# Household Investors' Belief Dispersion, Trading Volume, and Excess Returns

Dan Li                    Geng Li *

Federal Reserve Board          Federal Reserve Board

October 2016

## Abstract

We study how stock trading activities and returns are associated with household investors' belief dispersion regarding future macroeconomic conditions using more than 30 years of household survey data. We propose a novel approach of measuring household belief dispersions and document a robust positive relationship between this measure and the stock market trading volume, even after various series of professional analysts' belief dispersion are controlled for. Such a relationship is more pronounced for belief dispersion among individuals who are more likely to own stocks and for trading volume of stocks that are more visible to household investors and stocks that are more sensitive to macro-economic conditions. Furthermore, we find that stocks whose trading volume is most sensitive to household belief dispersion tend to carry a positive excess return during recessions, when the dispersion is high. Interestingly, such an excess return is not detected when we rank stocks by their trading volume sensitivity to professional analyst belief dispersion.

# 1  Introduction

Household investors directly own about 40% of outstanding equities in the U.S. and hold an additional 20% of outstanding equities through mutual funds.[1] Despite their substantial presence, we know little about the impact of household investors on the dynamics of the stock market. Much of the existing literature views household investors as less sophisticated, uninformed, trading randomly (see, for example, Campbell (2007) and Barber, Odean, and Zhu (2006)). Existing research largely focuses on household investors' suboptimal behaviors in the stock market, such as limited participation (Mankiw and Zeldes 1991, Vissing-Jørgensen 2002) and excessive trading (Odean 1999, Agnew, Balduzzi, and Sundén 2003).

This paper studies how belief dispersion on macroeconomic conditions among household investors affects trading volume in the stock market. Our results shed light on the extent to which household investors trade based on their beliefs—as opposed to trading randomly— and the extent to which the belief-based trading of household investors drives the trading volume in the broad stock market, relative to those of the professional investors.

In a standard representative agent model, no trading occurs because all investors are assumed to be identical. Trading may arise in models where investors have different beliefs (Basak 2000, Basak 2005).[2] The theoretical significance of disagreement among investors on trading volume has been appreciated at least since Varian (1985) and Karpoff (1986), who show that trading arises if investors interpret signals differently or if they interpret signals in the same way but start with different prior beliefs.[3] Empirical tests on these models have largely focused on how the trading of securities of individual firms is affected by disagreement among professional analysts, typically within a short period around earnings releases or other major corporate news announcements.

Our analysis brings household investors to the center of the stage and examines how dis-

---

[1]Flow of Funds Accounts published by the Federal Reserve Board.

[2]Other models have shown that trading arises among agents who have different endowment levels (Wang 1994), discount rates (Gollier and Zeckhauser 2005), and preferences (Dumas 1989, Jiang 1996).

[3]Throughout the paper, we will use "disagreement" and "belief dispersion" interchangeably.

agreement among them about the macro economy affects the stock market. Because beliefs of household investors (and the dispersion therein) are potentially correlated with those of professional analysts, we run a horse race between household-based and analysts-based belief dispersion measures. To further challenge the household-based dispersion measure, we consider a wide range of professional analyst surveys—Survey of Professional Forecasters, Blue Chip Economic Indicator Survey, and I/B/E/S analyst forecasts of earnings. Our results suggest that, even after controlling for all three analyst-based dispersion measures, the positive correlation between household belief dispersion and stock market turnover rates remains substantial and significant. Specifically, a one-standard-deviation increase in household belief dispersion implies an increase in monthly market-wide turnover of about one quarter of its standard deviation. Interestingly, the effects of professional belief dispersion on trading volume diminish in the presence of household belief dispersion.

We also contribute to the literature by speaking directly to the issue of causality, using split-sample analysis and event study. We first exploit the household socioeconomic and demographic information unique to our data and examine whether the sensitivity of trading volume to belief dispersion varies across subgroups of households with different propensities of investing in stocks. We find that trading volume is more sensitive to the belief dispersion among richer and more educated households, who are more likely to own stocks. In addition, belief dispersion has a larger impact on the trading volume of stocks that are more visible to household investors. Indeed, an event study suggests firms recently added to the S&P 500 index experience an increase in the sensitivity of the trading volume to household belief dispersion. Furthermore, consistent with the fact that our household beliefs are derived from surveys of their views on macroeconomic conditions, we find that household belief dispersion is more significantly related to the trading volumes of stocks that are more sensitive to macroeconomic factors, such as those with larger market capitalization, and those with higher book-to-market ratios (value stocks).

Our paper also makes a methodological innovation in measuring investor disagreements

when beliefs are reported in categorical, instead of numerical, values. We introduce the weighted Herfindahl index to measure dispersion of ranked categorical variables. By giving more weight to responses that are different from the current status-quo, we are able to reduce noise from "lazy" answers.[4] We show that our constructed series of household belief dispersion demonstrates significant counter-cyclicality. The pattern is more pronounced than what is demonstrated in the dispersion series of professional forecasts. Moreover, household belief dispersion peaks earlier in recessions compared to professional belief dispersion .

The baseline results hold in a number of robustness analysis. First, we adopt gross flow of equity mutual funds as an alternative measure of trading volume. We find that belief dispersion among households has a similar positive effect on fund flows, which represent indirect trading of stocks through mutual funds. Second, our results hold when we calculate trading volume from only two of the exchanges, NYSE and AMEX, to address the potential double counting problem with the NASDAQ trading volume (Anderson and Dyl 2005b). Third, we show that our results are not sensitive to the choice of detrending methods for the turnover series. Finally, we experiment with varying weights in constructing our belief dispersion measure—the weighted Herfindahl index—and with using standard deviations to measure belief dispersion.

We argue that households may possess information not available to professional analysts, as the population of consumers are sufficiently heterogeneous in terms of income level, geographic location, profession, personal financial and economic experience, and industry affiliation. The professional population, on the other hand, are much more homogeneous along these dimensions. Since these characteristics likely affect belief systematically, household surveys that use nationwide representative samples likely have an advantage in capturing the dispersion of opinions that are relevant for stock market trading. That said, it is important to note that we do not presume households necessarily have rational expectations. Neither

---

[4]Survey respondents may simply check the modal answer "unchanged" to avoid taking a stand or being asked follow-up questions.

does this paper test the rationality of household expectations. The focus of the paper is to link beliefs directly to trades, and furthermore, understand whose beliefs matter more for aggregate trading volume.

We complete our analysis by exploiting the short-panel structure of our data that allows us to directly measure investors' belief changes over time and to study how dispersion of belief changes affects trading activities—a question most of the earlier work has not addressed (with Barron (1995) and Bamber, Barron, and Stober (1997) being two notable exceptions).[5] Our results confirm that greater dispersion in belief changes is also associated with greater trading volume.

The paper proceeds as follows. Section 2 summarizes the literature on the theory and empirics regarding the relationship between investor belief heterogeneity and trading volume. Section 3 describes the data. Section 4 introduces our measures of belief dispersion. Section 5 presents the main empirical results and robustness analysis. Section 6 addresses the identification strategy for establishing a potential causal relationship between belief dispersion and trading volume. Section 7 extends our analysis to the effect of dispersion of belief changes on stock trading volume. We conclude in Section 8.

## 2   Related Literature

One of the surprising and elegant pieces of economic theory is the No-Trade theorem (Milgrom and Stokey 1982). The theorem states that in a speculative market composed of fully rational agents with identical prior beliefs, no trade will occur in equilibrium, even in the presence of asymmetric information. The prediction is obviously not meant to hold in reality, but it provides a starting point for any attempt to answer the question—why do people trade in financial markets?

---

[5]Dispersion of belief changes is often dubbed "belief jumbling," a term introduced in Karpoff (1986) and Bamber, Barron, and Stober (1997).

Tirole (1982) describes the conditions under which the No-Trade theorem does not hold: (1) there exist irrational traders, or noise traders who trade for liquidity reasons; (2) some investors trade for hedging or diversification purposes; and (3) agents have different prior beliefs. Our focus in this paper is the third condition.[6] This argument is, first and foremost, empirically sound. Considering "the glass is half full or half empty" argument, it speaks to the deep psychological roots of the dispersion of human optimism or pessimism. As Hong and Stein (2007) argue, "disagreement models uniquely hold the promise of being able to deliver a comprehensive joint account of stock prices and trading volume, which we consider to be one of the highest priorities for theoretical work in asset pricing."

Various theoretical papers have demonstrated how differences in beliefs can be linked to trading volume. Karpoff (1986) shows that both different interpretations of the same information and different prior beliefs can stimulate trading activities.[7] Hence, two different aspects of belief heterogeneity—(1) dispersion of prior beliefs and (2) dispersion of belief changes—can stimulate trade, we address both in this paper. By no means is this paper the first attempt at providing empirical evidence for Karpoff's theory. Rather, we are motivated by the gaps in the existing, albeit vast, empirical literature. We highlight several major ways in which our paper improves upon and extends the previous empirical work.

First, most existing empirical studies focus on belief dispersion among financial analysts. However, Dinh and Gajewski (2007) point out that such proxies can be inaccurate since

---

[6]Regarding the first possibility, a rapidly growing literature in psychology and behavioral finance has documented the behavioral biases of human beings in making financial decisions. Hirshleifer (2001) and Barberis and Thaler (2003) provide thorough reviews of earlier contributions. Outside of the school of behavioral finance, a large body of the literature investigates trading volume under Tirole's second assumption, allowing agents to have different endowments or different preferences. For example, Wang (1994) introduces both heterogeneous investment opportunities (endowments) and asymmetric information in a competitive market, and identifies a link between the nature of heterogeneity among investors and the dynamics of trading volume. The challenge that the heterogeneous endowment argument faces is that it can generate only one round of trade, after which no further trade will take place.

[7]For subsequently developed models with different prior beliefs, see Detemple and Murthy (1994); for models in which investors have different ways of updating their posterior beliefs, see Harris and Raviv (1993) and Kandel and Pearson (1995). More recently, Scheinkman and Xiong (2002) suggest investor overconfidence as a potential source of heterogeneous beliefs, a hypothesis that finds empirical support in Statman, Thorley, and Vorkink (2006)

they represent only a distinct group of economic agents, who are often more informed and more sophisticated. In addition, analysts' forecasts can be biased by their private incentives. Such bias may render their belief dispersion an inefficient proxy for traders' disagreement. For example, analysts' desire to win investment banking clients may lead them to adjust their forecasts to avoid earnings disappointments (Chan, Karceski, and Lakonishok 2003). Moreover, Hong, Kubik, and Solomon (2000) find that analysts, especially inexperienced ones, herd in their forecasts because of career concerns. These caution us in using and interpreting the forecasts made by professional analysts.[8]

In addition, previous studies have largely focused on earnings forecasts of individual firms, instead of beliefs about future macroeconomic conditions. However, expectations on future macroeconomic and business conditions play a pivotal role in shaping investors' strategies and portfolio choices as these variables tend to influence the level and riskiness of future dividend flows of all firms and the interest rate at which future dividends are discounted.

Furthermore, most prior studies examine specific events, such as corporate earnings releases, and measure beliefs and trading volumes over a short period of time around such events (Comiskey, Walkling, and Weeks 1987, Ziebart 1990, Lang and Litzenberger 1989). Nevertheless, agents do not trade only on their opinions about earnings releases. Investors' opinions about the economy and their perspectives on interest rates and employment should all be critical in forming their opinions about financial investment and trading strategies. In reality, agents receive new information on a continuous basis, especially information concerning the economy. As new information comes, investors update both their short- and long-term outlooks for the economy and financial markets. It is therefore an empirical question as to what extent each piece of information matters in generating trades. We address

---

[8]Effort has been made to characterize belief heterogeneity when beliefs are not measurable. Bessembinder, Chan, and Seguin (1996) consider the open interest of S&P 500 index futures a proxy for dispersion of traders' opinions regarding underlying values and find it positively related to trading volume. Goetzmann and Massa (2005) construct an indirect proxy of belief dispersion from age, income, and occupation information about 100,000 retail investors and find the proxy positively related to contemporaneous trading volume and stock return.

these concerns by constructing metrics of belief dispersion using data of self-reported expectations directly collected in a nationwide representative survey of consumers.

Finally, our paper is also related to growing literature that studies the financial decisions and trading behaviors of household or retail investors. For example, Amromin and Sharpe (forthcoming) examine stock market beliefs and portfolio choices of household investors, and Kelley and Tetlock (2013) study the role of retail investors in stock pricing. Using one year worth of data from a large discount broker from the Netherland, Hoffmann, Post, and Pennings (2013) present evidence that retail investors beliefs affect their trading activities.

# 3 Data Description

## 3.1 Surveys of Consumers

We use self-reported beliefs of households from the Thomson Reuters/University of Michigan Surveys of Consumers (SCA), which is the building block of the Consumer Sentiment Index. Introduced in the late 1940s, the Index has established itself as one of the most widely followed indicators of households' sentiments about current and future economic and business conditions.

The information content of the survey has long been recognized by the research community. Validation studies by the SCA staff show that the information collected in the SCA predicts the dynamics of the nationwide economy quite well. For example, the time series correlation between the index of consumer assessment of economic conditions and real GDP growth was 0.9. The correlation between the index of unemployment expectation and subsequent realized unemployment rate was 0.8. In addition, Carroll, Fuhrer, and Wilcox (1994) find the Index of Consumer Sentiment derived from the Survey and Consumers predicts future changes of household spending. More recently, Barsky and Sims (2012) document that the "Michigan Survey of Consumers have powerful predictive implications for the future paths of macroeconomic variables." Moreover, Ang, Bekaert, and Wei (2007) find that

the mean inflation projection of the same survey outperforms statistical time series or and term structure forecast models. These studies consistently demonstrate that, despite the substantial dispersion among household beliefs, their expectations on future macroeconomic conditions appear to be quite accurate on average.

Since 1978, the SCA has been conducting monthly surveys of a minimum of 500 consumers, the majority of whom were contacted within about two weeks.[9] Our sample covers the period from 1978 to 2011, containing nearly 35 years' worth of monthly consumer-level data. The SCA survey data have a few notable advantages over various surveys of professional forecasters. The size of the SCA cross-section is more than 500, significantly larger than the sizes of professional forecasts, which typically have less than 50 respondents. Therefore, belief dispersion can be calculated more reliably in the household surveys. In addition, collecting all answers within a short window of 2 weeks ensures that forecasts are made based on a roughly similar information set across respondents. Another critical and unique feature of the SCA is that it collects key information about the demographic characteristics and the economic status of sampled consumers, which informs us of the respondens' propensity to own stocks. Finally, the SCA features a short longitudinal structure. About 40 percent of the respondents were surveyed again six months after their first interview.[10] We exploit this feature later and study the effect of dispersion of belief changes on trading volume.

Each month, the SCA asks about 50 core questions broadly related to consumers' assessments of current economic conditions and their expectations about the future economic conditions of both their households and the country.[11] We focus on the dispersion of beliefs (and its changes) about future business conditions, personal financial conditions, unemploy-

---

[9]Although the survey started shortly after World War II, respondent-level data for the years before 1961 are not publicly available. For the period from 1961 to 1965, the respondent-level data are available only in February; for 1966, they are available in February and August; and for 1967 to 1977, the respondent-level data are available quarterly in February, May, August, and December.

[10]However, these consumers were not contacted again after the follow-up interview.

[11]From time to time, additional questions, known as the "riders", were added in special modules. These questions, though interesting and potentially closely related to stock market trading activities, are typically asked only for a limited number of months and are not asked at regular monthly frequency.

ment, and interest rates. The upper panel of Table 1 lists the survey variables that our study focuses on. Four of the five questions are about consumers' expectations in the near term, typically within the next 12 months. The only question regarding long-term expectations is about consumers' expectations of business conditions during the next five years.

Most SCA questions have categorical, instead of numerical, answers.[12] For example, when asked about unemployment expectations, consumers choose from three answers—"more unemployment," "about the same," and "less unemployment". Similarly, when asked about future business conditions, consumers choose from "better off," "same," and "worse off." Categorical answers are likely easier for a typical household to respond. It is also less affected by "wild answers." However, constructing dispersion measures from categorical answers is less straightforward. We address this challenge by introducing a new measure of belief dispersion, which is explained in detail in Section 4.

## 3.2   Forecasts of Professionals

Earlier research has documented that dispersion of beliefs regarding corporate earnings among business analysts can influence trading activities of individual stocks (Comiskey, Walkling, and Weeks 1987, Ziebart 1990, Lang and Litzenberger 1989). Conceivably, wider belief dispersion among professional forecasters regarding future macroeconomic conditions can also induce higher trading volume—a hypothesis we test. Should professional-based disagreement matter for stock trading volume, we are particularly interested in whether household-based belief dispersion has any net effects on trading volume beyond the extent to which their beliefs are correlated with those among professional forecasters. To study this, we collect professional forecasters' beliefs from the following three sources:

---

[12]The only exceptions are two questions about future inflation rates, for which consumers are asked to give numerical answers. We did not include inflation expectations in our study because, relative to dispersion of categorical responses, dispersion of numerical responses in consumer surveys are more prone to be influenced by "wild" answers. For example, some reported inflation expectations were as high as 50 percent per year. As a result, the cross-sectional standard deviations of inflation expectations in the SCA are much higher than those in the Survey of Professional Forecasters.

- Survey of Professional Forecasters

  The Survey of Professional Forecasters (SPF) is a survey of professional investors and analysts on their views about macroeconomic conditions.[13] The SPF began in late 1960s, and we use the later portion of the SPF data that overlaps with the SCA data, from 1978 to 2011. The numbers of respondents for the SPF in recent years vary between 30 and 50. In addition to survey participants, the SPF differs from the SCA in a two other aspects. First, the SPF is conducted quarterly whereas the SCA is done monthly. We interpolate the SPF data to a monthly frequency to facilitate comparison. Moreover, the SPF answers are numerical, unlike the categorical answers in the household survey. The middle panel of Table 1 describes the SPF variables our study focuses on—growth of GDP, industrial production and corporate profit, and unemployment.[14] The forecasters typically report to the SPF their level projections of an economic variable over the current and the next five quarters. We derived the implied annual growth forecast from the level forecast. For unemployment forecast, we use the annual averages.

- Blue Chip Economic Indicator Survey

  Blue Chip Economic Indicator Survey are survey data collected by a monthly newsletter published by Wolters Kluwer. Similar to the SPF, participants of the Blue Chip surveys are a mix of economists at large banks, consulting firms, or academic institutions. On average, the survey has around 50 participants each month. Our study focuses on the forecasts of the following variables—summarized in the lower panel of Table 1— GDP growth, industrial production growth, nonresidential investment growth, unemployment rate, short-term interest rate, and longer-term interest rate.[15] Our sample of the Blue Chip survey data covers the period from 1984 to 2011.

---

[13]The survey was conducted by the National Bureau of Economic Research before being transferred to the Federal Reserve Bank of Philadelphia in early 1990s.

[14]SPF began collecting forecasts of short- and longer-term interest rates in the third quarter of 1981. We do not include these forecasts in our baseline analysis to keep the sample period identical. Robustness tests that include the interest rates forecasts (using a shorter sample period) yields qualitatively similar results.

[15]Unlike the SPF, Blue Chip Economic Indicator Survey contains growth, instead of level, forecasts.

- I/B/E/S Analyst Forecasts of Corporate Earnings

  The I/B/E/S database contains monthly updates of financial analyst forecasts of near-term levels of earnings-per-share (EPS) for the firms they cover. Pinto (2010) constructs a series of analysts' earning forecasts dispersion as the coefficient of variation of analysts' EPS estimates averaged across companies, which we use in this study. The series begin in 1976, and we use the period from 1978 to 2011 that overlaps with our main SCA data. The earnings forecasts are not exactly about macroeconomic conditions, which are the focus of the SCA, SPF and Blue Chip surveys. However, it is likely that the aggregate measure of forecasts' dispersion on earnings contains relevant information about the belief dispersion of macroeconomic variables that are particularly relevant for the stock market.

## 3.3 Trading Volume and Control Variables

Our measure of trading volume is the monthly turnover rate of the aggregate U.S. stock market (the total number of shares traded in a period divided by the average total number of shares outstanding during that period). Normalizing trading volume with shares outstanding allows us to abstract from increases in volume that are due mainly to the growth of the economy and the stock market. The turnover measure has been used in various studies, such as Campbell, Grossman, and Wang (1993). Data on both the number of shares traded and shares outstanding are from the Center for Research in Security Prices (CRSP). In our baseline analysis, we aggregate the monthly trading volume and shares outstanding of all securities traded on the New York Stock Exchange (NYSE), American Stock Exchange (AMEX), and NASDAQ.

As shown in the upper panel of Figure 1, turnover rates in the U.S. stock market steadily increased during the SCA data sample period. Many explanations have been offered to explain this trend. For example, Smidt (1990) suggests that the long-run trend in equity turnover can be attributed to transaction cost changes. Some researchers also attribute

this trend to the increasing importance of high-frequency trading. The Dickey-Fuller test we conduct suggests that the series is trend stationary. We then remove the trend using various detrending methods. The middle panel of Figure 1 shows the cubic detrended series of turnover, which is used in our baseline analysis. The cubic detrending method leaves smaller residual autocorrelation than linear, quadratic or fourth-order polynomial detrending method.[16] The series has a zero mean and a standard deviation equal to 0.024.

As an alternative measure of market trading volume, we also use the gross flow of equity mutual funds. The data is from the Investment Company Institute. Monthly gross flows to the equity market are calculated as the sum of sales and redemption, normalized by total asset under management by equity mutual funds. The fund flow data start from January 1984 and are presented in the lower panel of Figure 1. Unlike the turnover rates, gross flows of equity mutual funds do not show any pronounced upward trend. One plausible explanation is that mutual fund flows reflect low-frequency trading that is not affected by the increasing importance of high-frequency trading, which likely contributed to the increase in stock turnovers.

We include the S&P 500 index return and the S&P 500 index volatility as control variables. Both variables are calculated from the CRSP data. In addition, we control for stock market liquidity, which is the Pastor-Stambaugh series (Pastor and Stambaugh 2003) from WRDS. A higher value of the measure indicates more liquid market conditions.

# 4  Measures of Belief Dispersion

## 4.1  Dispersion Measures—Weighted Herfindahl Index

To measure belief dispersion from the SCA surveys, which have categorical answers, we introduce a weighted negative Herfindahl index (WNHI). Our measure is a variant of the widely used Herfindahl index, commonly used as a measure of market concentration in mar-

---

[16]In our robustness analysis, we experiment with different polynomial trends and a detrending technique due to Baker and Stein (2004).

keting and industrial organization research (see, for example, Neumark and Sharpe (1992)). We adapt it to measure the opposite of concentration—dispersion. Recall that the standard Herfindahl index is defined as

$$HI = \sum_{i=1}^{N} p_i^2, \tag{1}$$

where $p_i$ is the share of the $i$-th element among $N$ elements. The Herfindahl index takes a value between $1/N$ and 1. A lower value of the index indicates greater dispersion.

The standard Herfindahl index treats each of the $N$ elements symmetrically, without taking into account the ordering among the elements. In other words, the distances between elements are equal. However, one important aspect of the SCA data is that different answers are naturally ranked, and hence the distance between answers matters. For example, a sample consisting of 50 percent survey responses of "better off" and 50 percent "worse off" will yield the same value of standard Herfindahl index as a sample consisting of 50 percent "better off" and 50 percent "about the same" answers, although opinions in the first sample are more dispersed. To explicitly account for such relative distances, we construct (for each survey month) a weighted negative Herfindahl index as

$$WNHI = -\sum_{i=1}^{N} \omega_i p_i^2, \tag{2}$$

where $\omega_i$ is a weight assigned to element $i$. We take the negative value of the index for expositional convenience to make higher value of the index indicate greater dispersion. We give lower weights to elements closer to the polars and higher weights to elements in the middle so that we have less negative WNHI, or greater measured dispersion, for belief distributions with more polar responses. Specifically, in our baseline analysis, we let the weights on the answers of "better off" and "worse off" be equal to one and the weight on the answer of "about the same" be equal to two.

The weighting scheme of the measure not only allows us to preserve the rankings among categorical answers, but also provides flexibility in giving different importance to answers that are different in informational values. Answers of "about the same" are potentially "lazy

answers" and reveal less information about the beliefs of the respondents.[17] We can adjust for this by giving a high $\omega_i$ to these answers. That said, our results do not rely on any particular choice of weights; we alter the weights in the robustness analysis, and show that the results are qualitatively preserved.[18]

## 4.2 Composite Dispersion Measure—A Principal Component Approach

Figure 2 presents the time series of belief dispersion, measured using $WNHI$, for each SCA question we study. Recall that higher values of the $WNHI$ (closer to zero) suggest more dispersed distribution of beliefs. Belief dispersions on different questions seem to follow a common pattern. In particular, consistent with Patton and Timmermann (2010), three of the five series of belief dispersion exhibit strong counter-cyclicality. The peaks in dispersion of expectations about near-term economic conditions (BEXP), interest rate, and unemployment largely coincide with recessions as defined by the National Bureau of Economic Research. However, the cyclical patterns in the belief dispersion for expectations about personal financial (PEXP) and longer-term business conditions (BUS5) are more muted. Moreover, it appears that beliefs about longer-term business conditions in the next five years (BUS5), are more dispersed than beliefs about shorter-term economic conditions in one year (BEXP).

Expectations on various macroeconomic indicators held by the same investor are likely correlated (people expecting lower unemployment also tend to expect better business conditions), potentially making dispersion of beliefs on these macroeconomic indicators also correlated. To summarize in a parsimonious manner the information contained in the five series of belief dispersion, following Buraschi and Whelan (2010), we compute the principal

---

[17]Choosing "about the same" answer is less cognitively consuming than answers that move away from the status quo. The so called "status quo bias" is extensively documented in the psycology and behavioral economics literature. See, for example, Samuelson and Zeckhauser (1988).

[18]One caveat associated with the weighting scheme is that the weighted Herfindahl index of a variable, as a measure of its dispersion, is not independent to the mean of this variable. However, as we are going to show in Section 5.4, our baseline results are qualitatively robust to alternative weighting choices within a rather broad range. In addition, the results are qualitatively unchanged to an alternative specification where the mean of a variable is explicitly being controlled for when constructing the dispersion index.

components of these series. Our subsequent analyses focus on the first principal component, which accounts for 50 percent of total variance. Each of the successive principal components explains no more than 20 percent of the total variance. As shown in the lower right panel of Figure 2, the first principal component exhibits pronounced counter-cyclicality. Heuristically, people may disagree more when greater uncertainty prevails. The cyclicality in our belief dispersion measures is broadly consistent with the cyclicality of economic uncertainty as documented in Bloom (2009).

The SPF and Blue Chip surveys provide numerical forecasts by professionals, which allow us to calculate belief dispersion as the cross-sectional standard deviation of the forecasts. As with the SCA, we compute the first principal components of the SPF and Blue Chip data, respectively. Finally, belief dispersion in the I/B/E/S data is summarized into a single series by taking a weighted average of firm-specific analysts' dispersion in their earnings forecasts (Pinto 2010).

## 4.3   Visual Comparison of Various Dispersion Measures

Figure 3 contrasts the belief dispersion series derived from the four data sources. First, we notice that all four series demonstrate counter-cyclicality to some extent, with the cyclical pattern being particularly pronounced in the household belief dispersion series (SCA). A closer examination of the chart also reveals that household belief dispersion tends to rise sharply just before recessions (over the past twenty years in particular), whereas dispersion of professional forecasters tends to peak towards the end of recessions. This suggests that some household investors may have information predicting business cycles that is not possessed by professional analysts.[19]

It is too early to declare household investors a winner of the race, as each series shows some idiosyncratic dynamics that could be crucial for stock market trading volume. Indeed, as shown in the upper panel of Table 2, correlations among these series can be quite low,

---

[19]In a recent working paper, Loh and Stulz (2014) also find that in bad times the accuracy of analysts' earnings forecasts is worse and that they disagree more.

albeit positive. The SCA and SPF series have the highest correlation coefficient—0.63, while the SCA and the Blue Chip series have the lowest correlation coefficient—0.20, suggesting that the household belief dispersion series contain information orthogonal to what is reflected in the disagreement among professional analysts.

## 5 Baseline Results and Robustness Analysis

### 5.1 Baseline Analysis

We estimate the following model for stock market turnover:

$$Turnover_m = \alpha + \rho Turnover_{m-1} + \beta WNHI_m + \gamma Mean_m^{ICE} + \delta_1 R_m + \delta_2 \sigma_m + \delta_3 LIQ_m$$
$$+ \eta Pre2007 + \sum_{i=1}^{11} \psi_i I_{i=m} + \varepsilon_m, \quad (3)$$

where $Turnover_m$ is the cubic polynomial-detrended turnover for month $m$. We include one lag, $Turnover_{m-1}$, of the dependent variable as a control variable, taking into account the auto-correlations exhibited in the detrended turnover series. We control for the mean levels of the expectation index, $ICE$. The index is constructed by the SCA staff as a summary of investors' expectations about economic fundamentals and is therefore likely related to stock market trading activities. $R_m$ is the contemporaneous gross return in the S&P 500 index. Many papers look at the relationship between stock returns and trading volume (see for example, Campbell, Grossman, and Wang (1993) and Llorente, Michaely, Saar, and Wang (2002)). We also control for stock market volatility $\sigma_m$ and stock market liquidity $LIQ_m$, two additional factors potentially affecting market trading volume.[20] In addition, considering the fact that market trading volume exploded during the financial crisis, and particularly so for stocks with the greatest level of institutional holdings (Chordia, Roll, and Subrahmanyam 2011), we include a dummy, $Pre2007$, that indicates if the year is before 2007. The dummy controls for possible shifts in trading volume since the financial crisis that

---

[20]We also considered a specification with forward looking stock market volatility—the VIX, which measures the implied volatility of S&P 500 index options. The results are little changed.

is not necessarily related to belief dispersion. Finally, we include a vector of monthly dummies to control for seasonal factors. Hong and Yu (2009) find that trading volume in summer vacation months is significantly lower than that in other months. In contrast, trading around year-end could be higher, partly driven by tax-related reasons. These seasonal fluctuations can be captured by the monthly dummies, denoted as $\sum_{i=1}^{11} \psi_i I_{i=m}$ in Equation (3).

In the above specification, the parameter of interest is $\beta$. Recall that we construct the $WNHI$ in such a way that higher (closer to zero) $WNHI$ indicates greater belief dispersion. Should greater belief dispersion indeed induce larger trading volume, we will observe $\beta > 0$ in Equation (3). Column Model 1.a of Table 3 reports the baseline results of the relationship between stock trading volume and belief dispersion among household investors. All standard errors are adjusted for auto-correlations and heteroskadasticity using the Newey-West method.[21] Our estimates show that greater belief dispersion among household investors is indeed associated with higher stock turnover rates, and the estimated $\beta$ coefficient is positive, sizeable, and highly statistically significant. Putting the point-estimate in perspective, if the dispersion among household investors increases one standard deviation, the detrended monthly turnover rate will increase 0.56 percentage point, about one quarter of the standard deviation of the detrended turnover rate.

Regarding the effects of the control variables on stock turnover rate, we find that contemporaneous stock market returns are positively related to turnover rates, with a one-percent higher return inducing an approximately 0.4 percent higher detrended turnover rate. In addition, we find that the mean level of the household expectation index, $Mean(ICE)$, has a positive effect on turnover rates. Consistent with the literature, we find S&P 500 index volatility (annualized) to be significantly positively related to stock market trading volume. Somewhat surprisingly, stock market liquidity, as in Pastor and Stambaugh (2003), is negatively related to stock market trading volume, potentially due to the fact that episodes of

---

[21]We use the Newey-West method with first-order autocorrelation. Allowing for higher orders of autocorrelation does not change the results qualitatively.

high trading volume caused by fire sales tend to coincide with time of low market liquidity.

## 5.2 Horse Race Against Professional Analysts

Columns with heading Model 1.b through Model 1.d of Table 3 explore the relationships between stock market turnover and various belief dispersion measures for professional analysts. The models are almost identical to the one in Equation (3), except that the households' belief dispersion, $WNHI$, is replaced with those of professional analysts. Perhaps not surprisingly, estimated coefficients for belief dispersion in all three models are positive and statistically significant. An increase of one standard deviation in the belief dispersion among $IBES$, Blue-Chip, and SPF analysts correspond to increase in detrended turnover rate of 9.5 percent, 12.1 percent and 10.9 percent, respectively. These numbers, though somewhat smaller than what is found for household belief dispersion, remain economically significant. To the best of our knowledge, this is the first exercise that documents a significant positive relationship between aggregate stock market trading volume and belief dispersion among professionals regarding macroeconomic conditions.

Since professional investors are typically deemed more sophisticated and informed than household investors, we are then interested in any incremental explanatory power of household belief dispersion over those among professional analysts. We do this by adding professional belief dispersions into our baseline model cumulatively. The results of the horse race are presented in columns Model 2 through Model 4 in Table 3. As shown in column Model 2, where the I/B/E/S dispersion series is added to Equation (3), the coefficient on household belief dispersion is little changed and remains statistically significant. Whereas the estimated coefficient for the I/B/E/S dispersion shrinks in magnitude compared to the result in Model 1.b and becomes statistically insignificant. In Model 3, we add Blue Chip dispersion to Model 2, and find that the size of the estimated coefficient for household belief dispersion somewhat increases and remains statistically significant. By contrast, the coefficient on Blue Chip dispersion series becomes smaller than in Model 1.c and less significant. Finally,

when all four belief dispersion series are included, the coefficient on household belief dispersion (shown in the column Model 4) becomes the only one among four dispersion measures that is sizeable and significant at the 99-percent level. In contrast, statistical significance disappeared for the estimated coefficients on I/B/E/S and the SPF dispersion.[22]

The horse race results highlight the additional information in household belief dispersion that has a material effect on the dynamics in stock market trading volume. Our results thus sufficiently challenge the traditional view of households as uninformed and insignificant participants in the stock market. To the extent that disagreement among such investors is significantly related to stock market trading volume, it is likely that, household investors collectively have information that professional analysts do not have and therefore contribute to price discovery and market efficiency.[23]

## 5.3 Belief Dispersion and Mutual Fund Flow

A substantial portion of stocks are held by households through equity mutual funds. We estimated that gross equity mutual fund flows account for as much as 10 percent of total stock trading volume. In addition, the time series variations in mutual fund flows account for 20 to 25 percent of those in the detrended stock market turn over series. We therefore hypothesize that greater disagreements among household investors may also leave footprints on equity mutual fund flows as they alter their exposures to market risks by changing their allocations to equity mutual funds. To test this hypothesis, we estimate Equation (3) with *Turnover* being replaced by the gross equity fund flow rate. The flow rate is calculated as the ratio between gross volume of equity mutual fund flows—the sum of new sales and

---

[22]Broadly speaking, the estimates of control variables are largely little changed across all columns in Table 3. There are only a few exceptions. First, the mean expectation coefficients are small and insignificant for Model 1.b (I/B/E/S) and Model 1.c (Blue Chip), and second, the coefficients of S&P return become less significant in columns Model 1.c, Model 3, and Model 4. The somewhat weaker effects of S&P returns on trading volume may be related to the lower sample frequencies (monthly) that our study uses, compared with the daily frequency used in the literature

[23]Examining market and limited orders placed by retail investors and the relationship between these orders and stock returns and firm news, (Kelley and Tetlock 2013) also argue that retail investors potentially contribute to market efficiency.

redemption—and the total asset under management by stock mutual funds.[24] The results, shown in Table 4, present several notable findings.

First, household belief dispersion is strongly positively related with gross equity fund flows. Model 1.a of Table 4 includes only the household belief dispersion variable. Professional dispersion series are added cumulatively in Model 2 through Model 4. Similar to the results for stock market turnover, we find equity fund flows to be positively related to household belief dispersion with a very high level of statistical significance. Second, the magnitude of household belief dispersion's effect on equity fund flows is significant across all models—a one-standard-deviation increase in household dispersion is associated with an increase in gross flow of one fifth of its standard deviation. Finally, disagreement among professional forecasters is either unrelated or negatively related to equity mutual fund flows. Somewhat surprisingly, most control variables, such as S&P return, stock market liquidity, and mean consumer expectation are unrelated to fund flows. Only the coefficient of S&P volatility is positively and statistically significant.

## 5.4 Robustness Analysis

In our baseline analysis, we made specific choices on model specifications, on the weights used in computing the dispersion measure $WNHI$, and on the construction and detrending of stock turnover rates. We now implement an array of robustness test to evaluate whether our results are sensitive to these choices. For all alternative specifications we run as robustness tests, we estimate the four models(Model 1.a, 2, 3, and 4) as in Table 3. The estimated $\beta$ coefficients from these alternative specifications are summarized in Table 5. The baseline results are presented in the top row of the table for comparison.

---

[24]The mutual fund literature have been focusing on net flow, which is gross purchase minus gross redemption. Empirical studies on net flows to mutual funds largely establish a positive relationship between net flows and stock returns (see, for example, Warther (1995)).

### 5.4.1 Controlling for Business Cycle Effects

To begin with, we note that the household belief dispersion series, as shown in figure 2, is highly cyclical. Arguably, households may trade stocks actively during economic downturns to free up liquidity for consumption smoothing purposes, to rebalance their portfolios, or (for many perhaps) to trim realized losses. It is therefore important to examine whether our results speak to an intrinsic association between households' belief dispersion and trading activities or merely pick up the business cycle effects. Accordingly, we experiment with including controls of contemporaneous and various lags (up to one year) of unemployment rate and GDP growth to better capture the cyclical position of the market and the economy.[25] In addition, recent research reveals that economic fundamentals appear to have only a limited impact on household expectations of future returns once lagged realized returns and price-dividend (P/D) ratios are controlled for (**?**). We therefore include the corresponding lagged returns and P/D ratios as additional controls.**For Dan: Is this the best way of motivating adding lagged returns and P/D ratios as additional controls—I was citing the comments of a referee. Also, we need to add something in the data section to describe the source of P/D ratios.**

We first add contemporaneous unemployment rate, GDP growth, and P/D ratios to the baseline specification, which has controlled for contemporaneous realized S&P returns. In the next experiment, we add to this augmented specification one quarter-lags of unemployment, GDP growth, realized S&P returns, and P/D ratio. We subsequently further add one year-lags of these four variables to this specification. As shown in Table 5, the estimated $\beta$ coefficients are qualitatively little changed from the baseline estimates. Moreover, as in the baseline regression, adding professional analysts' dispersion measures (models 2, 3, and 4) does not undermine the statistical and economic significance of $\beta$ as we are adding more complete controls of business cycle factors. Finally, we project each of the SCA, IBES, Blue

---

[25]GDP growth is available at a quarterly frequency. The three months in the same calendar quarter have the same value of GDP growth. We use the final BEA release for GDP growth.

Chip, and SPF dispersion measures on contemporaneous and one quarter- and one year-lag of unemployment, GDP growth, realized S&P returns, and P/D ratio, and replace the dispersion measures in Equation (3) with the respective residuals of these projections. The results, presented in row (4) of Table 5, indicate that even after filtering out the cyclical movements out of the dispersion series in a rather elaborated way, the residual component of household belief dispersion remain strongly associated with stock trading volume. This pattern holds even multiple indicators of professional analysts' belief dispersion are controlled for.

### 5.4.2 Alternative Measures of Household Belief Dispersion

We next experiment with assigning different weights, $\omega_i$ in Equation (2), to survey answers of "about the same" when we compute the weighted Herfindahl index. In our baseline analysis, we give a weight of 2 to such answers. We now experiment with smaller weights of 1.5 and 1, with the latter corresponding to an unweighted index. Two observations stand out from the results presented in the next panel of Table 5. First, the estimated $\beta$ coefficient remains sizeable and statistically significant for both weighting options, regardless whether professional analysts' disagreement is controlled for. Second, the magnitude and statistical significance of $\beta$ estimates of $\omega = (1,1.5,1)$ are consistently smaller than in our baseline specification, and even more so for $\omega = (1,1,1)$, suggesting the extent to which we penalize central answers in constructing the belief dispersion series does matter regarding the explanatory power of the series with respect to trading activities. Indeed, as shown in the next row, we also experimented with assigning numerical values -1, 0, 1 to categorical answers of "worse," "same," and "better," respectively and compute the standard deviation as an alternative measure of belief dispersion. We find the estimated $\beta$ coefficients lose even more significance. Thus, we argue that, on balance, the $WNHI$ we introduce serves as an informative and flexible way to better extract information regarding belief dispersion in categorical survey responses. That said, we also acknowledge that measuring household

belief dispersion is tricky and the $WNHI$ has its own potential limitations.

One caveat of our $WNHI$ measure is that the dispersion index of a variable is not independent to its mean. We examine whether our results are driven by such a correlation between the dispersion and the mean index. To do so, for each $WNHI$ of the five variables we consider—$PEXP$, $BEXP$, $BUS5$, $UNEMP$, and $RATES$—we regress the $WNHI$ series on the mean series of the respective variable, and construct the first principal component of the five residual series generated from these regressions. The first principal component series so constructed is highly correlated with the series in our baseline analysis, with a correlation coefficient equal about 0.95. Replacing this first principal component series in the baseline specification Equation 3, the results (shown at the bottom of this panel, labelled as $WHI$ de-meaned) are qualitatively little changed from the baseline analysis.

### 5.4.3 Alternative Sample Period, Market, and Detrending Specifications

Another observation from Figure **??** is that the variance of the turnover series increased after 2007. We re-estimated the model using only the pre-2007 sample to confirm that our results are not driven by this period of relatively high variances. As shown in the first row of the bottom panel, while the $\beta$ coefficients estimated using this sub-sample are smaller than those estimated from the entire sample, they remain statistically significant and economically sizeable, with and without including professional analysts' belief disperison. We then examine if our results are sensitive to any particular technique used to detrend the series of turnover rates. Specifically, we consider a linear trend, a quadratic trend, and a detrending algorithm used in Baker and Stein (2004), who propose a stochastic detrending technique of subtracting a lagged five-year mean from the current series. The results, presented in the same panel, reassure that our estimates are robust with respect to various detrending methods. Finally, we exclude trading in the NASDAQ exchange when constructing the turnover rate series to address concerns that NASDAQ inter-dealer trades are double counted ((Anderson and Dyl 2005a)). The results, presented at the bottom of the table, suggest that

our baseline results are not specific to the NASDAQ exchange.

## 5.5  Why Household Beliefs Matter More?

What is the source of additional information possessed by households but not by professional analysts? We argue that the SCA panel is more heterogeneous and potentially captures information from a broad scope, as it consists of over 500 households that straddle different geographic locations, professions, personal financial and economic experience, and industry affiliations. Many studies document how economic shocks can vary among people across these dimensions. For example, Souleles (2004) finds household forecast errors in the SCA to be correlated to their demographic characteristics, as aggregate shocks do not hit all people equally. Favara and Song (2014) argue that dispersion of income shocks is a good proxy for information dispersion among city residents. The dispersion of income shocks also varies across cities, which helps explain cross-sectional differences in house price volatilities across US cities. In addition, two recent episodes of recessions affect specific industries, technology, and financial services disproportionately. Households who are directly linked to those industries likely form systematically different expectations from people linked to other industries. Finally, even with the same information set, people can form different expectations based on their different personal experience. Kaustia and Knüpfer (2008) find that people with more experience in IPO subscription make better return in future IPO investment. Choi, Laibson, Madrian, and Metrick (2009) find personal experience to affect saving decisions.

Compared to the SCA household data, the cross-section of professional forecasters in Blue-Chip, SPF, and I/B/E/S likely come from a more homogeneous pool. They typically work for the same industry (financial industry), live in a smaller set of metropolitan areas (likely financial center cities), and likely share similar personal financial experience. Thus, professional forecasters may possess an incomplete information set.

# 6  Identification Strategy and Evidence for A Causal Relationship

Thus far we have shown a pronounced positive correlation between belief dispersion among household investors and stock market trading volume. Such a relationship is robust to various model specifications and, in particular, to a horse race with belief dispersion among professional analysts. We now explore whether our results speak to a causal relationship between household disagreement and stock market trading.

Omitted variable bias is one obstacle we need to deal with before claiming causality. For example, one can potentially argue that household disagreement on macroeconomic conditions may be related to the amount of sunshine people have, which may also affect the amount of trading we observe. Although neither of such relationships has been documented, we cannot *a priori* rule them out. After all, Hirshleifer and Shumway (2003) find sunshine to be related to stock return. We address these concerns in two ways. First, we split the sample of households into different demographic groups and compare the sensitivity of trading volume to the belief dispersions within these groups. The premise of our test is that if household belief dispersion indeed affects trading volume, it should be true that belief dispersion among households who are more likely to own stocks has a stronger impact on trading volume.

Second, household belief dispersion should have a greater impact on the trading volumes of stocks that are more visible to household investors. We conduct an event study for this. We focus on events that increase visibility of individual stocks to household investors, and compare the sensitivity of the stock's turnover to household belief dispersion before and after such events. Specifically, the events we study are stock index inclusions and exclusions. We predict that post index inclusion, stocks' turnovers are more sensitive to household belief dispersion.

Finally, because our focus is households' expectations on macroeconomic conditions, we

expect that the relationship between belief dispersion and trading volume to be more pronounced for stocks that are more sensitive to macroeconomic conditions. Empirical asset pricing literature generally finds that value stocks are riskier (larger consumption beta and market beta) in bad times (Lettau and Ludvigson (2001) and Petkova and Zhang (2005)). Indeed, our cross-sectional analysis suggests that household belief dispersion is more significantly correlated with the trading volumes of stocks that have large capitalization and high book-to-market ratios (value stocks).

## 6.1 Analysis by Demographic Characteristics

To begin with, we study the effect on trading volume for belief dispersion among subgroups of households that have different demographic and socioeconomic characteristics and therefore have different propensities to own stocks. Existing studies, such as Hong, Kubik, and Stein (2004), document that prime-age, more educated, white, and higher-income investors are more likely to hold stocks. Consequently, we construct demographic-specific $WNHI$ measures, as defined in Equation (2), using answers from subgroups of households that differ in dimensions such as age, educational attainment, race, and income, respectively. Specifically, we estimated $WNHI$ among consumer below versus above age 35, high school and below versus college and above, black versus white, below median income versus above median.

We then estimate a variation of Equation (3),

$$Turnover_m = \alpha + \rho Turnover_{m-1} + \beta^L WNHI_m^L + \beta^H WNHI_m^H + \gamma Mean_m^{ICE} + \delta_1 R_m + \delta_2 \sigma_m + \delta_3 LIQ_m$$
$$+ \eta Pre2007 + \sum_{i=1}^{11} \psi_i I_{i=m} + \varepsilon_m, \quad (4)$$

with the demographic-specific $WNHI$s. In the above regression, $WNHI^H$ and $WNHI^L$ denote the belief dispersion among the households with high and low likelihood of owning stocks, respectively. Observing that belief dispersion among prime-age, better educated, white, and higher-income consumers having a more pronounced effect on stock trading volume is consistent with a causal relationship between more dispersed beliefs and higher trading

26

volume.

We present in Table 6 the estimate of $\beta$ from the model that involves only household belief dispersion (Model 1.a) and the estimates from models that control for increasing number of analyst-based belief dispersion measures (Model 2, 3, 4). First, we note that the standard deviations of the sub-population $WNHI$ series are very similar, with the exception that $WNHI$ among white consumers is about 15 percent more volatile than that among black consumer. Figure **??** plots the $WNHI$ series computed among consumers younger than 35 and consumers 35 and older. The two series appear to fluctuate within a similar band and share similar cyclical trends.

Examining the estimated $\beta^L$ and $\beta^H$ coefficients, we find that $\beta^H$ coefficients corresponding to consumers older than 35 are positive and statistically significant, whereas $\beta^H$ coefficients corresponding to younger consumers have imprecisely estimated small negative values. Moreover, across all model estimated, the differences between the estimated $\beta^L$ and $\beta^H$ values are statistically significant. Belief dispersion among consumers with only high school or below education does not appear to affect stock trading volume much, whereas the coefficients estimated for belief dispersion among people who have some college education all have significant positive values. Here the F-tests for $\beta^L = \beta^H$ are statistically significant for model 1.a and model 2, but not so but models 3 and 4. Very similar patterns can be found when we contrast $WNHI$ for consumers whose income is below and above the median. That said, the contrast between the effects of white and black consumers' belief dispersion is less stark. While the coefficients estimated for the $WNHI$ among black consumers, who are on average less likely to own stocks, are smaller and less statistically significant than those for white consumers, the differences are generally not statistically significant.

Finally, we divide the SCA sample by imputed stock ownership. Specifically, we estimate a linear probability model of stock ownership using data from the Survey of Consumer Finances and use the estimated coefficient to impute the stock ownership likelihood in the SCA sample. We then split the sample at the median of the imputed likelihood and compute

the $WNHI$ for each subsample, respectively. The estimated $\beta^L$ and $\beta^H$ coefficients are reported at the bottom of the table. We continue to find that the coefficients of the belief dispersion among consumers who are more likely to own stocks are consistently significant, whereas the coefficients estimated for consumers with below-median likelihood of owning stocks are much smaller and statistically insignificant.

To summarize, our results consistently show that the belief dispersion among consumers who are more likely to own stocks has a stronger association with stock trading volume, which is consistent with a causal relationship between belief dispersion and stock trading volume.

## 6.2  Evidence from S&P 500 Index Inclusion and Exclusion

One way for individual stocks to gain visibility among household investors is to be included in widely used indices, such as the S&P 500 index. We use events of index inclusion and exclusion to examine the changes in sensitivities of trading volume to household belief dispersion after changes in visibility. If household investors gravitate more towards highly visible stocks, such as those included in the S&P 500 index, we expect the sensitivity of trading volume to the belief dispersion among household investors to increase after index inclusion.

We use the S&P 500 index composition history file in CRSP to come up with a list of 728 index inclusion events, and 252 index exclusion events during the period of 1978-2011. The monthly turnover of these stocks around the index change events are calculated. The event window is chosen to be [-6,+6] months, and the event month, month 0, is removed. Since we are interested in time series variations, firm fixed-effects are included in the model. The panel setting in this exercise distinguishes itself from our baseline model where only aggregate stock market trading volume is studied.

We find that after being included in the S&P index, a stock's trading volume becomes more sensitive to household belief dispersion(Column 1 to 4 of Table 7), as suggested by the significantly positive coefficient for the term "$PostEvent \times HouseholdDisp$", which is an

interaction term between the post event indicator and household belief dispersion. Trading volume's sensitivity to belief dispersion is not lowered immediately after index deletion (Column 5 to 8), where the interaction term becomes insignificant. The asymmetric results are consistent with similar asymmetric patterns of stock abnormal returns around index inclusion and exclusion events (Chen, Noronha, and Singal 2004), which is explained by changes in investor awareness around index change events. Our results support the idea that stocks gain visibility to retail investors after being included in the index, but do not lose that visibility after being removed from the index. The event study results also suggest that our baseline results are not spurious; it holds true even in a panel setting, and is suggestive of a causal relationship between belief dispersion and trading volume.

## 6.3  Cross-Sectional Implications

If dispersion of beliefs regarding future macroeconomic conditions drives stock trading activities, one would expect to observe the relationship between belief dispersion and trading volume to be more pronounced for stocks more sensitive to fluctuations in the broad economy. We explore such a potential cross-sectional variation in two ways—stocks of different market capitalization and different book-to-market ratios. Specifically, we hypothesize that trading volumes of large cap stocks and value stocks are more sensitive to belief dispersion among household investors regarding their expectations on future macroeconomic conditions. We estimate Equation (3) with the stock market turnover rates being replaced with turnover rates of companies in the top-, medium-, and bottom-terciles of market cap and book-to-market ratio distribution, respectively.

The results, controlling for all three belief dispersion series of professional analysts, are reported in Table 8, and they largely confirm our hypothesis. The $\beta$ coefficient estimated for large-cap companies is similar to our baseline result (the left column), whereas household belief dispersion does not appear to materially affect trading volume of stocks of small- or medium-cap companies. Interestingly, Da, Engelberg, and Gao (2011) find that a direct

measure of retail investors' attention to stocks—the Google search volume on the stock ticker—is significantly positively related to the size of the stock. Thus, our results also speak to the notion that household belief dispersion has a greater effect on the trading volume of stocks they follow more closely.

Regarding the distribution of book-to-market ratio, we note two observations in the table. First, household belief dispersion is the only one among the four belief dispersion measures we consider that appears to have consistent statistically significant effects on trading volume of stocks across the book-to-market ratio distribution. Second, in line with our hypothesis, the effect of household belief dispersion on trading volume is the strongest for for stocks with the highest book-to-market ratio.

# 7 Trading Sensitivity and Excess Returns

To Be Added

# 8 Conclusion

This paper presents strong evidence that disagreement among household investors drives trades in the stock market. We find robust and pronounced evidence that stock turnovers and equity mutual fund flows are both positively related to the dispersion of household beliefs about the future macroeconomic outlook. Notably, household belief dispersion's effects on stock trading volume remain substantial and statistically significant even after controlling for up to three professional analyst-based dispersion measures, while those professional belief dispersion measures' impact on volume becomes statistically insignificant. In addition, our paper introduces the weighted Herfindahl index as a measure of belief dispersion for categorical responses provided in household surveys, a methodological innovation to the literature.

Exploiting some unique features of our consumers survey data, we present substantial

evidence that are consistent with a causal relationship between household belief dispersion and stock trading volume. In particular, we find that stock turnovers are more sensitive to belief dispersion among consumers who have greater propensities to own stocks. The impact of belief dispersion on trading activities is more prominent for stocks that are more visible to retail investors, such as stocks that are recently included in the S&P 500 index.

Consistent with Campbell (2007) and Kelley and Tetlock (2013), our results indicate that household investors are strong drivers of trading volumes in the U.S. stock market. Household investors' collective involvement in the stock market may also contribute to market efficiency and price discovery. Using the household belief dispersion index we introduce here, a recent study by Duarte, Navarro-Staicos, and Rosa (2014) finds that household belief dispersion is an important factor in explaining the low volatility in asset prices prior to the financial crisis. A natural avenue of future research is to explore broadly the pricing implications of household beliefs.

# References

Agnew, Julie, Pieluigi Balduzzi, and Annika Sundén, 2003, Portfolio choice and trading in a large 401(k) plan, *American Economic Review* 93, 193–215.

Amromin, Gene, and Steven A. Sharpe, forthcoming, From the horse's mouth: How do investor expectations of risk and return vary with economic conditions?, *Management Science*.

Anderson, Anne-Marie, and Edward Dyl, 2005a, Market structure and trading volume, *Jounal of Financial Research* 28, 115–131.

Anderson, Anne-Marie, and Edward A. Dyl, 2005b, Market structure and trading volume, *Journal of Financial Research* 28, 115–131.

Ang, Andrew, Geert Bekaert, and Min Wei, 2007, Do macro variables, asset markets, or surveys forecast inflation better?, *Jounal of Monetary Economics* 54, 1163–1212.

Baker, Malcolm, and Jeremy Stein, 2004, Market liquidity as a sentiment indicator, *Journal of Financial Markets* 7, 271–299.

Bamber, Linda Smith, Orie E. Barron, and Thomas L. Stober, 1997, Trading volume and different aspects of disagreement coincident with earnings announcements, *The Accounting Review* 72, 575–597.

Barber, Brad, Terrance Odean, and Ning Zhu, 2006, Do noise traders move markets, *Working Paper*.

Barberis, Nicholas, and Richard Thaler, 2003, *A survey of behavioral finance* vol. 1 of *Handbook of the Economics of Finance* . pp. 1053–1128 (Elsevier).

Barron, Orie E., 1995, Trading volume and belief revisions that differ among individual analysts, *Accounting Review* 70, 581–597.

Barsky, Robert B., and Eric R. Sims, 2012, Information, animal spirits, and the meaning of innovations in consumer confidence, *American Economic Review* 102, 1343–1377.

Basak, Suleyman, 2000, A model of dynamic equilibrium asset pricing with heterogeneous beliefs and extraneous risk, *Journal of Economic Dynamics and Control* 24, 63–95.

——— , 2005, Asset pricing with heterogeneous beliefs, *Journal of Banking & Finance* 29, 2849–2881.

Bessembinder, Hendrik, Kalok Chan, and Paul J. Seguin, 1996, An empirical examination of information, differences of opinion, and trading activity, *Journal of Financial Economics* 40, 105–134.

Bloom, Nick, 2009, The impact of uncertainty shocks, *Econometrica* 77, 623–685.

Buraschi, Andrea, and Paul Whelan, 2010, Term structure models with differences in beliefs, *Working Paper*.

Campbell, John, 2007, Who are the noise traders, *Economist's View* 61.

Campbell, John Y., Sanford J. Grossman, and Jiang Wang, 1993, Trading volume and serial correlation in stock returns, *Quarterly Journal of Economics* 108, 905–939.

Carroll, Christopher D, Jeffrey C Fuhrer, and David W Wilcox, 1994, Does consumer sentiment forecast household spending? if so, why?, *American Economic Review* 84, 1397–1408.

Chan, Louis K. C., Jason Karceski, and Josef Lakonishok, 2003, Analysts' conflict of interest and biases in earnings forecasts, Discussion paper, National Bureau of Economic Research, Inc.

Chen, Honghui, Gregory Noronha, and Vijay Singal, 2004, The price response to s&p 500 index additions and deletions: Evidence of asymmetry and a new explanation, *Journal of Finance* 59, 1901–1930.

Choi, James J., David Laibson, Brigitte C. Madrian, and Andrew Metrick, 2009, Reinforcement learning and savings behavior, *The Journal of Finance* 64, 2515–2534.

Chordia, Tarun, Richard Roll, and Avanidhar Subrahmanyam, 2011, Recent trends in trading activity and market quality, *Journal of Financial Economics* 101, 243 – 263.

Comiskey, Eugene E., Ralph A. Walkling, and Michael A. Weeks, 1987, Dispersion of expectations and trading volume, *Journal of Business Finance & Accounting* 14, 229–239.

Da, Zhi, Joseph Engelberg, and Pengjie Gao, 2011, In search of attention, *Journal of Finance* 66, 1461–1499.

Detemple, Jerome, and Shashidhar Murthy, 1994, Intertemporal asset pricing with heterogeneous beliefs, *Journal of Economic Theory* 62, 294–320.

Dinh, Thanh Huong, and Jean-Franois Gajewski, 2007, An experimental study of trading volume and divergence of expectations in relation to earnings announcement, CIRANO Working Papers 2007s-24 CIRANO.

Duarte, Fernando, Juan Navarro-Staicos, and Carlo Rosa, 2014, What can we learn from prior periods of low volatility?, *Liberty Street Economics*.

Dumas, Bernard, 1989, Two-person dynamic equilibrium in the capital market, *Review of Financial Studies* pp. 157–188.

Favara, Giovanni, and Zheng Song, 2014, House price dynamics with dispersed information, *Journal of Economic Theory* 149, 350–382.

Goetzmann, William N., and Massimo Massa, 2005, Dispersion of opinion and stock returns, *Journal of Financial Markets* 8, 324–349.

Gollier, Christian, and Richard Zeckhauser, 2005, Aggregation of heterogeneous time preferences, *Journal of Political Economy* 113, 878–896.

Harris, M, and A Raviv, 1993, Differences of opinion make a horse race, *Rev. Financ. Stud.* 6, 473–506.

Hirshleifer, David, 2001, Investor psychology and asset pricing, *Journal of Finance* 56, 1533–1597.

——— , and Tyler Shumway, 2003, Good day sunshine: Stock returns and the weather, *The Journal of Finance* 58, 1009–1032.

Hoffmann, Arvid O.I., Thomas Post, and Joost M.E. Pennings, 2013, Individual investor perceptions and behavior during the financial crisis, *Journal of Banking & Finance* 37, 60–74.

Hong, Harrison, Jeffrey D. Kubik, and Amit Solomon, 2000, Security analysts' career concerns and herding of earnings forecasts, *RAND Journal of Economics* 31, 121–144.

Hong, Harrison, Jeffrey D. Kubik, and Jeremy C. Stein, 2004, Social interaction and Stock-Market participation, *The Journal of Finance* 59, 137–163.

Hong, Harrison, and Jeremy C. Stein, 2007, Disagreement and the stock market, *Journal of Economic Perspectives* 21, 109–128.

Hong, Harrison, and Jialin Yu, 2009, Gone fishin': Seasonality in trading activity and asset prices, *Journal of Financial Markets* 12, 672–702.

Jiang, Wang, 1996, The term structure of interest rates in a pure exchange economy with heterogeneous investors, *Journal of Financial Economics* 41, 75–110.

Kandel, Eugene, and Neil D Pearson, 1995, Differential interpretation of public signals and trade in speculative markets, *Journal of Political Economy* 103, 831–72.

Karpoff, Jonathan M, 1986, A theory of trading volume, *Journal of Finance* 41, 1069–87.

Kaustia, Markku, and Samuli Knüpfer, 2008, Do investors overweight personal experience? evidence from ipo subscriptions, *The Journal of Finance* 63, 2679–2702.

Kelley, Eric K., and Paul C. Tetlock, 2013, How wise are crowds? insights from retail orders and stock returns, *Journal of Finance* 68, 1229–1265.

Lang, Larry H.P., and Robert H. Litzenberger, 1989, Trading volume and changes in heterogeneous expectations, .

Lettau, Martin, and Sydney Ludvigson, 2001, Resurrecting the (C)CAPM: A Cross-Sectional Test When Risk Premia Are Time-Varying, *Journal of Political Economy* 109, 1238–1287.

Llorente, Guillermo, Roni Michaely, Gideon Saar, and Jiang Wang, 2002, Dynamic volume-return relation of individual stocks, *Review of Financial Studies* 15, pp. 1005–1047.

Loh, Roger K., and René M. Stulz, 2014, Is sell-side research more valuable in bad times, *NBER Working Paper 19778*.

Mankiw, Gregory, and Steven Zeldes, 1991, The consumption of stockholders and nonstockholders, *Journal of Financial Economics* 29, 97–112.

Milgrom, Paul, and Nancy Stokey, 1982, Information, trade and common knowledge, *Journal of Economic Theory* 26, 17–27.

Neumark, David, and Steven A Sharpe, 1992, Market structure and the nature of price rigidity: Evidence from the market for consumer deposits, *The Quarterly Journal of Economics* 107, 657–80.

Odean, Terrance, 1999, Do investors trade too much?, *American Economic Review* 89, 1279–1298.

Pastor, Lubos, and Robert F. Stambaugh, 2003, Liquidity risk and expected stock returns, *Journal of Political Economy* 111, 642–685.

Patton, Andrew J., and Allan Timmermann, 2010, Why do forecasters disagree? lessons from the term structure of cross-sectional dispersion, *Journal of Monetary Economics* 57, 803 – 820.

Petkova, Ralitsa, and Lu Zhang, 2005, Is value riskier than growth?, *Journal of Financial Economics* 78, 187 – 202.

Pinto, Eugenio, 2010, Forecasting investment with earnings expectations, *unpublished memo.*

Samuelson, William, and Richard Zeckhauser, 1988, Status quo bias in decision making, *Journal of Risk and Uncertainty* 1, 7–59.

Scheinkman, José, and Wei Xiong, 2002, Overconfidence, short-sale constraints, and bubbles, *Princeton Economic Theory Working Papers.*

Souleles, Nicholas S., 2004, Expectations, heterogeneous forecast errors, and consumption: Micro evidence from the michigan consumer sentiment surveys, *Journal of Money, Credit and Banking* 36, pp. 39–72.

Statman, Meir, Steven Thorley, and Keith Vorkink, 2006, Investor overconfidence and trading volume, *Rev. Financ. Stud.* 19, 1531–1565.

Tirole, Jean, 1982, On the possibility of speculation under rational expectations, *Econometrica* 50, 1163–81.

Varian, Hal R, 1985, Divergence of opinion in complete markets: A note, *Journal of Finance* 40, 309–17.

Vissing-Jørgensen, Annette, 2002, Limited stock market participation and the elasticity of intertemporal substitution, *Journal of Political Economy* 110, 825–853.

Wang, Jiang, 1994, A model of competitive stock trading volume, *Journal of Political Economy* 102, 127–168.

Warther, Vincent A., 1995, Aggregate mutual fund flows and security returns, *Journal of Financial Economics* 39, 209 – 235.

Ziebart, David A., 1990, The association between consensus of beliefs and trading activity surrounding earnings announcements., *Accounting Review* 65, 477–488.

**Table 1:** Description of Expectation Variables in SCA, SPF, and Bluechip

| Variable | Description |
|---|---|
| <td colspan="2" align="center">Surveys of Consumers</td> |
| $PEXP$ | Now looking ahead—do you think that a year from now you (and your family living there) will be better off financially, or worse off, or just about the same as now? |
| $BEXP$ | A year from now, do you expect that in the country as a whole business conditions will be better, or worse than they are at present, or just about the same? |
| $BUS5$ | Looking ahead, which would you say is more likely—that in the country as a whole well have continuous good times during the next five years or so, or that we will have periods of widespread unemployment or depression, or what? |
| $UNEMP$ | How about people out of work during the coming 12 months—do you think that there will be more unemployment than now, about the same, or less? |
| $RATEX$ | Do you think will happen to interest rates for borrowing money during the next 12 months—will they go up, stay the same, or go down? |
| <td colspan="2" align="center">Survey of Professional Forecasters</td> |
| GDP growth | log(ngdp5) - log(ngdp1), log difference between the GDP level projection of the quarter of the survey and the quarter one year later. |
| Industrial production growth | log(indprod5) - log(indprod1), log difference between the industrial production level projection of the quarter of the survey and the quarter one year later. |
| Corporate profit growth | log(cprof5) - log(cprof1), log difference between the corporate profit level projection of the quarter of the survey and the quarter one year later. |
| Unemployment | (unemp2 + unemp3 + unemp4 + unemp5) / 4, average of the projected unemployment over the next four quarters. |
| <td colspan="2" align="center">Blue Chip Economic Indicator Survey</td> |
| GDP growth | Projected annual GDP growth |
| Industrial production growth | Projected annual industrial production growth |
| Investment growth | Projected annual non-residential investment growth |
| Unemployment | Projected unemployment |
| Short-term interest rate | Projected 3-month Treasury bills rate |
| Longer-term interest rate | Projected 10-year Treasury notes rate over the next year |

**Table 2:** Overview of Dispersion Measures in Four Data Sources

The upper panel presents pair-wise correlations for all four belief dispersion measures. The "SCA Household Disp" refers to belief dispersion among households surveyed in the Reuters/University of Michigan Surveys of Consumers. The "SPF Professional Disp" refers to belief dispersion among professional forecasters surveyed in the Survey of Professional Forecasters, "Blue-Chip Disp" refers to belief dispersion among respondents surveyed by the Blue Chip Economic Indicators, and "IBES Analyst Disp" refers to the series of corporate earnings belief dispersion among professional analysts that was constructed by Pinto (2010). All measures are monthly time series. The "SPF Professional Disp" is interpolated from quarterly surveys. The lower panel of the table presents standard deviations and sample period of the four series.

| Variables | SCA Household Disp | SPF Disp | IBES Analyst Disp | Blue-Chip Disp |
|---|---|---|---|---|
| SCA Household Disp | 1.00 | | | |
| SPF Disp | 0.63 | 1.00 | | |
| IBES Analyst Disp | 0.31 | 0.33 | 1.00 | |
| Blue-Chip Disp | 0.20 | 0.55 | 0.56 | 1.00 |

| | Std. Dev. (pcts) | Sample Period | N |
|---|---|---|---|
| SCA Household Disp | 1.55 | Jan. 1978–Dec. 2011 | 408 |
| SPF Professional Disp | 1.63 | Jan. 1978–Dec. 2011 | 408 |
| IBES Analyst Disp | 1.42 | Jan. 1978–Dec. 2011 | 408 |
| Blue-Chip Disp | 1.81 | Jul. 1984–Dec. 2011 | 330 |

**Table 3:** Turnovers and Belief Dispersion

The table reports turnover's responses to the belief dispersion among households surveyed in SCA and among professional forecasters in SPF, Blue-Chip, and IBES. The model is described in Equation (3). The dependent variable $Turnover$ is measured monthly and is quoted in percentage points. $Turnover$ is also trend-adjusted using cubic detrending. Independent variables are described in Section 5.1. Model 1.a through Model 1.d include one of the four belief dispersion measures. Model 2 through Model 4 add professional dispersion measures to household dispersion measures cumulatively. Numbers in parentheses are Newey-West adjusted standard errors. ***, ** and * denote statistical significance at 1%, 5% and 10% levels, respectively.

| | Model 1.a | Model 1.b | Model 1.c | Model 1.d | Model 2 | Model 3 | Model 4 |
|---|---|---|---|---|---|---|---|
| SCA Household Disp | 0.421*** | | | | 0.413*** | 0.470*** | 0.463*** |
| | (0.116) | | | | (0.114) | (0.125) | (0.124) |
| IBES Disp | | 0.180** | | | 0.163* | 0.207 | 0.153 |
| | | (0.088) | | | (0.088) | (0.173) | (0.176) |
| Blue-Chip Disp | | | 0.218*** | | | 0.162** | 0.148* |
| | | | (0.065) | | | (0.079) | (0.081) |
| SPF Disp | | | | 0.196** | | | 0.135 |
| | | | | (0.081) | | | (0.158) |
| | | | | | | | |
| Lag Turnover | 0.488*** | 0.505*** | 0.482*** | 0.501*** | 0.467*** | 0.404*** | 0.399*** |
| | (0.085) | (0.085) | (0.097) | (0.088) | (0.083) | (0.090) | (0.091) |
| Mean Expectation | 0.038*** | 0.010 | 0.011 | 0.021*** | 0.037*** | 0.054*** | 0.053*** |
| | (0.010) | (0.006) | (0.010) | (0.008) | (0.010) | (0.014) | (0.014) |
| S&P Return | 0.414*** | 0.347** | 0.321* | 0.337** | 0.382** | 0.335* | 0.327* |
| | (0.149) | (0.152) | (0.184) | (0.154) | (0.151) | (0.177) | (0.178) |
| S&P Volatility | 7.381*** | 6.967** | 7.421** | 7.272** | 6.816** | 7.071** | 6.848** |
| | (2.795) | (3.063) | (3.374) | (2.941) | (2.847) | (3.140) | (3.096) |
| Stock Liquidity | −3.054 | −3.772* | −3.528 | −3.129 | −3.311 | −3.032 | −3.103 |
| | (2.071) | (2.189) | (2.331) | (2.161) | (2.049) | (2.150) | (2.134) |
| Unemployment Rate | −0.103 | −0.028 | −0.053 | −0.039 | −0.170** | −0.236** | −0.258** |
| | (0.075) | (0.070) | (0.111) | (0.070) | (0.082) | (0.119) | (0.124) |
| Pre 2007 | −0.829* | 0.028 | −0.535 | −0.322 | −0.846* | −1.578*** | −1.553*** |
| | (0.441) | (0.385) | (0.430) | (0.417) | (0.435) | (0.501) | (0.490) |
| | | | | | | | |
| Constant | −3.793*** | −3.511*** | −2.237 | −3.091*** | −4.068*** | −4.778** | −4.125** |
| | (1.010) | (1.039) | (1.628) | (1.012) | (1.018) | (1.930) | (2.045) |
| Monthly Fixed-effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Adj. R-Squared | 0.566 | 0.544 | 0.558 | 0.544 | 0.570 | 0.594 | 0.593 |
| N | 407 | 407 | 330 | 407 | 407 | 330 | 330 |

**Table 4:** Mutual Fund Flow and Belief Dispersion

The table reports stock mutual fund flow's responses to the belief dispersion among households surveyed in SCA and among professional forecasters in SPF, Blue-Chip, and IBES. The model is described in Equation (3). The dependent variable is mutual fund gross flow as a percentage of total asset under management by stock mutual funds, and is quoted in percentage points. Independent variables are described in Section 5.1. Model 1.a through Model 1.d include one of the four belief dispersion measures. Model 2 through Model 4 add professional dispersion measures to household dispersion measures cumulatively. Numbers in parentheses are Newey-West adjusted standard errors. ***, ** and * denote statistical significance at 1%, 5% and 10% levels, respectively.

|  | Model 1.a | Model 1.b | Model 1.c | Model 1.d | Model 2 | Model 3 | Model 4 |
|---|---|---|---|---|---|---|---|
| SCA Household Disp | 0.057** | | | | 0.071** | 0.074*** | 0.084*** |
|  | (0.026) | | | | (0.028) | (0.028) | (0.028) |
| IBES Disp | | −0.065** | | | −0.079** | −0.049 | 0.005 |
|  | | (0.032) | | | (0.034) | (0.039) | (0.046) |
| Blue-Chip Disp | | | −0.052*** | | | −0.037 | −0.026 |
|  | | | (0.020) | | | (0.023) | (0.021) |
| SPF Disp | | | | −0.115*** | | | −0.121** |
|  | | | | (0.039) | | | (0.051) |
|  | | | | | | | |
| Lag Fund Flow | 0.530*** | 0.543*** | 0.501*** | 0.520*** | 0.530*** | 0.494*** | 0.473*** |
|  | (0.066) | (0.063) | (0.064) | (0.063) | (0.061) | (0.061) | (0.062) |
| Mean Expectation | 0.006* | −0.000 | 0.004 | 0.001 | 0.004 | 0.007** | 0.009*** |
|  | (0.003) | (0.003) | (0.003) | (0.003) | (0.003) | (0.003) | (0.003) |
| S&P Return | 0.007 | 0.035 | 0.020 | 0.042 | 0.040 | 0.029 | 0.036 |
|  | (0.083) | (0.076) | (0.080) | (0.071) | (0.076) | (0.078) | (0.073) |
| S&P Volatility | 2.003*** | 2.380*** | 2.486*** | 2.696*** | 2.355*** | 2.518*** | 2.831*** |
|  | (0.565) | (0.577) | (0.567) | (0.576) | (0.545) | (0.538) | (0.538) |
| Stock Liquidity | −0.465 | −0.477 | −0.476 | −0.395 | −0.413 | −0.400 | −0.309 |
|  | (0.537) | (0.557) | (0.553) | (0.557) | (0.550) | (0.550) | (0.547) |
| Unemployment Rate | 0.008 | 0.055** | 0.069*** | 0.075*** | 0.040 | 0.061** | 0.080*** |
|  | (0.022) | (0.025) | (0.022) | (0.026) | (0.025) | (0.024) | (0.026) |
| Pre 2007 | −0.099 | 0.099 | 0.163* | 0.096 | −0.018 | 0.045 | 0.015 |
|  | (0.107) | (0.088) | (0.089) | (0.086) | (0.100) | (0.101) | (0.098) |
|  | | | | | | | |
| Constant | 1.435*** | 1.737*** | 1.057*** | 1.062*** | 1.729*** | 1.262*** | 0.662 |
|  | (0.364) | (0.377) | (0.359) | (0.374) | (0.384) | (0.423) | (0.491) |
| Monthly Fixed-effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Adj. R-Squared | 0.596 | 0.597 | 0.600 | 0.605 | 0.605 | 0.609 | 0.616 |
| N | 335 | 335 | 330 | 335 | 335 | 330 | 330 |

**Table 5:** Robustness of Baseline Results

This table reports coefficient $\beta$ as in Equation (3) estimated from variants of the baseline model. $\beta$ measures the sensitivity of stock market turnover to households' belief dispersion. Column titles "Model 1.a", "Model 2","Model 3","Model 4" correspond to model specifications in identically titled columns in Table 3. Where "Model 1.a" includes only "SCA Household Disp", and Model 2 through 4 add professional dispersion measures from SPF, Blue-Chip, and IBES cumulatively. The dependent variable $Turnover$ is measured monthly and is quoted in percentage points. In the baseline model, the dependent variable is trend-adjusted using cubic detrending, and the household belief dispersion measure is calculated using the $WNHI$ as defined in Equation (2) where the weights for answers of "better off", "about the same" and "worse off" are 1, 2 and 1, respectively. Details about the $WNHI$ measure are described in Section 4.1. The upper part of this table deviates from the baseline model by changing the way the dependent variable ($Turnover$) is calculated. The lower part of the table deviates from the baseline model by changing the way household dispersion variable is calculated. Numbers in parentheses are Newey-West adjusted standard errors. ***, ** and * denote statistical significance at 1%, 5% and 10% levels, respectively.

| Variation Type | Model 1.a | Model 2 | Model 3 | Model 4 |
|---|---|---|---|---|
| Baseline | 0.421*** | 0.413*** | 0.470*** | 0.463*** |
| (cubic detrending, $WNHI, w = (1, 2, 1)$) | (0.116) | (0.114) | (0.125) | (0.124) |
| **Additional controls of business cycle effects, lags of returns and P/D ratios** | | | | |
| (1) baseline + Contemporaneous GDP growth, | 0.435*** | 0.436*** | 0.493*** | 0.491*** |
| unemployment rates, and P/D ratios | (0.112) | (0.107) | (0.127) | (0.127) |
| (2) (1) + Quarter-lagged GDP growth, | 0.409*** | 0.424*** | 0.450*** | 0.451*** |
| unemployment rates, S&P returns, | (0.110) | (0.106) | (0.128) | (0.128) |
| and P/D ratios | | | | |
| (3) (2) + Year-lagged GDP growth, | 0.400*** | 0.424*** | 0.460*** | 0.460*** |
| unemployment rates, S&P returns, | (0.110) | (0.107) | (0.130) | (0.131) |
| and P/D ratios | | | | |
| (4) Using residuals of projecting | 0.214** | 0.227** | 0.343*** | 0.376*** |
| dispersion measures on contemporaneous, | (0.105) | (0.103) | (0.128) | (0.127) |
| quarter-, and year-lagged GDP growth, | | | | |
| unemployment rates, S&P returns, and P/D ratios | | | | |
| **Variations in belief dispersion measures** | | | | |
| WHI  $w = (1, 1.5, 1)$ | 0.295*** | 0.266*** | 0.346*** | 0.343*** |
| std.  dev. $= 1.52$ | (0.087) | (0.091) | (0.116) | (0.116) |
| WHI  $w = (1, 1, 1)$ | 0.261*** | 0.228** | 0.270** | 0.264* |
| std.  dev. $= 1.41$ | (0.086) | (0.093) | (0.137) | (0.138) |
| Standard Deviation | 0.122 | 0.082 | 0.259** | 0.253** |
| std.  dev. $= 1.42$ | (0.082) | (0.081) | (0.109) | (0.109) |
| WHI  de-meaned | 0.234*** | 0.205** | 0.288*** | 0.285*** |
| std.  dev. $= 1.41$ | (0.084) | (0.085) | (0.108) | (0.108) |
| **Variations in sample period and dependent variable** | | | | |
| Using pre-2007 sample | 0.129*** | 0.121** | 0.154** | 0.162** |
| | (0.049) | (0.049) | (0.063) | (0.063) |
| Linear Detrending | 0.426*** | 0.447*** | 0.230* | 0.174 |
| | (0.106) | (0.109) | (0.123) | (0.120) |
| Quadratic Detrending | 0.359*** | 0.261*** | 0.324*** | 0.312*** |
| | (0.092) | (0.096) | (0.120) | (0.120) |
| Baker-Stein Detrending | 0.311*** | 0.352*** | 0.458*** | 0.477*** |
| | (0.110) | (0.115) | (0.139) | (0.138) |
| Excluding NASDAQ | 0.293** | 0.242* | 0.343** | 0.350** |
| | (0.120) | (0.124) | (0.150) | (0.151) |

**Table 6:** Turnover Sensitivities To Belief Dispersion Among Demographic Groups

This table reports coefficient $\beta$ as in Equation (3) estimated from different subsamples of the data based on demographic information. $\beta$ measures the sensitivity of stock market turnover to households' belief dispersion. Column titles "Model 1.a", "Model 2","Model 3","Model 4" correspond to model specifications in identically titled columns in Table 3. "Model 1.a" includes only "SCA Household Disp", and Model 2 through 4 add professional dispersion measures from SPF, Blue-Chip, and IBES cumulatively. The dependent variable $Turnover$ is measured monthly and is quoted in percentage points. Numbers in parentheses are Newey-West adjusted standard errors. ***, ** and * denote statistical significance at 1%, 5% and 10% levels, respectively.

| Subgroups | Std. dev. $WHI$ | Model 1.a | Model 2 | Model 3 | Model 4 |
|---|---|---|---|---|---|
| by age | | | | | |
| Age < 35 | 1.42 | -0.033 | -0.022 | -0.039 | -0.031 |
| | | (0.099) | (0.100) | (0.106) | (0.108) |
| Age $\geq$ 35 | 1.45 | 0.397*** | 0.373*** | 0.441*** | 0.430*** |
| | | (0.106) | (0.106) | (0.112) | (0.112) |
| $p$-value < 0.1 | | yes | yes | yes | yes |
| by education | | | | | |
| High school and below | 1.51 | 0.032 | 0.037 | 0.055 | 0.048 |
| | | (0.099) | (0.100) | (0.122) | (0.123) |
| Some college and above | 1.49 | 0.432*** | 0.422*** | 0.423** | 0.423** |
| | | (0.146) | (0.146) | (0.176) | (0.175) |
| $p$-value < 0.1 | | yes | yes | no | no |
| by income | | | | | |
| Below median | 1.47 | 0.030 | 0.045 | 0.097 | 0.098 |
| | | (0.089) | (0.091) | (0.105) | (0.105) |
| Above median | 1.45 | 0.404*** | 0.374*** | 0.382*** | 0.376*** |
| | | (0.123) | (0.127) | (0.137) | (0.137) |
| $p$-value < 0.1 | | yes | yes | no | no |
| by race | | | | | |
| Black | 1.22 | 0.141* | 0.162** | 0.140 | 0.145 |
| | | (0.072) | (0.073) | (0.092) | (0.091) |
| White | 1.53 | 0.233** | 0.225** | 0.300*** | 0.298*** |
| | | (0.093) | (0.090) | (0.114) | (0.112) |
| $p$-value < 0.1 | | no | no | no | no |
| by likelihood to hold stocks | | | | | |
| Below median | 1.52 | 0.067 | 0.076 | 0.154 | 0.154 |
| | | (0.101) | (0.101) | (0.114) | (0.114) |
| Above median | 1.47 | 0.406*** | 0.383*** | 0.338** | 0.330** |
| | | (0.123) | (0.127) | (0.143) | (0.142) |
| $p$-value < 0.1 | | yes | yes | no | no |

**Table 7:** Turnover Sensitivities To Belief Dispersion Before and After Index Addition and Deletion

This table documents results from our event study of firm-specific stock turnover's sensitivity to aggregate household belief dispersion before and after the stock's inclusion into or deletion from the S&P 500 index. The "Event" is index addition for column (1) through (4), and index deletion for column (5) through (8). Stock turnover 6 months before and 6 months after index inclusion/deletion events are considered. The month of the event is removed from the sample. Column titles "Model 1.a", "Model 2", "Model 3", "Model 4" correspond to model specifications in identically titled columns in Table 3. Where "Model 1.a" includes only "SCA Household Disp", and Model 2 through 4 add professional dispersion measures to household dispersion measures cumulatively. *PostEvent* equals one if the month of trading is after the month when the stock is either added into or deleted from the S&P 500 index, and zero otherwise. The index composition history file is from CRSP. Firm fixed effects are included in all regressions. The dependent variable *Turnover* is measured monthly and is quoted in percentage points. Numbers in parentheses are standard errors clustered at firm level. ***, ** and * denote statistical significance at 1%, 5% and 10% levels, respectively. Other control variables are not listed in the table.

| | Addition | | | | Deletion | | | |
|---|---|---|---|---|---|---|---|---|
| | Model 1.a | Model 2 | Model 3 | Model 4 | Model 1.a | Model 2 | Model 3 | Model 4 |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| SCA Household Disp | 0.077*** | 0.074*** | 0.078*** | 0.085*** | 0.120** | 0.136*** | 0.136*** | 0.143*** |
| | (0.023) | (0.023) | (0.022) | (0.026) | (0.051) | (0.049) | (0.047) | (0.053) |
| Post Event | 0.003 | −0.004 | 0.001 | 0.044 | −0.030 | 0.012 | 0.012 | 0.026 |
| | (0.032) | (0.033) | (0.031) | (0.042) | (0.133) | (0.131) | (0.131) | (0.162) |
| Post Event * Household Disp | 0.044** | 0.037** | 0.039** | 0.050* | 0.020 | 0.056 | 0.056 | 0.058 |
| | (0.019) | (0.018) | (0.018) | (0.028) | (0.082) | (0.083) | (0.082) | (0.105) |
| IBES Analyst Disp | | 0.047* | 0.074* | −0.017 | | −0.251*** | −0.252*** | −0.266** |
| | | (0.027) | (0.038) | (0.034) | | (0.077) | (0.093) | (0.108) |
| SPF Professional Disp | | | −0.072 | 0.075* | | | 0.003 | 0.048 |
| | | | (0.056) | (0.041) | | | (0.101) | (0.146) |
| Blue-Chip Disp | | | | −0.012 | | | | −0.046 |
| | | | | (0.014) | | | | (0.048) |
| Constant | 1.283*** | 0.959** | 0.876* | 1.576*** | 1.063 | 2.800*** | 2.808*** | 2.926** |
| | (0.333) | (0.431) | (0.452) | (0.431) | (0.844) | (1.017) | (0.998) | (1.152) |
| Month Effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Firm Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Adj. R-Square | 0.847 | 0.847 | 0.848 | 0.850 | 0.591 | 0.593 | 0.593 | 0.583 |
| N | 7,974 | 7,974 | 7,974 | 6,875 | 2,573 | 2,573 | 2,573 | 2,365 |

**Table 8:** Turnovers and Belief Dispersion, by Market Cap and Book-to-Market Ratio Tercile

This table compares the sensitivity of turnover of stocks in different terciles based on either market cap or book-to-market to the belief dispersion among households. The model is described in Equation (3). The dependent variable *Turnover* is measured monthly and is quoted in percentage points. It is also trend adjusted using cubic detrending. Independent variables are described in Section 5.1. Large, medium, and small cap are defined as the top 33%, 33%-66% and bottom 33% stocks in terms of the market capitalization distribution as of the end of the month. High B2M, medium B2M, and low B2M are defined as the top 33%, 33%-66% and bottom 33% stocks in terms of the book-to-market ratio distributions as of the end of the month. Numbers in parentheses are Newey-West adjusted standard errors. ***, ** and * denote statistical significance at 1%, 5% and 10% levels, respectively.

| | All | Large Cap | Medium Cap | Small Cap | High B2M | Medium B2M | Low B2M |
|---|---|---|---|---|---|---|---|
| SCA Household Disp | 0.415*** | 0.404*** | 0.145 | 0.098 | 0.417*** | 0.157* | 0.160* |
| | (0.097) | (0.125) | (0.099) | (0.088) | (0.157) | (0.094) | (0.092) |
| IBES Analyst Disp | 0.106 | 0.138 | 0.240 | 0.185 | 0.438* | 0.165 | −0.253* |
| | (0.149) | (0.195) | (0.148) | (0.130) | (0.245) | (0.146) | (0.141) |
| Blue-Chip Disp | 0.133* | 0.200* | −0.060 | −0.180** | 0.099 | 0.133* | 0.259*** |
| | (0.081) | (0.105) | (0.081) | (0.074) | (0.131) | (0.079) | (0.078) |
| SPF Professional Disp | 0.021 | 0.237 | −0.047 | 0.098 | 0.464* | 0.290* | 0.157 |
| | (0.168) | (0.221) | (0.166) | (0.149) | (0.280) | (0.168) | (0.162) |
| | | | | | | | |
| Lagged Turnover | 0.420*** | 0.449*** | 0.611*** | 0.707*** | 0.569*** | 0.364*** | 0.303*** |
| | (0.048) | (0.047) | (0.041) | (0.034) | (0.043) | (0.051) | (0.051) |
| Mean Expectation | 0.054*** | 0.040** | 0.038*** | 0.039*** | 0.037* | 0.024* | 0.005 |
| | (0.014) | (0.017) | (0.013) | (0.012) | (0.022) | (0.013) | (0.013) |
| S&P Return | 0.320 | 0.108 | 0.910*** | 1.161*** | 0.630* | 0.219 | 0.052 |
| | (0.212) | (0.275) | (0.233) | (0.220) | (0.357) | (0.207) | (0.201) |
| S&P Volatility | 7.675*** | 10.774*** | 0.256 | −0.778 | 6.765*** | 7.332*** | 7.553*** |
| | (1.396) | (1.832) | (1.364) | (1.264) | (2.175) | (1.351) | (1.313) |
| Stock liquidity | −3.314** | −4.331** | −3.093* | 0.163 | −3.881 | −2.859* | −2.410 |
| | (1.608) | (2.078) | (1.816) | (1.703) | (2.705) | (1.560) | (1.524) |
| Pre 2007 | −1.076*** | −2.319*** | −1.223*** | −0.366 | −2.839*** | −2.259*** | −1.931*** |
| | (0.393) | (0.536) | (0.397) | (0.338) | (0.687) | (0.414) | (0.388) |
| | | | | | | | |
| Constant | −6.134*** | −4.446** | −4.093*** | −3.437** | −4.497* | −2.470 | 0.915 |
| | (1.619) | (2.053) | (1.560) | (1.398) | (2.557) | (1.523) | (1.480) |
| Month Effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Adj. R-Square | 0.586 | 0.705 | 0.482 | 0.615 | 0.752 | 0.683 | 0.570 |
| N | 330 | 330 | 330 | 330 | 330 | 330 | 330 |

**Table 9:** Turnovers and Dispersion of Household Belief Changes

The table reports turnover's responses to dispersion of changes in beliefs among SCA respondents. The model is similar to Equation (3), except that belief dispersion variable $WNHI$ is replace with dispersion of belief change variable $DBC$. The definition and calculation of $DBC$ is described in Section 4.2. The dependent variable $Turnover$ is quoted in percentage points and trend-adjusted using cubic detrending. In column (1), the dependent variable is the turnover at the last month of the belief change window. In column (2), the dependent variable is the total turnover over the six months window when belief change happened. Other independent variables are described in Section 5.1. Numbers in parentheses are Newey-West adjusted standard errors. ***, ** and * denote statistical significance at 1%, 5% and 10% levels, respectively.

|  | Turnover (1) | Turnover 6 month (2) |
|---|---|---|
| Disp of Belief Changes | 0.167*** | 1.135*** |
|  | (0.062) | (0.359) |
|  |  |  |
| Lag Turnover | 0.509*** | 0.520*** |
|  | (0.041) | (0.055) |
| Mean Expectation | 0.016** | 0.014* |
|  | (0.007) | (0.008) |
| S&P Return | 0.388** | 5.386** |
|  | (0.174) | (2.646) |
| S&P Volatility | 7.698*** | 11.702*** |
|  | (1.244) | (1.974) |
| Stock liquidity | -3.456** | -5.085 |
|  | (1.490) | (3.719) |
| Pre 2007 | -0.200 | 1.161 |
|  | (0.278) | (1.859) |
| Constant | -3.189*** | -28.088*** |
|  | (0.664) | (5.102) |
|  |  |  |
| Month Effects | Yes | Yes |
| Adj. R-Square | 0.546 | 0.657 |
| N | 401 | 396 |

**Figure 1:** Stock Market Turnover and Aggregate Flows to Stock Market Mutual Funds

This figure plots monthly time series of our proxies for household trading activities in the US stock market. The top panel shows the turnover rates, the middle panel shows the turnover rates after cubic detrending and the bottom panel shows the aggregate flows to stock market mutual funds. All three variables are in fractions. Turnover rate is defined as the combined number of shares traded in NYSE, AMEX and NASDAQ in a given month divided by the average total number of shares outstanding during the same month. Mutual fund flow is defined as the sum of outflow and inflow as a fraction of total asset under management by equity mutual funds in the same month. The date range is from 1978 to 2011 for turnover rates and from 1984 to 2011 for total fund flow variable. Shaded areas correspond to NBER recessions.
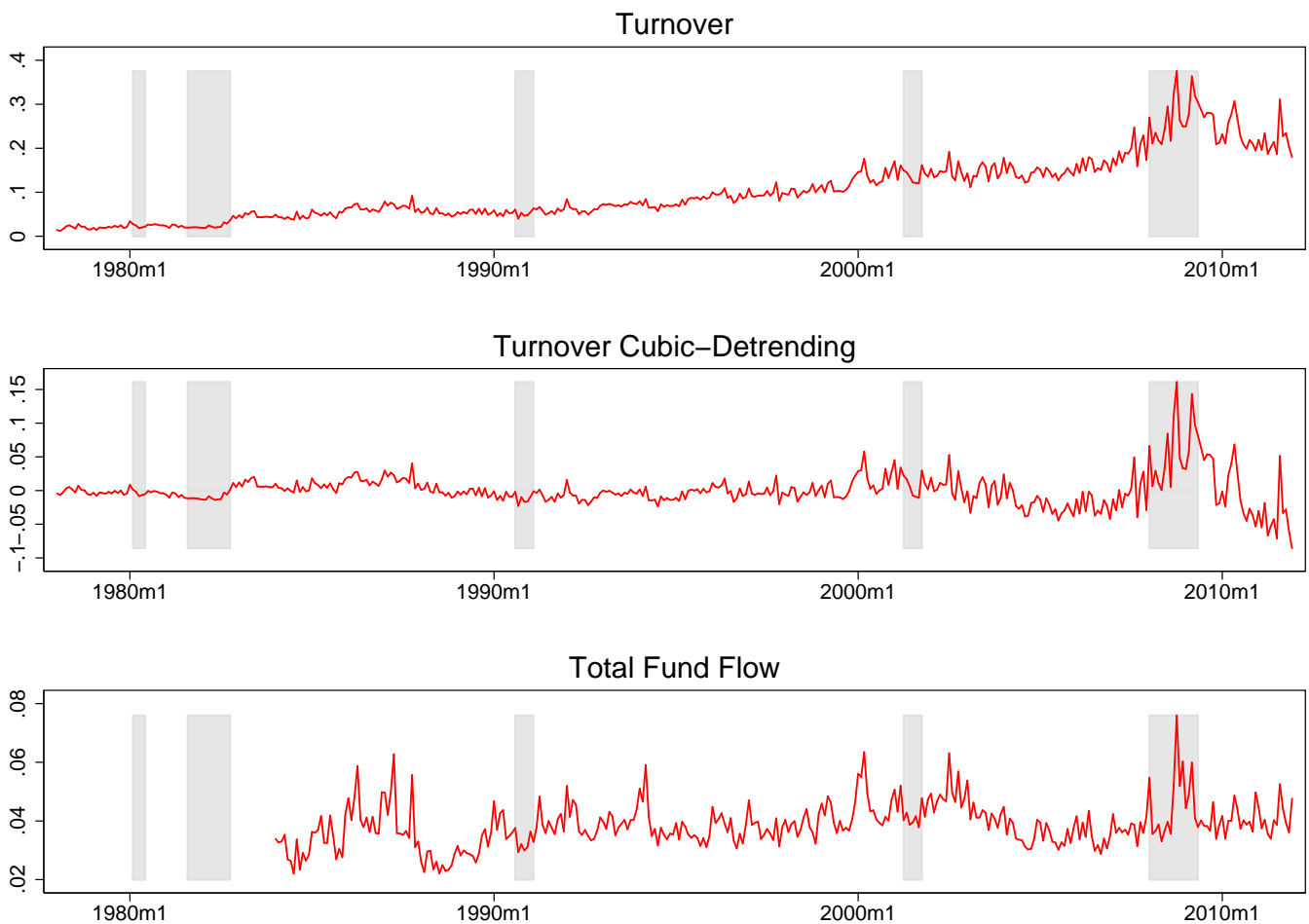
**Figure 2:** Monthly Belief Dispersion from Household Surveys

This figure plots time series of dispersion of beliefs on five expectation variables in the SCA. The five expectation variables are unemployment, interest rates, short-term business conditions (BEXP), personal financial conditions (PEXP), and long-term business conditions (BUS5). The last panel plots the first principle component of the five dispersion series. Belief dispersion is measured using weighted negative Herfindahl index (WNHI) described in Equation (2). Larger values indicate higher dispersion. Shaded areas are NBER recession periods.
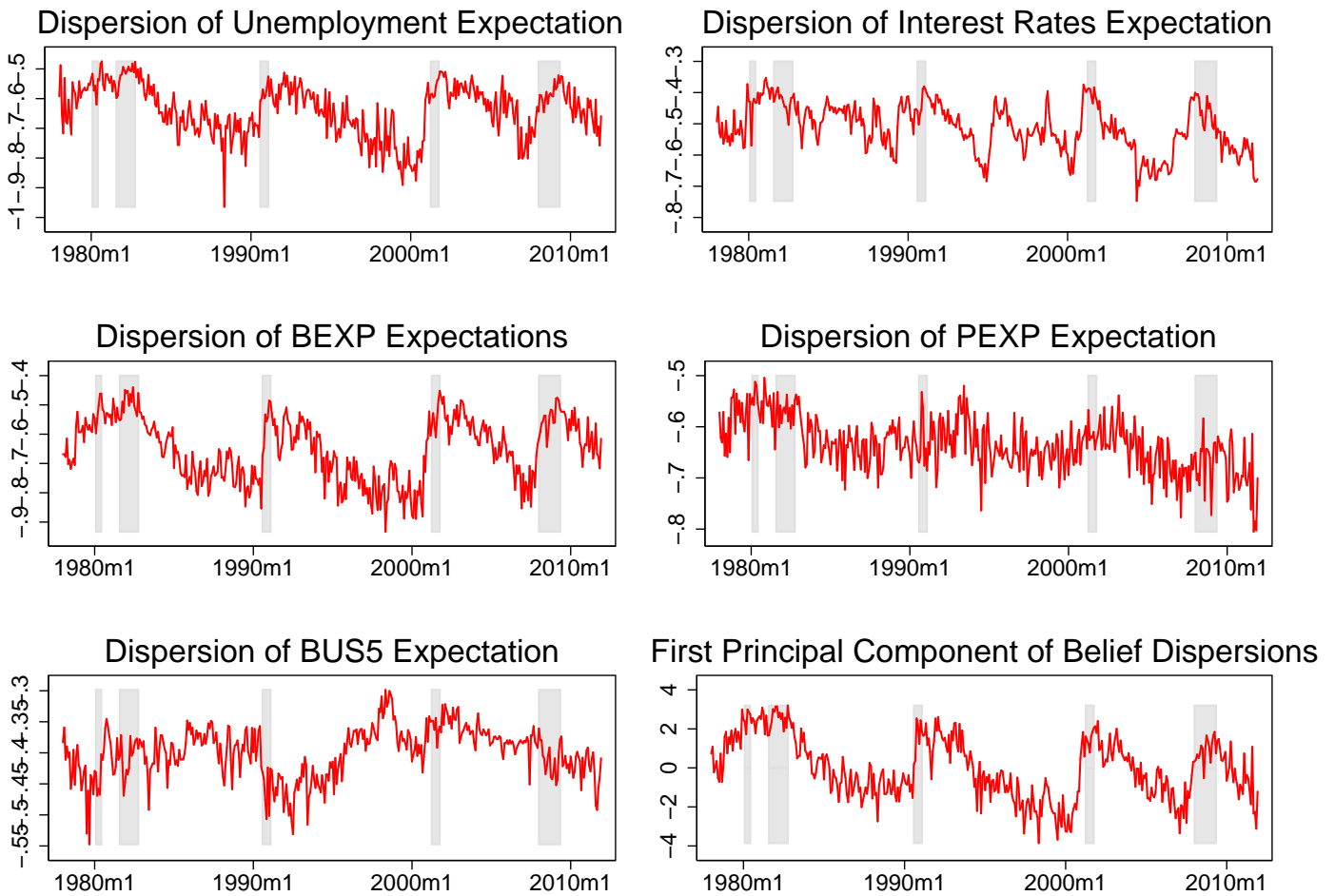
**Figure 3:** Comparison of Household Belief Dispersion Measure to Professional Belief Dispersion Measures

This figure compares belief dispersion among households in the SCA survey to the belief dispersions among professional forecasters in the SPF, IBES and Blue Chip surveys. Detailed description of the construction of these dispersion series is given in Section 4. Data frequency is monthly. The date range is from Jan. 1978 to Dec. 2011 for all time series except the Blue Chip time series, which is only available from Jul. 1984. Shaded areas correspond to NBER recessions.
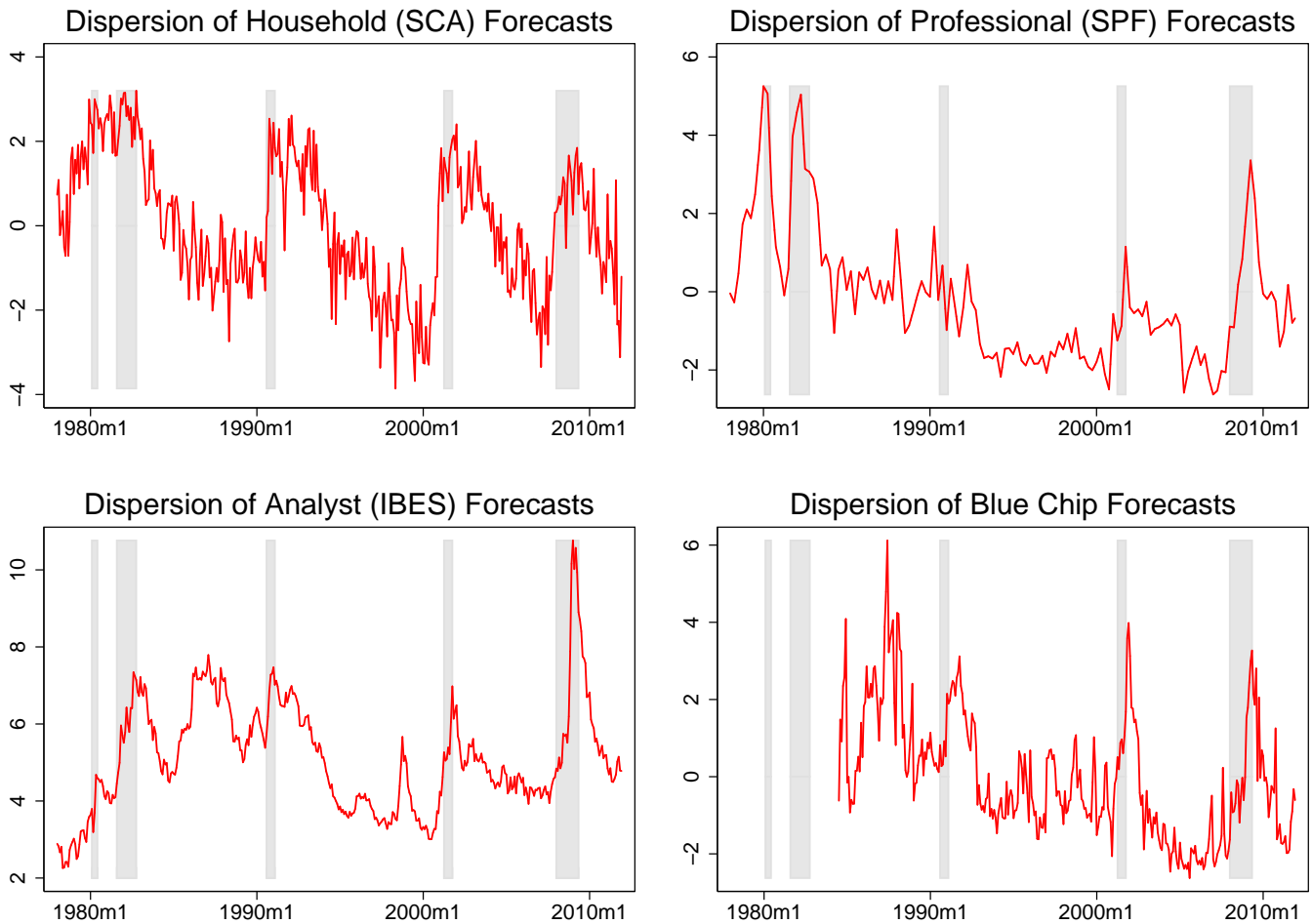
# A   Trading Volume and Dispersion of Belief Changes

Karpoff (1986) argues that, beside belief dispersion itself, dispersion in belief changes over a given period of time also affects trading activities. However, the theory has been tested far less frequently than the effect of the static belief dispersion dispersion. To test this hypothesis, similar to Bamber, Barron, and Stober (1997), we construct a measure of dispersion in belief changes that captures the differences in household belief dynamics.

In the SCA surveys, about 40 percent of the consumers are surveyed again six months after they were originally surveyed. We can thus track belief changes of the same household over a six-month interval.[26]  When constructing a dispersion measure for belief changes, a similar problem arises.  Recall that all of the SCA questions we study have categorical answers—whether the future of the economy will be better, worse, or about the same; whether an economic indicator, such as the interest rate and the unemployment rate, will go up, go down, or remain the same.

To measure dispersion of changes in categorial beliefs, in a similar way we measure dispersion in static beliefs, we first classify all possible belief changes into the following five types. For example, if a respondent's answer to a question six months ago was that "things will be worse", and the answer now is "things will be better", we consider the person having a "significantly improved outlook".  Other changes of outlook are defined accordingly, as listed below. To reflect the fact that different types of belief changes are naturally ordered, we also assign different weights to them. The weights are chosen so that greater weights are assigned to no change in beliefs.

- significantly improved outlook: "worse"→"better" $\omega = 1$
- moderately improved outlook: "worse"→"the same" and "the same"→ "better" $\omega = 1$
- moderately deteriorated outlook:  "the same"→"worse" and "better"→ "the same" $\omega = 3$
- significantly deteriorated outlook: "better"→"worse" $\omega = 3$
- no change "better"→"better", "the same"→"the same" and "worse"→"worse" $\omega = 9$

The measure of dispersion of belief changes (DBC) is thus defined as below,

$$DBC = -\sum_{j=1}^{5} \omega_j p_j^2,\tag{5}$$

where $j$ corresponds to each of the five possible belief change outcomes.

To estimate the effects of dispersion of belief changes on trading volume, we replace the dispersion measure $WNHI$ in Equation (3) with the measure of dispersion of belief changes, $DBC$, defined above.[27] We study the effect of $DBC$ on stock turnover in month $m + 6$ and in the six month period between $m$ and $m+6$. All control variables are adjusted accordingly

---

[26]One caveat of the data is that, for each consumer, only one observation of belief change is available because he will not be contacted again after the second interview.

[27]Assigning smaller weights to the status quo answers yields qualitatively similar but statistically less significant results.

to be consistent with the time period for which the turnover rate refers to. The results are presented in Table 9 and are broadly consistent with the predictions of Karpoff (1986), that the extent to which people's belief revisions differ from each other also affects trading volume.[28] Indeed, $\beta$-coefficients in both columns of Table 9 are sizeable and statistically significant.

---

[28]Because it is largely infeasible to construct belief revision measures using the surveys of professional analysts, we do not run a horse race as we do for belief dispersion.