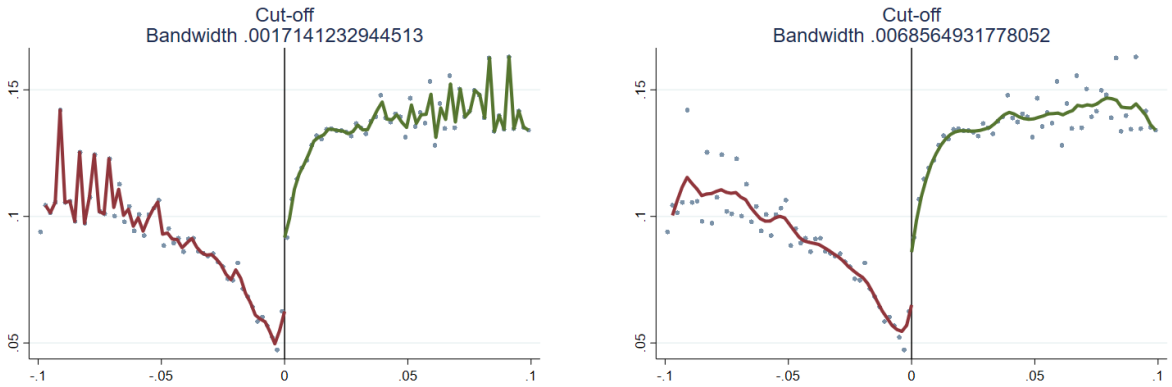


Figure C.11: Local polynomial, optimal bandwidth, HP 1 - 30

Notes: This figure plots the results of a local linear polynomial regression estimated on transactions with holding periods between 1 and 30 days. The bandwidth has been chosen with -rd stata command, using Imbens and Kalyanaraman's algorithm.



(a) Optimal Bandwidth/2

(b) Optimal Bandwidth*2

Figure C.12: Local polynomials, HP 1 - 30

Notes: This figure plots the results of a local linear polynomial regression estimated on transactions with holding periods between 1 and 30 days. The bandwidth has been chosen with -rd stata command, using Imbens and Kalyanaraman's algorithm. Panel (a) shows the results using half the optimal bandwidth for the estimation. Panel (b) shows the results using twice the optimal bandwidth.

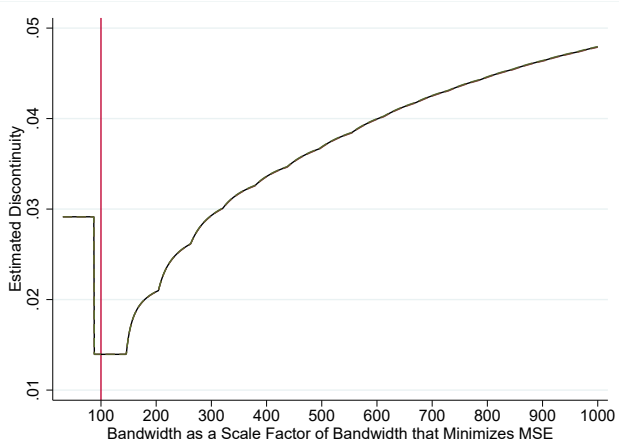


FIGURE C.13: ESTIMATED DISCONTINUITY FROM A LOCAL LINEAR REGRESSION, AS A FUNCTION OF THE SIZE OF THE BANDWIDTH, HP 1 - 30

Notes: This figure shows the discontinuity at 0 estimated with a local linear regression, as a function of the size of the bandwidth. Bandwidth 100 corresponds to the optimal bandwidth chosen with the -rd command. The other bandwidths correspond to scale factors of the optimal one. Confidence intervals are included as dashed lines around the estimates. The sample is restricted to transactions with holding periods between 1 and 30 days.

D Theoretical Extensions

Here we extend the model from Section 4 of the main text to highlight some additional details. We first discuss explicitly separating realization utility from other concerns, potentially including tax concerns or dynamic concerns. We then address dynamic concerns directly using an approach similar to that of Barberis and Xiong's model of realization utility.

D.1 Separating Realization Utility from Other Concerns

As before the individual decides whether to buy or hold some asset. We now express the difference in utility between selling and holding as

$$u_i(1) - u_i(0) = v(x) + g(x) + \varepsilon_i \tag{4}$$

where $v(x)$ is a prospect theory utility function with a reference point at $x = 0$, and $g(x)$ captures all dependence between the rate of return and motives to sell other than the jolt from prospect theory. The main theoretical result in the paper in this modified model implies that either $v(x)$ or $g(x)$ contains a discontinuity at 0, but we explicitly separate the prospect theory concerns in $v(x)$ with other concerns. To add structure,

we could use a consider prospect theory payoffs of the following form:

$$v(x) = \begin{cases} \nu(x) & \text{if } x > 0 \\ -\lambda\nu(-x) - \gamma & \text{if } x < 0, \end{cases} \tag{5}$$

with sub-utility function $\nu(x)$ having $\nu' > 0$ and $\nu'' \leq 0$ - the so-called "diminishing sensitivity" property - and with $\nu(0)$ normalized to 0. Most of the literature uses payoff functions of this form to justify observed reference dependence, except that much of the literature supposes that $\gamma = 0$. With $\gamma = 0$ there is a kink in the utility function at $x = 0$, with the magnitude of the kink governed by $\lambda > 1$ (e.g. Imas (2016); O'Donoghue and Sprenger (2018)). A sizable literature studies the behavioral implications of such a kink in the utility function. Comparatively fewer papers consider models with a discontinuity in utility, i.e. where $\gamma > 0$.

The discussion of our main result in the paper supposed that $v(x) + g(x)$ at $x = 0$ comes through $v(x)$. If $g(x)$ must be continuous, the only way to generate a discontinuity like the one we observe empirically is therefore to have a discontinuity in $v(x)$, i.e. a large and significant value of $\gamma > 0$ in equation (5). To see how γ influences the discontinuity at $x = 0$ take the limit as $x \to 0$ from the left and from the right of $\Pr(a_i = 1|x)$. From equations (5) and (3), we have that the size of this discontinuity will be:

$$\lim_{x \to 0^+} \Pr(a_i = 1|x) - \lim_{x \to 0^-} \Pr(a_i = 1|x) = F(\gamma) - F(0) \tag{6}$$

Note that $F(\gamma) > F(0)$, so the model predicts a discrete increase in $\Pr(a = 1|x)$ as x goes from just below zero to just above 0, which is consistent with the data.

Furthermore, there is no good reason to believe that $g(x)$ is discontinuous at $x = 0$. We think of $g(x)$ as a reduced form function capturing all non-behavioral motives to sell depending on x, everything apart from the contemporaneous realization-based payoff from prospect theory. Fundamentally, we argue that no classically rational individual would evaluate the prospect of realizing an extremely small tiny gain as discretely better than realizing an extremely tiny loss. We discuss the specific dynamic concerns that could influence $g(x)$ below, accounting for forward-looking preferences. In Section 4 of the main text we alluded to these dynamic concerns, as well as the fact that gains and losses sometimes receive different effective tax treatments – which leads to a kink in $g(x)$ at $x = 0$. These factors can shape the distribution of $\Pr(a_i = 1|x)$ around 0 but they cannot introduce a discontinuity at zero.

Finally, we note that in principle, one might also propose other *behavioral* reasons for a discontinuity in the payoff function at 0, separately from reference dependence. For example, one possibility would be that the individual's beliefs about the future performance of the asset might be discretely more pessimistic

conditional on a small loss relative to a small gain due to some bias in beliefs. This would actually cause the individual to be *more* likely to realize the loss in order to reallocate wealth away from the asset, which is the opposite from what we observe.

D.2 Dynamics

There are 4 periods indexed by $t = 0, 1, 2, 3$. The agent acquires the asset exogenously in period 0. The agent then makes a sell versus hold decision in each subsequent period, and if the agent still holds the asset in the final period, then they realize the gain or loss in that period. Thus, there are two periods in which the agent makes a non-trivial sell or hold decision, periods 1 and 2. The rate of return as of each period is x_t, which evolves according to $x_{t+1} = x_t + \eta$ for a random innovation η with a stationary distribution. We assume the distribution of η is continuous with positive support over all $\eta \in \mathbb{R}$.

We first show that our above results are unaffected by the presence of dynamic concerns. First consider period 2, where the agent is only deciding between selling in the current period and selling in period 3. The agent should account for the fact that selling the asset in period 3 instead of 2 will generate a payoff in period 3.[27] Comparing utility for selling in period two ($a_2 = 1$) versus holding in period 2 and selling in period 3 ($a_2 = 0$) we have:

$$\Delta u(a_2, x_2) = u(1) - u(0) = v(x_2) + \tilde{g}(x_2) + \varepsilon_{i2} - \beta E[v(x_2 + \eta) + \tilde{g}(x_2 + \eta) + \varepsilon_{i3}|x_2, \varepsilon_{i2}] \tag{7}$$

where $v(x)$ is a prospect theory payoff like above, $\tilde{g}(x)$ captures all other dependence between x and the contemporaneous payoff from realizing the gain or loss, β discounts the payoff in the future, and ϵ_{it} captures all idiosyncratic motives to sell in a given period, which may be correlated across periods.

The reduced form payoff function summarized by equation 4 nests this payoff function with

$$g(x) \equiv \tilde{g}(x_2) - \beta E[v(x_2 + \eta) + \tilde{g}(x_2 + \eta)|x_2] \tag{8}$$

$$\varepsilon_i \equiv \varepsilon_{i2} - \beta E[\varepsilon_{i3}|\varepsilon_{i2}].$$

As such Proposition 1 on the importance of discontinuities in utility for discontinuities in behavior – $p(a_i = 1|x)$ – continues to hold in the dynamic model. Next we will show that provided that $\tilde{g}(x)$ is continuous in x, $g(x)$ is still continuous in x at x_0, even when there is a discontinuity in $v(x)$ at $x = 0$. To see why this is the case, note that we can re-write $g(x)$ as follows by the law of iterated expectations and equation 5:

[27] One might wonder whether agents are forward-looking, especially when it comes to the prospect theory component of their preferences. The model where agents are not forward-looking is obviously already nested by the static model presented in equations (4) through (6), so we will focus here on the case where agents *are* forward-looking.

$$g(x) \equiv \tilde{g}(x_2) - \beta\{E[\tilde{g}(x_2+\eta)|x_2] + E[v(x_2+\eta) + |x_2, x_2+\eta > 0]\Pr(x_2+\eta > 0|x_2) \\ + E[v(-\lambda(x_2+\eta)) - \gamma|x_2, x_2+\eta < 0]\Pr(x_2+\eta < 0|x_2)\}$$

Simplifying this further we have

$$g(x) \equiv \tilde{g}(x_2) - \beta\{E[\tilde{g}(x_2+\eta)|x_2] + E[v(x_2+\eta)|x_2, x_2+\eta > 0]\Pr(x_2+\eta > 0|x_2) \\ + E[v(-\lambda(x_2+\eta))|x_2, x_2+\eta < 0]\Pr(x_2+\eta < 0|x_2) \\ - \gamma\Pr(x_2+\eta < 0|x_2)\}.$$

The first two terms of this expression are obviously continuous at $x_2 = 0$ because $\tilde{g}(x)$ is assumed to be continuous. Both expectations, $E[v(x_2+\eta)|x_2, x_2+\eta > 0]$ and $E[v(-(x_2+\eta))|x_2, x_2+\eta < 0]$ are also continuous in x_2 because the sub-utility function $v(x)$ is continuous in x_2. The probability of being in the gain domain in Period 3, $\Pr(x_2+\eta > 0|x_2)$ is also continuous in x_2 by the conditions we put on the distribution of η. As sums and products of continuous functions are continuous, the expression for $g(x)$ above is continuous. Intuitively, the reason that $g(x)$ is continuous despite discontinuities of v is that the innovation to future returns acts to smooth out the discontinuity in $v(x)$ when calculating expectations for the future.

In summary, we can nest period 2 behavior by the static model presented in equations (4) through (6), without introducing discontinuities into $g(x)$ that would cause the logical argument presented above to break down. We next show that the same basic argument continues to hold even if we examine decision-making in period 1 rather than period 2. This decision is more complicated because the agent is forward-looking and might decide to sell or hold in period 2. In period 1, the agent sells based on the following comparison:

$$\Delta u(a_1, x_1) = u(1) - u(0) = v(x_1) + \tilde{g}(x_1) + \varepsilon_i 1 \\ - \Pr(\Delta u(a_2, x_1+\eta_2) > 0|x_1)\beta E[v(x_1+\eta_2) + \tilde{g}(x_1+\eta_2) + \varepsilon_i 2|\varepsilon_{i1}, x_1, \Delta u(a_2, x_1+\eta_2) > 0] \\ + \Pr(\Delta u(a_2, x_1+\eta_2) < 0|x_1)\beta^2 E[v(x_1+\eta_2+\eta_3) + \tilde{g}(x_1+\eta_2+\eta_3) + \varepsilon_{i3}|\varepsilon_{i1}, x_1, \Delta u(a_2, x_1+\eta_2) < 0]$$

This is a complicated expression but we know by the same logic as above: the only possible discontinuity in x_1 in this expression comes from the first term, $v(x_1)$.

References

Allen, E. J., Dechow, P. M., Pope, D. G., and Wu, G. (2016). Reference-Dependent Preferences: Evidence from Marathon Runners. *Management Science*, 63(6):1657–1672.

Barber, B. M., Lee, Y.-T., Liu, Y.-J., and Odean, T. (2007). Is the Aggregate Investor Reluctant to Realise Losses? Evidence from Taiwan. *European Financial Management*, 13(3):423–447.

Barber, B. M. and Odean, T. (2000). Trading is Hazardous to Your Wealth: The Common Stock Investment Performance of Individual Investors. *The Journal of Finance*, 55(2):773–806.

Barber, B. M. and Odean, T. (2004). Are Individual Investors Tax Savvy? Evidence From Retail and Discount Brokerage Accounts. *Journal of Public Economics*, 88(1-2):419–442.

Barber, B. M. and Odean, T. (2013). The Behavior of Individual Investors. In *Handbook of the Economics of Finance*, volume 2, pages 1533–1570. Elsevier.

Barberis, N. (2018). Psychology-Based Models of Asset Prices and Trading Volume. In Bernheim, B. D., DellaVigna, S., and Laibson, D., editors, *Handbook of Behavioral Economics*, volume 1 of *Handbook of Behavioral Economics: Applications and Foundations*, pages 79 – 175. North-Holland.

Barberis, N. and Xiong, W. (2009). What Drives the Disposition Effect? An Analysis of a Long-Standing Preference-Based Explanation. *The Journal of Finance*, 64(2):751–784.

Barberis, N. and Xiong, W. (2012). Realization Utility. *Journal of Financial Economics*, 104(2):251–271.

Ben-David, I. and Hirshleifer, D. (2012). Are Investors Really Reluctant to Realize Their Losses? Trading Responses to Past Returns and the Disposition Effect. *The Review of Financial Studies*, 25(8):2485–2532.

Brown, P., Chappel, N., da Silva Rosa, R., and Walter, T. (2006). The Reach of the Disposition Effect: Large Sample Evidence Across Investor Classes. *International Review of Finance*, 6(1-2):43–78.

Chang, T. Y., Solomon, D. H., and Westerfield, M. M. (2016). Looking for Someone to Blame: Delegation, Cognitive Dissonance, and the Disposition Effect. *The Journal of Finance*, 71(1):267–302.

Constantinides, G. M. (1984). Optimal Stock Trading with Personal Taxes: Implications for Prices and the Abnormal January Returns. *Journal of Financial Economics*, 13(1):65–89.

Cooper, M., McClelland, J., Pearce, J., Prisinzano, R., Sullivan, J., Yagan, D., Zidar, O., and Zwick, E. (2016). Business in the United States: Who Owns It, and How Much Tax Do They Pay? *Tax Policy and the Economy*, 30(1):91–128.

Diecidue, E. and Van De Ven, J. (2008). Aspiration Level, Probability of Success and Failure, and Expected Utility. *International Economic Review*, 49(2):683–700.

Fishburn, P. C. (1977). Mean-Risk Analysis with Risk Associated with Below-Target Returns. *American Economic Review*, 67(2):116–26.

Gelman, A. and Imbens, G. (2019). Why High-Order Polynomials Should Not Be Used in Regression Discontinuity Designs. *Journal of Business & Economic Statistics*, 37(3):447–456.

Grinblatt, M. and Han, B. (2005). Prospect Theory, Mental Accounting, and Momentum. *Journal of Financial Economics*, 78(2):311–339.

Grinblatt, M. and Keloharju, M. (2001). What Makes Investors Trade? *The Journal of Finance*, 56(2):589–616.

Hirshleifer, D. (2015). Behavioral Finance. *Annual Review of Financial Economics*, 7(1):133–159.

Hoopes, J., Langetieg, P., Nagel, S., Reck, D., Slemrod, J., and Stuart, B. (2016). Who Sold During the Crash of 2008-9? Evidence from Tax-Return Data on Daily Sales of Stock. Working Paper 22209, National Bureau of Economic Research.

Imas, A. (2016). The Realization Effect: Risk-Taking after Realized versus Paper Losses. *American Economic Review*, 106(8):2086–2109.

Ivković, Z., Poterba, J., and Weisbenner, S. (2005). Tax-Motivated Trading by Individual Investors. *American Economic Review*, 95(5):1605–1630.

Kahneman, D. and Tversky, A. (1979). Prospect Theory: An Analysis of Decisions and the Psychology of Choice. *Econometrica*, 47:313–327.

Kaustia, M. (2010). Prospect Theory and the Disposition Effect. *The Journal of Financial and Quantitative Analysis*, 45(3):791–812.

Köszegi, B. and Rabin, M. (2006). A Model of Reference-Dependent Preferences. *The Quarterly Journal of Economics*, 121(4):1133–1165.

Köszegi, B. and Rabin, M. (2007). Reference-Dependent Risk Attitudes. *American Economic Review*, 97(4):1047–1073.

Köszegi, B. and Rabin, M. (2009). Reference-Dependent Consumption Plans. *American Economic Review*, 99(3):909–36.

Liao, J., Peng, C., and Zhu, N. (2020). Price and Volume Dynamics in Bubbles.

Meng, J. (2010). The Disposition Effect and Expectations as Reference Point.

Odean, T. (1998). Are Investors Reluctant to Realize Their Losses? *The Journal of Finance*, 53(5):1775–1798.

Odean, T. (1999). Do Investors Trade Too Much? *American Economic Review*, 89(5):1279–1298.

O'Donoghue, T. and Sprenger, C. (2018). Reference-Dependent Preferences. In Bernheim, B. D., DellaVigna, S., and Laibson, D., editors, *Handbook of Behavioral Economics - Foundations and Applications 1*, volume 1 of *Handbook of Behavioral Economics: Applications and Foundations 1*, pages 1 – 77. North-Holland.

Payne, J. W. (2005). It Is Whether You Win or Lose: The Importance of the Overall Probabilities of Winning or Losing in Risky Choice. *Journal of Risk and Uncertainty*, 30(1):5–19.

Poterba, J. M. and Weisbenner, S. J. (2001). Capital Gains Tax Rules, Tax-loss Trading, and Turn-of-the-Year Returns. *The Journal of Finance*, 56(1):353–368.

Rees-Jones, A. (2017). Quantifying Loss-Averse Tax Manipulation. *The Review of Economic Studies*, 85(2):1251–1278.

Shefrin, H. and Statman, M. (1985). The Disposition to Sell Winners Too Early and Ride Losers Too Long: Theory and Evidence. *The Journal of Finance*, 40(3):777–790.

Stiglitz, J. E. (1983). Some Aspects of the Taxation of Capital Gains. *Journal of Public Economics*, 21(2):257–294.

Wang, H., Yan, J., and Yu, J. (2017). Reference-Dependent Preferences and the Risk–Return Trade-Off. *Journal of Financial Economics*, 123(2):395–414.