THE ECONOMIC JOURNAL

The Economic Journal, **132** (*January*), 299–325 https://doi.org/10.1093/ej/ueab059 Published by Oxford University Press on behalf of Royal Economic Society 2021.
Advance Access Publication Date: 19 July 2021

WHO SELLS DURING A CRASH? EVIDENCE FROM TAX RETURN DATA ON DAILY SALES OF STOCK*

Jeffrey L. Hoopes, Patrick Langetieg, Stefan Nagel, Daniel Reck, Joel Slemrod and Bryan A. Stuart

Using United States tax return data containing the universe of individual taxable stock sales from 2008 to 2009, we examine which individuals increased their sale of stocks following episodes of market tumult. We find that the increase was disproportionately concentrated among investors in the top 1% and top 0.1% of the overall income distribution, retired individuals and individuals at the very top of the dividend income distribution. Our estimates suggest that, following the day when Lehman Brothers collapsed, taxpayers in the top 0.1% sold \$1.7 billion more in stocks than individuals in the bottom 75%. This difference is equal to 89% of average daily sales by taxpayers in the top 0.1%.

Periods of turmoil in stock markets—such as September 2008 in the wake of the Lehman Brothers bankruptcy or during the COVID-19 crisis in March 2020—are associated with abnormally high intra-day price volatility, high trading volume and large declines in prices. Market commentary often characterises 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 certain 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. Empirical evidence on individual investors’ reaction to aggregate shocks has so far remained elusive, mainly because of a lack of appropriate data.[1] As a consequence, it is unclear what dimensions of heterogeneity give rise to trading volume and

* Corresponding author: Jeffrey L. Hoopes, Kenan-Flagler Business School, University of North Carolina at Chapel Hill, 27599, USA. Email: hoopes@unc.edu

This paper was received on 9 April 2020 and accepted on 7 July 2021. The Editor was Heski Bar-Isaac.

The data and codes for this paper are available on the Journal website. They were checked for their ability to reproduce the results presented in the paper. The authors were granted an exemption to publish parts of their data because access to these data is restricted. However, the authors provided a simulated or synthetic dataset that allowed the Journal to run their codes. The synthetic/simulated data and the codes for the parts subject to exemption are also available on the Journal website. They were checked for their ability to generate all tables and figures in the paper, however, the synthetic/simulated data are not designed to reproduce the same results.

We thank John Friedman, Terry Odean, Andrei Shleifer, Tyler Shumway, Clemens Sialm, Chris Williams, Stefan Zeume, and participants at the University of Michigan Public Finance seminar, the National Tax Association meetings, the NBER Behavioral Finance Meeting, the US 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. During the writing of this paper, some authors were 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, Federal Reserve Bank of Philadelphia, or Federal Reserve System.

[1] Anecdotal accounts of investor trading during such episodes and evidence from small subsets of the investor population point in different directions. 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 (2004) documents that Wall Street professionals tended to blame the 1929 stock market crash on ‘panic’ selling by small investors. In contrast, Young (2020) documents that,

asset reallocation. Understanding the nature of heterogeneous selling behaviour is also necessary for predicting the tax revenue consequences of market turmoil, especially with respect to the tax liability generated by realised net capital gains.

In this paper, we study anonymised administrative data from the US Internal Revenue Service (IRS) consisting of billions of third-party reports on all sales of stock in taxable individual accounts, matched with anonymised individual tax returns. These data allow us to investigate, at a daily frequency, which individuals sold stocks and mutual funds during the tumultuous market events of 2008 and 2009, a period that includes the market upheaval following the bankruptcy of Lehman Brothers in September 2008 and the subsequent decline in turbulence. We measure market tumult with lagged daily changes in the volatility index (VIX), a widely used measure of (risk-adjusted) expected market volatility. Because we focus on stock sales and, hence, on changes in investors' stock holdings, which should be triggered by a variation in the market environment, our tumult measure is based on changes in the VIX.

In our analysis of heterogeneity in investors' selling response, we focus on investor characteristics that are known to be associated with the average risky asset share in individual investors' portfolios (see, e.g., the survey by Curcuru *et al.*, 2009). These characteristics—which may proxy for differences in risk tolerance, background risk and beliefs—could also play a role in explaining individuals' responses to a tumult shock. For example, a fall in stock prices combined with a rise in risk may trigger rebalancing trades. As Kimball *et al.* (2020) highlight, in this situation investors with high risk tolerance would be sellers of risky assets in equilibrium. Coming into the tumultuous period with the highest risky asset exposure, they are the ones who would require the highest compensation in terms of expected returns to refrain from selling. Along similar lines, investors that are on average more optimistic would likely be sellers in equilibrium.

Empirically, we find that when the VIX rises, selling volume increases the most for taxpayers with high levels of taxable income, dividend income and private business income—characteristics that earlier literature has linked to willingness to bear risk and optimism (Heaton and Lucas, 2000; Carroll, 2001; Wachter and Yogo, 2010; Das *et al.*, 2020). Strikingly, this phenomenon is concentrated at the very top of the income distribution: investors in the top 1%, and even the top 0.1%, have a much greater propensity to sell during times of market tumult than do investors in the rest of the income distribution. Our estimates suggest that, following the day of the Lehman Brothers collapse, taxpayers in the top 0.1% of the income distribution sold \$1.7 billion more than those in the bottom 75%, while taxpayers in the top 1% but not the top 0.1% sold \$860 million more. These increases amount to 89% and 27% of average daily sales by taxpayers in the top 0.1% and top 1% but not top 0.1%, respectively. Moreover, because the top 1% holds about half of total stock market wealth (Wolff, 2010), this means that the high-income half of the wealth-weighted individual investor population appear as sellers in response to tumult. This result challenges the perceived wisdom in the business press that 'small investors' are most prone to 'panic', but it is not inconsistent with standard portfolio choice models.[2]

Investors with the highest level of consumption commitments also should be 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

among the clients of Vanguard, a big investment manager, wealthier investors were more likely to pull out of the equity market during the turmoil in March 2020.

[2] Of course, factors other than optimism and risk tolerance could lead high-income and high-wealth individuals to be more responsive to market tumult. For example, tax-loss selling, volatility timing, overconfidence (Graham *et al.*, 2009), or greater attention (Sicherman *et al.*, 2016), and less influence of the 'disposition effect' (Shefrin and Statman, 1985) may also help explain our results, as we discuss.

receipt of social security income as proxies. For young investors, the presence of labour income may imply that the net consumption commitment is low or even negative, especially if coupled with a potentially flexible response of labour 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. Our estimates suggest that, following the day of the Lehman Brothers collapse, taxpayers that receive social security sold \$112 million more than those who do not, an amount equal to 4% of average daily sales by social security recipients. Lower opportunity costs of paying attention to the market for retired investors are a potential additional contributor to this effect. We do not find a separate effect of age when controlling for social security income.

The data we analyse are on several levels substantially better than the data sets that have been studied heretofore to address heterogeneous responses to aggregate shocks—e.g., data from investor surveys (Shiller, 1987; Hudomiet *et al.*, 2011; Guiso *et al.*, 2013), non-randomly selected samples of portfolio holdings data (Dorn and Weber, 2013; Hoffmann *et al.,* 2013; Weber *et al.,* 2013; Barrot *et al.,* 2016) and annual administrative data from Sweden (Calvet *et al.*, 2009). Our data let us investigate, for the first time, all taxable sales by the population of US individuals at a daily frequency. This extremely large sample size allows us to analyse heterogeneity in selling, including for the extreme upper tails of the income distribution, in ways that are not possible with small survey or brokerage account data sets.

An important limitation of the data set is that it 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 (IRA). To investigate the implications of this, we analyse data from the 2007–9 panel of the Survey of Consumer Finances, 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 larger 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 of Consumer Finances are strongly related to total net sales, suggesting that our analysis of sales in taxable accounts is informative about total asset holdings.

1. 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. To be sure, there could be many explanations for why sensitivity to tumult might vary with specific observables; the purpose of the discussion in this section is to help us interpret the heterogeneity that we observe using standard portfolio choice theory.

We start by analysing 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. These perceptions of prospective risk and return are subjective and not necessarily rational. The investor receives a stream of income from riskless background wealth (e.g., labour 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 labour income and is negative otherwise. We label X the net committed consumption.

The investor solves the portfolio problem in each period assuming that 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), the optimal share of wealth invested in the stock portfolio is

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

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

$$\mu_j = v_j - p, \tag{2}$$

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

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	–

Note: 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.

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

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

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 need 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),[3] plugging in the optimal solution (1) for α_j, and then taking the derivative with respect to the 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 summarises the results. In this analysis, we assume that expected returns from the risky asset exceed the risk-free return, $\mu_j > r_F$. 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 labour 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.[4] 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 alters

[3] 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. If all investors react similarly to the tumult-driven price change, then our empirical strategy, which compares the differential response of different types of investors, will not reflect this response. On the other hand, if different types of investors respond differently to the tumult-driven price change, this will be reflected in our estimates.

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

the risky asset share unless the investor rebalances. The last row in Table 1 shows the rebalancing behaviour 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 risk aversion and non-myopic investors such as Kimball *et al.* (2020) 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 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 US data, that wealthier households take higher risk in their wealth portfolio. Bach *et al.* (2020) find that individuals in the top 1% 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% of the income distribution. Das *et al.* (2020) find that high-income individuals are more optimistic about stock market returns. Where the cut-off 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% 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 pay-offs 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 positive 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 labour income as a share of wealth compared with younger individuals (Heaton and Lucas, 2000). Thus, net of their labour income stream, consumption commitments are more likely to be positive for older taxpayers.

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.

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.

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 characterise the expected relationship between taxpayer characteristics and sales behaviour.

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 are 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.

Other factors. High-income investors may be less subject to the ‘disposition effect’, which describes the behavioural tendency to sell stocks with capital gains and hold stocks with losses (Shefrin and Statman, 1985). Increases in volatility are generally associated with decreases in prices and investors may find it advantageous to harvest losses to offset capital gains. Tax-loss selling could be particularly common among high-income investors. Although we cannot rule out that tax-loss harvesting plays an important role, existing evidence suggests that tax-loss harvesting mostly affects behaviour at the end of the tax year (Hoopes *et al.*, 2015), when investors reallocate their portfolio to minimise taxes, while in other periods factors like the (countervailing) disposition effect appear to be more important (Ivković *et al.*, 2005). High-income investors may also be more likely to follow a volatility-timing strategy in which risk exposure falls when volatility rises (Moreira and Muir, 2017). Finally, high-income individuals may trade more often because they perceive themselves to be more knowledgeable (Graham *et al.*, 2009).

In summary, the theories discussed above suggest that selling behaviour 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.

2. Data

The primary source of data is third-party information assembled by brokerages that itemises all transactions made during a tax year, which is provided to the IRS and taxpayers on Form 1099-B. We analyse data from all Form 1099-Bs for trades occurring between 1 January 2008 and 31 December 2009. We match the Forms 1099-B to data from individual income tax returns (Form 1040) and social security information. This allows us to observe wages and salaries, dividends, interest payments, retirement benefits, and net income from self-employment. In addition to income sources, we learn age and gender, number of dependents, whether the

Table 2. *Descriptive Statistics.*

	Number of transactions (millions)	Dollar volume (billions of USD)
Panel A: Sample selection		
All 1099-Bs in 2008–9	1,433	\$37,181
Eliminate non-trading and partial days	1,428	\$37,101
Eliminate negative and trades over \$2 billion	1,411	\$36,284
Individual taxpayers	870	\$9,575
Taxpayers age over 17	861	\$9,548
Stocks and stock mutual funds	274	\$6,794
Panel B: Aggregate relationship between log sales volume and market tumult		
Dependent variable: log 1099-B sales volume, USD		
Change in log VIX	6.831	
	(1.316)	
Observations (days)	498	
Observations (individuals)	24.9 million	
R^2	0.21	

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 31 December 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). In Panel B, regression includes 15 one-day lagged log VIX changes from $t-16$ to $t-1$, and we report the sum of these coefficients. We estimate Newey–West SEs (allowing for a 15-day lag) and report the SE on the sum of the coefficients in parentheses.

filing household takes a mortgage interest deduction, and the ZIP (postal) code of the filing address.[5]

Panel A of 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 (such as might result from the sale of the businesses by its founders or major shareholders), we are left with 1.41 billion 1099-Bs and \$36 trillion of volume.[6] 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.[7] 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

[5] The 1099-B provides the identity (anonymous to the researcher) of the taxpayer, so we can determine individual attributes of the taxpayer. However, characteristics that come from Form 1040, such as income, number of dependents, etc., are computed at the household level and then matched to all adults with 1099-Bs from that household.

[6] 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. Our sample contains many large trades: 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.

[7] 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 *et al.* (2016) provide evidence about the ultimate owners of pass-through entities, suggesting that they are substantially more concentrated among high earners.

taxpayer or spouse on at least one Form 1040 from 2000 to 2007. This leaves us 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 assembled by the Statistics of Income Division of the IRS in 2008 (Wilson and Liddell, 2013).[8] Online Appendix A contains additional details on the data.

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 measures the implied volatility of stock prices based on option contracts sold on the S&P 500 stock index with a one-month maturity. 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 e.g., Adrian and Shin, 2010; Longstaff, 2010; Nagel, 2012).[9] For purposes of presentation, we divide the VIX by 100 throughout and make any transformations on this re-scaled variable, and often analyse the logarithm of the VIX. Unless noted otherwise, we examine behaviour only on full trading days.

In Panel A of Figure 1 we plot the VIX, in logs and levels, at a daily frequency from 2008 to 2009. Until mid-2008, the VIX was low relative to levels seen later during the crisis. Starting in the second week of September 2008, VIX increased dramatically, from 0.23 on 8 September to 0.80 on 27 October.[10] Panel B displays VIX from September to November 2008. On the day of the Lehman Brothers bankruptcy (15 September 2008), the VIX increased by 24%.[11] 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 VIX occurred on 29 September, 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 US House of Representatives rejected legislation proposed by the Department of Treasury regarding the purchase of troubled assets. On 14 October, 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 18 November, executives of three large US auto companies testified before Congress and requested TARP funds, triggering an increase in VIX that began to turn around only on 21 November. The VIX peaked on 20 November (at 0.81), and then began to decrease toward pre-crisis levels.

[8] 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. 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 are 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, whereas the 1099-B will capture all trades of public equities.

[9] Our qualitative conclusions about heterogeneity in investor response to market tumult are preserved if we use lagged negative market returns as an alternative measure of market tumult.

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

[11] This narrative is based on https://www.stlouisfed.org/financial-crisis/full-timeline.

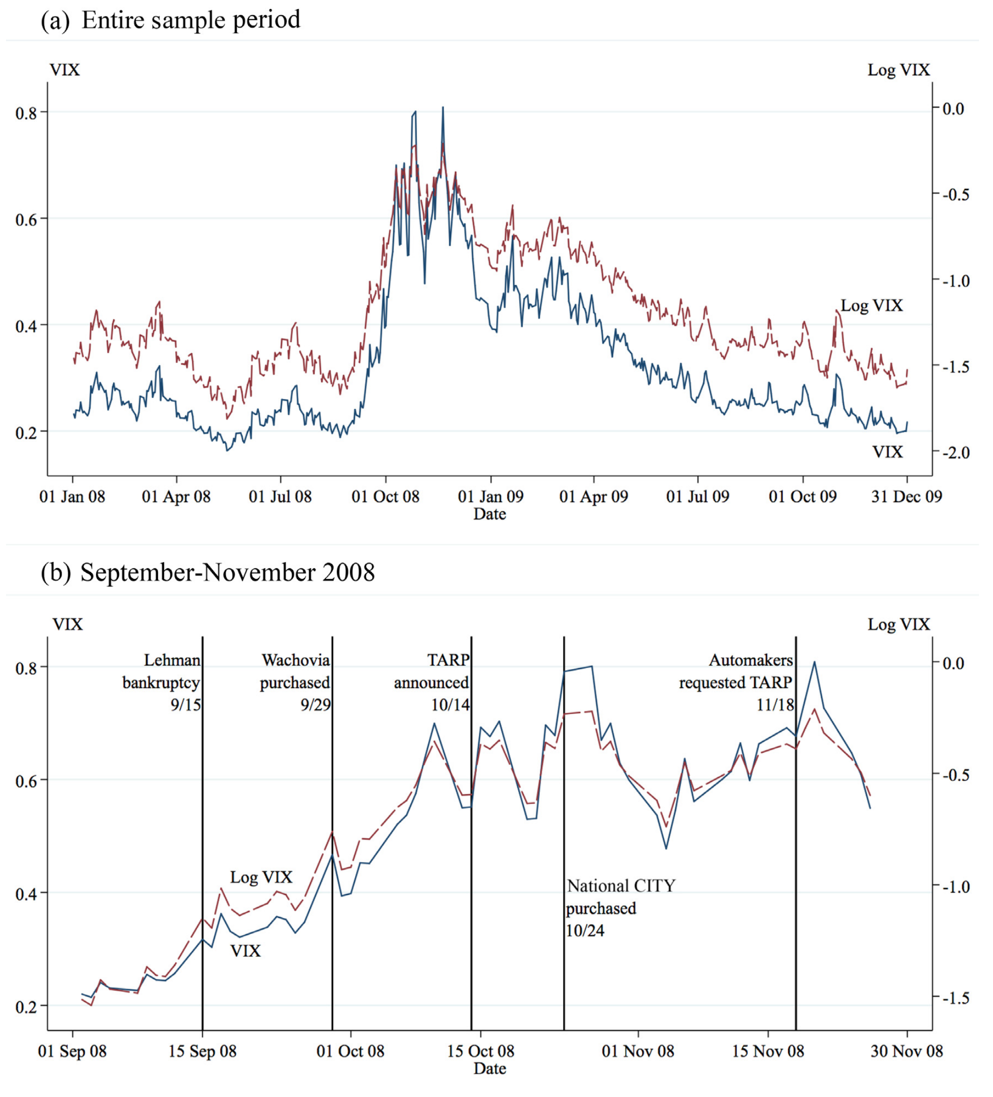

Fig. 1. *Stock Market Volatility (VIX S&P 500) Over Time.*

3. Evidence on Investor Selling Behaviour

We begin with an analysis at the aggregate level of the individual taxpayer population. Panel B of Table 2 presents a regression of the log of the aggregate value of stocks sold on changes in lagged log VIX. 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. Investors might not always respond instantaneously to these shocks, and the speed of response might vary across individuals. To flexibly capture the total effect of a change in volatility, we include several daily lags of log VIX changes and we report the sum of the coefficients along with the SE for this sum of coefficients. We test for the number of lags that are significant which leads us to include

15 daily lags of log VIX changes.[12] The table shows that market tumult is associated with a substantial amount of additional stock sales during the following days. Over our sample period, a 10% increase in VIX from day $t-16$ to $t-1$ is associated with additional sales amounting to about 98% of typical daily sales volume.[13]

The positive relationship between log VIX changes and the volume of trades is consistent with the well-known relationship between market-wide trading volume and volatility (see the survey by Karpoff, 1987), but here we document for the first time that such a relationship in daily data exists also for individual taxpayer selling volume. 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 behaviour. For now, we focus on gross sales, measured using IRS data, and in Subsection 3.3 consider the relationship between gross and net sales.

3.1. *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:

$$y_{it} = \boldsymbol{x}_i' \boldsymbol{\gamma_t} + \delta_t + \varepsilon_{it} \tag{4a}$$

$$\boldsymbol{\gamma}_t = \boldsymbol{\gamma}_0 + \boldsymbol{\beta} \boldsymbol{\Delta V_{t-1}} + \boldsymbol{\xi}_t, \tag{4b}$$

where y_{it} is a measure of individual i's sales on date t, which depend on a vector $\boldsymbol{x}_i$ of K individual characteristics in a time-varying fashion, as captured by the vector of coefficients $\boldsymbol{\gamma}_t$. Projecting $\boldsymbol{\gamma}_t$ on $\boldsymbol{\Delta V_{t-1}}$, a vector of 15 one-day changes in log VIX from day $t-16$ to $t-1$, we decompose this time variation into a part related to changes in volatility and an orthogonal residual $\boldsymbol{\xi_t}$, as shown in (4b). Our main interest centres on estimating the $K \times 15$ matrix $\boldsymbol{\beta}$, which describes how the sensitivity of sales to changes in volatility varies with the observed characteristics in $\boldsymbol{x_i}$. Because we include a time fixed effect δ_t in (4a), we identify $\boldsymbol{\beta}$ 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 $\boldsymbol{\Delta V_{t-1}}$ that we examine in Table 2. This is similar to a difference-in-differences specification, where we identify $\boldsymbol{\beta}$ using cross-sectional variation in investor characteristics and time-series variation in volatility, while controlling for time-invariant and demographic-invariant differences in investor behaviour.[14] By adding up the 15 coefficients

[12] Table A.3 in the Online Appendix reports the results of regressing the logarithm of total 1099-B sales on 5, 10, 15 or 20 daily lags of log VIX changes. The sum of the coefficients stabilises after including 15 daily lags, and lags 16–20 are not statistically significant, which leads us to choose 15 lags as our preferred specification throughout.

[13] Our regression identifies how changes in VIX from days $t-16$ to $t-1$ affect sales on day t. However, we calculate the cumulative effects of a change in VIX by adding up responses across all relevant days. For example, suppose that the VIX rises by 0.67% (=10/15) during each of the 15 days leading up to the end of day $t-1$. On day t, sales go up by $6.8 \times 0.67\% = 4.56$ log points. But, this captures only part of the response to this episode of tumult, which is spread out over days $t-14$ to $t+14$. The rise by 0.67% on day $t-15$ is followed by a cumulative effect of $6.8 \times 0.67\% = 4.56$ log points on days $t-14$ to t, the rise by 0.67% on day $t-14$ is followed by a cumulative effect of $6.8 \times 0.67\% = 4.56$ log points on days $t-13$ to $t+1, \ldots$, and the rise by 0.67% on day $t-1$ is followed by a cumulative effect of $6.8 \times 0.67\% = 4.56$ log points on days t to $t+14$. Hence, the cumulative effect across all days is $15 \times 4.56\% = 68$ log points (which translates to 98%). Moreover, the timing of the 10% rise of the VIX within the lagged time window from day $t-16$ to day $t-1$ does not matter for the cumulative effect. For example, if the entire 10% rise in the VIX occurs on day $t-1$, with zero changes prior to $t-1$, the cumulative effect over the days t to $t+14$ is given by the sum of coefficients times 10%, which is $6.8 \times 10\% = 68$ log points of average sales volume on a single day.

[14] The term $\boldsymbol{x}_i' \boldsymbol{\gamma_t}$ is akin to a group fixed effect. Because we are interested in the differential responsiveness of investors at the group-level, there is little to gain by adding individual fixed effects to equation (4a).

for each characteristic in $\boldsymbol{\beta}$, we capture the overall response to a change in tumult. Below, we show that our results are robust to alternative approaches.

We estimate equations (4a) and (4b) for our sample period of 2008 and 2009. The vector $\boldsymbol{x}_i$ includes indicators for specified percentiles of AGI (adjusted gross income) (averaged from 2000 to 2007), age ranges (as of 31 December 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 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 three challenges. First, most individuals do not trade on any given day. We address this by analysing 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 where zero-sales observations are excluded.[15]

The second challenge is that residuals ε_{it} are likely to be cross-sectionally correlated across individuals, and the third challenge is that combining equations (4a) and (4b) would require estimating a regression with about 70 billion observations.[16] We address both issues with a two-step procedure inspired by 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 $\boldsymbol{\gamma}_t$. 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 $\boldsymbol{\gamma}_t$ as the dependent variable, we then estimate (4b) as the second stage.[17] Based on the estimates of the elements of $\boldsymbol{\beta}$ 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 VIX. If members of this group raise their sales more than other groups following high $\boldsymbol{\Delta V_{t-1}}$, the group's element of $\boldsymbol{\gamma}_t$ will be high following days when VIX went up and the corresponding entry of $\boldsymbol{\beta}$ will therefore be positive; if the magnitude of this group's propensity to sell relative to other groups does not vary with $\boldsymbol{\Delta V_{t-1}}$, the element of $\boldsymbol{\gamma}_t$ corresponding to this group's membership indicator will be constant across time and hence the group's entry in $\boldsymbol{\beta}$ will be zero. The Fama–MacBeth method accounts for arbitrary cross-sectional correlation in the residuals ε_{it} and we address auto-correlation in the second-stage residuals by estimating Newey–West SEs that include lags up to 15 days.

[15] We also exclude individuals who never realise 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).

[16] There are about 140 million taxpayers per year and 500 days in our sample. Excluding those who never sell in a given year leaves us with fewer than a billion observations, which, when estimated at the day-level, is more reasonably handled in the IRS computing environment.

[17] Because the daily first-stage estimates of $\boldsymbol{\gamma}_t$ are subject to estimation error $\boldsymbol{x}_i'\varepsilon_{it}$, our sample version of (4b) includes this estimation error in the residual in addition to $\boldsymbol{\xi}_t$. 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 $\boldsymbol{\gamma}_t$ is by construction orthogonal to a time-series variable like $\boldsymbol{\Delta V_{t-1}}$, so we can consistently estimate $\boldsymbol{\beta}$ with this two-step approach even if unobserved time-series factors affect sales volume in (4a).

Table 3. *Heterogeneity in Propensity to Sell Stock: Linear Probability Model Estimates.*

	Coefficient on characteristic given no change in log VIX (1)	Interaction between characteristic and change in log VIX (2)
AGI in [75, 95)	−0.077	−0.711
	(0.008)	(0.745)
AGI in [95, 99)	0.182	0.041
	(0.021)	(1.778)
AGI in [99, 99.9)	1.539	7.398
	(0.064)	(4.098)
AGI in [99.9, 100]	3.724	16.948
	(0.138)	(8.084)
Age 30–9	0.008	−0.068
	(0.001)	(0.051)
Age 40–9	0.003	−0.050
	(0.000)	(0.026)
Age 50–9	−0.000	−0.021
	(0.000)	(0.012)
Age 60–9	−0.003	−0.007
	(0.000)	(0.011)
Age 70+	−0.008	−0.031
	(0.000)	(0.016)
Average dividends in [75, 95)	−0.479	4.518
	(0.038)	(1.919)
Average dividends in [95, 99)	−0.092	7.825
	(0.046)	(3.423)
Average dividends in [99, 99.9)	0.980	8.349
	(0.070)	(5.944)
Average dividends in [99.9, 100]	0.937	7.543
	(0.075)	(6.399)
Receipt of partnership/S corp income indicator	0.347	0.999
	(0.019)	(1.062)
Social Security receipt indicator	0.022	1.589
	(0.013)	(0.628)
Male	−0.011	−0.196
	(0.002)	(0.071)
Married	0.404	0.908
	(0.008)	(0.470)
Trading days per year	0.310	1.404
	(0.009)	(0.449)
Intercept	0.956	−6.273
	(0.071)	(2.976)

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 15 one-day lagged log VIX changes, and we report the sum of these coefficients. Coefficients are scaled by 100 to facilitate interpretation. We estimate Newey–West SEs (allowing for a 15-day lag) and report the SE on the sum of the coefficients in parentheses. There are 24.9 million individuals in the sample used to estimate equation (4a).

3.2. *Results on Stock Sales*

Table 3 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 $\boldsymbol{\gamma}_0$ and the row sums of the $\boldsymbol{\beta}$ matrix, respectively, for the specification including all trading days in our sample period. From the estimates of $\boldsymbol{\gamma}_0$ in column (1) of Table 3, we see that sales occur more frequently in the highest income groups. For instance, holding other characteristics fixed, the fraction of individuals in the top 0.1% 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%—the 'left-out' category among the income groups. This difference is substantively large, as on average 1.7% of all individuals in our sample trade on a given day. We also see 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. The cumulative effect, including any delayed reactions, equals the percentage change in VIX times the sum of the 15 coefficients we report (scaled by 100) in column (2) (see Note 13). 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 (or 35%), which is the approximate increase that occurred in the 15-day period around the Lehman Brothers collapse on 15 September 2008 (see also Figure 1). Figure 2(a) plots the results from multiplying the estimated effects in column (2) of Table 5 by 0.3, thus depicting heterogeneity in the probability of selling due to a large but realistic increase in volatility.

The regression results indicate that individuals in the top 0.1% of the income distribution respond substantially more than others to increases in VIX. A 35% increase in VIX over the previous 15 days increases the probability that an individual in the top 0.1% sells by 5 percentage points cumulatively relative to individuals in the bottom 75% of the income distribution. This gap is about one-third larger than the difference in average daily trading propensities between individuals in the top 0.1% and bottom 75%, controlling for other characteristics, that we obtained from the $\boldsymbol{\gamma}_0$ estimates in column (1). For individuals in the top 1% but not the top 0.1%, this estimated difference is 2.2 percentage points, which is roughly 45% bigger than the difference in average daily trading propensities between these groups shown in column (1). As discussed in the introduction, a higher selling propensity of high-income individuals in response to a rise in volatility and falling stock prices can be rationalised if high-income proxies for risk tolerance, relative optimism and, possibly, sophistication as an investor. This finding is very different from the popular story of the panic-prone small investor.

Portfolio choice theory also suggests retirement should be a strong predictor of response to market tumult, because consumption commitments in retirement are high relative to future income. For our sample period, we find that individuals with social security income, proxying for retirement, are significantly more likely to sell in response to an increase in volatility than individuals not receiving social security income: a 35% increase in VIX increases the probability of selling for individuals receiving social security payments by about 0.5 percentage points cumulatively relative to those who do not receive such payments, an effect equal to 28% of the total daily sales propensity for this group (see Table A.6 in the Online Appendix). Notably, the estimated effect of age conditional on receiving social security income is very small. Thus, retirement status rather than age seems to affect investment behaviour 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% of the dividend income distribution are as large as estimates for the top 1% 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

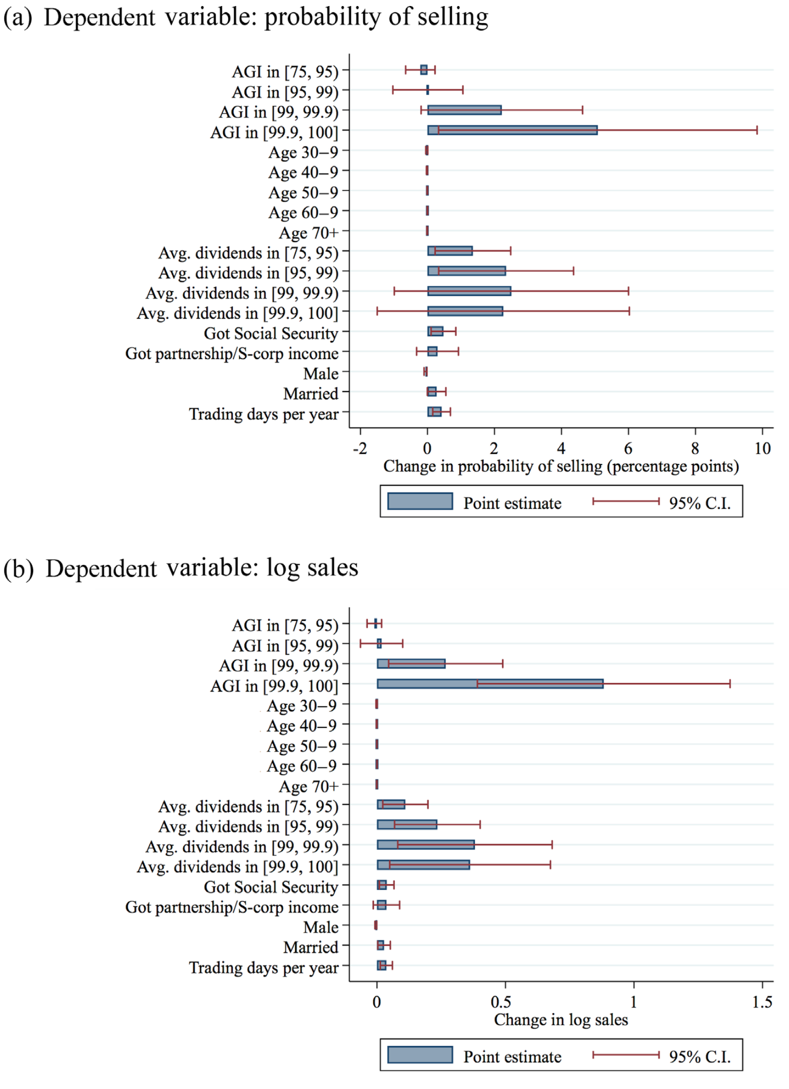

Fig. 2. *Predicted Changes in Sales for a 35% Increase in VIX.*

Notes: This figure plots main effects of interest, measuring sensitivity of sales to market tumult by investor characteristics. Panel (a) uses the regression coefficients in column (2) of Table 3, and Panel (b) uses the 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 (equal to 35%), which is approximately the change occurring following the Lehman Brothers collapse in September 2008. We also plot 95% confidence intervals based on the SEs from the regressions (see Tables 3 and 4).

point estimates are sizeable. Likewise, the total number of days in which an individual traded in 2007 is significantly and positively related to VIX sensitivity: ten more trading days is associated with a 4.2 percentage point increase in the daily sales probability following a 35% increase in VIX. 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. Our results are robust to using alternative measures of market volatility.[18]

Table 4 presents results using log sales as the dependent variable in equation (4a). The columns of Table 4 present estimates exactly as in Table 3, and we continue to scale coefficients by 100. 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 individuals and individuals who traded on a larger number of days in 2007.

Figure 2(b) plots the cumulative change in log sales for a 35% increase in VIX implied by the interaction estimates, exactly as in Figure 2(a). In our sample period, a 35% increase in VIX over the previous 15 days increases mean sales by individuals in the top 0.1% of the AGI distribution cumulatively by 0.88 log points, or 141% of their typical daily sales volume, relative to individuals in the bottom 75% of the income distribution. For individuals in the top 1% but not the top 0.1%, mean sales increase by 31%. While these effects are large, average daily sales volume is over 50 times as large for individuals in the top 0.1% as for individuals in the bottom 75% and over 10 times as large for individuals in the top 1% (see Table A.6 in the Online Appendix). Given the average daily sales volume, these estimates imply that individuals in the top 0.1% increased their mean sales cumulatively by \$10,340 (= \$11,750 × 0.88) more than individuals in the bottom 75%, holding other investor characteristics equal. Because there are 163,000 taxpayers in the top 0.1% in 2008, this amounts to a difference of \$1.7 billion in total sales following a one-time increase in VIX of 35%. Similar calculations imply that total sales by taxpayers in the top 1% but not the top 0.1% increased by \$860 million more than total sales by taxpayers in the bottom 75%.

The response of high-income taxpayers to changes in VIX is sizeable. On an average day, taxpayers in the bottom 75% sell \$2.3 billion of stocks in total, taxpayers in the top 1% but not the top 0.1% sell \$3.2 billion of stock, and taxpayers in the top 0.1% sell \$1.9 billion (see Table A.6 in the Online Appendix). The \$1.7 billion increase of taxpayers in the top 0.1% is equal to 89% of their average daily sales and 13% of average daily sales among all taxpayers in the Form 1040 data. Relative to all sales of stock, the \$1.7 billion increase is considerably smaller (slightly under 1%). However, almost three-quarters of market-level sales are made by high-frequency trading firms (see the Online Appendix), and a sizeable share of the remaining sales are made by other investment firms. Consequently, the response of individual investors is meaningfully large, but not large enough to dominate the market.[19]

A 35% increase in VIX over the 15 previous days increases mean sales by individuals in the top 0.1% of the dividend distribution cumulatively by 0.36 log points, or 43%, relative to individuals in the bottom 75% of the dividend distribution. Scaled up, this amounts to a difference of \$123 million in daily sales volume. While this effect is smaller than that for AGI, it is still substantial. The similarly scaled estimate for social security recipients, relative to those who do not receive social security income, is \$112 million in daily sales.

[18] In Tables A.4 and A.5 in the Online Appendix, we present alternative estimates using negative lagged daily market index returns as the measure of tumult. This alternative specification yields qualitatively similar results. This is not surprising as during this period large increases in VIX coincide with strongly negative market returns. Figure A.2 in the Online Appendix plots the time series of $\boldsymbol{\gamma}_t$ coefficients from equation (4a). Visual examination of the characteristics associated with high-tumult sensitivity in Table 3 shows noticeable spikes in $\boldsymbol{\gamma}_t$ on the important dates in the financial crisis depicted in Figure 1, such as the Lehman collapse and the announcement of TARP. We therefore conclude that our results likely do not depend on using any particular measure of tumult.

[19] The average day sees \$13.5 billion of sales in the Form 1040 data and \$234 billion of sales market wide.

Table 4. *Heterogeneity in Propensity to Sell Stock: Log Sales Model Estimates.*

	Coefficient on characteristic given no change in log VIX (1)	Interaction between characteristic and change in log VIX (2)
AGI in [75, 95)	−0.438	−3.421
	(0.050)	(4.807)
AGI in [95, 99)	2.033	5.975
	(0.161)	(13.973)
AGI in [99, 99.9)	13.916	89.128
	(0.529)	(37.763)
AGI in [99.9, 100]	45.609	294.086
	(1.158)	(83.604)
Age 30–9	0.089	−0.456
	(0.008)	(0.414)
Age 40–9	0.046	−0.351
	(0.004)	(0.222)
Age 50–9	0.013	−0.147
	(0.002)	(0.090)
Age 60–9	−0.012	−0.010
	(0.002)	(0.080)
Age 70+	−0.053	−0.249
	(0.002)	(0.130)
Average dividends in [75, 95)	−3.882	36.820
	(0.304)	(14.949)
Average dividends in [95, 99)	−1.666	78.321
	(0.377)	(28.325)
Average dividends in [99, 99.9)	8.171	127.061
	(0.598)	(51.093)
Average dividends in [99.9, 100]	7.449	120.739
	(0.635)	(53.185)
Receipt of partnership/S corp. income indicator	3.179	12.266
	(0.160)	(8.726)
Social Security receipt indicator	−0.026	12.628
	(0.105)	(4.845)
Male	−0.051	−1.532
	(0.013)	(0.589)
Married	3.333	9.176
	(0.074)	(4.136)
Trading days per year	2.452	12.164
	(0.077)	(4.013)
Intercept	6.620	−49.826
	(0.536)	(22.603)

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 15 one-day lagged log VIX changes, and we report the sum of these coefficients. Coefficients are scaled by 100 to facilitate interpretation. We estimate Newey–West SEs (allowing for a 15-day lag) and report the SE on the sum of the coefficients in parentheses. There are 24.9 million individuals in the sample used to estimate equation (4a).

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 what we reported in Table 3. Overall, these results suggest that heterogeneity on the extensive margin closely resembles heterogeneity on the intensive margin, conditional on having positive sales.[20]

[20] Figure A.3 in the Online Appendix plots the time series of $\boldsymbol{\gamma}_t$ coefficients from equation (4a) for the log sales model. The results clearly echo the results in Table 4 regarding which individuals sell more during periods of market tumult.

Table 5. *Gross Sales and Changes in Dividend Income.*

	Dependent variable: dividends in $t+1$ minus dividends in $t-1$			
	$t = 2008$ (1)	$t = 2009$ (2)	$t = 2008$ (3)	$t = 2009$ (4)
Gross sales	−0.0065 (0.000041)	−0.0068 (0.000051)	−0.0049 (0.000086)	−0.043 (0.000094)
Gross sales × AGI percentile [75, 95)			−0.0004 (0.00011)	−0.0008 (0.00013)
Gross sales × AGI percentile [95, 99)			−0.0005 (0.00012)	−0.0012 (0.00014)
Gross sales × AGI percentile [99, 99.9)			−0.0006 (0.00013)	−0.0017 (0.00015)
Gross sales × AGI percentile [99.9, 100]			−0.0004 (0.00022)	−0.0017 (0.00027)
AGI group fixed effects	No	No	Yes	Yes
R^2	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 winsorise all variables at the 1% and 99% levels to eliminate the effect of outliers (many of which are obvious data errors) on the estimates. Heteroscedasticity-robust SEs in parentheses.

3.3. *Gross and Net Sales in Discount Brokerage Data*

The interpretation of the evidence from gross sales volume in taxable accounts depends on whether net sales across all accounts respond similarly. 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 correspond to 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. Here we present evidence that the gross selling we observe in the Form 1099-B data is in fact quite informative about net selling.

3.3.1. *Gross versus net sales: changes in dividend income*

First, 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.[21] If gross sales are associated with net sales (i.e., gross sales minus purchases), then an 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 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.[22] We restrict the sample to individuals receiving dividends in year $t-1$, and we winsorise gross sales and dividend income changes at the 1 and 99% levels to eliminate the effect on the results of some obvious data errors.[23]

Table 5 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 interpret the

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

[22] 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.

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

meaning of these estimates, consider \$1 of gross sales on some day. For the average individual, we expect the dividend yield to be somewhere near the S&P 500 dividend yield of 2%. Thus, dividing the coefficient estimates in Table 5 of 0.0065 and 0.0068 by 0.02 we obtain the result that \$1 of gross sales corresponds approximately to net sales of \$0.33 that are not invested back into stocks within one year.[24]

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 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 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 these reasons, 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 and net sales does not vary across groups. 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 5 report the results of this test for AGI groups: we interact the specification in columns (1) and (2) with the AGI groups used in Subsection 3.2. The presence of negative estimated 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% 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 finding 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.

3.3.2. *Gross and net sales in discount brokerage data*

Next, we analyse the Barber and Odean (2000) data set of daily trades in a discount brokerage account from 1991 to 1996.[25] In the brokerage account data, we can observe both gross and net sales. 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 such assets as American depositary receipts, Canadian stocks, REITs (real estate investment trusts), 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

[24] This number can be used to approximately convert the gross sales changes that we report above into changes in net sales. For example, we estimate that individuals in the top 0.1% of the AGI distribution increased their gross sales by \$1.7 billion following a one-time increase in VIX of 35%. Scaled by one-third, our estimates suggest that net sales for this group rose by \$0.6 billion.

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

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.[26]

We first estimate our regressions with aggregate time series from Table 2B 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 15 as in Tables 3 and 4. 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 changes in log VIX. The estimate in column (1) implies that a 10% change in VIX from day $t - 3$ to t is associated with a rise in sales volume on day t of about 20%, 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 1% of the portfolio value is associated with a net sale of 0.34%. Notably, this is nearly identical to the estimate implied by our examination of dividend receipt in the tax data in Table 5. The adjusted R^2 of approximately 27% 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 behaviour 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 behaviour 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).

[26] 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.

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

	Dependent variable:					
	Log gross sales volume, taxable accounts (1)	Net sales volume (% of portfolio value), taxable accounts (2)	Gross sales volume (% of portfolio value), taxable accounts (3)	Net sales volume (% of portfolio value), taxable accounts (4)	Log gross sales volume, non-taxable accounts (5)	Log gross sales volume, taxable accounts, large–small portfolios (6)
Change in log VIX	1.753		0.917	0.315	2.231	1.152
	(0.463)		(0.131)	(0.149)	(0.652)	(0.320)
Contemp. gross sales volume		0.342				
(% of portfolio value)		(0.029)				
Adjusted R^2	0.0263	0.2685	0.0429	0.0111	0.0164	0.0101
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. The dependent variable in the last column is the log gross sales volume of large accounts (above the 80th percentile by the total value of all positions at the end of the previous month) minus the log gross sales volume of all other accounts. We estimate Newey–West SEs (allowing for a 10-day lag) and report the SE on the sum of the coefficients in parentheses.

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 how taxable accounts restricted to customers with large portfolios, defined as those above the 80th portfolio value percentile, differ from those of customers with smaller portfolios. Portfolio value is an imperfect way to approximate the high-AGI sample in the tax return data. The dependent variable in this regression is the log gross sales volume of large accounts (above the 80th percentile) minus the log gross sales volume of smaller accounts (below the 80th percentile). We obtain a statistically significant positive coefficient of 1.152. Thus, consistent with our IRS data analysis, large accounts/higher-income investors react substantially more strongly to VIX changes than other investor groups.[27] 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).

3.3.3. *Taxable and non-taxable accounts*

Using data from the 2007–9 panel of the Survey of Consumer Finances (SCF), we next provide further suggestive evidence that our inability to observe activity in non-taxable accounts does not confound the qualitative results from the tax return data. 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 wealth in taxable accounts separately, 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.[28] Using this data, we construct measures of (*a*) equities held in taxable accounts, (*b*) equities in all accounts, (*c*) net sales or purchases of equities in taxable accounts, and (*d*) net sales or purchases of equities in all accounts.[29]

How might our inability to observe non-taxable accounts influence our results? The percentage change in an individual's overall equity holdings sold in response to an uptick in volatility (our principal parameter of interest) depends on (*i*) the percentage change in their taxable equities, (*ii*) the share of overall equities held in taxable accounts, and (*iii*) the relative intensity of their stock trading in taxable accounts. Our main results suggest that (*i*) is higher for certain groups, like those with higher income. A comparison of (*i*) 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 3 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 to 2007 and, due to power concerns, we group the top 0.1% of the income distribution with the rest of the top 1%. 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 their wealth in taxable accounts, perhaps due to the limits on contributions to tax-deferred retirement accounts. These

[27] Moreover, consistent with our analysis of IRS data, the response is stronger for the top decile (above the 90th percentile) than the second-highest decile.

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

[29] 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.

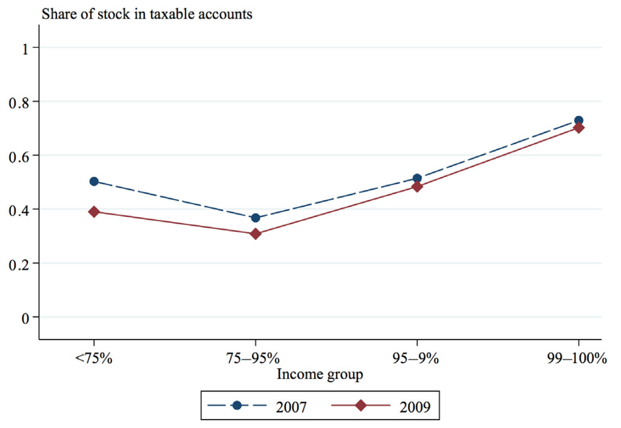

Fig. 3. *Share of Wealth in Taxable Accounts by Income Group.*
Notes: The data for this analysis come from the 2007–9 panel of the Survey of Consumer Finances. Income groups are defined based on income reported in the 2007 wave of the survey. We calculate the shares by dividing the total equities (stocks and mutual funds) held in taxable accounts in each group by the total equities held in any account in the same group.

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 column (2) of Table 6, but here we 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. Unfortunately, the SCF does not contain higher-frequency information on changes in stock holdings.

Table 7. *Net Sales in Taxable Accounts and Total Net Sales, 2007–9.*

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 × HH income percentile [75, 95)		0.146
		(0.139)
Net taxable purchases × HH income percentile [95, 99)		0.030
		(0.140)
Net taxable purchases × HH income percentile [99,100]		0.015
		(0.153)
AGI group fixed effects	No	Yes
R^2	0.772	0.728
Observations	3,857	3,857

Notes: Household (HH) income percentiles are based on the SCF. Heteroscedasticity-robust SEs, adjusted for multiple imputations in the survey data, are reported in parentheses.

Table 7 reports the results of the regression. The slope coefficient is 0.925; 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 small relative to the overall effect.[30] 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 because of data limitations it does not directly analyse the response of equity holdings in various accounts to tumult, but rather the overall variation in equity holdings.

In addition to this empirical evidence, the existing literature suggests a limited role for active trading in non-taxable accounts (Tang *et al.*, 2012). Agnew *et al.* (2003) examine 7,000 retirement accounts from 1994 to 1998 and find 'very little portfolio reshuffling' (p. 193), suggesting that many traders in retirement accounts are buy and hold traders. Further, Mitchell *et al.* (2006, Abstract) summarise their analysis of non-taxable trading data: 'Almost all participants (80%) initiate no trades, and an additional 11% makes only a single trade, in a two-year period'. They find an annual turnover due to participant trading of about 9%. For comparison, the average portfolio turnover in the Barber–Odean brokerage account data (Barber and Odean, 2000) is 75% per year. In a sample that includes the financial crisis, Tang *et al.* (2012) find that among 2.25 million participants in 401(k) plans, only about 0.68 million made at least one trade during a more than three-year period. Like in Mitchell *et al.* (2006), these numbers imply that only roughly 10% of plan participants make at least one trade per year.

4. Conclusion

As the explosion of stock market volatility in March 2020 makes clear, extreme stock market events are not a relic of the past, although little is known about the micro-structure of trading during these times. In this paper we study in detail the previous episode of exceptional volatility, during 2008–9, by analysing anonymised administrative data from the Internal Revenue Service

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

consisting of billions of third-party reports on all sales of stock in United States taxable individual accounts, with the goal of understanding which individuals sold during this period. On many dimensions, this data set is 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 characterise the sellers by the demographic information available on income tax returns and matched social security records.

Overall, our results show that there is substantial heterogeneity in investors' responses to market tumult. We find that individuals at the top 1%, and even the top 0.1%, of the income distribution are much more responsive to market tumult than individuals at the bottom 75%, 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. Evidence from several auxiliary data sets shows that these results are likely informative about net sales in taxable and non-taxable accounts and not just the gross sales in taxable accounts that we can observe directly.

These results challenge the popular notion that 'small' investors are more prone to 'panic'. In addition to providing evidence on financial market activity that draws widespread attention, our results are useful for disciplining portfolio choice models. We use these models to motivate our empirical analysis, but much work remains on integrating different selling motivations into a common framework. Incorporating these sources of heterogeneity into asset pricing models with time-varying risk would be an interesting avenue for further research. Further research could also address the heterogeneous response to tumult in more recent periods of market tumult, notably during the coronavirus pandemic.

University of North Carolina at Chapel Hill, USA
Internal Revenue Service, USA
University of Chicago Booth School of Business, USA
London School of Economics, UK
University of Michigan, USA
Federal Reserve Bank of Philadelphia, USA

Additional Supporting Information may be found in the online version of this article:

Online Appendix
Replication Package

References

Adrian, T. and Shin, H.S. (2010). 'Liquidity and leverage', *Journal of Financial Intermediation*, vol. 19(3), pp. 418–37.
Agnew, J., Balduzzi, P. and Sundén, A. (2003). 'Portfolio choice and trading in a large 401(k) plan', *American Economic Review*, vol. 93(1), pp. 193–215.
Bach, L., Calvet, L.E. and Sodini, P. (2020). 'Rich pickings? Risk, return, and skill in the portfolios of the wealthy', *American Economic Review*, vol. 110(9), pp. 2703–47.
Barber, B.M. and Odean, T. (2000). 'Trading is hazardous to your wealth: The common stock investment performance of individual investors', *The Journal of Finance*, vol. 55(2), pp. 773–806.
Barrot, J.N., Kaniel, R. and Sraer, D. (2016). 'Are retail traders compensated for providing liquidity?', *Journal of Financial Economics*, vol. 120(1), pp. 146–68.
Behr, P. and Vise, D.A. (1987). 'Wave of selling sweeps across international borders', *The Washington Post*, 20 October.
Bricker, J., Bucks, B.K., Kennickell, A., Mach, T.L. and Moore, K. (2011). 'Drowning or weathering the storm? Changes in family finances from 2007 to 2009', Working Paper 16985, National Bureau of Economic Research.

Calvet, L.E., Campbell, J.Y. and Sodini, P. (2009). 'Measuring the financial sophistication of households', *The American Economic Review*, vol. 99(2), pp. 393–8.

Campbell, J.Y. and Viceira, L.M. (2002). *Strategic Asset Allocation: Portfoliio Choice for Long-Term Investors*, Oxford: Oxford University Press.

Carroll, C.D. (2001). 'Portfolios of the rich', in (L. Guiso, M. Haliassos and T. Jappelli, eds.), *Household Portfolios*, pp. 389–430, Cambridge, MA: MIT Press.

Chai, J., Horneff, W., Maurer, R. and Mitchell, O.S. (2011). 'Optimal portfolio choice over the life cycle with flexible work, endogenous retirement, and lifetime payouts', *Review of Finance*, vol. 15(4), 875–907.

Chan, Y. L. and Kogan, L. (2002). 'Catching up with the Joneses: Teterogeneous preferences and the dynamics of asset prices', *Journal of Political Economy*, 110(6), 1255–85.

Cooper, M., McClelland, J., Pearce, J., Prisinzano, R., Sullivan, J., Yagan, D., Zidar, O. and Zwick, E. (2016). 'Business in the United States: Who owns it and how much tax do they pay?', *Tax Policy and the Economy*, vol. 30(1), pp. 91–128.

Curcuru, S., Heaton, J., Lucas, D. and Moore, D. (2009). 'Heterogeneity and portfolio choice: Theory and evidence', in (Y. Ait-Sahalia and L.P. Hansen, eds.), *Handbook of Financial Econometrics*, pp. 337–82, Oxford: North-Holland.

Das, S., Kuhnen, C. and Nagel, S. (2020). 'Socioeconomic status and macroeconomic expectations', *Review of Financial Studies*, vol. 33(1), pp. 395–432.

Dorn, D. and Weber, M. (2013). 'Individual investors' trading in times of crisis: Going it alone or giving up?', Working Paper, University of California, San Diego.

Fama, E.F. and MacBeth, J.D. (1973). 'Risk, return, and equilibrium: Empirical tests', *Journal of Political Economy*, vol. 81(3), pp. 607–36.

Graham, J.R., Harvey, C.R. and Huang, H. (2009). 'Investor competence, trading frequency, and home bias', *Management Science*, vol. 55(7), pp. 1094–106.

Guiso, L., Sapienza, P. and Zingales, L. (2013). 'The determinants of attitudes toward strategic default on mortgages', *The Journal of Finance*, vol. 68(4), pp. 1473–515.

Heaton, J. and Lucas, D. (2000). 'Portfolio choice and asset prices: The importance of entrepreneurial risk', *The Journal of Finance*, vol. 55(3), pp. 1163–98.

Hoffmann, A.O.I., Post, T. and Pennings, J.M.E. (2013). 'Individual investor perceptions and behavior during the financial crisis', *Journal of Banking & Finance*, vol. 37(1), pp. 60–74.

Hoopes, J.L., Reck, D.H. and Slemrod, J. (2015). 'Taxpayer search for information: Implications for rational attention', *American Economic Journal: Economic Policy*, vol. 7(3), pp. 177–208.

Hudomiet, P., Kézdi, G. and Willis, R.J. (2011). 'Stock market crash and expectations of American households', *Journal of Applied Econometrics*, vol. 26(3), pp. 393–415.

Ivković, Z., Poterba, J. and Weisbenner, S. (2005). 'Tax-motivated trading by individual investors', *American Economic Review*, vol. 95(5), pp. 1605–30.

Karpoff, J.M. (1987). 'The relation between price changes and trading volume: A survey', *The Journal of Financial and Quantitative Analysis*, vol. 22(1), pp. 109–26.

Kimball, M.S., Shapiro, M.D., Shumway, T. and Zhang, J. (2020). 'Portfolio rebalancing in general equilibrium', *Journal of Financial Economics*, vol. 135(3), pp. 816–34.

Longstaff, F.A. (2010). 'The subprime credit crisis and contagion in financial markets', *Journal of Financial Economics*, vol. 97(3), pp. 436–50.

Malmendier, U. and Nagel, S. (2011). 'Depression babies: Do macroeconomic experiences affect risk taking?', *The Quarterly Journal of Economics*, vol. 126(1), pp. 373–416.

Mitchell, O.S., Mottola, G.R., Utkus, S.P. and Yamaguchi, T. (2006). 'The inattentive participant: Portfolio trading behavior in 401(k) plans', Working Paper, Michigan Retirement Research Center, University of Michigan.

Moreira, A. and Muir, T. (2017). 'Volatility-managed portfolios', *The Journal of Finance*, vol. 72(4), pp. 1611–44.

Moskowitz, T.J. and Vissing-Jørgensen, A. (2002). 'The returns to entrepreneurial investment: A private equity premium puzzle?', *American Economic Review*, vol. 92(4), pp. 745–78.

Nagel, S. (2012). 'Evaporating liquidity', *Review of Financial Studies*, vol. 25(7), pp. 2005–39.

Santos, T. and Veronesi, P. (2017). 'Habits and leverage', Working Paper w22905, National Bureau of Economic Research.

Shefrin, H. and Statman, M. (1985). 'The disposition to sell winners too early and ride losers too long: Theory and evidence', *The Journal of Finance*, vol. 40(3), pp. 777–90.

Shiller, R.J. (1987). 'Investor behavior in the October 1987 stock market crash: Survey evidence', Working Paper 2446, National Bureau of Economic Research.

Sicherman, N., Loewenstein, G., Seppi, D.J. and Utkus, S.P. (2016). 'Financial attention', *Review of Financial Studies*, vol. 29(4), pp. 863–97.

Tang, N., Mitchell, O. and Utkus, S. (2012). 'Trading in 401(k) plans during the financial crisis', in (O. Mitchell, R. Maurer and M.J. Warshawsky, eds.), *Reshaping Retirement Security: Lessons from the Global Financial Crisis*, pp. 101–19, Oxford: Oxford University Press.

Traflet, J. (2004). 'Spreading the ideal of mass shareownership: Public relations and the NYSE', *Essays in Economic & Business History*, vol. 22(1), pp. 257–73.

Wachter, J.A. and Yogo, M. (2010). 'Why do household portfolio shares rise in wealth?', *The Review of Financial Studies*, vol. 23(11), pp. 3929–65.

Weber, M., Weber, E.U. and Nosić, A. (2013). 'Who takes risks when and why: Determinants of changes in investor risk taking', *Review of Finance*, vol. 17(3), pp. 847–83.

Wilson, J. and Liddell, P. (2013). 'Sales of capital assets reported on individual tax returns, 2008–2009', *Statistics of Income Bulletin*, spring, pp. 59–111.

Wolff, E.N. (2010). 'Recent trends in household wealth in the United States: Rising debt and the middle-class squeeze—an update to 2007', Working Paper 589, Levyhu Economics Institute of Bard College.

Young, J.A. (2020). 'US household trading: Coronavirus market volatility', Research note, Vanguard.