Who Sold During the Crash of 2008-9?
Evidence from Tax-Return Data on Daily Sales of Stock

Jeffrey Hoopes, University of North Carolina at Chapel Hill

Patrick Langetieg, Internal Revenue Service

Stefan Nagel, University of Chicago Booth School of Business

Daniel Reck, London School of Economics

Joel Slemrod, University of Michigan

Bryan Stuart, George Washington University

November 30, 2017

Abstract: Based on an analysis of the universe of individual taxable stock sales from 2008 to 2009, using population tax-return data, we examine which individuals increased their sale of stocks following episodes of market tumult. The increase was disproportionately concentrated among investors in the top 1 and top 0.1 percent of the overall income distribution, retired individuals, and individuals at the very top of the dividend income distribution. We show that these results are plausibly consistent with standard portfolio choice models. Supplementary analysis suggests that the gross sales observable in tax return data are informative about overall net sales.

We thank John Friedman, Terry Odean, Clemens Sialm, Andrei Shleifer, Tyler Shumway, Chris Williams, Stefan Zeume, and participants at the University of Michigan Public Finance seminar, the 2015 National Tax Association meetings, the NBER Behavioral Finance Meeting, the U.S. Treasury Office of Tax Analysis, the University of North Carolina Annual Tax Symposium, and the University of Michigan Econ-Finance Day for helpful discussion and comments. We also thank John Guyton, Barry Johnson, Michael Strudler, and Janette Wilson of the Research, Applied Analytics, and Statistics Division of the Internal Revenue Service for help with using the IRS administrative data. All data work for this project involving confidential taxpayer information was done at IRS facilities, on IRS computers, by IRS employees, and at no time was confidential taxpayer data ever outside of the IRS computing environment. Professors Hoopes, Reck, and Stuart are IRS employees under an agreement made possible by the Intragovernmental Personnel Act of 1970 (5 U.S.C. 3371-3376). The views expressed here are those of the authors alone, and do not necessarily reflect the views of the Internal Revenue Service. Send correspondence to jslemrod@umich.edu.

1. Introduction

Periods of turmoil in stock markets—such as September 2008 in the wake of the Lehman Brothers bankruptcy—are associated with large declines in prices, abnormally high intra-day price volatility, and high trading volume. Market commentary often characterizes these periods as "sell-offs." As always, there is a buyer for every seller, so investors as a group cannot all be sellers. What may be happening instead during these instances is that some investors sell out, leading to a reallocation of asset ownership among heterogeneous investors. Little is known, however, about the characteristics of the investors that are prone to sell in the midst of market turmoil. While a substantial literature explores the selling behavior of individual investors in response to stock-level variables such as past returns or firm-specific news, empirical evidence on individual investors' reaction to aggregate shocks has so far remained largely elusive due to lack of data.

In this paper, we study administrative data from the Internal Revenue Service (IRS), consisting of billions of third-party reports on all sales of stock in United States taxable individual accounts, to investigate, at a daily frequency, which individuals sold stocks and mutual funds during the tumultuous market events of 2008 and 2009. We combine this information with data extracted from the universe of (anonymized) individual tax returns filed with the IRS to match asset sales reported for capital gains taxation purposes with income and other demographic information on each taxpayer. While we do not observe asset purchases in these tax records, we present indirect evidence based on dividend receipts and a supplementary brokerage account data set suggesting that individuals with high levels of gross sales are also, to a substantial extent, net sellers of stocks.

Our objective is to analyze the demographics of who sold during tumultuous periods. Rational portfolio choice theories and behavioral hypotheses suggest a number of channels through which the dramatic change in market conditions during market tumult could lead to a heterogeneous response and trading. Understanding which types of investors sell in response to tumult offers insights into which of these channels are driving trade at high frequencies.

We focus our analysis on 2008 and 2009, a period that includes the dramatic increase in market tumult following the bankruptcy of Lehman Brothers in September 2008 and the subsequent decline in tumult. In our baseline analyses, we measure market tumult with lagged daily changes in the VIX index. The VIX index is a measure of (risk-adjusted) expected market volatility, and it is commonly used as a proxy for market tumult and as a crisis indicator (see Adrian and Shin 2010; Longstaff 2010; Nagel 2012). Because we focus on stock sales, and hence on a *change* in investors' stock holdings, which should be triggered by a *change* in the market environment, our tumult measure is based on changes in the VIX index. During our sample period, stock market returns are typically strongly negative on days when the VIX rises, and we interpret VIX changes as a broad proxy of market turmoil rather than a narrow measure of volatility.[1]

Looking first at aggregate numbers, we find that the total selling volume of all individual taxpayers rises strongly in response to a rise in the VIX. Our main interest, however, centers on heterogeneity. To understand this heterogeneity, we start with a standard static portfolio choice framework in which investors differ in their asset return expectations, risk aversion, and consumption commitments (net of labor income). We then consider how investors react to a rise in volatility, fall in stock prices, and a rise in the present value of committed consumption—due

---

[1] We obtain similar results with negative lagged daily market index returns as an alternative proxy for tumult instead of VIX.

to a fall in interest rates or a heightened sensitivity of investors to the need to fund future consumption commitments—that usually occur in market tumult. While stock prices and expected returns adjust in equilibrium to clear the market, investors who require a relatively greater increase in expected returns than others in response to the change in market conditions will be sellers of stocks.

One result that naturally emerges in this setting is that optimistic and risk-tolerant investors are sellers. Coming into the tumultuous period with the highest risky asset exposure, they are also the ones who would require the highest compensation in terms of expected returns to refrain from selling. Empirically, we proxy for optimism and risk tolerance with a high level of taxable income, dividend income, and private business income—characteristics that earlier literature has linked to willingness to bear risk. We find that all three of these proxies are strongly positively related to increased selling volume in response to a rise in the VIX. Strikingly, this phenomenon is concentrated at the very top of the income distribution: investors in the top 1 percent, and even the top 0.1 percent, have a much greater propensity to sell during times of market tumult than do investors in the rest of the income distribution. Because the top 1 percent hold about half of total stock market wealth, this means that the high-income half of the wealth-weighted individual investor population appear as sellers in response to tumult. This result is in conflict with the popular notion that "small investors" may be most prone to "panic," but it is not inconsistent with standard portfolio choice models.[2]

In addition to the risk-sharing motives captured by simple portfolio choice models, tax-loss selling and attempts at volatility-timing could further contribute to tumult-induced selling by

[2] For example, in the crash of 1987 it was reported by the press that small investors were pulling out of the market—"Mutual funds reported record redemptions by panicked customers—primarily small investors—demanding cash" (Behr and Vise 1987). Traflet (2012) documents that Wall Street professionals tended to blame the 1929 stock market crash on "panic" selling by small investors. The notion that small investors is not, however, universally held (see, for example, Barboza 1998).

high-income taxpayers.[3] Taxpayers with high income may be most likely to undertake these relatively sophisticated strategies. To the extent that high-income investors are more prone to overconfidence (Graham et al. 2009) and pay greater attention to their portfolios (Sicherman et al. 2016), this could further reinforce this effect.

Our baseline portfolio choice framework further suggests that investors with the highest level of consumption commitments also should be most prone to selling stocks when the present value of these commitments rises or investors become more sensitive to these commitments. While we cannot measure net commitments directly, we investigate age and receipt of social security income as proxies. For young investors, the presence of labor income may imply that the net consumption commitment is low or even negative, especially if coupled with a potentially flexible response of labor supply. In contrast, for older investors, and especially those already in retirement, the present value of the net consumption commitment stream may be large and positive. Consistent with this prediction, we find that investors who receive social security income are more prone to sell in response to a rise in the VIX. Lower opportunity costs of paying attention to the market for retired investors are a potential additional contributor to this effect. Controlling for social security income, we do not find a clear separate effect of age.

Overall, there is substantial heterogeneity in selling behavior in response to market tumult, which supports models that emphasize such heterogeneity to explain investor behavior and asset prices. Our finding that those who are likely to be greater risk takers and more sophisticated investors were more likely to be sellers of stock following tumultuous days in the stock market may be somewhat surprising. However, as we have argued, standard portfolio choice theory provides a good explanation of this finding.

---

[3] Moreover, recent work by Moreira and Muir (2016) suggests that temporarily reducing stock market exposure following a burst of high volatility may in fact be a utility-improving timing strategy.

The data we analyze are, in a number of ways, substantially better than the data sets that have been studied up to now to address this set of questions. Existing studies of investor responses to market movements and changes in risk use data either from investor surveys (Guiso et al. 2013; Hudomiet et al. 2011; Shiller 1987), from non-randomly selected samples of portfolio holdings data (Dorn and Weber 2013; Hoffmann et al. 2013; Weber et al. 2012; Barrot et al. 2016), administrative data from Sweden that are available only at annual frequency (Calvet et al. 2009), or data on institutional investor portfolios (Ben-David, Franzoni, and Moussawi (2012) and Cella et al. (2013) provide evidence for the crisis in 2008 and Brunnermeier and Nagel (2004) and Griffin et al. (2011) for the Nasdaq crash in 2000). Our data let us investigate, for the first time, *all* taxable sales by the population of U.S. individuals at a daily frequency. This extremely large sample size allows us to analyze heterogeneity in selling, including for the upper tails of the income distribution, in ways that are not possible with small survey or brokerage account data sets.

To be sure, the data set is not perfect. The data set covers reported taxable sales, but not purchases of stocks and mutual funds. Additional analyses show, however, that there is a strong relationship between gross selling, which we observe, and net selling (i.e., sales minus purchases), which we do not observe. We document this relationship in two separate ways. First, we examine data from a discount brokerage that reports both gross and net sales (Barber and Odean 2000), and find a very strong positive relationship between gross and net sales. Furthermore, net sales of brokerage customers rise with changes in the VIX in similar ways as the gross sales of taxpayers in our data. Second, in the IRS data, we examine changes in dividend income reported on individual tax returns. Here we find a strong negative relationship between gross sales in a given year (e.g., 2008) and the change in dividend income from the previous year

(2007) to the subsequent year (2009); this is consistent with gross sales being correlated with net sales and therefore a decline in stockholding. Despite coming from different sources and different time periods, the quantitative results based on analysis of both discount brokerage data and changes in dividend income on tax returns are highly consistent with one another, suggesting that $1 of gross sales corresponds to about $0.33 in net sales.

Another shortcoming of the data set is that we do not observe sales in non-taxable accounts, such as Individual Retirement Accounts. To investigate the implications of this, we analyze data from the 2007-2009 panel of the Survey of Consumer Finance, which contains data on wealth in taxable and non-taxable accounts, including pensions and trusts. We find that the share of wealth in taxable accounts is relatively higher for individuals at the top of the income distribution and for older individuals. These facts rule out the concern that our main findings are driven by higher-income and/or retired people holding a disproportionately small share of their equities in taxable accounts, and thus having relatively less of these assets to sell in tumultuous times. Additionally, net sales in taxable accounts between the 2007 and 2009 waves of the survey are strongly related to total net sales, suggesting that our analysis of sales in taxable accounts is informative about total asset holdings.

2. Heterogeneous Response to Market Tumult: Theoretical Background

In our empirical analysis, we examine the heterogeneity of who sold in response to market tumult along several demographic dimensions of observable taxpayer characteristics. In this section, we discuss how these observable characteristics relate to underlying drivers of investors' asset allocation and trading decisions.

2.1 Heterogeneity in beliefs, risk aversion and consumption commitments

We start by analyzing the potential role of heterogeneous beliefs, risk aversion, and consumption commitments within a simple portfolio choice model. This model explains how such heterogeneity could lead to different reactions to market tumult.

Consider an investor endowed with wealth W who can invest in two traded assets: a one-period risk-free asset with log return r_F and a broad stock portfolio with log expected return μ and volatility σ^2. The investor receives a stream of income from riskless background wealth (e.g., labor income). At the same time, the investor has also committed to a stream of future consumption. We denote with X the difference in present value, with risk-free discounting, between the two streams, so that X is positive if the level of committed consumption exceeds the labor income and is negative otherwise. We label X the net committed consumption.

The investor solves the portfolio problem each period assuming returns are independent and identically distributed over time. With constant relative risk aversion γ, log-normal returns, and a second-order Taylor approximation of the investor's first-order condition (Campbell and Viceira 2002, Ch. 6), we obtain the optimal share of wealth invested in the stock portfolio to be

$$\alpha = \left(\frac{\mu - r_F}{\gamma\sigma^2}\right)\left(1 - \frac{X}{W}\right). \qquad (1)$$

The first term in parentheses is the standard myopic mean-variance risky asset share. The second term in parentheses reflects the fact that the investor puts an amount X into the risk-free asset to ensure that the future stream of net committed consumption can be financed with certainty. The remainder of wealth is then invested according to the standard mean-variance portfolio rule.

We use this simple portfolio choice model to consider investor heterogeneity and to ask which types of investors are most prone to sell when investors' views about the return distribution change in a tumultuous period. First, we allow for heterogeneity in beliefs about expected stock returns. The log expected returns of investor j is

$$\mu_j = v_j - p \qquad (2)$$

where v_j is the log expected value at the end of the investor j's time horizon and p is the current log price of the risky asset. Investors are heterogeneous in their views about v_j, which results in differences in expected returns. Furthermore, we consider heterogeneity in risk aversion, γ_j, and in the net committed consumption level X_j.

Now suppose that, going into a new period, these heterogeneous investors hold their optimal risky asset shares α_j. Going into this new period, asset prices may change, which would mechanically change the risky asset share if investors remained passive without trading. We denote this passive risky asset share with $\alpha_j(p)$. Moreover, investors' perception of the return distribution changes. This change in perception as well as the desire to rebalance following a price change can prompt investors to seek a risky asset share that differs from $\alpha_j(p)$ unless the current risky asset price and, consequently, the expected return adjusts to make them willing to continue to hold $\alpha_j(p)$. Based on (1) and (2), the current price would have to adjust to

$$p_j = v_j - r_F - \left(\frac{\alpha_j(p)\ \gamma_j \sigma^2}{1 - \frac{X_j}{W_j}} \right). \qquad (3)$$

Heterogeneity implies that different investors may need a different p_j to remain satisfied with their $\alpha_j(p)$. Because there can be only one price in equilibrium, there will be trade. Those who need a higher p_j to stick to their passive risky asset share will be sellers in equilibrium. Those who are happy with a lower p_j to remain at $\alpha_j(p)$ will be buyers. By taking the derivative of (3) with respect to the parameter that changes when entering into a tumultuous period (σ, X, and p),[4] plugging in the optimal solution (1) for α_j, and then taking the derivative with respect to the

[4] In equilibrium, the changes in σ^2 and X may cause a simultaneous change in p, but for the sake of analytical clarity we consider the first-order effect of a change in p separately.

investor characteristic of interest (e.g., γ), we can determine how the investor characteristic affects the desired p_j and hence the propensity to sell.

Table 1 summarizes the results. In this analysis, we assume $\mu_j > r_F$. The first parameter change associated with market tumult we consider is a rise in volatility σ. As the table shows, investors who are optimistic about the expected returns are sellers when volatility rises. Neither heterogeneity in risk aversion nor the net consumption commitment level generate heterogeneity in selling when volatility changes.

The second change we consider is a rise in the present value of the net consumption commitment stream. This rise could result from a tumult-induced downward revision in investors' anticipated labor income stream or an upward revision in their consumption commitments. The investor then needs to invest more into the risk-free asset to ensure that sufficient funds will be available to match the revised net consumption commitment stream. If tumult causes investors to be more sensitive or attentive to their consumption commitments than they are in other periods, this would generate similar effects.[5] The strength of this response is heterogeneous. As Table 1 shows, investors that come into the tumultuous period with high X_j are sellers. Moreover, optimists are relatively more likely to be sellers.

Finally, the rise in volatility in tumultuous periods is often accompanied by a fall in stock prices, consistent with the relationship discussed in (3). Such a price change mechanically changes the risky asset share unless the investor rebalances. The last row in Table 1 shows the rebalancing behavior of investors in response to this price change. Risk-tolerant investors sell in response to a fall in stock prices, while investors with high risk aversion buy. The result in our model is consistent with similar results in fully specified equilibrium models with heterogeneous

[5] Santos and Veronesi (2017) consider state-dependent habit sensitivity in a model with habit preferences, which is a closely related phenomenon.

risk aversion and non-myopic investors such as Kimball et al. (2015) and Chan and Kogan (2002). The intuition is the same as in our simplified setting: investors with higher risk aversion or more pessimistic beliefs hold a lower share of wealth in stocks. A drop in stock prices shrinks their stock holdings by a much greater percentage than it shrinks their wealth. As a consequence, they have a greater desire to rebalance and buy stocks than risk-tolerant or optimistic investors whose stock holdings have shrunk, relative to their decline in wealth, to a smaller degree or not at all.

Empirically, we cannot measure optimism, risk aversion, and net consumption commitments directly. However, there are several taxpayer characteristics in our data that are likely correlated with these unobserved factors. To proxy for higher risk tolerance and optimism, which predict a higher propensity to sell in response to tumult, we use the following characteristics:

Income. Prior empirical evidence indicates that richer individuals take more financial risk. Heaton and Lucas (2000) and Wachter and Yogo (2010) observe, examining U.S. data, that wealthier households take higher risk in their wealth portfolio. Bach, Calvet, and Sodini (2017) find that individuals in the top 1 percent of the income distribution in Sweden take much higher systematic risks. Risk tolerance and optimistic beliefs both appear to contribute to this willingness to take higher risk. Carroll (2001) finds higher self-reported risk tolerance of individuals in the top 1 percent of the income distribution. Das, Kuhnen, and Nagel (2017) find that high-income individuals are more optimistic about stock market returns.

Where the cutoff is, in equilibrium, between buyers and sellers depends on the wealth distribution. Trade has to aggregate to zero in dollars, i.e., in wealth-weighted terms. Before the start of the Great Recession, the top 1 percent by household income owned, directly or indirectly through delegated investments, about half of all publicly traded stock (Wolff 2010). To the extent that high-income individuals are less risk-averse and more optimistic than lower income groups, one might then see higher sales concentrated at the very top of the income distribution.

Dividend income. Higher stock holdings might indicate low risk aversion or optimism about payoffs from stock investments. We use dividend income as a proxy for stock holdings, as do Heaton and Lucas (2000) in their analysis of tax data.

Private business ownership. Private business wealth is very risky (Moskowitz and Vissing-Jørgensen 2002). Heaton and Lucas (2000) find that private business owners have a much greater share of their wealth invested in equity (including private and public equity) than non-business owners. Overall, the prior evidence is consistent with risk tolerant and/or optimistic individuals being attracted to private business ownership. We measure whether taxpayers receive income from partnerships or S-corporations before the rise in market tumult.

We use the following characteristics to proxy for high positive consumption commitments, which predict a higher propensity to sell in response to tumult.

Age. Older individuals have a smaller present value of future labor income as a share of wealth compared with younger individuals (Heaton and Lucas 2000). Thus, net of their labor income stream, consumption commitments are more likely to be positive for older taxpayers.

Receipt of social security income. Social security receipt is particularly likely to occur in retirement. Moreover, retired individuals have largely lost the option to flexibly adjust their work effort in response to shocks (Chai et al. 2011). We use receipt of social security income as a proxy for retirement.

Mortgage interest deduction. Presence of a mortgage interest deduction indicates the presence of committed consumption of housing services.

In addition to the motives for trading that arise from risk sharing in standard portfolio theory, these demographic characteristics could also capture a number of additional effects that we discuss next. As a result, Table 1 might not fully characterize the expected relationship between taxpayer characteristics and sales behavior.

2.2 Volatility timing

Some investors may perceive volatility timing opportunities during tumult. Moreira and Muir (2016) suggest that temporarily reducing stock market exposure following a burst of high volatility may be a utility-improving timing strategy. This strategy is related to the portfolio policies analyzed above, but it is possible that relatively sophisticated investors would be most likely to attempt such volatility timing. The level of financial sophistication is likely positively related to income. Overconfidence in market-timing abilities could reinforce this behavior. Graham, Harvey, and Huang (2009) document that high-income investors are more prone to overconfidence.

2.3 Belief updating

Malmendier and Nagel (2011) provide evidence that young individual investors are more sensitive to recent returns when forming expectations about future returns or allocating portfolios. Hence, young investors could be the ones most likely to sell. Combined with the retirement effect discussed above, this could lead to a U-shaped relationship in age.

2.4 Attention

Investors may be heterogeneous in the degree of attention paid to their portfolios. This is particularly relevant in our study, as we examine reactions at a daily frequency. The reaction to market tumult of a less than perfectly attentive investor is not necessarily instantaneous, and may not occur at all if the tumult subsides before the individual pays attention. Sicherman et al.

(2016) find that investors with large portfolios log in to their online accounts more often. This suggests that high-income taxpayers could be more attentive to the market. Moreover, it seems plausible that retired investors have a lower opportunity cost of attention and hence are more likely to pay attention.

2.5 Tax-loss harvesting

Tax-loss harvesting is another potential reason for selling. However, even if tumult is associated with a fall in price and therefore with a potential for realization of short-term capital losses that may be beneficial for tax purposes, it is not clear that these losses should be realized immediately. Constantinides (1984) shows that immediate realization of short-term losses is optimal only with zero transaction costs. With transaction costs, optimal tax-loss harvesting requires a gradual realization of short-term losses towards the end of the year. [6] Grinblatt and Moskowitz (2004, 543) argue that "it is common folk wisdom that investors pay attention to the tax implications of their portfolios at the end of the year." Similarly, Hoopes, Reck, and Slemrod (2015) find evidence of taxpayers acquiring information at year-end relevant to tax loss harvesting. Nevertheless, tax-loss selling is potentially one of the reasons why investors sell following tumult. This channel is likely most relevant for high-income investors.

In summary, the theories discussed above suggest that selling behavior could differ based on a number of demographic characteristics. Some of these predictions, e.g., that risk-tolerant and optimistic investors are more likely to be sellers in tumultuous markets, may be counter-intuitive. Whether such heterogeneity really exists and, in particular, whether it is evident at very short time horizons such as the daily frequency that we examine here, is an open empirical question that we turn to next.

---

[6] See Dai et al. (2015) for further reasons for optimal deferral of loss realization, especially for high-income taxpayers.

3. Data

This section describes the confidential administrative data and publicly available data we use, and provides summary statistics and match rates across different data sources.

3.1 Tax Return Data

We use two types of tax-return data on U.S. individuals trading in taxable accounts. The primary source of data is third-party information assembled by brokerages and provided to the IRS and taxpayers on Form 1099-B.[7] We use data from all Form 1099-Bs for trades occurring between January 1, 2008 and December 31, 2009. For any covered financial asset sold in this period, Form 1099-B provides the proceeds from the sale and date, the Committee on Uniform Security Identification Procedures (CUSIP) number identifying the asset, and an anonymized version of the taxpayer identification number (TIN), which for an individual seller is a Social Security number, along with several less relevant items.[8]

The second source of data is demographic information from individual income tax returns (Form 1040) and Social Security information. This information includes age and gender (from Social Security records), number of dependents, whether the individual takes a mortgage interest deduction, and the ZIP code of the filing address. We also observe several income components, including wages and salaries, dividends, interest payments, retirement benefits, and net income from self-employment, many of which are supported by third-party information.

Because we only observe activity in individual taxable accounts, if individuals' propensity to sell assets during times of turmoil systematically differs between non-taxable retirement accounts and taxable accounts, our results will be limited in scope, because they are

[7] For a current year 1099-B, see www.irs.gov/pub/irs-pdf/f1099b.pdf.

[8] For some assets acquired after January 1, 2011, Form 1099-B also lists the date of acquisition, the cost basis, the capital gain or loss, and whether the capital gain is short-term or long-term. We do not use this information in this paper.

only informative about the latter. Furthermore, we observe gross sales, but not purchases, so we cannot provide direct evidence on net sales; we do, however, observe annual dividend income, which is related to stock ownership. We address these issues at length in Section 5.

3.2 Match Rates and Aggregate Statistics

Table 2 provides details of our data selection process and sample statistics. We start with the population of 1.43 billion 1099-Bs filed for tax years 2008 and 2009, representing $37 trillion in total trading volume. This comprised about 22 million distinct taxpayers (individuals and institutions) in 2008, and about 21 million distinct taxpayers in 2009. After eliminating non-trading days and partial trading days, negative trade amounts and seemingly erroneous and very large trades, we are left with 1.41 billion 1099-Bs and $36 trillion of volume.[9] Next, we keep only sales related to individual taxpayers, substantially reducing our sample to 870 million transactions and $9.6 trillion in volume; the excluded trades are executed mostly by entities such as partnerships, corporations, and trusts.[10] Of these 1099-Bs that have a valid Social Security Number as a TIN (individual taxpayers), we discard trades entered into by minors (those under 18), leaving 861 million 1099-Bs in the sample, representing $9.5 trillion in volume. Although many different assets are subject to 1099-B reporting, we focus on stocks and stock mutual funds, which we refer to as stocks for simplicity, represented by 273 million 1099-Bs and $6.8 trillion in trading volume. Finally, because our main income measure derives from average income from 2000 to 2007, we retain only transactions in 2008 and 2009 for taxpayers who appear as the taxpayer or spouse on at least one Form 1040 from 2000 to 2007. This leaves us

---

9 Specifically, we discard data from a trivial number of 1099-Bs (under 10) that are clearly errors (single sales of stock in the tens of billions of dollars) and several large sales apparently related to a single event in a single state. Many large trades remain in our sample; in our sample there are over 13,000 sales over $10 million and over 140 sales over $100 million. We verified as valid by hand a random set of these transactions.

10 If a demographic group is unusually likely to execute trades through such entities, we might misstate the relative sensitivity of these groups' overall sales. Cooper at al. (2015) provide evidence about the ultimate owners of pass-through entities, suggesting that they are substantially more concentrated among high earners.

with a final sample of $6.8 trillion in trading volume across 2008 and 2009—$3.7 trillion in 2008, and $3.1 trillion in 2009. Our total trading volume of $3.7 trillion in 2008 compares to the estimate of $2.2 trillion from the Sales of Capital Assets (SOCA) sample in 2008 (Wilson and Liddell 2013).[11] Additional summary statistics on the final sample are presented in Appendix Table A.1.

On average, the sales volume reported on individual Form 1099-B's amounts to about 6 percent of total equity market sales volume as reported by CRSP. Given estimates that about 73 percent of U.S. equity trading in our sample period is done by computer-driven, high-frequency (HF) traders,[12] a 6 percent coverage rate implies that our data covers the sales volume of a substantial fraction—about 22 percent—of non-HF trading. Because HF traders typically close their positions at the end of each day, they are better viewed as intermediaries that hold temporary positions rather than investors that add risk-bearing capacity to the market. For a study like ours that focuses on who ultimately bears stock market risk, the sales volume of HF traders is not particularly relevant. The non-HF part of sales volume of course also includes the sales volume of mutual funds, hedge funds, and other non-HF institutional investors. We therefore likely capture a substantial part of individual investors' sales volume. Table A.2 shows that although individual 1099-B sales volume increases significantly with our measure of market tumult, our coverage of total trading volume, as measured by the ratio of individual 1099-B transactions to CRSP sales volume, is slightly decreasing in tumult. This could be driven by institutional investors selling more during periods of tumult than individual investors.

---

[11] See https://www.irs.gov/pub/irs-soi/08in03soca.xls. A number of factors might account for the difference between the universe of 1099-B transactions and the sample in the SOCA data assembled by the Statistics of Income Division of the IRS. For 2008, SOCA estimates are based on a sample of 58,521 taxpayers (Wilson and Liddell 2013). Based on conversations with IRS staff, we believe that the data in the SOCA is based on when a return is filed, as opposed to when a trade is executed. Further, the SOCA study only records a limited number of short-term trades (500) per taxpayer, due to the costliness of transcribing Schedule D data.

[12] See MacKenzie (2009), which references estimates by the Tabb Group, a consulting firm.

Unfortunately, we cannot observe institutional trading directly to test this hypothesis further. See the online Appendix for more details.

3.3 Market Turmoil and the Financial Crisis

To proxy for market tumult, we use the Chicago Board Options Exchange Volatility Index (VIX), obtained from the Center for Research in Security Prices (CRSP). The VIX index measures the implied volatility of stock prices based on option contracts sold on the S&P 500 stock index with a one-month maturity.[13] Because it is based on option prices, it is a forward-looking measure of investor uncertainty. It reflects the expected S&P 500 stock-index return volatility at a one-month horizon as well as the risk premium that investors are willing to pay to insure against shocks to volatility over this horizon. The VIX is widely used in academic studies as a measure of tumult in stock markets and the financial system more generally (see, for example, Adrian and Shin (2010), Longstaff (2010), and Nagel (2012)).[14] For purposes of presentation, we divide the VIX by 100 throughout and make any transformations on this re-scaled variable, and often analyze the logarithm of the VIX. Unless noted otherwise, we examine behavior only on full trading days.[15]

Figure 1 plots the evolution of the VIX at a daily frequency. In Panel A, we plot the VIX, in logs and levels, from 2008 to 2009. Until mid-2008, the VIX was low relative to levels seen during the crisis. Starting in the second week of September 2008, the VIX increased

---

[13] The VIX calculated based on the S&P 500 is highly correlated with alternatives such as the VIX based on the Dow Jones Industrial Average or the NASDAQ, with rank correlations in excess of 0.95 between each pair of these measures over our sample period.

[14] The VIX is, to be sure, not the only reasonable measure of market tumult, one alternative being the lagged negative market return. Our qualitative conclusions about investor heterogeneity in their response to market tumult are preserved if we use this alternative measure.

[15] For our sample period, the half-trading days are 7/3/2008, 11/28/2008, 12/24/2008, 12/26/2008, 7/2/2009, 11/27/2009, and 12/24/2009. The market is fully closed on weekends and holidays.

dramatically, from 0.23 on September 8 to 0.80 on October 27.[16] Panel B displays the VIX from September to November 2008. On the day of the Lehman Brothers bankruptcy (September 15, 2008), the VIX increased by 24 percent.[17] The following day, American International Group (AIG) avoided bankruptcy after receiving an $85 billion loan from the Federal Reserve Bank of New York. The next major increase in the VIX occurred on September 29, the day on which Citigroup agreed to purchase Wachovia, the Federal Open Market Committee (FOMC) expanded swap lines with several other central banks, and the U.S. House of Representatives rejected legislation proposed by the Department of Treasury regarding the purchase of troubled assets. On October 14, the Treasury Department announced the Troubled Asset Relief Program (TARP), and the VIX increased considerably on the following day. Ten days later, the VIX reached a new peak, when National City Bank was purchased by PNC. Almost a month later, on November 18, executives of three large U.S. auto companies testified before Congress and requested TARP funds, triggering an increase in the VIX that began to turn around only on November 21. The VIX peaked on November 20 (at 0.81), and then began to decrease toward pre-crisis levels.

4. Evidence on Investor Selling Behavior

This section presents the main results on investor heterogeneity in the propensity to sell stock during periods of high market tumult as measured by the VIX index. To establish a baseline, we begin with an analysis at the aggregate level of the individual taxpayer population before we examine, in our main estimation, the heterogeneity around this baseline.

Table 3 presents a regression where we add up gross sales across all taxpayers in the IRS data each day. We then regress the log of the aggregate value of stocks sold on changes in lagged

---

[16] Roughly speaking, a VIX value of 23 (scaled to 0.23) means that option prices imply that a one standard deviation movement in the S&P 500 over the next month is 6.6 percent (=$23/\sqrt{12}$) of the current index level or 23 percent annualized.

[17] This narrative is based on the account in https://www.stlouisfed.org/financial-crisis/full-timeline.

log VIX. We use changes in VIX because stock sales presumably reflect a largely unanticipated *change* in an investor's desired stock holdings that is caused by an unexpected *change* in the conditions perceived by the investor. Because the level of the VIX is highly persistent at a daily frequency, changes in VIX measured over a short period are quite close to unexpected innovations. Because investors might not always respond instantaneously to these shocks, we include ten lags of log VIX changes and we report the sum of the ten coefficients along with the standard error for this sum of coefficients. In this way, we allow for a delayed sales reaction. We show in the Appendix that alternative measures of tumult yield similar results. The table shows that market tumult induces individual taxpayers to engage in a substantial amount of stock sales during the following days. Over our sample period, a 10 percent increase in the VIX from day *t-11* to *t-1* is associated with additional sales amounting to roughly 50 percent of typical daily sales volume.[18]

To understand the reasons for taxpayers' stock sales during market tumult, we now look at the data in a more disaggregated way and study heterogeneity in selling behavior. At the aggregate market level, there is a buyer for every seller and hence gross sales equal gross purchases, with zero net sales. However, at a more disaggregated level, when we compare heterogeneous investor types, gross sales can be indicative of net sales. We demonstrate this explicitly in Section 5 with an alternative data set for one particular subgroup of investors—

[18] The timing of the 10 percent rise of the VIX within the lagged time window from day t-11 to day t-1 does not matter for the cumulative effect. For example, if the VIX rises by 1 percent during each of the 10 days leading up to the end of day *t*-1, sales on day *t* are estimated to rise by the sum of coefficients reported in Table 3 times 1 percent, i.e., 4.9 × 1% = 4.9%; sales on day *t*-1 and day *t*+1 rise by 0.9 × 4.9 × 1%; sales on day *t*-2 and *t*+2 rise by 0.8 × 4.9 × 1%; and so on, which yields a total cumulative effect of 49 percent of daily sales volume. Alternatively, if the entire 10 percent rise in the VIX occurs on day *t*-1, with zero changes prior to *t*-1, the cumulative effect over the days *t* to *t*+9 is given by the sum of coefficients times 10 percent, which is again 4.9 × 10% = 49 percent of daily sales volume.

discount brokerage customers—for which we observe both gross sales and net sales and we find a strong relationship between them. For now, we focus on gross sales, measured using IRS data.

4.1 Empirical Specification

To estimate the heterogeneous sensitivity of stock sales to changes in the log VIX, we estimate individual-level regressions based on the following specification,

$$y_{it} = x'_i\gamma_t + \delta_t + \varepsilon_{it} \qquad (4a)$$

$$\gamma_t = \gamma_0 + \beta\Delta V_{t-1} + \xi_t, \qquad (4b)$$

where y_{it} is a measure of individual i's sales on date t, and these stock sales relate to a vector x_i of K individual characteristics in a time-varying fashion, as captured by the vector of coefficients γ_t. Projecting γ_t on ΔV_{t-1}, a vector of ten one-day changes in log VIX from day *t-11* to *t-1*, we decompose this time-variation into a part related to changes in volatility and an orthogonal residual ξ_t, as shown in (4b). Our main interest centers on estimating the $K \times 10$ matrix β, which describes how the sensitivity of sales to changes in volatility varies with the observed characteristics in x_i. Because we include a time fixed effect δ_t in (4a), we identify β purely from cross-sectional differences in the sales-VIX sensitivity between individuals with different characteristics, not from the time-series correlation of aggregate stock sales and ΔV_{t-1} that we examined in Table 3 above. By adding up the ten coefficients for each characteristic in β, we capture the overall response to a change in tumult. In Appendix Tables A.3 and A.4, we present alternative estimates using a single lagged one-day change in log VIX as the measure of tumult, rather than ten one-day changes. This alternative specification yields qualitatively similar results.

We estimate equations (4a) and (4b) for our sample period of 2008 and 2009. The vector x_i includes indicators for specified percentiles of AGI (averaged from 2000 to 2007), age ranges (as of December 31, 2008), percentiles of dividend income (averaged from 2000 to 2007), along with indicators for contemporaneous receipt of social security income, whether the individual received positive income from partnerships or S corporations in any year from 2000 to 2007, gender, and marital status (married or unmarried, based on filing status), and the number of days an individual traded in the year 2007, which measures how actively the individual trades in general.

Our strategy for estimating and interpreting the parameters in equations (4a) and (4b) must address two challenges. First, most individuals do not trade on any given day, so that how we handle zero-sales observations is important. We address this by analyzing both a linear probability model (LPM) specification, in which y_{it} is an indicator for whether individual i had any sales on date t, and a continuous specification, in which y_{it} is the logarithm of the dollar value of sales by individual i on date t. In the latter case, zero-sales observations are excluded from the estimation sample.[19] We also exclude individuals who never realize any taxable capital gains during the given year. Some of these excluded individuals may well own stock and could be sellers in other years, but many of them are likely to not directly participate in the stock market (in taxable accounts). For this reason, we do not find it meaningful to include these individuals for which the dependent variable is always zero. Further, including all 140 million taxpayers would expand the sample size enormously to more than 70 billion observations, creating additional computational difficulties. Excluding those who never sell in a given year leaves us with fewer than a billion observations, which, when estimated at the day level (explained later), is more reasonably estimable with our IRS computing resources.

Second, the residuals ε_{it} are likely to be cross-sectionally correlated among groups of millions of individuals. To deal with both problems, we use a two-step procedure inspired by

[19] Using an inverse hyperbolic sine transformation of sales as y_{it}, which resembles the log sales specification for positive values but permits the inclusion of zero-sales observations, generates very similar results.

Fama and MacBeth (1973). In the first stage, we run cross-sectional, within-day regressions of the sales variable on individual characteristics to estimate (4a) and get a time series of daily estimates of the vector γ_t. These estimates tell us, for example, whether individuals in the highest income group were more or less inclined than members of other groups to sell on a given day. Using these daily estimates of γ_t as the dependent variable, we then estimate (4b) as the second stage.[20] Based on the estimates of the elements of β obtained in this second stage, we can infer, for example, whether a greater tendency to sell by members of the highest income group is more pronounced on days following a rise in the VIX index. If members of this group raise their sales more than other groups following high ΔV_{t-1}, the group's element of γ_t will be high following days when the VIX went up and the corresponding entry of β will therefore be positive; if the magnitude of this group's propensity to sell relative to other groups does not vary with ΔV_{t-1}, the element of γ_t corresponding to this group's membership indicator will be constant across time and the hence the group's entry in β will be zero.

As in standard asset pricing and corporate finance applications of the Fama-MacBeth method, cross-sectional correlation in ε_{it} is properly accounted for because any such correlation raises the estimation error and hence the time-series variation of the daily first-stage estimates of γ_t, which in turn boosts the variance of the residuals in the second-stage regressions and thus the standard error of the β estimates. Auto-correlation in ε_{it} would show up as autocorrelation in the estimation error of γ_t and hence autocorrelation in the second-stage residuals. We address this issue by estimating Newey-West standard errors in the second-stage regression that allow for autocorrelation in the residual up to a maximum of 10 days.

[20] Because the daily first-stage estimates of γ_t are subject to estimation error $x_i'\varepsilon_{it}$, our sample version of (4b) includes this estimation error in the residual in addition to ξ_t. However, because we have time dummies in the first stage regression (i.e., a different intercept each day), the estimation error in the elements of γ_t is by construction orthogonal to a time-series variable like ΔV_{t-1}, so we can consistently estimate β with this two-step approach even if unobserved time-series factors affect sales volume in (4a).

4.2 Results

Table 4 presents results on heterogeneity in investors' tendency to sell. All coefficients are scaled by 100 to facilitate interpretation. Columns 1 and 2 present the estimates of γ_0 and the row sums of the β matrix, respectively, for the specification including all trading days in our sample period.[21] From the estimates of γ_0 in column 1 of Table 4, we see that, perhaps unsurprisingly, sales occur more frequently in the highest-income groups. For instance, holding other characteristics fixed, the fraction of individuals in the top 0.1 percent of the AGI distribution who sell on a given day is on average 3.7 percentage points higher than for individuals in the bottom 75 percent – the "left-out" category among the income groups. This difference is substantively large, as on average 1.7 percent of all individuals in our sample trade on a given day. We also see modestly higher sales probabilities for individuals with higher average dividend income, individuals who previously received partnership or S corporation income, and married individuals. A few other characteristics attract statistically significant coefficients that are very small in magnitude.

Our principal interest is column 2, which presents estimates of the heterogeneous response of the probability of selling to stock market volatility. As in Table 3, the cumulative effect including any delayed reactions equals the percentage change in the VIX times the sum of the 10 coefficients we report (scaled by 100) in column 2 (see footnote 18). To facilitate interpretation of these results, we also calculate the implied change in the cumulative probability of selling from an increase in log VIX of 0.3, which is roughly the increase that occurred in the 10-day period around the Lehman Brothers collapse on September 15, 2008 (see also Figure 1). Figure 2 plots the results from multiplying the estimated effects in column 2 of Table 4 by 0.3,

[21] Appendix Table A.3 presents results from using a single lagged change in log VIX, and Appendix Table A.4 presents results using lagged negative market returns as our proxy of tumult.

thus depicting heterogeneity in the probability of selling due to a large but realistic increase in volatility. The estimates demonstrate substantial heterogeneity in the response of sales to changes in VIX by individual characteristics.

The regression results indicate that individuals in the top 0.1 percent of the AGI distribution respond substantially more to increases in the VIX. A 30 percent increase in the VIX over the previous 10 days increases the probability that an individual in the top 0.1 percent sells by roughly 4 percentage points cumulatively relative to individuals in the bottom 75 percent of the income distribution. This gap equals roughly the difference in average daily trading propensities between individuals in the top 0.1 percent and bottom 75 percent, controlling for other characteristics, that we obtained from the γ_0 estimates in column 1. For individuals in the top 1 percent but not the top 0.1 percent, this estimated difference is 1.9 percentage points, which is slightly bigger than the difference in average daily trading propensities between these groups shown in column 1. As we discussed in Section 2, a higher selling propensity of high-income individuals in response to a rise in volatility and falling stock prices can be rationalized if high income proxies for risk tolerance, relative optimism and, possibly, sophistication as an investor. Even so, it is surprising to see that selling activity surged within the very highest ranks of the income distribution. This is very different from the popular story of the panic-prone small investor.

As discussed above, portfolio choice theory suggests retirement should be a strong predictor of response to market tumult. For our sample period, we find that individuals with social security income, which proxies for retirement, are significantly more likely to sell in response to an increase in volatility than are individuals not receiving social security income. The magnitude of this effect suggests that a 30 percent increase in VIX increases the probability of selling for individuals receiving social security by about 0.4 percentage points cumulatively relative to those who do not receive social security. The estimates of the effect of the age indicators, conditional on receiving social security income, are very small in magnitude. Thus, retirement status rather than age seems to determine investment behavior in times of tumult.

Observed characteristics associated with capital ownership and financial sophistication strongly predict sensitivity to increases in VIX. Holding other variables, including AGI, constant, dividend income is positively associated with volatility sensitivity. The effect sizes for the top 5 percent of the dividend income distribution are comparable in magnitude to estimates for the top 1 percent of the AGI distribution, suggesting that even holding AGI fixed, high-wealth individuals are more responsive to increases in VIX. The effects are statistically insignificant for the top two groups, although the point estimates are sizable. Likewise, the total number of days in which an individual traded in the year 2007 is significantly and positively related to VIX sensitivity. Finally, the receipt of partnership and S corporation income is positively related to volatility sensitivity. Gender and marital status have a relatively small association with sensitivity to changes in VIX, although some coefficients are statistically significant.

Another way to see the estimated relationship between the likelihood that an individual with given characteristics sells and market tumult is to simply plot the γ_t coefficients from equation (4a). These coefficients measure the effect of some characteristic on selling on date *t,* holding other characteristics fixed. If certain characteristics are associated with sensitivity to tumult, we should expect that the component of γ_t for those characteristics to spike following the significant events of the financial crisis (compare to equation (4b), where we explicitly measure tumult using the change in log VIX). We plot these coefficients for AGI and dividend income in Figure 3. To compare the coefficients in a meaningful way, we normalize the coefficients by subtracting their mean over the full sample period, which serves a similar purpose as the inclusion of γ_0 in the second stage equation (4b). For the same groups that Table 4 shows to be especially likely to sell in response to tumult, we see noticeable spikes in γ_t on the important dates in the financial crisis depicted in Figure 1, such as the Lehman collapse and the announcement of TARP. Recall that the analysis contained in Figure 2 suggested that the spike in volatility occurring with the Lehman collapse should be associated with a 4 percentage point increase in sales by the top 0.1 percent income group. In Figure 3, we observe a spike in the coefficient for the top income groups of about 3.5 percentage points, which is similar in magnitude.[22] Thus, Figure 3 constitutes visual evidence corroborating our findings regarding tumult sensitivity, using a method that does not depend on any specific measure of tumult. Appendix Figure A.3 plots these coefficients for other characteristics studied in Table 4; for these characteristics, the effects are much smaller or zero, and thus the time series of coefficients do not appear to be strongly responsive to the events of the financial crisis.

Table 5 presents results using log sales as the dependent variable in equation (4a).[23] The columns of Table 5 present estimates exactly as in Table 4, and we continue to scale coefficients by 100. Figure 4 plots the cumulative change in log sales for a 30 percent increase in VIX implied by the interaction estimates in order to facilitate interpretation, exactly as in Figure 2. As with the LPM results, sales amounts are typically larger for individuals with higher AGI and dividend income, individuals receiving partnership and S corporation income, married

[22] To be concrete, the regression results generating Figure 2 are based on changes in VIX over 10 days, not a one-day change like the one after the Lehman collapse. However, the spike in VIX due to Lehman Brothers dwarfs the other changes in the days before and after. The regression also included time fixed effects that are not used in Figure 3; one can observe visually in Figure 3 that the spike in sales probabilities in high-income groups after the Lehman collapse will be much larger than the fixed effect for that day. Given all this, it is reassuring that the implied results from the regression agree with the non-parametric plot of coefficients in Figure 3.

[23] Appendix Table A.5 presents results from using a single lagged change in log VIX, and Appendix Table A.6 presents results using lagged negative market returns as our proxy of tumult.

individuals, and individuals who traded on a larger number of days in 2007. In our sample period, a 30 percent increase in the VIX over the previous 10 days increases mean sales by individuals in the top 0.1 percent of the AGI distribution cumulatively by 0.65 log points, or about 91 percent of their typical daily sales volume, relative to individuals in the bottom 75 percent of the income distribution. For individuals in the top 1 percent, but not the top 0.1 percent, mean sales increase by 20 percent. While these effects are large, average daily sales volume is over 50 times as large for individuals in the top 0.1 percent as for individuals in the bottom 75 percent and over 10 times as large for individuals in the top 1 percent (see Appendix Table A.5). Given the average daily sales volume, these estimates imply that individuals in the top 0.1 percent increased their mean sales cumulatively by $7,638 (=11,750 x 0.65) more than individuals in the bottom 75 percent, holding other investor characteristics equal. Because there are 163,000 taxpayers in the top 0.1 percent in 2008, this amounts to a difference of $1.2 billion in total sales following a one-time increase in VIX of 30 percent. Similar calculations imply that total sales by taxpayers in the top 1 percent but not the top 0.1 percent increased by $620 million more than total sales by taxpayers in the bottom 75 percent.

With a few exceptions, heterogeneity in the response of sales amounts in response to volatility is qualitatively similar to heterogeneity in the propensity to sell that we reported in Table 4. Overall, these results suggest that heterogeneity on the extensive margin closely resembles heterogeneity on the intensive margin, conditional on having positive sales.

Figure 5 depicts time series plots of the first stage coefficients for AGI and dividend income groups in the log sales specification, similarly to Figure 3. Once again, we observe that the first-stage coefficients noticeably spike following the salient events of the financial crisis for the characteristics that, in our regression analysis, we found to be strongly and significantly related to tumult sensitivity. We therefore conclude that our results likely do not depend on using any particular measure of tumult. Figure A.4 in the Appendix plots the other first stage coefficients.

5. Evidence that Gross Sales Are Informative about Net Sales

So far we have interpreted the evidence from gross sales volume in taxable accounts as indicative of a likely similar (although quantitatively perhaps accentuated) behavior of net sales across all accounts. As the purchase of a security is not a taxable event, we do not observe purchases of securities in our data. Thus, it is possible that the sales we observe do not represent net sales as investors exit the market or reduce their net exposure to it, but rather individuals reconfiguring their portfolio while maintaining the same broad exposure to the equity market as a whole. This section presents evidence that the gross selling we observe in 1099-B data can be used to infer net selling. To examine whether gross sales track net sales, we use detailed daily trading data from discount brokerage accounts, and we examine the evolution of dividend income in tax-return data. To examine whether sales in taxable accounts track total sales of equities, we use detailed wealth data in the Survey of Consumer Finance.

5.1 Gross and Net Sales in Discount Brokerage Data

We begin by analyzing the Barber and Odean (2000) data set of daily trades in a discount brokerage account from 1991 to 1996.[24] In the brokerage account data, we can observe both gross and net sales (i.e., gross sales minus purchases). We eliminate option trades and trades in fixed-income mutual fund shares. The resulting sample largely comprises trades in domestic common stock and equity mutual fund shares, but it also includes small amounts of trades in

[24] We thank Terry Odean for allowing us to access these data.

such assets as ADRs, Canadian stocks, REITs, and preferred shares. This sample contains roughly 1,000 individual trades per day.

For each trading day, we calculate two aggregate sales numbers for the whole brokerage account sample. The first is the amount of net sales, which is simply the aggregate dollar amount (positive for sales, negative for purchases) added across all brokerage customers each day. The second is the amount of gross sales, which includes only the dollar amount of sales across all brokerage customers. The latter gross sales number corresponds to the sales numbers that we get from the tax-return data. We further observe the aggregate value of brokerage customers' portfolios at the beginning of each month (including all assets, not just stocks and stock mutual funds), and we create versions of gross and net sales expressed as a percentage of this aggregate portfolio value.[25]

We first estimate our baseline regressions from Table 3 with the brokerage account data. Column 1 in Table 6 shows the results from a regression of log gross sales volume in taxable accounts—the equivalent to 1099-B sales volume in the tax-return data—on the contemporaneous change in the log VIX index and two lags. We report the sum of these coefficients in the table. We include the contemporaneous change in log VIX here because brokerage account customer sales—unlike taxpayer sales in the tax-return data—are strongly related to the contemporaneous change and not just lagged changes in log VIX. But the effects of lagged changes also quickly die out in the brokerage account data. For this reason, we include only two lags here, not ten as in Table 3. These differences are to be expected: discount-brokerage customers are more likely to react to same-day news and trade more actively than the average taxpayer. As Table 6 shows, brokerage account customer sales are strongly related to

[25] We take the absolute value of each position in the calculation of the portfolio value; that is, short positions enter with a positive value. We do this because we want to scale trading activity variables with the gross size of an investor's portfolio rather than the net equity of the portfolio.

changes in log VIX. The estimate in column 1 implies that a 10 percent change in VIX from day *t-3* to *t* is associated with a rise in sales volume on day *t* of about 18 percent, but the effect is less persistent during the following days than in the tax-return data because only shorter lags of VIX are involved here. For our purposes, the relevant take-away is that the tax return data and the brokerage account data tell a similar story about the relationship between changes in log VIX and gross sales volume.

Column 2 of Table 6 presents the most important piece of evidence from the brokerage account data. Here we use net sales (which we do not observe in the tax-return data) as the dependent variable and gross sales (which we do observe in the tax-return data) as the explanatory variable, both expressed as a percentage of the portfolio value. The results show that there is a very strong relationship between these two variables. A gross sale of one percent of the portfolio value is associated with a net sale of 0.34 percent. The adjusted R-squared of approximately 27 percent also indicates that there is a strong relationship between gross sales and net sales.

Columns 3 and 4 compare regressions on log VIX changes with gross sales and net sales as dependent variables, both expressed as a percentage of portfolio value. A comparison of the estimates from these two regressions can help us understand to what extent a rise of gross sales in times of market tumult also implies a rise in net sales. Both gross sales and net sales are associated with the log VIX change, with the coefficient estimates on gross sales being about three times as large as with net sales. The estimates reported in columns 3 to 4 together suggest that the behavior of gross sales from the tax-return data is informative about the unobserved net sales.

Unlike the tax-return data, the brokerage account data also contains trades in non-taxable (IRA and Keogh) accounts. This allows us to check whether in tumultuous times the behavior of investors in non-taxable accounts is fundamentally different. We find that they are not. The results reported in column 5 are quite similar to the results for taxable accounts in column 1. Thus, it seems that the results from our analysis of taxable trades in tax-return data could also carry over to some extent to non-taxable accounts.

Finally, column 6 looks at the taxable accounts restricted to customers with large portfolios, defined as those above the 80th portfolio value percentile. This is an imperfect way to approximate the high-AGI sample in the tax-return data. Based on the point estimates, the relationship with the VIX index changes is slightly stronger than in column 1, but the difference is not statistically significant and the magnitude of the difference is much smaller than in our AGI-based sample splits in the tax-return data. Part of the reason could be that the value of the brokerage-house portfolio is not as good a measure of the level of affluence as is AGI in the tax-return data. Moreover, the brokerage customers are a rather special selected sample that likely differs from average taxpayers on a number of dimensions. We also repeated the regressions in columns 3 and 4 with the large-portfolio sample (not reported). We find that the estimated coefficients on log VIX changes are slightly higher than those reported in columns 3 and 4.

5.2 Changes in Dividend Income

Next we present evidence that annual gross sales by a given individual are associated with decreases in dividend income reported on that individual's tax return (Form 1040 Schedule B). Intuitively, one can think of qualified dividend income as a rough proxy for the amount of stocks held in an individual's portfolio.[26] If gross sales are associated with net sales, then an

[26] A qualified dividend is one that is taxed at the preferential lower tax rate. Regular dividends paid out to shareholders of for-profit U.S. corporations are generally qualified.

individual's portfolio should contain less stock after a year of high gross sales, and thus the individual's dividend income should decrease.

We run simple regressions of the change in dividend income from year *t-1* to year *t+1* on gross sales in year *t*, where *t* is either 2008 or 2009.[27] We restrict the sample to individuals receiving dividends in year *t-1*, and we winsorize gross sales and dividend income changes at the 1 and 99 percent levels to eliminate the effect on the results of some obvious data errors.[28]

Table 7 reports the results of this analysis. In columns 1 and 2, we document a statistically significant relationship between gross sales and decreases in dividend income. To assess whether the coefficient we estimate is reasonable and consistent with the analysis of the discount brokerage data in Table 6, consider $1 of gross sales on some day. The earlier analysis suggests that $1 of gross sales corresponds on average to $0.33 of net sales on the same day. Suppose that the $0.33 reallocated from stocks on that day is not invested back into stocks within one year. Then the decrease in dividend income will be roughly $0.33 times the dividend yield in the individual's portfolio. For the average individual, we expect the dividend yield to be somewhere near the S&P 500 dividend yield of 2 percent. In this case the drop in gross sales would be about $0.33*0.02 = $0.0066. This number is nearly identical to our estimated coefficients in Table 7, which are 0.0065 and 0.0068.

This exercise relies on a number of assumptions. Our interpretation requires that changes in dividend yields from year *t-1* to year *t+1* should be reasonably unrelated to gross sales, and to the share of gross sales that pass through to net sales. For example, the first condition fails if individuals disproportionately sell dividend-paying stocks, and the second fails if individuals sell

[27] We have also estimated regression specifications with transformed versions of the same dependent and independent variables, including logarithmic specifications and those in which all variables are scaled by adjusted gross income. In all instances, the qualitative results are the same. We prefer the specifications reported here because their interpretation is relatively straightforward.

[28] The results are nearly identical if we also exclude individuals with zero gross sales in the given year.

high-dividend-paying stocks while buying low-dividend paying stocks. While this analysis is an imperfect test of the relationship between gross and net sales for the reasons described above, we believe the most plausible explanation for the strong negative association between gross sales in year *t* and changes in dividend income from year *t-1* to year *t+1* is that gross sales are associated with net sales, especially given that the magnitudes of the estimated coefficients so closely align with this interpretation.

We also use changes in dividend income to test an implicit assumption made above, that the relationship between gross sales to net sales does not vary across groups. Specifically, if this implicit assumption is satisfied, the relationship between dividend income changes and gross sales should be roughly constant across groups. Columns 3 and 4 of Table 7 report the results of this test for AGI groups: we interact the specification in columns 1 and 2 with the AGI groups used in Section 4. The negative coefficients on the interaction between gross sales and high-AGI group membership suggests that individuals in the higher-AGI groups have a *higher* rate of pass-through from gross to net sales than people in the bottom 75 percent of the income distribution. While there may be heterogeneity in pass-through rates, heterogeneity of the kind suggested by these results would actually *strengthen* our interpretation of the results in Section 4 that high-income groups disproportionately sold out of the stock market during the financial crisis. The interpretation of the regressions in columns 3 and 4 in terms of pass-through rates from gross to net sales is subject to similar caveats about dividend yields described in the previous paragraph.

5.3 Taxable and Non-Taxable Accounts

Using data from the 2007-2009 panel of the Survey of Consumer Finances (SCF), we next provide suggestive evidence that our inability to observe activity in non-taxable accounts does not confound the qualitative results described in Section 4. The SCF contains detailed information on wealth for 3,857 households interviewed in late 2007 and late 2009, and the survey deliberately oversamples high-wealth individuals (see Bricker et al (2011) for an overview). Importantly, the data allow us to examine separately wealth in taxable accounts, which includes directly held stock, mutual funds, and hedge funds, and wealth in non-taxable accounts, where the latter includes tax deferred retirement accounts, trusts, other managed assets, and annuities.[29] Using this data, we construct measures of 1) equities held in taxable accounts, 2) equities in all accounts, 3) net sales or purchases of equities in taxable accounts, and 4) net sales or purchases of equities in all accounts.[30]

How might our inability to observe non-taxable accounts influence our results? The percent change in an individual's overall equity holdings sold in response to an uptick in volatility (our principal parameter of interest) depends on (1) the percent change in their taxable equities, (2) the share of overall equities held in taxable accounts, and (3) the relative intensity of their stock trading in taxable accounts. Our main results suggest that (1) is higher for certain groups, like those with higher income. A comparison of (1) alone, however, could be misleading if higher-income individuals hold a smaller share of wealth in taxable accounts and/or they execute more of their equity sales in their taxable accounts.

Figure 6 plots the share of wealth held in taxable accounts by income. We use similar group definitions as elsewhere in the paper, but because of data limitations we use income in 2007 rather than average AGI from 2000-2007 and, due to power concerns, we group the top 0.1 percent of the income distribution with the rest of the top 1 percent. Examining this figure rules out the first potential pitfall, that higher-income individuals hold a smaller share of wealth in their taxable accounts. Indeed, the opposite is true: high-income people hold a higher share of

[29] The data do not include wealth held by foundations controlled by an individual.

[30] To be comparable with the IRS data, we consider a transaction in the SCF to be taxable if it would lead to a reported sale on a 1099-B linked to an individual taxpayer.

their wealth in taxable accounts, perhaps due to the limits on contributions to tax-deferred retirement accounts. These facts on their own suggest that the heterogeneity across income groups in responses to market tumult is likely higher than what we document.

Although our results are clearly not driven by differences in the share of wealth held in taxable accounts, it could still be the case that higher-income individuals conduct much more of their volatility-driven net sales in taxable accounts, while lower-income individuals mix their activity between taxable and non-taxable accounts. This could cause our results to be misleading, as the overall sales of lower-income individuals would be higher than what we measure, and maybe not that different from the high-income individuals.

To address this second problem, we regress across the individuals in the SCF the change in total stock holdings between 2007 and 2009 on the change in stock holdings in taxable accounts, with and without an interaction with income group indicators. When calculating these changes, we adjust the stock holdings in 2009 for the change in the Wilshire 5000 Total Market Stock Index between the survey dates in 2007 and 2009. The remaining change in stock holdings equals approximately the amount of stocks bought or sold. This exercise is similar in spirit to the regression comparing gross and net sales in Table 6 (column 2), but we here compare net taxable sales and total net sales. If individuals conduct all their trading in taxable accounts and no trading in non-taxable accounts, or if trading in non-taxable accounts is uncorrelated with trading in taxable accounts, the slope coefficient in such a regression would be approximately one: a dollar in net taxable sales is associated with a dollar in total sales. If selling in taxable and non-taxable accounts is positively correlated, this coefficient would be larger than one. Another possibility is that individuals tend to sell in taxable accounts when they buy in non-taxable accounts, in which case the coefficient would be less than one. The main caveat to this approach is that not all variation in net sales in these data is a response to market tumult, although to be sure a large amount of activity between 2007 and 2009 was driven by the tumult of the financial crisis.

Table 8 reports the results of the regression. The slope coefficient is 0.92; this estimate is statistically different from zero ($p < 0.001$) but not from one ($p \approx 0.37$). When we include interactions for income groups in column 2, we find that the group interaction terms are all statistically insignificant, and the point estimates are relatively small relative to the overall effect.[31] Thus, we find no evidence that the relative trading activity in taxable versus non-taxable accounts confounds our main results. These results also rule out that gross sales in taxable accounts over this period were primarily due to shifting assets from taxable to non-taxable accounts (in which case the estimated coefficient would be zero). To be sure, this exercise is suggestive rather than dispositive, as due to data limitations it does not directly analyze the response of equity holdings in various accounts to tumult, but rather the overall variation in equity holdings.

6. Conclusion

In this paper we have investigated which types of individuals sell stocks during periods of stock market turmoil. We use administrative data from the Internal Revenue Service consisting of billions of third-party reports on all sales of stock in United States taxable individual accounts, to understand which individuals sold during the tumultuous market events of 2008 and 2009. On many dimensions, this data set is vastly superior to the kinds of data that have been brought to bear heretofore on related questions. The unique advantage of our data is that we can identify for each sale exactly who sold securities on each day, and characterize the sellers by the demographic information available on income tax returns and matched Social Security records.

[31] If we include interactions for age groups, the point estimates for interactions are also small and statistically insignificant.

Overall, our results show that there is substantial heterogeneity in investors' responses to market tumult. We find that individuals at the top 1 percent, and even the top 0.1 percent, of the income distribution are much more responsive to market tumult than individuals at the bottom 75 percent, even accounting for the higher amounts of sales overall of individuals in top income groups. We also find that retired individuals, and individuals likely to hold high amounts of wealth in stocks, according to various proxies (i.e., large amounts of dividend income), are disproportionately responsive to tumult.

Our examination of a model of portfolio choice suggests that, while perhaps surprising to some readers, these results are not inconsistent with standard portfolio choice theory. Prior literature suggests that high-income and high-wealth taxpayers have higher risk tolerance and greater optimism, which the theory predicts should not only lead to higher risky asset holdings on average, but also a greater propensity to sell risky assets in response to tumult. Perceived market-timing skills or higher attention of financially more sophisticated investors could also play a role. Retired individuals face consumption commitments and have little labor income to shelter themselves from negative shocks to income from risky assets, which theory suggests should lead them to sell in response to tumult.

Our analysis thus uncovers important dimensions of investor heterogeneity that contribute to the elevated level of trading activity during tumultuous periods in the stock market. An explanation of price movements during these episodes likely needs to take into account these shifts between investor groups in their willingness to bear risk. Incorporating these sources of heterogeneity into asset pricing models with time-varying risk would be an interesting avenue for further research.

References

Adrian, T., and H. S. Shin. 2010. Liquidity and Leverage. *Journal of Financial Intermediation* 19 (3). Risk Transfer Mechanisms and Financial Stability: 418–437.

Anderson, A. M., and E. A. Dyl. 2007. Trading Volume: NASDAQ and the NYSE. *Financial Analysts Journal* 63 (3): 79–86.

Bach, L., L. E. Calvet, and P. Sodini. 2017. Rich Pickings? Risk, Return, and Skill in the Portfolios of the Wealthy.

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

Barboza, D. 1998. The Market Turmoil: The Investors; Small Stockholders Again Refuse to Panic. *The New York Times*, August 6, sec. Business Day.

Behr, P., and D. A. Vise. 1987. Wave of Selling Sweeps Across International Borders. *The Washington Post*, October 20.

Ben-David, I., F. Franzoni, and R. Moussawi. 2012. Hedge Fund Stock Trading in the Financial Crisis of 2007–2009. *Review of Financial Studies* 25 (1): 1–54.

Bricker, J., B. K. Bucks, A. Kennickell, T. L. Mach, and K. Moore. 2011. Drowning or Weathering the Storm? Changes in Family Finances from 2007 to 2009.

Calvet, L. E., J. Y. Campbell, and P. Sodini. 2009. Measuring the Financial Sophistication of Households. *The American Economic Review* 99 (2): 393–398.

Campbell, J. Y., and L. M. Viceira. 2002. *Strategic Asset Allocation: Portfolio Choice for Long-Term Investors*. Clarendon Lectures in Economic.

Carroll, C. D. 2001. Portfolios of the Rich. In *Household Portfolios*, edited by L. Guiso, M. Haliassos, and T. Jappelli. Cambridge, MA: MIT Press.

Cella, C., A. Ellul, and M. Giannetti. 2013. Investors' Horizons and the Amplification of Market Shocks. *Review of Financial Studies*: hht023.

Chai, J., W. Horneff, R. Maurer, and O. S. Mitchell. 2011. Optimal Portfolio Choice over the Life Cycle with Flexible Work, Endogenous Retirement, and Lifetime Payouts. *Review of Finance* 15 (4): 875–907.

Chan, Y. L., and L. Kogan. 2002. Catching Up with the Joneses: Heterogeneous Preferences and the Dynamics of Asset Prices. *Journal of Political Economy* 110 (6): 1255–1285.

Constantinides, G. M. 1984. Optimal Stock Trading with Personal Taxes. *Journal of Financial Economics* 13 (1): 65–89.

Cooper, M., J. McClelland, J. Pearce, R. Prisinzano, J. Sullivan, D. Yagan, O. Zidar, and E. Zwick. 2015. Business in the United States: Who Owns it and How Much Tax Do They Pay?

Dai, M., H. Liu, C. Yang, and Y. Zhong. 2015. Optimal Tax Timing with Asymmetric Long-Term/Short-Term Capital Gains Tax. *The Review of Financial Studies* 28 (9): 2687–2721.

Das, S., C. Kuhnen, and S. Nagel. 2017. Socioeconomic Status and Macroeconomic Expectations. *Working Paper*.

Dorn, D., and M. Weber. 2013. Individual Investors' Trading in Times of Crisis: Going It Alone or Giving Up? *Working Paper*.

Fama, E. F., and J. D. MacBeth. 1973. Risk, Return, and Equilibrium: Empirical Tests. *Journal of Political Economy* 81 (3): 607–636.

Graham, J. R., C. R. Harvey, and H. Huang. 2009. Investor Competence, Trading Frequency, and Home Bias. *Management Science* 55 (7): 1094–1106.

Griffin, J. M., J. H. Harris, T. Shu, and S. Topaloglu. 2011. Who Drove and Burst the Tech Bubble? *The Journal of Finance* 66 (4): 1251–1290.

Grinblatt, M., and T. J. Moskowitz. 2004. Predicting Stock Price Movements from Past Returns: The Role of Consistency and Tax-Loss Selling. *Journal of Financial Economics* 71 (3): 541–579.

Guiso, L., P. Sapienza, and L. Zingales. 2013. The Determinants of Attitudes toward Strategic Default on Mortgages. *The Journal of Finance* 68 (4): 1473–1515.

Heaton, J., and D. Lucas. 2000. Portfolio Choice and Asset Prices: The Importance of Entrepreneurial Risk. *The Journal of Finance* 55 (3): 1163–1198.

Hoffmann, A. O. I., T. Post, and J. M. E. Pennings. 2013. Individual Investor Perceptions and Behavior During the Financial Crisis. *Journal of Banking & Finance* 37 (1): 60–74.

Hoopes, J. L., D. H. Reck, and J. Slemrod. 2015. Taxpayer Search for Information: Implications for Rational Attention. *American Economic Journal: Economic Policy* 7 (3): 177–208.

Hudomiet, P., G. Kézdi, and R. J. Willis. 2011. Stock Market Crash and Expectations of American Households. *Journal of Applied Econometrics* 26 (3): 393–415.

K. Brunnermeier, M., and S. Nagel. 2004. Hedge Funds and the Technology Bubble. *The Journal of Finance* 59 (5): 2013–2040.

Karpoff, J. M. 1987. The Relation Between Price Changes and Trading Volume: A Survey. *The Journal of Financial and Quantitative Analysis* 22 (1): 109–126.

Kimball, M. S., M. D. Shapiro, T. Shumway, and J. Zhang. 2015. Portfolio Rebalancing in General Equilibrium. *Working Paper*.

Longstaff, F. A. 2010. The Subprime Credit Crisis and Contagion in Financial Markets. *Journal of Financial Economics* 97 (3): 436–450.

Mackenzie, M. 2009. High-Frequency Trading Under Scrutiny. *Financial Times*, July 28.

Malmendier, U., and S. Nagel. 2011. Depression Babies: Do Macroeconomic Experiences Affect Risk Taking? *The Quarterly Journal of Economics* 126 (1): 373–416.

Moreira, A., and T. Muir. 2016. Volatility Managed Portfolios. *Working Paper*.

Moskowitz, T. J., and A. Vissing-Jørgensen. 2002. The Returns to Entrepreneurial Investment: A Private Equity Premium Puzzle? *American Economic Review* 92 (4): 745–778.

Nagel, S. 2012. Evaporating Liquidity. *Review of Financial Studies* 25 (7): 2005–2039.

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

Santos, T., and P. Veronesi. 2017. Habits and Leverage. *Working Paper*.

Shiller, R. J. 1987. *Investor Behavior in the October 1987 Stock Market Crash: Survey Evidence*. Working Paper. National Bureau of Economic Research.

Sicherman, N., G. Loewenstein, D. J. Seppi, and S. P. Utkus. 2016. Financial Attention. *Review of Financial Studies* 29 (4): 863–897.

Traflet, J. 2012. Spreading the Ideal of Mass Shareownership:: Public Relations and the NYSE. *Essays in Economic & Business History* 22: 257–273.

Wachter, J. A., and M. Yogo. 2010. Why Do Household Portfolio Shares Rise in Wealth? *The Review of Financial Studies* 23 (11): 3929–3965.

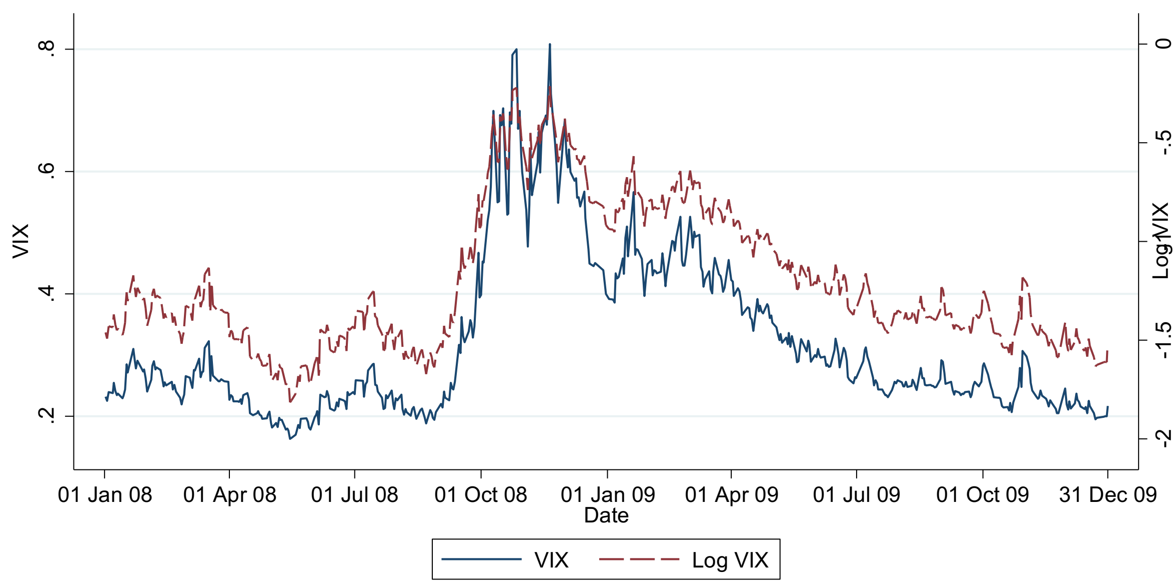
Weber, M., E. U. Weber, and A. Nosić. 2012. Who Takes Risks When and Why: Determinants of Changes in Investor Risk Taking. *Review of Finance* 17: 847–883.

Wilson, J., and P. Liddell. 2013. *Sales of Capital Assets Reported on Individual Tax Returns, 2008–2009*. Statistics of Income Bulletin, Spring 2013. Internal Revenue Service.

Wolff, E. N. 2010. Recent Trends in Household Wealth in the United States: Rising Debt and the Middle-Class Squeeze - An Update to 2007.

Figure 1. Stock Market Volatility (VIX S&P 500) Over Time

Panel A. Entire Sample Period



Panel B. September-November 2008

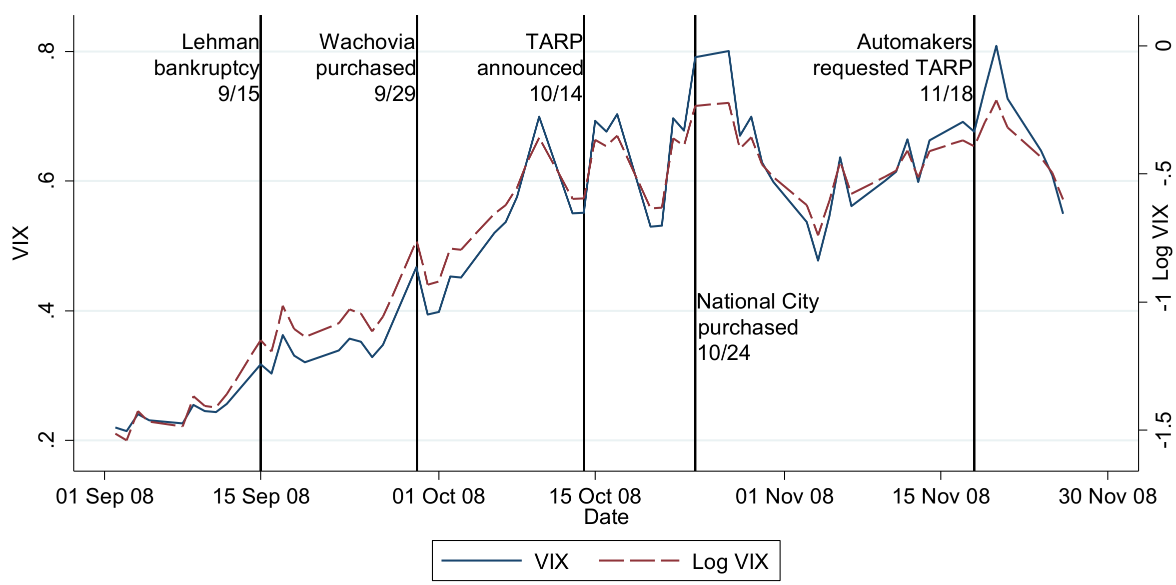
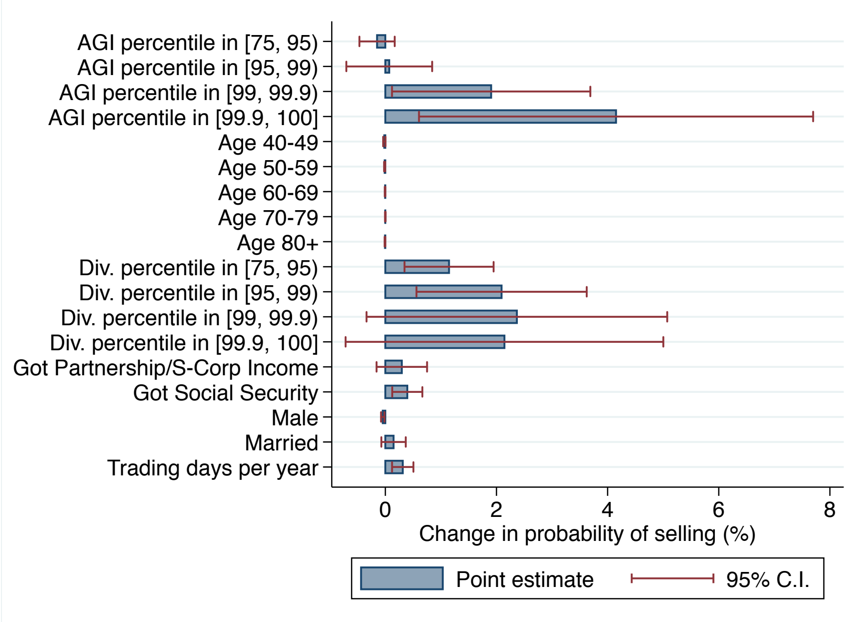


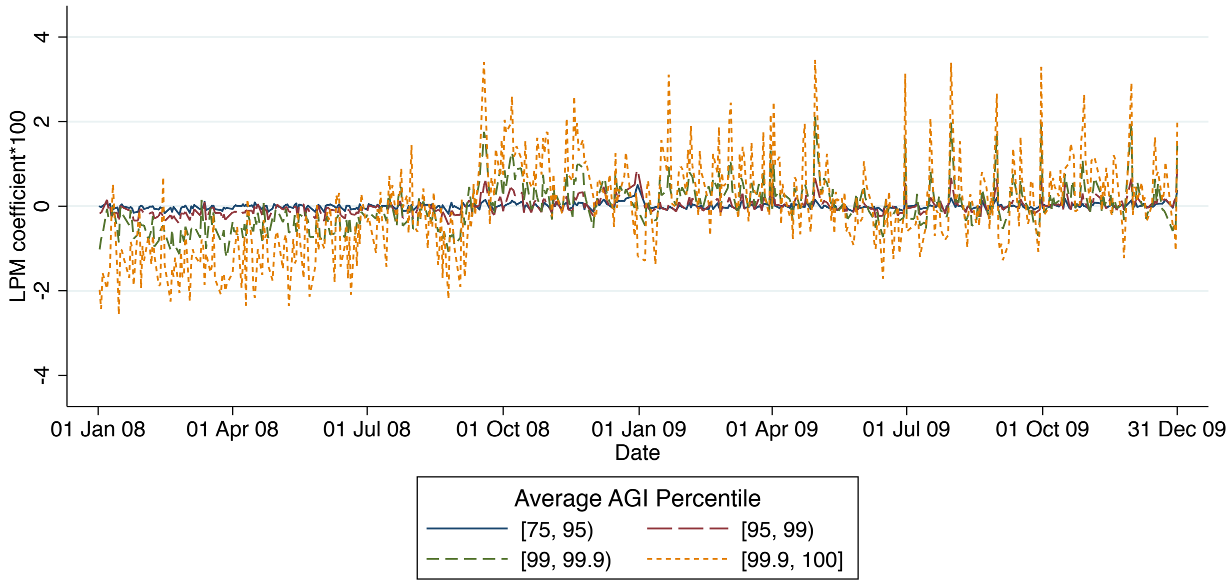
Figure 2. Predicted Changes in the Probability of Selling for a 30 Percent Increase in VIX



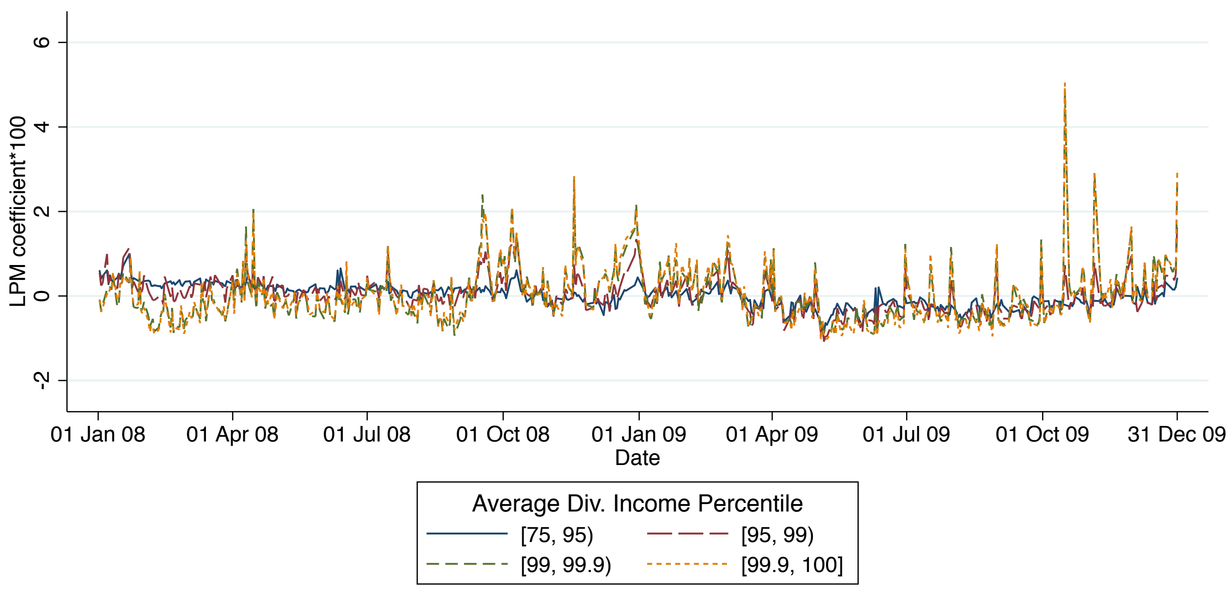
Notes: This figure plots main effects of interest, measuring sensitivity of sales to market tumult by investor characteristics. The figure uses the regression coefficients in column 2 of Table 4. We multiply these coefficients by 0.3 to calculate the implied change in the probability of selling from an increase in log VIX of 0.3, which is approximately the change occurring following the Lehman Brothers collapse in September 2008. We also plot 95% confidence intervals based on the standard errors from the regressions.

Figure 3. Plots of First-Stage Coefficients: Linear Probability Model Estimates

Panel A. AGI

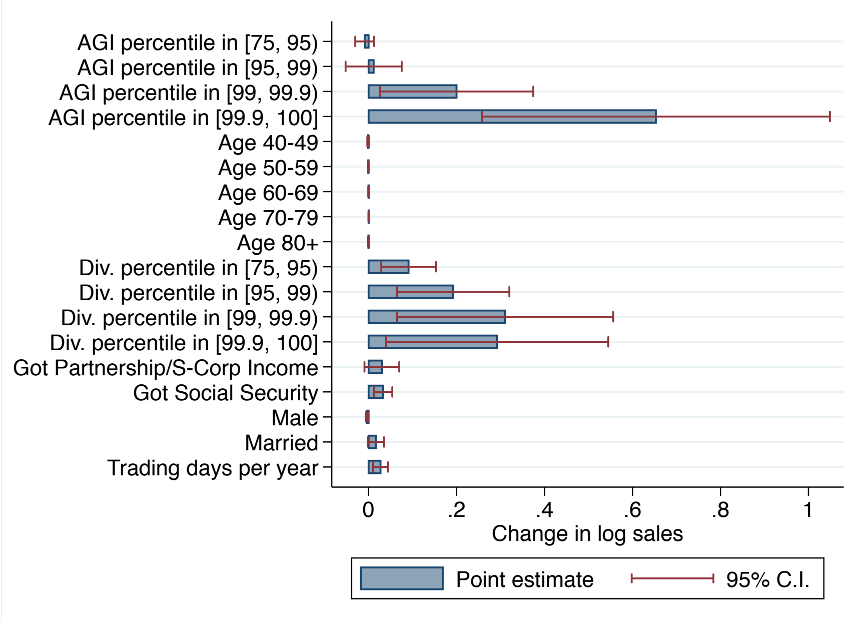


Panel B. Dividend Income



Note: These figures plot the time series of coefficients on selected characteristics for the probability of trading on a given day, resulting from the estimation of the γ_t parameters in equation (4a). We normalize by the mean value of γ_t for a given independent variable, which allows us to compare coefficients across observable characteristics, similarly to how we include the constant, γ_0, terms in the second stage regression, equation (4b).

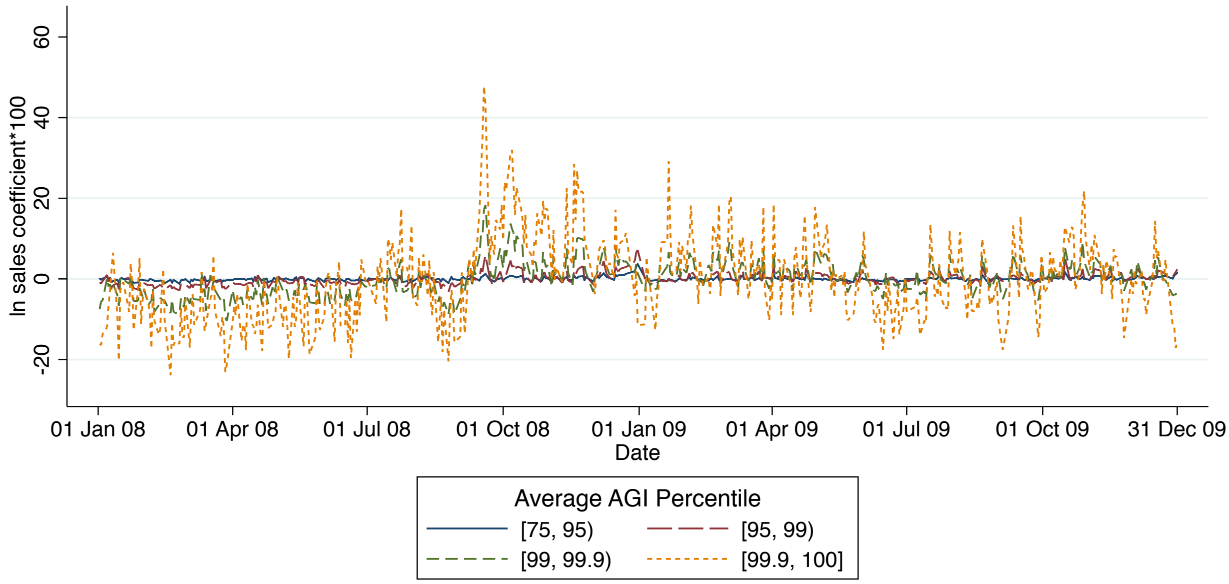
Figure 4. Predicted Changes Log Sales for a 30 Percent Increase in VIX



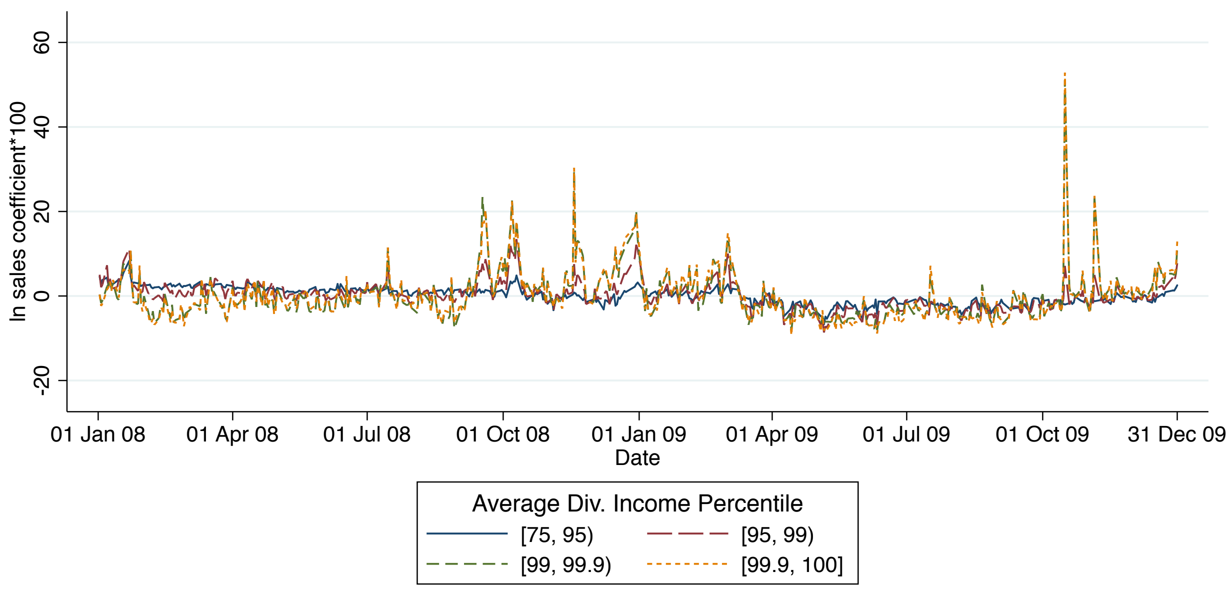
Notes: This figure plots main effects of interest, measuring sensitivity of sales to market tumult by investor characteristics. The figure uses the regression coefficients in column 2 of Table 5. Regression coefficients are not scaled by 100 here, so the units of the x axis are simply the change in log sales. We multiply these coefficients by 0.3 to calculate the implied change in log sales from an increase in log VIX of 0.3, which is approximately the change occurring following the Lehman Brothers collapse in September 2008. We also plot 95% confidence intervals based on the standard errors from the regressions.

Figure 5. Plots of First-Stage Coefficients: Log Sales Model Estimates

Panel A. AGI

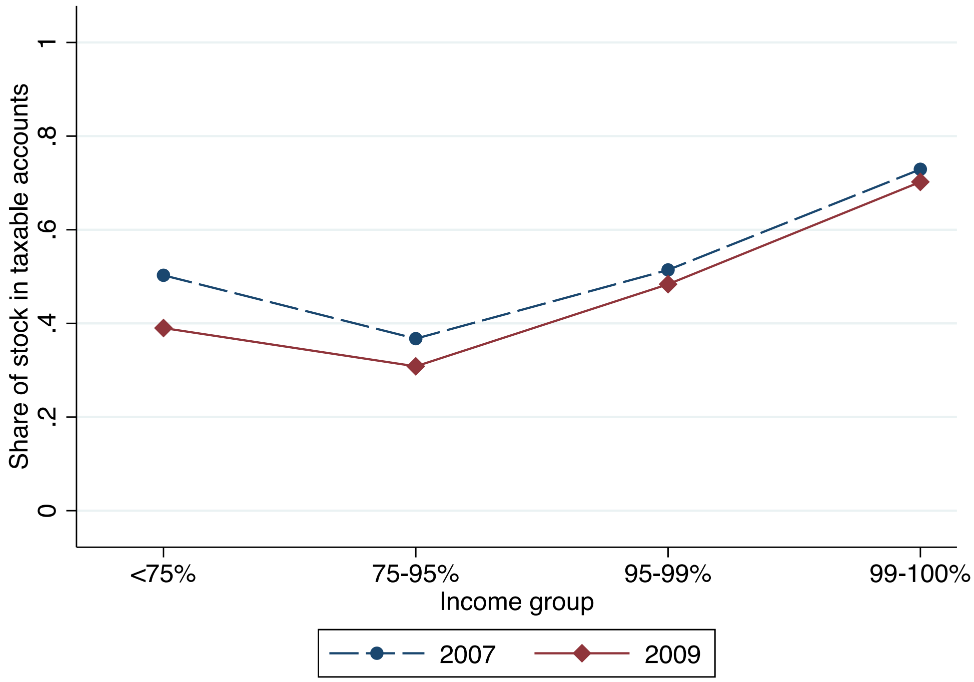


Panel B. Dividend Income



Note: These figures plot the time series of coefficients on selected characteristics for the log sales on a given day, resulting from the estimation of the γ_t parameters in equation (4a). We normalize by the mean value of γ_t for a given independent variable, which allows us to compare coefficients across observable characteristics, similarly to how we include the constant, γ_0, terms in the second stage regression, equation (4b).

Figure 6. Share of Wealth in Taxable Accounts by Income Group



Note: The data for this analysis come from the 2007-2009 panel of the Survey of Consumer Finance. Income groups and age groups are defined based on income and age reported in the 2007 wave of the survey. We calculate the shares by dividing the total equities (stocks and mutual funds) held in taxable accounts in each group by the total equities held in any account in the same group.

Table 1. Heterogeneous Investors: Who sells?

	In response to indicated change in σ, X, and p, investors with higher value of characteristic are more or less likely to sell:		
	Belief v_j	Risk aversion γ_j	Commitment level X_j
σ rises	more	--	--
X_j rises	more	--	more
p falls	more	less	--

Notes: The table shows which types of investors demand the lowest price to stay at their previous risky asset share in response to a rise in volatility, σ, a rise in the level of the net consumption commitment, X_j, or a fall in the log risky asset price p.

Table 2. Sample Selection

Sample Restriction	Transactions	Dollar Volume
All 1099-Bs in 2008-2009	1,432,614,704	$ 37,180,571,687,408
Non-trading and partial days eliminated	1,427,880,785	$ 37,101,171,401,512
Eliminate negative and trades over $2 billion	1,411,432,043	$ 36,283,816,476,155
Individual Taxpayers	870,141,589	$ 9,574,862,035,508
Taxpayers age over 17	861,220,943	$ 9,547,946,596,246
Stocks and Stock Mutual Funds	273,524,098	$ 6,793,511,903,794

Notes: Full trading days are defined as days with positive CRSP trading volume, less days marked as partial trading days. Age of the taxpayer is determined as of December 31, 2008. Stocks are defined as assets where the first two characters of US_CFI_CODE from the cusip.issue database on WRDS are ES (common equity) or EP (preferred shares).

Table 3. Overall Sales Response to Market Tumult

Dependent variable: log 1099-B sales volume, USD	
Change in log VIX	4.964***
	(1.115)
Observations	498
R-squared	0.16

Notes: Regressions include ten one-day lagged log VIX changes from t-11 to t-1, and we report the sum of these coefficients. Newey-West standard errors (10-day lag) in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 4. Heterogeneity in Propensity to Sell Stock: Linear Probability Model Estimates

	Coefficient on characteristic given no change in log VIX	Interaction between characteristic and change in log VIX
	(1)	(2)
AGI in [75, 95)	-0.077***	-0.495
	(0.008)	(0.540)
AGI in [95, 99)	0.182***	0.236
	(0.018)	(1.313)
AGI in [99, 99.9)	1.539***	6.349**
	(0.054)	(3.035)
AGI in [99.9, 100]	3.724***	13.84**
	(0.118)	(6.031)
Age 40-49	0.008***	-0.0596*
	(0.001)	(0.0355)
Age 50-59	0.003***	-0.0395**
	(0.000)	(0.0179)
Age 60-69	-0.000	-0.0147
	(0.000)	(0.00970)
Age 70-79	-0.003***	-0.00118
	(0.000)	(0.00902)
Age 80+	-0.008***	-0.0204*
	(0.000)	(0.0117)
Average Dividends in [75, 95)	-0.479***	3.823***
	(0.032)	(1.363)
Average Dividends in [95, 99)	-0.092**	6.975***
	(0.040)	(2.605)
Average Dividends in [99, 99.9)	0.980***	7.891*
	(0.062)	(4.601)
Average Dividends in [99.9, 100]	0.937***	7.146
	(0.066)	(4.862)
Receipt of Partnership/S-Corp Income Indicator	0.347***	0.989
	(0.016)	(0.773)
Social Security Receipt Indicator	0.022*	1.319***
	(0.011)	(0.461)
Male	-0.011***	-0.154***
	(0.001)	(0.0526)
Married	0.404***	0.494
	(0.007)	(0.373)
Trading days per year	0.310***	1.045***
	(0.008)	(0.327)
Intercept	0.955***	-4.952**
	(0.060)	(2.098)

Notes: There are 498 daily observations. This table reports estimates of equation (4b), using an indicator for whether the individual sold on a given day as the outcome variable in (4a). Columns 1 and 2 are estimated from the same regression. Regressions include ten one-day lagged log VIX changes, and we report the sum of these coefficients. Coefficients are scaled by 100 to facilitate interpretation. Newey-West standard errors (10-day lag) in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 5. Heterogeneity in Propensity to Sell Stock: Log Sales Model Estimates

	Coefficient on characteristic given no change in log VIX	Interaction between characteristic and change in log VIX
	(1)	(2)
AGI in [75, 95)	-0.438***	-2.956
	(0.045)	(3.661)
AGI in [95, 99)	2.033***	3.817
	(0.141)	(10.85)
AGI in [99, 99.9)	13.915***	66.66**
	(0.459)	(29.66)
AGI in [99.9, 100]	45.607***	217.7***
	(1.032)	(67.29)
Age 40-49	0.089***	-0.423
	(0.007)	(0.291)
Age 50-59	0.046***	-0.274*
	(0.004)	(0.153)
Age 60-69	0.013***	-0.0919
	(0.002)	(0.0751)
Age 70-79	-0.012***	0.0269
	(0.001)	(0.0679)
Age 80+	-0.053***	-0.152
	(0.002)	(0.0957)
Average Dividends in [75, 95)	-3.882***	30.27***
	(0.257)	(10.55)
Average Dividends in [95, 99)	-1.666***	64.14***
	(0.333)	(21.70)
Average Dividends in [99, 99.9)	8.170***	103.5**
	(0.538)	(41.76)
Average Dividends in [99.9, 100]	7.448***	97.45**
	(0.569)	(42.95)
Receipt of Partnership/S-Corp Income Indicator	3.179***	10.05
	(0.137)	(6.731)
Social Security Receipt Indicator	-0.027	10.95***
	(0.089)	(3.565)
Male	-0.051***	-1.206***
	(0.011)	(0.435)
Married	3.333***	5.487*
	(0.065)	(3.156)
Trading days per year	2.452***	9.021***
	(0.065)	(2.859)
Intercept	6.620***	-40.17**
	(0.450)	(16.06)

Notes: There are 498 daily observations. This table reports estimates of equation (4b), using log sales for a given individual on a given day as the outcome variable in (4a). Columns 1 and 2 are estimated from the same regression. Regressions include ten one-day lagged log VIX changes, and we report the sum of these coefficients. Coefficients are scaled by 100 to facilitate interpretation. Newey-West standard errors (10-day lag) in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 6. Gross and Net Sales in the Barber-Odean Discount Brokerage Data

	Dependent variable:					
	Log gross sales volume, taxable accounts	Net sales volume (% of portfolio value), taxable accounts	Gross sales volume (% of portfolio value), taxable accounts	Net sales volume (% of portfolio value), taxable accounts	Log gross sales volume, non-taxable accounts	Log gross sales volume, taxable accounts, large portfolios
	(1)	(2)	(3)	(4)	(5)	(6)
Change in log VIX	1.753**		0.917***	0.315**	2.231***	1.937***
	(0.463)		(0.131)	(0.149)	(0.652)	(0.554)
Contemp. gross sales volume		0.342***				
(% of portfolio value)		(0.029)				
Adj. R^2	0.0263	0.2685	0.0429	0.0111	0.0164	0.0195
Observations	1,496	1,475	1,474	1,474	1,496	1,474

Notes: Regressions include three one-day log VIX changes from t-3 to t (i.e., including the contemporaneous change from t-1 to t) and we report the sum of these coefficients. Large portfolios in the last column are those above the 80th percentile by the total value of all positions at the end of the previous month. Newey-West standard errors (10 day lag) in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 7. Gross Sales and Changes in Dividend Income

	Dependent Variable: Dividends in t+1 minus Dividends in t-1			
	t = 2008	t = 2009	t = 2008	t = 2009
	(1)	(2)	(3)	(4)
Gross Sales	-0.0065***	-0.0068***	-0.0049***	-0.043***
	(0.000041)	(0.000051)	(0.000086)	(0.000094)
Gross Sales x AGI percentile [75,95]			-0.0004***	-0.0008***
			(0.00011)	(0.00013)
Gross Sales x AGI percentile [95,99]			-0.0005***	-0.0012***
			(0.00012)	(0.00014)
Gross Sales x AGI percentile [99,99.9]			-0.0006***	-0.0017***
			(0.00013)	(0.00015)
Gross Sales x AGI percentile [99.9,100]			-0.0004***	-0.0017***
			(0.00022)	(0.00027)
AGI group fixed effects	NO	NO	YES	YES
R-square	0.0941	0.0648	0.1147	0.0869
Observations	1,888,174	1,885,874	1,888,174	1,885,874

Notes: For computational reasons, regressions are estimated on a random 10% sample of taxpayers who have positive dividends in year t-1. We winsorize all variables at the 1 and 99th percentile to eliminate the effect of outliers (many of which are obvious data errors) on the estimates. White standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 8. Net Sales in Taxable Accounts and Total Net Sales, 2007-2009

Dependent Variable: Net Purchases (Negative Sales) from 2007 to 2009		
	(1)	(2)
Net Taxable Purchases (or Sales)	0.925***	0.908***
	(0.075)	(0.123)
Net Taxable Purchases x HH income percentile [75,95]		0.146
		(0.139)
Net Taxable Purchases x HH income percentile [95,99]		0.030
		(0.140)
Net Taxable Purchases x HH income percentile [99,100]		0.015
		(0.153)
AGI group fixed effects	NO	YES
R-square	0.772	0.728
Observations	3,857	3,857

Notes: Household (HH) income percentiles are based on the SCF. Heteroskedasticity-robust standard errors, adjusted for multiple imputations in the survey data, are reported in parentheses *** p<0.01, ** p<0.05, * p<0.1.

Online Appendix. Aggregate Measures of Sales Volume

To gain a sense of the usefulness and limitations of the data we examine in the main part of the analysis, Figure A.1 compares total sales volume reported on 1099-Bs to total sales volume as measured by CRSP.[32] Panel A plots logged trading volume over time, and Panel B depicts the ratio of 1099-B to CRSP sales volume. This share increases modestly throughout 2009. We also see large increases in the matched 1099-B shares during the very last days of 2008 and 2009. This phenomenon is likely at least partly attributable to the well-known tendency of individual investors to rebalance their portfolio at year-ends, including their tendency to "harvest" capital losses for tax purposes (Hoopes et al. 2015; Poterba and Weisbenner 2001). Opportunities for loss harvesting were especially abundant at the end of 2008 and 2009 because of the crisis.

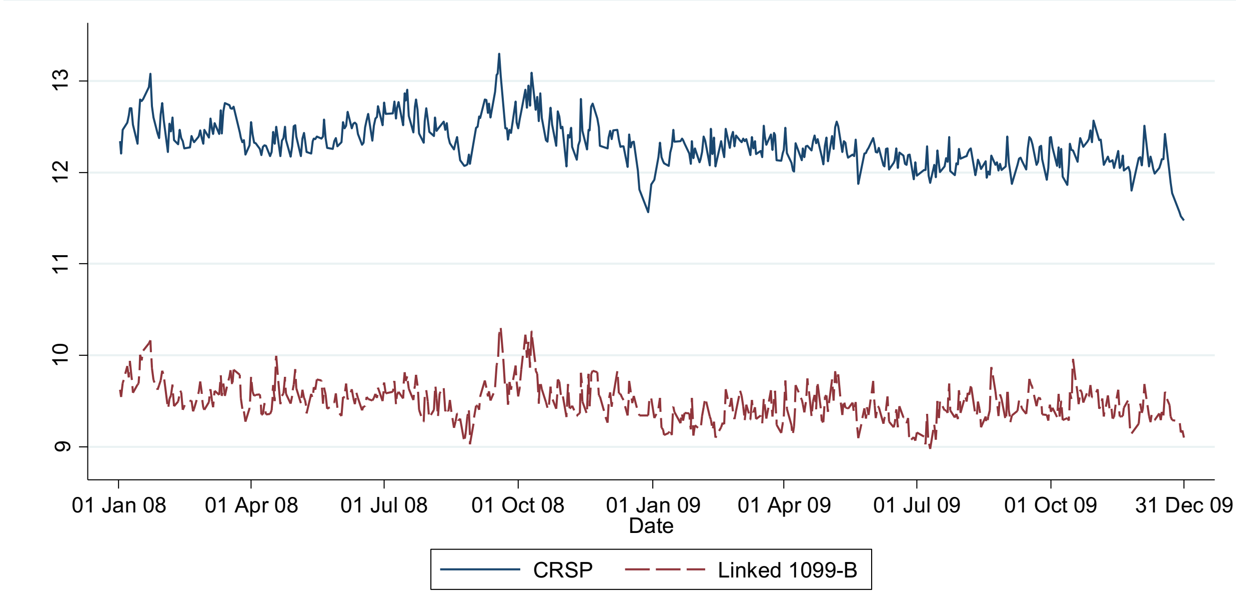
Figure A.2 examines the number of transactors and the sales volume per transactor. Panel A plots the sales volume and the number of individual investors selling stocks (the number of transactors) from our 1099-B data. The transactors on a given day ranges from 294,000 to 773,000. The two series, total sales volume and number of transactors, exhibit a strong correlation (0.69) during our sample period. Panel B displays the average daily sales volume per transactor, which ranges from \$16,000 to \$46,000 dollars. The increase in sales volume after an increase in VIX appears to be driven by both an increase in sales volume per transactor and an increase in the number of transactors, though the latter channel is somewhat more important.[33]

[32] The positive relationship between market sales volume and volatility is well documented in the literature (i.e., Karpoff 1987). In our sample period, a ten percent increase in log VIX over ten days is associated with a 116 percent increase in market sales volume. Sales volume from matched Form 1099-B's is slightly less strongly associated: a ten percent increase in log VIX over ten days is associated with a 64 percent increase in 1099-B sales volume. Appendix Table A.1 provides summary statistics and Appendix Table A.2 provides details of these regressions.

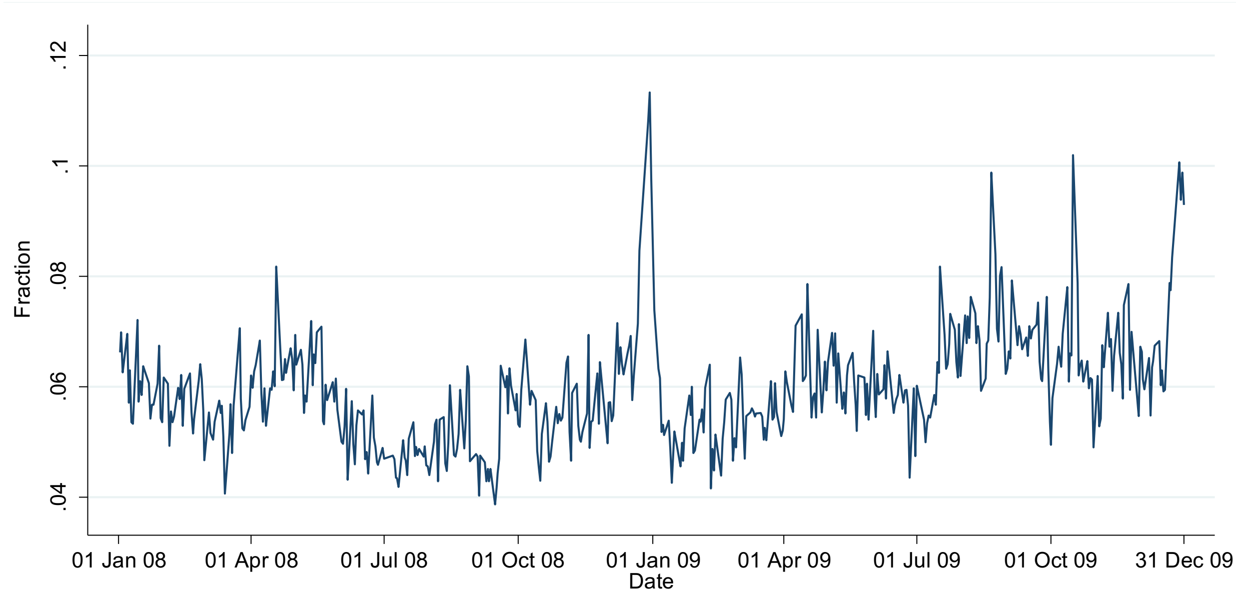
[33] Columns 4 and 5 of Table A.2 show that a ten percent increase in log VIX over ten days corresponds to a 20 percent increase in the number of transactors and a 37 percent increase in the average volume per transactor.

Figure A1. Sales and Trading Volume from Matched 1099-B's and CRSP

Panel A. Log Sales Volume, Matched 1099-B's and Universe of Stocks on CRSP



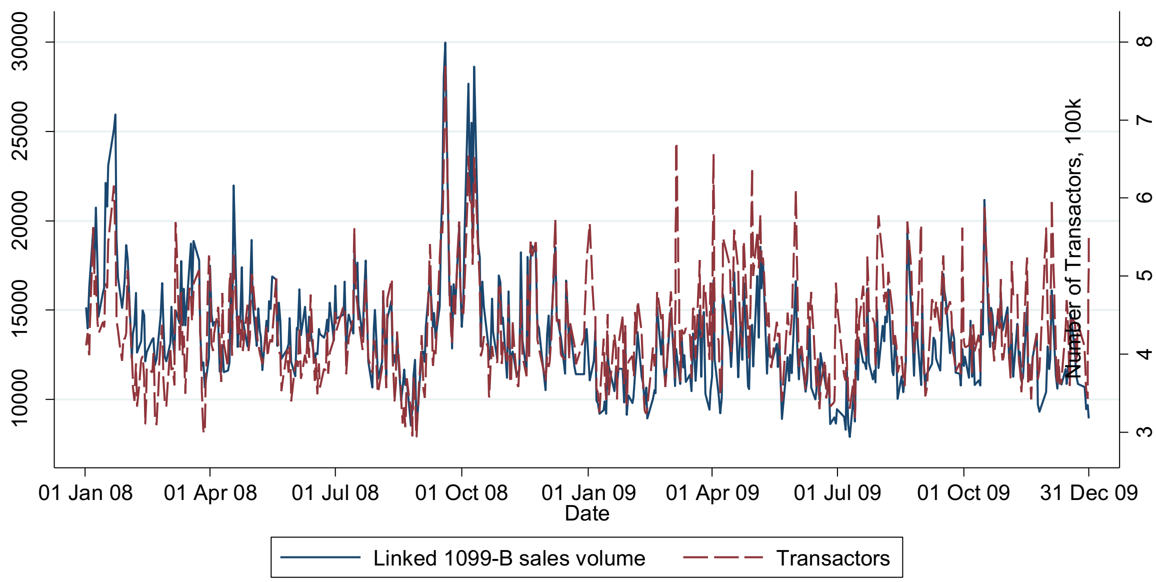
Panel B. Fraction of Total Sales Volume Represented by Matched 1099-B's



Notes: Total sales volume equals ending price per share times number of shares traded for the universe of stocks on CRSP. We divide the trading volume of NASDAQ stocks by two because it is a dealer market (see Anderson and Dyl 2007). The mean and median of the depicted share is 0.059.

Figure A2. Sales Volume and Number of Transactors

Panel A. Sales Volume and Number of Transactors, Matched Form 1099-B's



Panel B. Average Sales Volume per Transactor, Matched Form 1099-B's

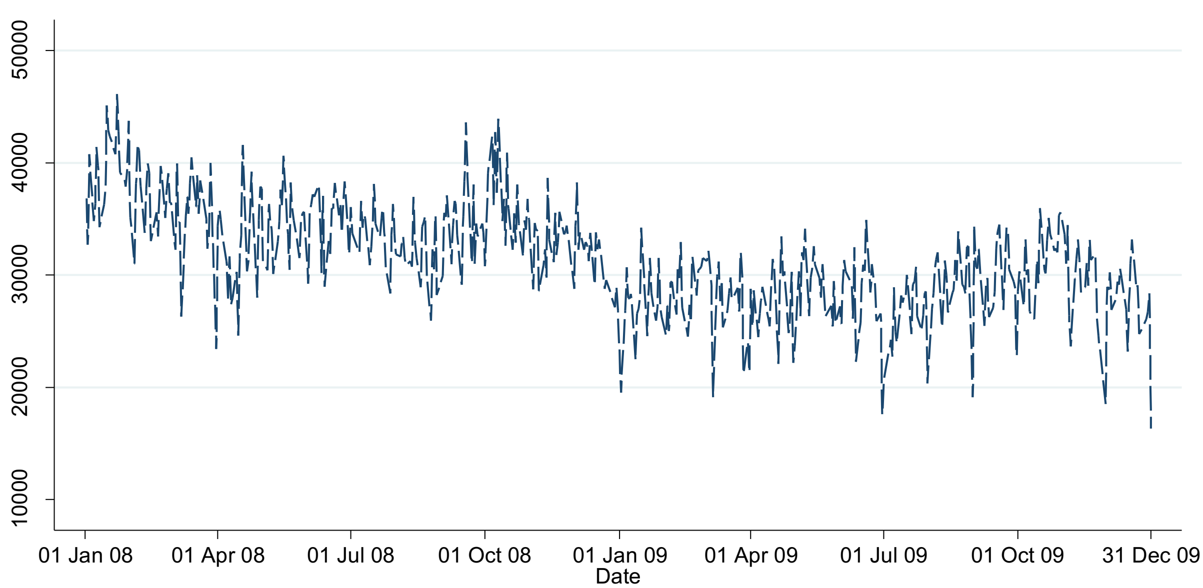
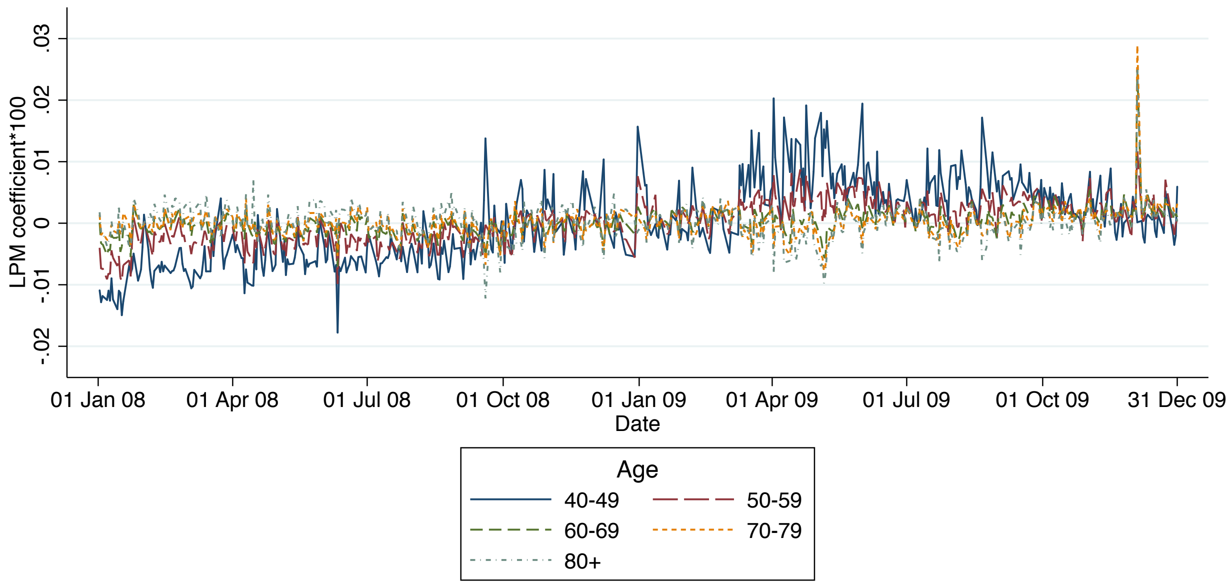
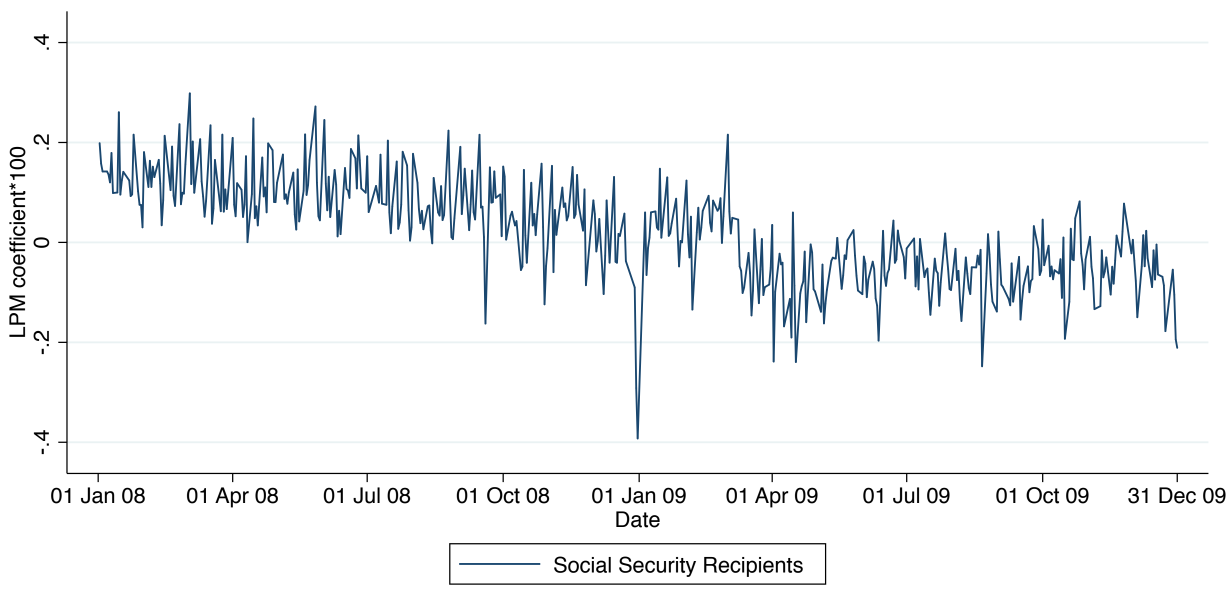


Figure A3. Additional Plots of First-Stage Coefficients: Linear Probability Model Estimates

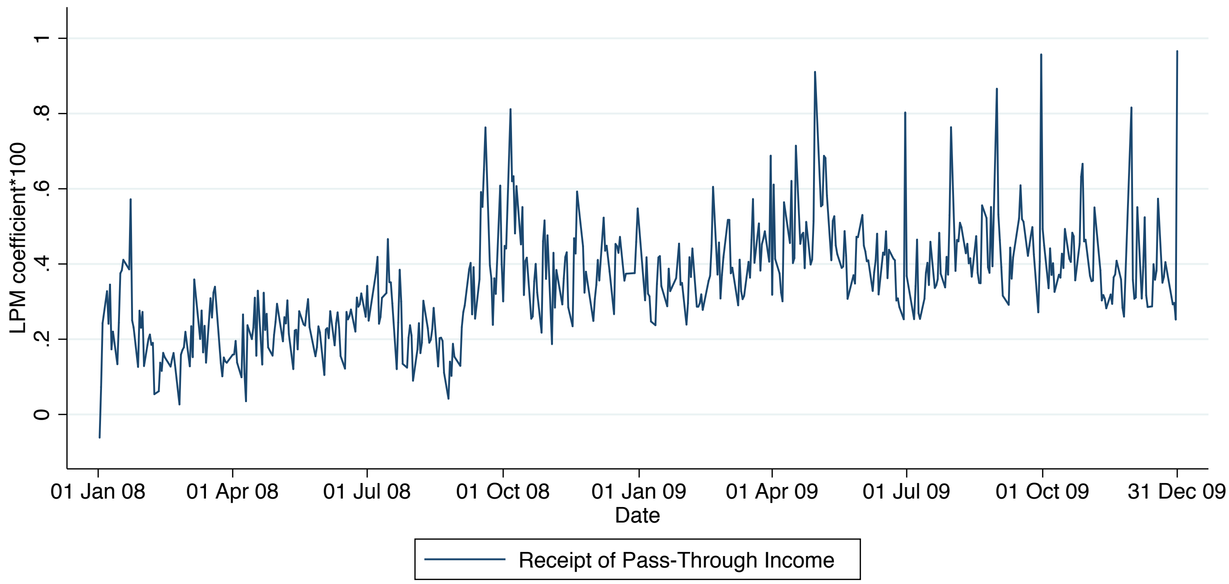
Panel A. Age



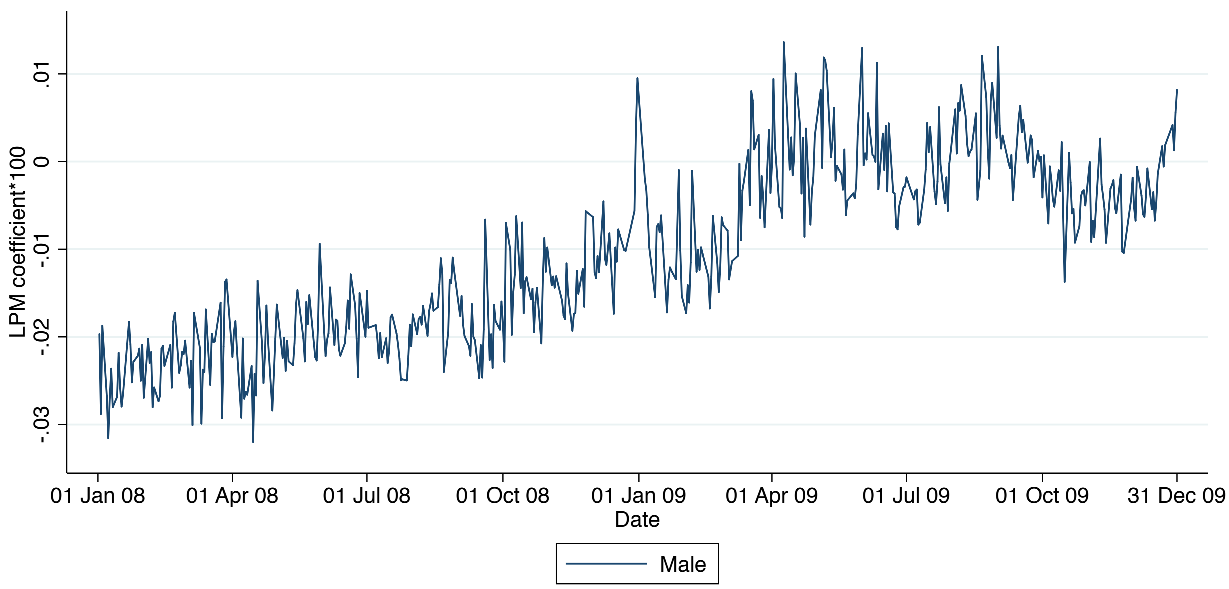
Panel B. Social Security



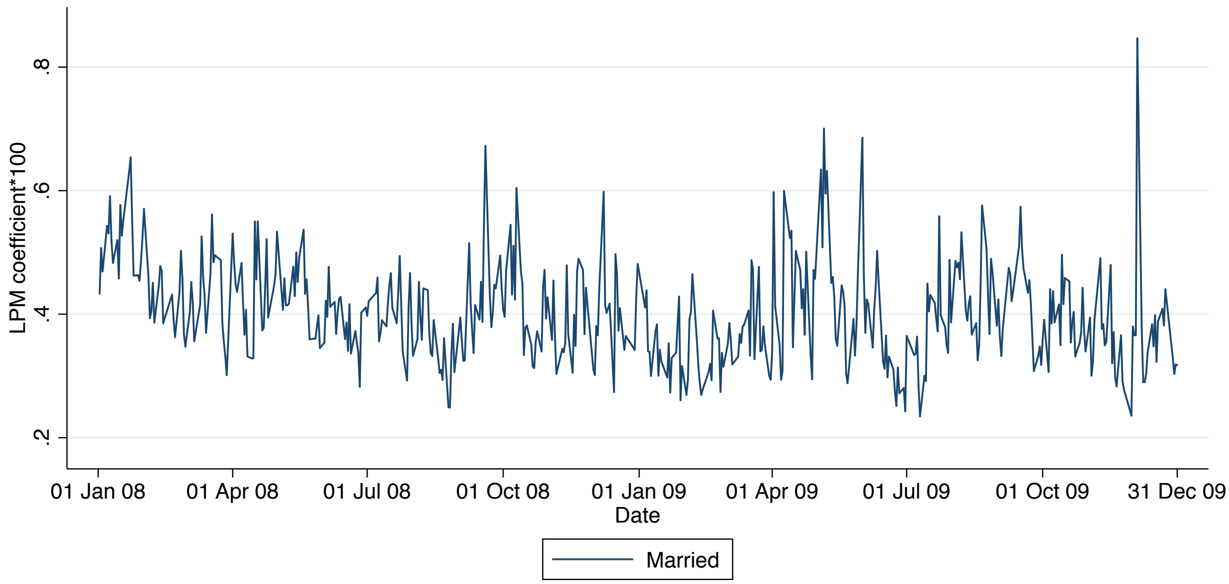
Panel C. Partnership/S-Corp Income



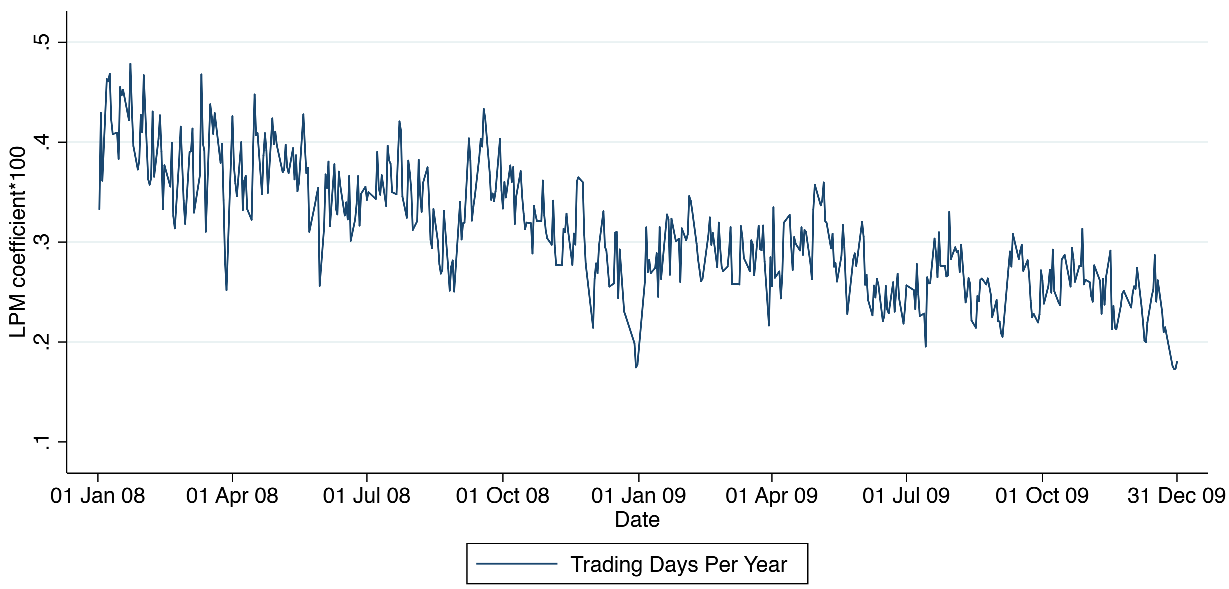
Panel D. Male



Panel E. Married



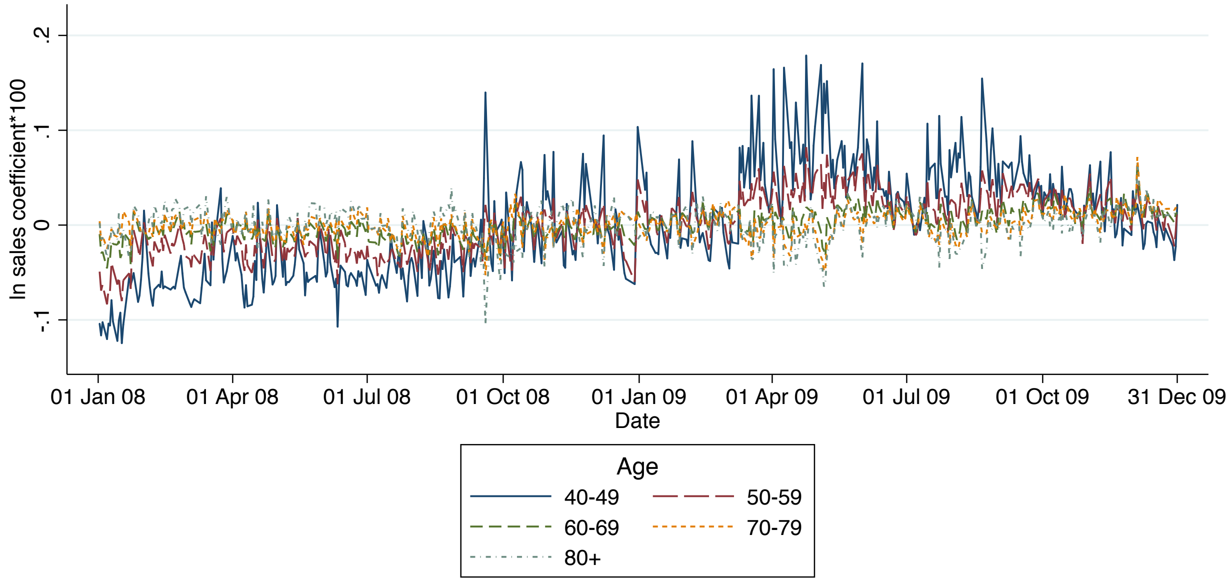
Panel F. Trading Days per Year



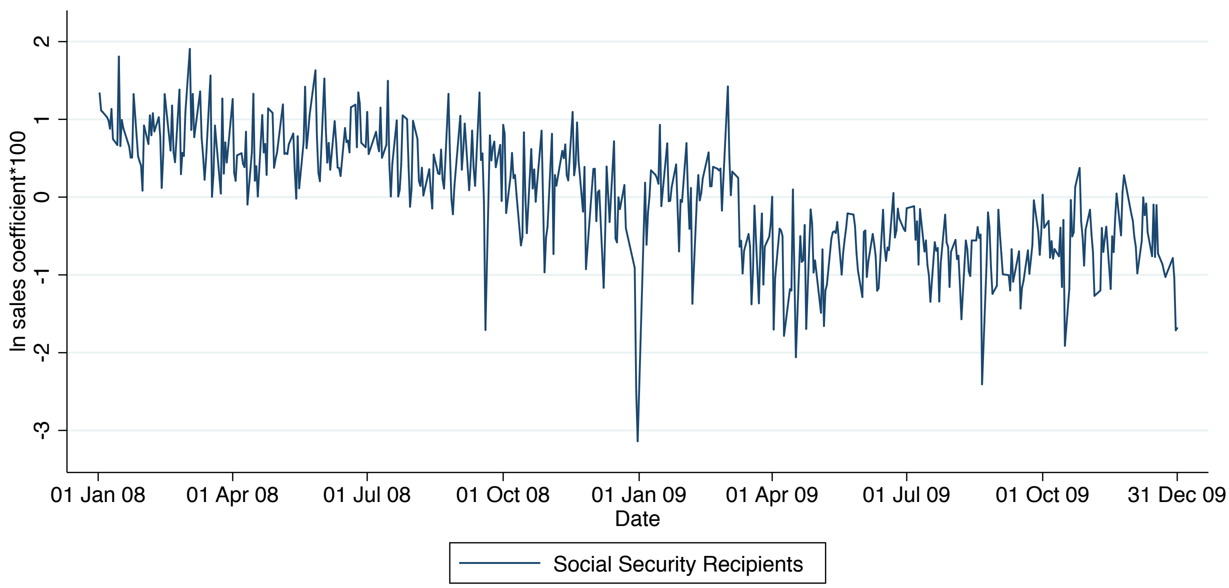
Note: These figures plot the time series of coefficients on given characteristics for the probability of trading on a given day, resulting from the estimation of the γ_t parameters in equation (4a). In Panels A, B, and C, we normalize by the mean value of γ_t for a given dependent variable, which allows us to compare coefficients across observable characteristics, similarly to how we include the constant, γ_0, terms in the second stage regression, equation (4b).

Figure A4. Additional Plots of First-Stage Coefficients: Log Sales Model Estimates

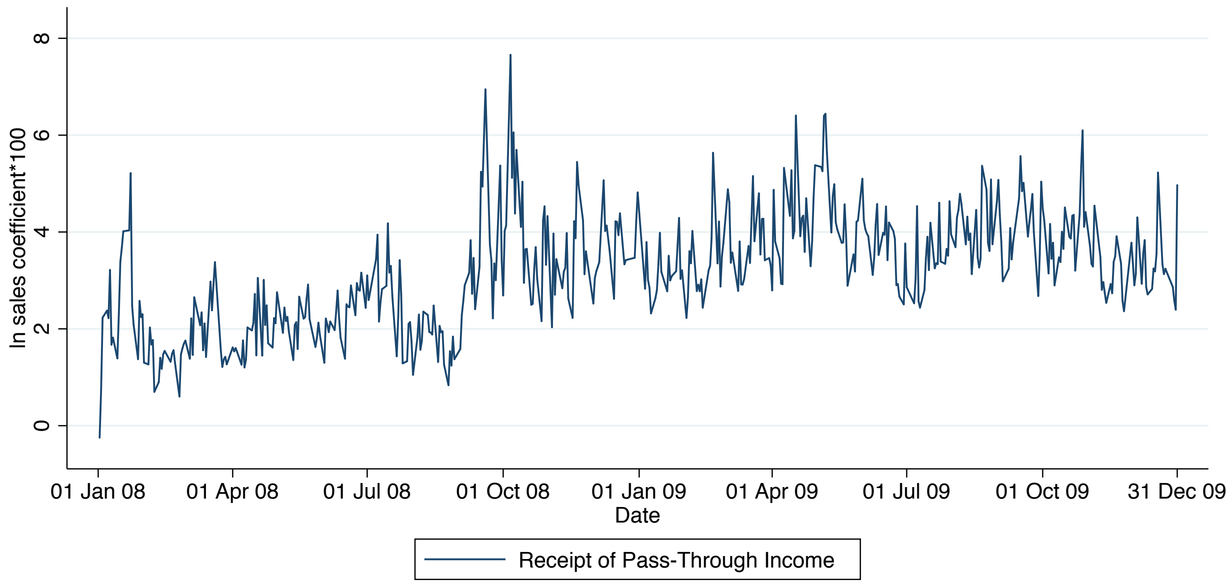
Panel A. Age



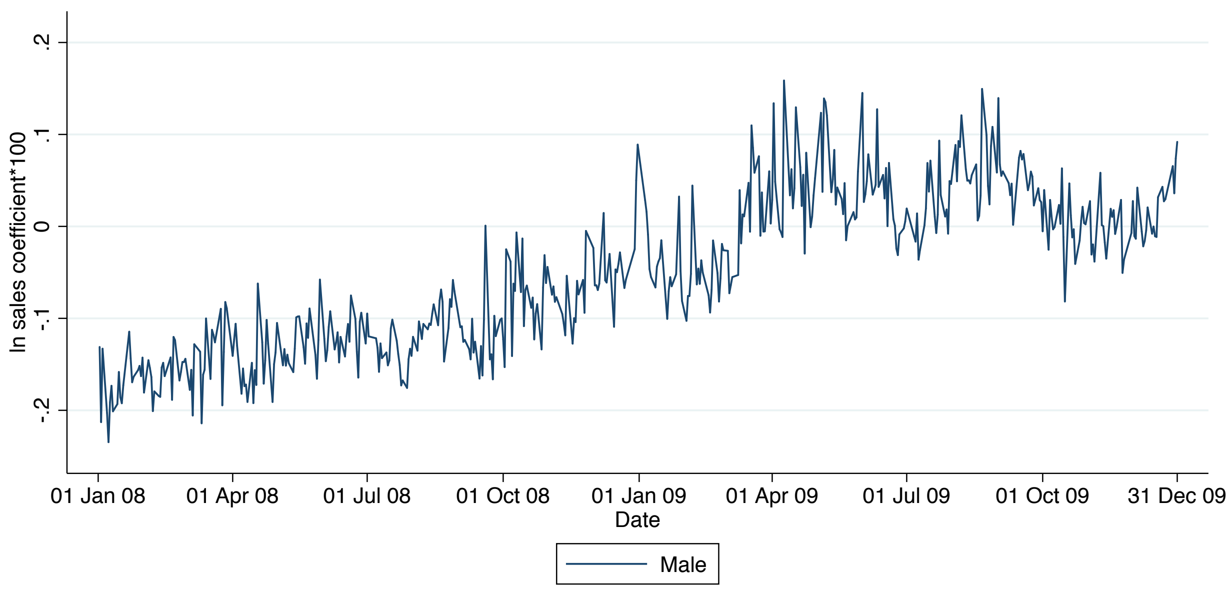
Panel B. Social Security



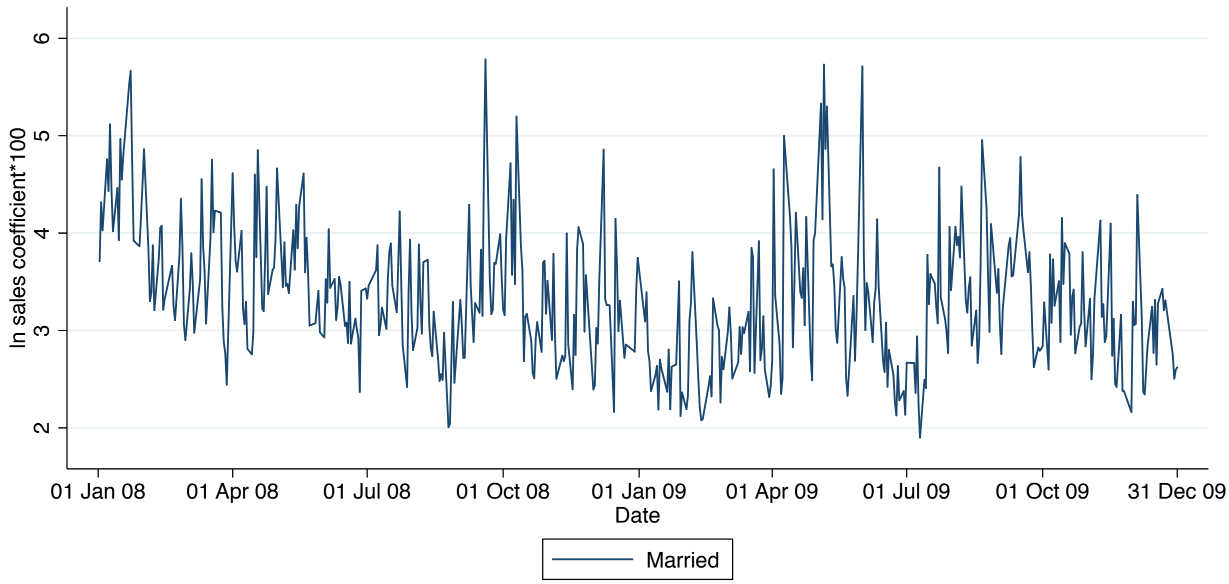
Panel C. Partnership/S-Corp Income



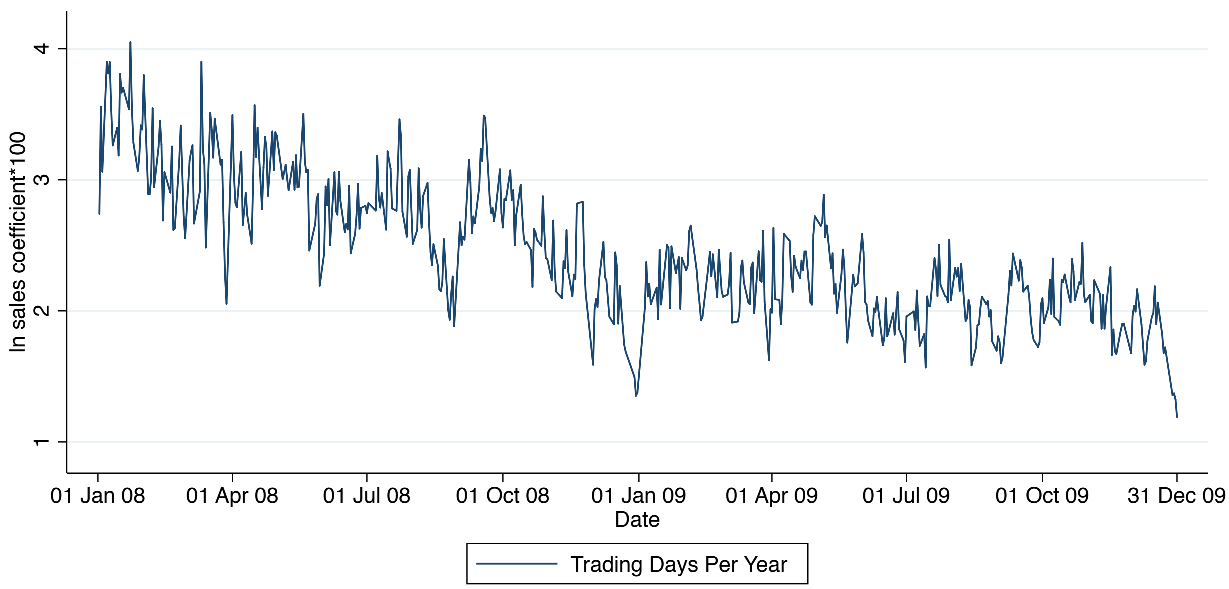
Panel D. Male



Panel E. Married



Panel F. Trading Days per Year



Note: These figures plot the time series of coefficients on given characteristics for the probability of trading on a given day, resulting from the estimation of the γ_t parameters in equation (4a). In Panels A, B, and C, we normalize by the mean value of γ_t for a given dependent variable, which allows us to compare coefficients across observable characteristics, similarly to how we include the constant, γ_0, terms in the second stage regression, equation (4b).

Table A.1. Aggregate Summary Statistics

	Mean	SD	Minimum	Maximum
Log sales volume, CRSP, millions USD	12.330	0.250	11.477	13.300
Log sales volume, 1099-B, millions USD	9.491	0.206	8.976	10.308
Ratio, 1099-B to CRSP volume	0.059	0.010	0.039	0.113
Log number of transactors, 100k	1.444	0.151	1.078	2.045
Log average volume per transactor, USD	10.350	0.153	9.700	10.738
1 day lagged difference, log VIX	-0.00008	0.068	-0.283	0.296
1 day lagged difference, VIX	-0.00002	0.029	-0.174	0.165
10 day lagged difference, log VIX	-0.00028	0.162	-0.370	0.700
10 day lagged difference, VIX	-0.00007	0.065	-0.230	0.352
Log VIX	-1.208	0.359	-1.814	-0.212
VIX	0.321	0.133	0.163	0.809

Table A.2. Total Sales Volume, Coverage Rates, Number of Transactors, and Market Tumult

	Dependent variable:				
	Log sales volume, CRSP, millions USD	Log sales volume, 1099-B, millions USD	Ratio of 1099-B to CRSP sales volume	Log number of transactors, 100k	Log average volume per transactor, USD
	(1)	(2)	(3)	(4)	(5)
Sum of 10 1-day lagged differences of log VIX	7.699***	4.964***	-0.172***	1.849**	3.115***
	(1.129)	(1.115)	(0.053)	(0.789)	(0.610)
Observations	498	498	498	498	498
R-squared	0.26	0.16	0.08	0.06	0.11

Notes: The dependent variable in column 4 is the natural log of the number of transactors divided by 100,000. Newey-West standard errors with a maximum of 10-day lag orders of autocorrelation are reported in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Table A.3. Heterogeneity in Propensity to Sell Stock: Linear Probability Model Estimates, One-Day Lagged Change in Log VIX

	Coefficient on characteristic given no change in log VIX	Interaction between characteristic and change in log VIX
	(1)	(2)
AGI in [75, 95)	-0.077***	-0.138
	(0.008)	(0.104)
AGI in [95, 99)	0.182***	-0.094
	(0.018)	(0.154)
AGI in [99, 99.9)	1.538***	0.414
	(0.055)	(0.316)
AGI in [99.9, 100]	3.724***	0.529
	(0.120)	(0.748)
Age 40-49	0.008***	-0.014***
	(0.001)	(0.005)
Age 50-59	0.003***	-0.007*
	(0.000)	(0.004)
Age 60-69	-0.000	-0.001
	(0.000)	(0.003)
Age 70-79	-0.003***	0.002
	(0.000)	(0.003)
Age 80+	-0.008***	0.000
	(0.000)	(0.003)
Average Dividends in [75, 95)	-0.479***	0.305**
	(0.033)	(0.143)
Average Dividends in [95, 99)	-0.092**	0.608**
	(0.043)	(0.246)
Average Dividends in [99, 99.9)	0.980***	0.580
	(0.063)	(0.406)
Average Dividends in [99.9, 100]	0.936***	0.392
	(0.067)	(0.438)
Receipt of Partnership/S-Corp Income Indicator	0.347***	0.126
	(0.016)	(0.081)
Social Security Receipt Indicator	0.022*	0.242***
	(0.012)	(0.057)
Male	-0.011***	-0.016***
	(0.001)	(0.005)
Married	0.404***	-0.008
	(0.007)	(0.054)
Trading days per year	0.310***	0.065**
	(0.008)	(0.031)
Intercept	0.955***	-0.319
	(0.061)	(0.297)

Notes: There are 498 daily observations. This table reports estimates of equation (4b), using an indicator for whether the individual sold on a given day as the outcome variable in (4a). Columns 1 and 2 are estimated from the same regression. Regressions include a single lagged log VIX change. Coefficients are scaled by 100 to facilitate interpretation. Newey-West standard errors (10-day lag) in parentheses *** p<0.01, ** p<0.05, * p<0.1

Table A.4. Heterogeneity in Propensity to Sell Stock: Linear Probability Model Estimates, Using Lagged Negative Market Return as Measure of Tumult

	Coefficient on characteristic given zero market return (1)	Interaction between characteristic and negative market return (2)
AGI in [75, 95)	-0.077***	1.821
	(0.007)	(1.777)
AGI in [95, 99)	0.182***	0.149
	(0.018)	(3.417)
AGI in [99, 99.9)	1.536***	-11.98
	(0.055)	(9.353)
AGI in [99.9, 100]	3.719***	-28.85
	(0.119)	(18.26)
Age 40-49	0.008***	0.312**
	(0.001)	(0.129)
Age 50-59	0.003***	0.154**
	(0.000)	(0.0603)
Age 60-69	-0.000	0.0230
	(0.000)	(0.0314)
Age 70-79	-0.003***	-0.0398
	(0.000)	(0.0309)
Age 80+	-0.008***	-0.00447
	(0.000)	(0.0456)
Average Dividends in [75, 95)	-0.482***	-17.56***
	(0.030)	(4.135)
Average Dividends in [95, 99)	-0.097***	-30.43***
	(0.036)	(4.835)
Average Dividends in [99, 99.9)	0.974***	-35.25***
	(0.059)	(7.753)
Average Dividends in [99.9, 100]	0.930***	-35.57***
	(0.063)	(8.004)
Receipt of Partnership/S-Corp Income Indicator	0.347***	0.544
	(0.016)	(2.566)
Social Security Receipt Indicator	0.021**	-5.806***
	(0.011)	(1.494)
Male	-0.011***	0.655***
	(0.001)	(0.170)
Married	0.404***	-0.0285
	(0.007)	(1.196)
Trading days per year	0.309***	-2.583***
	(0.008)	(0.823)
Intercept	0.960***	26.36***
	(0.056)	(8.071)

Notes: There are 498 daily observations. This table reports estimates of equation (4b), using an indicator for whether the individual sold on a given day as the outcome variable in (4a). Columns 1 and 2 are estimated from the same regression. Regressions include ten one-day lagged negative market returns, and we report the sum of these coefficients. Coefficients are scaled by 100 to facilitate interpretation. Newey-West standard errors (10-day lag) in parentheses *** p<0.01, ** p<0.05, * p<0.1

Table A.5. Heterogeneity in Propensity to Sell Stock: Log Sales Model Estimates, One-Day Lagged Change in Log VIX

	Coefficient on characteristic given no change in log VIX	Interaction between characteristic and change in log VIX
	(1)	(2)
AGI in [75, 95)	-0.438***	-0.694
	(0.045)	(0.462)
AGI in [95, 99)	2.033***	-0.503
	(0.139)	(0.980)
AGI in [99, 99.9)	13.915***	4.181
	(0.475)	(2.723)
AGI in [99.9, 100]	45.605***	10.374
	(1.127)	(7.843)
Age 40-49	0.089***	-0.113***
	(0.007)	(0.030)
Age 50-59	0.046***	-0.053***
	(0.004)	(0.018)
Age 60-69	0.013***	-0.001
	(0.002)	(0.011)
Age 70-79	-0.012***	0.023**
	(0.001)	(0.010)
Age 80+	-0.053***	0.008
	(0.002)	(0.014)
Average Dividends in [75, 95)	-3.882***	2.767***
	(0.265)	(0.993)
Average Dividends in [95, 99)	-1.667***	5.880***
	(0.360)	(1.950)
Average Dividends in [99, 99.9)	8.169***	7.005*
	(0.575)	(3.854)
Average Dividends in [99.9, 100]	7.446***	5.892
	(0.600)	(4.093)
Receipt of Partnership/S-Corp Income Indicator	3.179***	1.267**
	(0.136)	(0.641)
Social Security Receipt Indicator	-0.026	2.006***
	(0.091)	(0.389)
Male	-0.051***	-0.116***
	(0.012)	(0.037)
Married	3.334***	-0.113
	(0.066)	(0.413)
Trading days per year	2.452***	0.593**
	(0.068)	(0.265)
Intercept	6.619***	-3.988***
	(0.457)	(1.496)

Notes: There are 498 daily observations. This table reports estimates of equation (4b), using an indicator for whether the individual sold on a given day as the outcome variable in (4a). Columns 1 and 2 are estimated from the same regression. Regressions include a single lagged log VIX change. Coefficients are scaled by 100 to facilitate interpretation. Newey-West standard errors (10-day lag) in parentheses *** p<0.01, ** p<0.05, * p<0.1

Table A.6. Heterogeneity in Propensity to Sell Stock: Log Sales Model Estimates, Using Lagged Negative Market Return as Measure of Tumult

	Coefficient on characteristic given zero market return (1)	Interaction between characteristic and negative market return (2)
AGI in [75, 95)	-0.437***	6.537
	(0.045)	(8.506)
AGI in [95, 99)	2.031***	-9.965
	(0.140)	(26.29)
AGI in [99, 99.9)	13.889***	-149.2*
	(0.469)	(85.23)
AGI in [99.9, 100]	45.517***	-517.9***
	(1.068)	(179.8)
Age 40-49	0.090***	2.676**
	(0.007)	(1.091)
Age 50-59	0.046***	1.363***
	(0.004)	(0.526)
Age 60-69	0.013***	0.234
	(0.002)	(0.198)
Age 70-79	-0.012***	-0.393***
	(0.001)	(0.152)
Age 80+	-0.053***	-0.0271
	(0.002)	(0.285)
Average Dividends in [75, 95)	-3.907***	-148.4***
	(0.237)	(31.03)
Average Dividends in [95, 99)	-1.715***	-284***
	(0.287)	(39.21)
Average Dividends in [99, 99.9)	8.099***	-409***
	(0.495)	(77.21)
Average Dividends in [99.9, 100]	7.377***	-406.7***
	(0.522)	(78.88)
Receipt of Partnership/S-Corp Income Indicator	3.179***	-1.629
	(0.139)	(22.24)
Social Security Receipt Indicator	-0.034	-46.32***
	(0.083)	(11.53)
Male	-0.050***	5.078***
	(0.011)	(1.426)
Married	3.333***	-1.954
	(0.066)	(10.73)
Trading days per year	2.448***	-21.48***
	(0.066)	(7.486)
Intercept	6.657***	218.3***
	(0.420)	(56.52)

Notes: There are 498 daily observations. This table reports estimates of equation (4b), using log sales for a given individual on a given day as the outcome variable in (4a). Columns 1 and 2 are estimated from the same regression. Regressions include ten one-day lagged negative market returns, and we report the sum of these coefficients. Coefficients are scaled by 100 to facilitate interpretation. Newey-West standard errors (10-day lag) in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Table A.7. Summary Statistics by Investor Characteristic

	Average daily trading probability (1)	Average daily sales volume (2)	Number of taxpayers (3)
AGI in [0, 75)	1.1%	215	10,721,345
AGI in [75, 95)	1.4%	329	8,161,864
AGI in [95, 99)	2.4%	738	4,107,748
AGI in [99, 99.9)	5.9%	2,455	1,304,253
AGI in [99.9, 100]	12.8%	11,750	162,705
Age 18-39	0.9%	123	2,802,096
Age 40-49	1.6%	594	3,130,001
Age 50-59	1.9%	805	4,353,954
Age 60-69	2.1%	671	5,030,189
Age 70-79	2.1%	570	4,274,936
Age 80+	1.6%	332	4,864,961
Average Dividends in [0, 75)	1.0%	296	5,090,317
Average Dividends in [75, 95)	1.1%	324	11,569,151
Average Dividends in [95, 99)	2.5%	615	6,133,883
Average Dividends in [99, 99.9)	5.8%	2,434	1,549,438
Average Dividends in [99.9, 100]	5.9%	2,255	150,355
Receipt of Partnership/S-Corp Income	3.2%	1,249	5,542,750
No Receipt of Partnership/S-Corp Income	1.3%	328	18,924,712
Receipt of Social Security	1.6%	366	8,104,695
No Receipt of Social Security	1.8%	609	16,684,810
Male	2.0%	735	15,118,879
Female	1.4%	238	9,738,620
Married	1.9%	639	15,118,879
Not Married	1.4%	361	9,620,332

Notes: Table displays the average daily trading probability, average daily sales volume (including zeroes), and the number of taxpayers in our sample from 2008 to 2009. The counts in column 3 imply slightly different total numbers of taxpayers because some characteristics change over time (e.g., a taxpayer could be married in 2008 but not in 2009).