On the Effects of Private Information on Volatility

Anne Opschoor[1]
Michel van der Wel[1]
Dick van Dijk[1]
Nick Taylor[2]

1 Erasmus University Rotterdam, and Tinbergen Institute;

2 Cardiff Business School, Cardiff University, UK.

On the Effects of Private Information on Volatility

Anne Opschoor   Michel van der Wel[*]   Dick van Dijk   Nick Taylor

This version: May 9, 2011[†]

Abstract

We study the impact of private information on volatility in financial markets. We develop a comprehensive framework to investigate this link while controlling for the effects of both public information (such as macroeconomic news releases) and private information on prices and the effects of public information on volatility. Using a high-frequency 30-year U.S. Treasury bond futures data set, we find that private information variables, such as order flow and bid-ask spread, are statistically and economically significant explanatory variables for volatility. Private information is more important than public information, with the effect of a shock to order flow on volatility being four times larger than the effect of a surprise in the most influential macroeconomic news announcement. Moreover, we document an interaction between public and private information effects on volatility, with the impact of order flow on volatility depending positively on the dispersion of analysts' expectations about macroeconomic announcements. Finally, we find that the effect of private information on volatility is larger during contractions than during expansions.

Keywords: Information, macroeconomic announcements, order flow, Treasury futures, heterogeneity.

JEL: G14, E44.

1 Introduction

One of the most fundamental questions in financial economics is what drives asset prices and volatility. Both quantities are believed to change due to the arrival of new information. In this context it is useful to distinguish between public and private information. Public information concerns news that becomes available to all market participants at the same point in time, for example in the form of announcements of important macroeconomic variables. As all investors will be equally informed at the same time, the arrival of public information typically causes an immediate change in asset prices and volatility.[1] Private information refers to news that is distributed asymmetrically amongst market participants. The presence of such private information may be revealed by the trading process and changes in the asset price itself, as informed investors may buy or sell the asset based on their privately held knowledge. Other investors observe the trading process and make inferences on this private information, giving rise to further price adjustments.[2]

The effect of public information on prices and volatility and the effect of private information on prices have been established for many financial markets. Andersen, Bollerslev, Diebold and Vega (2007), for example, document the link between public information releases and prices and volatility in foreign exchange, Treasury and equity markets. Based on high-frequency intraday data they demonstrate that surprises in macro announcements (i.e. the difference between the actual release and the consensus market expectation) affect the conditional mean and volatility of exchange rates, treasuries and stocks. Recent literature also documents a relation between private information and prices for equity (see, e.g., Hasbrouck (1991)), foreign exchange (see, e.g., Evans and Lyons (2008)), and the Treasury market (see, e.g., Brandt and Kavajecz (2004), Green (2004) and Pasquariello and Vega (2007)).

The goal of this paper is to examine whether private information also influences the volatility of assets. We investigate this issue empirically for the 30-year U.S. Treasury bond

---

[1]The literature on the effects of public information on asset prices and volatility dates back to French and Roll (1986) and Cutler, Poterba and Summers (1989), and includes important contributions by Ederington and Lee (1993), Berry and Howe (1994), Fleming and Remolona (1999), Balduzzi, Elton and Green (2001), Boyd, Hu and Jagannathan (2005), Anderson, Bollerslev, Diebold and Vega (2003, 2007), Faust, Rogers, Wang and Wright (2007), Bartolini, Goldberg and Sacarny (2008), and Brenner, Pasquariello and Subrahmanyam (2009), among others.

[2]This idea was first formalized by Kyle (1985) and Glosten and Milgrom (1985).

futures. We conduct our analysis in a comprehensive framework that allows us to control for the effects of both public and private information on prices and the effects of public information on volatility. Our results demonstrate that proxies for the level of private information, such as order flow and bid-ask spread, do indeed affect volatility. Specifically, a larger degree of information asymmetry increases the uncertainty surrounding Treasury futures prices. This positive relation between private information and volatility is significant in both statistical and economic terms. For example, the average effect of order flow on daily volatility is around 80 basis points. Interestingly, we find that the effect of private information on volatility is considerably larger than the effect of public information in the form of macroeconomic news announcements. Among the 24 announcements that are included in the analysis, Nonfarm Payroll Employment is found to have the strongest impact on volatility. A one standard deviation surprise in an announcement of this variable increases daily volatility by only 17 basis points, which is more than four smaller than the effect of a comparable shock in order flow. The evident implication of our findings is that the level of private information should not be ignored when modeling and forecasting volatility.

We provide two extensions to analyze the robustness of our main result that private information affects volatility. First, we find that the effect of private information depends on the state of the economy, in the sense that the effect of the bid-ask spread on volatility is higher during contractions. Second, the effect of private information on volatility depends on the heterogeneity of analysts' expectations concerning upcoming public information releases. Order flow is found to influence volatility to a greater extent in times characterized by a high level of dispersion of beliefs.

We use a high-frequency data set consisting of transaction prices and volumes for the 30-year U.S. Treasury bond futures. The Treasury futures market is highly liquid and generates an average monthly trading volume of 5.2 million contracts (based on 2009 data). The 30-year U.S. Treasury bond futures is the most actively traded long-term interest contract in the world. Trading in Treasuries with a maturity of 30-years takes place almost solely in the futures market, hence we focus on this venue. Modeling returns and volatility of treasuries has the advantage that public information is aptly captured by macroeconomic news releases and thus easily tracked and measured. The availability of analysts' forecasts for each macro announcement provides an estimate of the market consensus, which in turn can be used to

construct the surprise component of each news release.

The idea of private information in the Treasury market needs some further clarification. The conventional interpretation of private information, as typically applied to individual stocks and corporate bonds, is advance knowledge of firm-related news, concerning earnings announcements, new investment projects, or changes in management, among others. Obviously, this interpretation does not apply to treasuries. Green (2004) provides several alternatives for the interpretation of private information in the Treasury market, such as information about endowments or the interpretation of macroeconomic news due to differential information processing skills. Despite the different interpretation, we can use similar variables to proxy private information in the Treasury market as commonly used for equities. We consider two such variables in our empirical analysis. First, we use order flow or net buying pressure (defined as the difference between the volume in buyer-initiated transactions and the volume in seller-initiated transactions) as a measure that could reveal heterogeneous information to the market. Second, we use the bid-ask spread. This latter proxy is more noisy than the former as it may include other liquidity determinants beyond private information (such as trader inventory effects).

Apart from uncovering the effect of private information on volatility as such, our study also contributes to the methodology for analyzing information effects on asset returns and volatility. We develop a modeling framework that allows us to simultaneously assess the effects of public and private information on return and volatility. Specifically, we propose a model specifying the dynamics of both returns and volatility in such a way that both equations can be estimated jointly. We split volatility in two (multiplicative) components, following the Spline-GARCH model of Engle and Rangel (2008). One component describes the effects of private and public information, while the other captures short-run GARCH-type behavior. This joint modeling approach extends and improves upon the specification of Andersen, Bollerslev, Diebold and Vega (2003, 2007). In their set-up, equations for return and volatility effects are estimated separately. Thus, parameter uncertainty of the return equation is neglected when estimating the volatility equation and, consequently, the parameters of both equations are not estimated efficiently. In addition, the approach could suffer from negative fitted values of the time-varying conditional volatility. In our framework, both return and volatility equations are estimated jointly by means of (quasi) maximum

likelihood, such that the parameters are estimated efficiently. Also, by construction, our model automatically avoids negative values of the volatility.

Our paper relates and contributes to three strands of the literature on information effects on return and volatility. The first identifies private information effects on volatility through the private information effects on returns. Evans and Lyons (2008) use a variance decomposition of the conditional mean of foreign exchange returns, which implies that they establish a link between the unconditional volatility and order flow. They show that the arrival of macro news can account for more than 30% of the daily variance of exchange rates, with 70% attributed to the direct effect of news. Likewise, He, Lin, Wang and Wu (2009) start with a structural model for returns and obtain a model-implied variance of price changes. In contrast to these two papers, in our work the term volatility implies a parameterization of the conditional volatility. Jiang and Lo (2011) use a Markov-Switching framework to identify the probability of private information flow, also from the return equation, and find that volatility is high when this probability is high controlling for other liquidity (shock) variables. Second, our work relates to the analysis of returns, volatility, liquidity and order flow, as in Chordia, Sarkar and Subrahmanyam (2005). One of their main findings is that volatility shocks are informative for predicting shifts in liquidity, highlighting the interrelation between the two. We focus on the reverse relationship and extend their work by investigating specifically the role of private information on volatility. We take order flow as the private information measure to be able to rely on economic theory that predicts causality from order flow to prices (as in Kyle (1985) and Glosten and Milgrom (1985)). Alternatively, one could look at the relation between volume and volatility as in Andersen (1996). Finally, our analysis of private information effects on volatility relates to the literature on the effects of buying- and selling-pressure in options markets (see, e.g., Bollen and Whaley (2004) and Garleanu, Pedersen and Poteshman (2009)). To take out price differences due to option contract specification (such as strike and maturity) the focus in this literature is typically on volatility rather than option prices directly.

The remainder of the paper is organized as follows. Section 2 discusses the data and presents summary statistics. Section 3 describes the joint modeling framework for analyzing the effects of public and private information on returns and volatility. Corresponding empirical results are also provided in this section. Section 4 reports extensions of the main

analysis, allowing for the effects of public and private information to depend on the level of disagreement among analysts concerning upcoming public information releases or on the state of the economy. Section 5 concludes.

2 Data and summary statistics

We combine two data sets to study the public and private information sources of volatility. In the following three subsections we discuss the 30-year U.S. Treasury bond futures data, the variables that we use as proxies for the presence of private information in the market, and the public information variables as constructed from expectations and announcements of macroeconomic fundamentals. In the final subsection, we provide a preliminary idea about the effects of public and private information on the Treasury futures returns and volatility, focusing on the announcements of Nonfarm Payroll Employment.

2.1 U.S. Treasury bond futures data

We employ a high-frequency data set of intraday transaction prices and volumes of 30-year U.S. Treasury bond futures contracts over the period from January 1, 2004 until December 31, 2009. The contract initially trades on the Chicago Board of Trade (CBOT), and after the merger of the Chicago Mercantile Exchange (CME) with the CBOT in 2007 on the CME. During our sample period trading takes place both on the trading floor and electronically. Floor trading occurs in a pit through open-outcry from 8:20 a.m. to 3:00 p.m. Eastern Standard Time (EST). Electronic trading occurs through Globex from 6.30 p.m. on Sunday to 5.00 p.m. on Friday. Since volume has gradually shifted from pit trading to electronic trading we focus on the volume generated through the electronic venue. In addition, we restrict ourselves to day trading (that is, from 8:00 a.m. to 5:00 p.m.) as this is where volume concentrates. The 30-year Treasury futures trade in the March quarterly cycle, that is, contracts mature in March, June, September or December. At each point in time the next three consecutive contracts can be traded.[3] We construct a single time series of transaction prices and volumes using the most nearby contract, which is the most intensely

---

[3]The last trading day of a given contract is the seventh business day preceding the last business day of the delivery month.

traded and is a close substitute for the underlying spot instrument. We roll over to the next contract when its daily tick volume exceeds the daily tick volume of the most nearby contract. This generally occurs between five to three days before expiration of the nearest-to-maturity contract. Our data set records the timestamp (in seconds), price and volume for each transaction.[4] We aggregate the data to 5-minute intervals.[5] For computing returns we use the last observed transaction price in each intraday interval. Overnight returns are excluded.

2.2 Private information variables

It is generally not possible to directly measure private information. In our study we use two variables that are commonly used to proxy private information, namely order flow and the bid-ask spread.

Order flow, or net buying pressure, is defined as the difference between the volume in trades initiated by a buyer and the volume in trades initiated by a seller during a certain interval. When there is a large positive order flow this could indicate investors are initiating trades based on having (the belief that) private information that indicates that the price of the asset is relatively low, and vice versa. Thus, order flow could have an impact on prices, and as such a large (positive or negative) value may mark a situation with high information asymmetry amongst participants.

We follow Pasquariello and Vega (2007) by using the unanticipated portion of aggregate order flow, as microstructure imperfections can cause lagged effects in the observed order flow, see Hasbrouck (2004a). Taking N_t as the number of trades in interval t and $v_{t,j}$ as the volume of the j-th trade in this interval, we first calculate our measure of order flow, denoted OF_t, as

$$OF_t = \sum_{j=1}^{N_t} q_{t,j} v_{t,j}, \tag{1}$$

where $q_{t,j} = 1$ if the j-th trade is initiated by a buyer and -1 if it is initiated by a seller. We standardize the order flow by scaling it by the sample standard deviation. Related to the definition of order flow OF_t as given above, this measure is also referred to as the 'signed'

---

[4]The data was obtained from Tickdata Inc., see `http://www.tickdata.com`.

[5]As a robustness check, we repeat our analysis also with 15-minute intervals. The main results do not change.

trading volume. In a second step, we estimate an appropriate ARMA model to remove the 'expected' order flow. The residuals of this model represent the aggregate unanticipated net order flow OF_t^* over each time interval t.[6] Our empirical analysis is based on unanticipated order flow, however we refer to order flow and unanticipated order flow interchangeably.

Our data set does not include the 'sign' variable $q_{t,j}$, indicating whether a trade is initiated by a buying or selling party. We estimate the sign of each trade by employing the method put forward by Van der Wel, Menkveld and Sarkar (2009), which is a likelihood-based implementation of the Bayesian algorithm suggested by Hasbrouck (2004b). [7]

Our second measure for private information is the bid-ask spread during a certain interval. This proxy may be slightly more noisy, as it could also include inventory costs and order processing costs beyond effects due to asymmetric (private) information.[8] A high bid-ask spread may imply a low level of liquidity, which could be due to a high level of information asymmetry. As our data set only comprises transaction prices, we need to estimate the bid-ask spread. We do so by taking the difference between the (volume-weighted) buy price and the average sell price during a particular interval, see also Manaster and Mann (1996). We remove spread estimates that are negative and trim observations that exceed the 99.5th percentile of the empirical distribution to avoid having noisily measured spreads dominating our results.

2.3   Public information variables

An advantage of studying the Treasury market is that public information affecting the prices and volatility is readily identified: these are the scheduled releases of macroeconomic variables. We use data on expectations and announcements of 24 key U.S. macro variables. This data is obtained from Econoday,[9] and feeds the information that is published on Bloomberg. The data on the actual released values comes from Haver Analytics. We derive our proxy for the market expectation from analysts' forecasts, which stem from Market News Interna-

---

[6]We perform the analysis also by using order flow as defined in (1). This does not affect the main results.

[7]For robustness, we also use the tick-test (Lee and Ready, 1991), which determines the sign based on the relation of the current trade price to the previous trade price. If a trade occurs at a price higher (lower) than the previous trade a trade is called an up-tick (down-tick), and is assumed to be initiated by the buying (selling) party. The main results do not change when we use the bid-ask spread based on the tick-test.

[8]See for example Biais, Glosten and Spatt (2005), who survey the literature on the analysis of the price formation and trading process.

[9]See http://www.econoday.com/.

tional and Thomson Financial. In order to obtain the analysts' forecasts the aforementioned companies hold a survey amongst a number of analysts for the announcements that are to come in the following week. The data set records the median of these forecasts, which provides an estimate of the market consensus. In addition, we have the lowest and highest analysts' forecasts for each announcement from June 2007 onwards.

We follow the existing literature (see, e.g., Balduzzi, Elton and Green (2001)) by considering the surprise in each announcement. This surprise is constructed as the difference between the actual released value and the consensus analysts' forecast. Since units of measurement vary widely across macroeconomic variables, we standardize the surprises by dividing by their sample standard deviation. Hence the surprise $S_{k,t}$ in the announcement of variable k at time t is

$$S_{k,t} = \frac{A_{k,t} - M_{k,t}}{\sigma_k}, \tag{2}$$

where $A_{k,t}$ denotes the announced value, $M_{k,t}$ is the median of analyst forecasts, and σ_k is the sample standard deviation of their difference.

Table 1 lists the 24 macroeconomic variables in our data set. Most of these are released at a monthly frequency. The exceptions are GDP, which is announced quarterly, and initial claims for unemployment insurance, which is released on a weekly basis. We have in total 1,825 announcements during the sample period 2004-2009. In our sample period of 1,519 days, 1,002 days contain at least one announcement release. The actual number of surprises that we use in our analysis is somewhat smaller than the total number of releases as the consensus forecast is unavailable for some announcements or the market is closed on the day of the announcement. On average, for each (monthly) macro variable we are able to use 70 of the 72 announcements. Table 1 also shows the number of available minimum and maximum analysts' forecasts for each variable. This number is considerably smaller than the number of surprises, as they are available only since June 2007.

Table 1: Macroeconomic announcements

This table describes scheduled macroeconomic announcements from 2004 to 2009. The data is from Econoday, the table is modeled after Anderson, Bollerslev, Diebold and Vega (2003, Table 1, p.43). The number of surprises is equal to the total number of announcements over the sample period minus missing consensus and missing trading days.

		Time(ET)	Number of announcements	Surprises in our sample	Number of dispersions
		Quarterly			
1	GDP advance	8:30 a.m.	24	24	10
2	GDP preliminary	8:30 a.m.	24	24	11
3	GDP final	8:30 a.m.	24	24	12
		Monthly			
Real Activity					
4	Nonfarm payroll employment[a]	8.30 a.m.	72	71	31
5	Retail sales[b]	8.30 a.m.	72	71	31
6	Industrial production[c]	8.30 a.m.	72	70	31
7	Capacity utilization[c]	9.15 a.m.	72	70	31
8	Personal income	9.15 a.m.	72	72	33
9	Consumer credit[d]	8.30 a.m.	72	65	31
Consumption					
10	Personal consumption exp.	3.00 p.m.	72	72	33
11	New home sales Investment[e]	8.30 a.m.	72	70	32
Investment					
12	Durable goods orders	10.00 a.m.	72	72	32
13	Construction spending	8.30 a.m.	72	72	31
14	Factory orders	10.00 a.m.	72	72	31
15	Business inventories[f]	10.00 a.m.[g]	72	70	30
Goverment Purchases					
16	Government budget deficit[h]	2.00 p.m.[i]	72	65	29
Net Exports					
17	Trade balance	8.30 a.m.	72	72	31
Prices					
18	Producer price index[j]	8.30 a.m.	72	69	31
19	Consumer price index	8.30 a.m.	72	72	32
Forward-looking					
20	Consumer confidence index	10.00 a.m.	72	72	33
21	NAPM index	10.00 a.m.[k]	72	72	33
22	Housing starts	8.30 a.m.	72	72	30
23	Index of leading indicators[l]	8.30 a.m.	73	71	32
		Weekly			
24	Initial unemployment claims[m]	8.30 a.m.	312	309	135

[a] No consensus on May 3, 2007; [b] No consensus on May 11, 2007; [c] No consensus on May 16, 2007 and no trading day at April 14, 2006. [d] No consensus on December 6, 2006; January 8 and February 7 2007; Market closed in afternoon on October 10, 2005; October 6, 2006; April 6 and October 5, 2007. [e] No consensus on July 27, 2006 and May 24, 2007. [f] No consensus on May 11 and September 14, 2007. [g] Earlier part of sample often at 8:30. [h] No consensus on October 14, 2004; October 12, 2006; October 11, 2007 and April 10, 2008. Market closed in afternoon on November 10, 2005 and January 12, 2006. No trading day on April 10, 2009. [i] Announcement at 1:30 on October 14, 2005; at 1:00 on October 14, 2008 and at 3:30 on October 16, 2009. [j] No consensus on March 18, 2004; April 1, 2004 and May 11, 2007. [k] Latest part of sample at 9:45 (since January 2007). [l] Announcements and no consensus on both May 17 and 21, 2007. [m] No consensus on May 10, 17 and 24, 2007.

## 2.4 Public and private information effects: A first impression

Figure 1 gives a first impression of the effect of public and private information variables on returns and volatility. The chart considers patterns in the volatility, bid-ask spread and order flow during all (72) days with scheduled Nonfarm Payroll Employment announcements at 8:30 during our sample period. For this purpose, we compute the average of these variables across these announcement days (where we use the absolute value of order flow) as well as corresponding averages across 72 randomly selected trading days without any macro announcement. Figure 1 depicts the average volatility, spread and order flow during these announcement and non-announcement days. The graphs provide two important insights. First, the patterns in volatility and order flow as shown in Panels A and B encourage our research question whether there is an effect of private information on volatility. Volatility spikes during the interval after the announcement time, reflecting the initial reaction of the market to the announcement. After the announcement, volatility does not immediately revert to its pre-announcement level. Unanticipated order flow also increases during the first minutes following the announcement and remains at a higher level for some time after the announcement. Second, the bid-ask spread peaks around the announcement time (8:30), showing the relationship found by Fleming and Remolona (1999) during announcement days. They argue that the high spread is a result of controlling inventory risk at the time of extreme price volatility. After the announcement the spread reverts to its normal level faster than volatility and order flow.

3 Modeling the response of Treasury futures to public and private information

In this section we describe the methodology that we adopt to examine the effects of private information on the volatility of the 30-year Treasury bond futures. We recognize that it is crucial to control for the effects of public and private information on returns as well as the effects of public information on volatility. To accomplish this, we propose a model specifying the dynamics of both returns and volatility in such a way that both equations can be estimated jointly. We develop our general modeling framework in different stages. First,

Figure 1: Order flow, bid-ask spread and volatility during announcement and non-announcement days

This figure shows time series of the volatility, the bid-ask spread and order flow of 30-year U.S. Treasury bond futures on announcement days and non-announcement days. Each series is obtained by the average over 72 trading days where no announcements were made (dotted line), and 72 trading days when an 8:30 Nonfarm Payroll Announcement was made (solid line). Panel A denotes the intraday volatility, measures by the absolute returns. Panel B depicts the unanticipated order flow on each 5-minute interval, whereas panel C denotes the bid-ask spread.
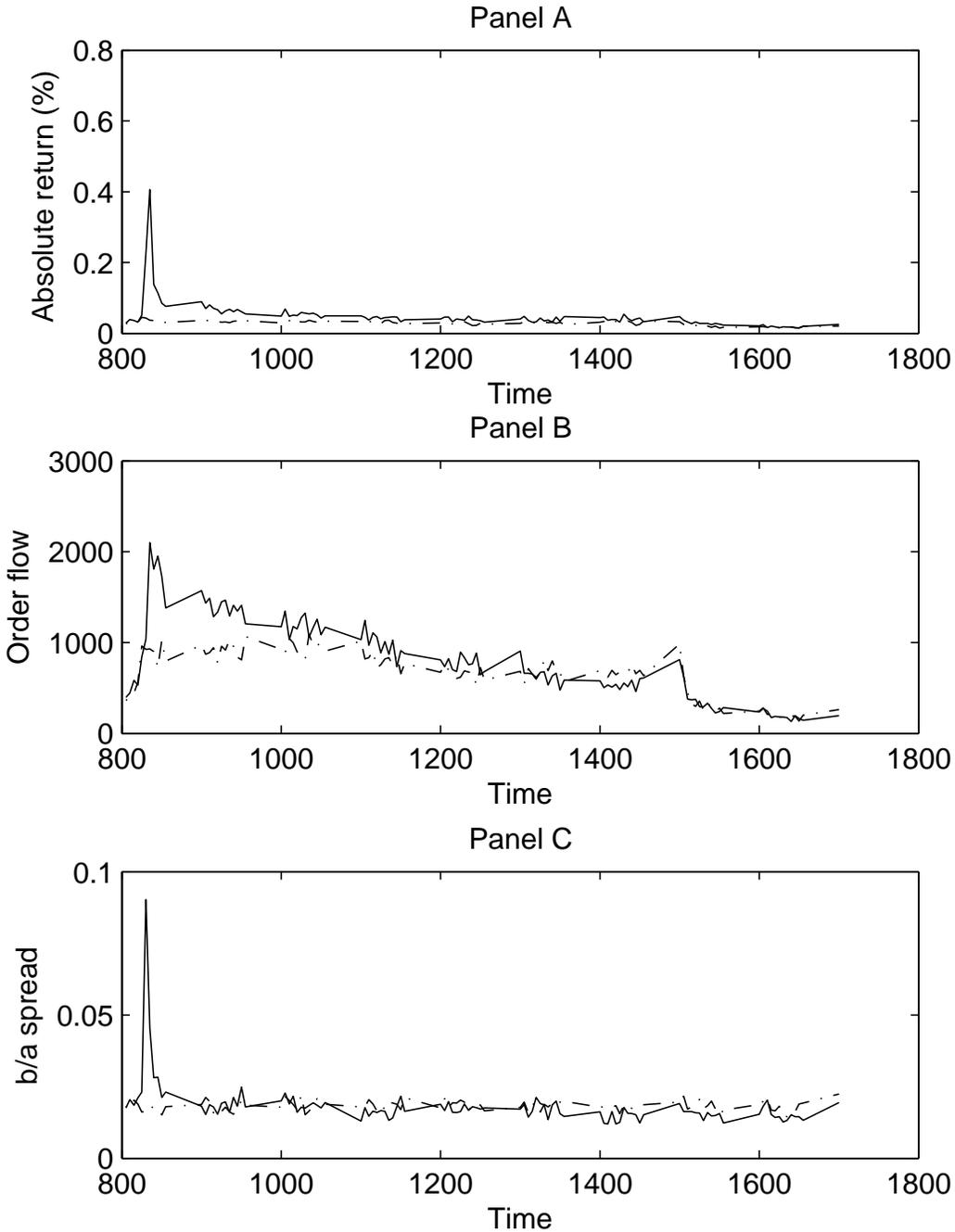
we consider the model equation for the 5-minute intraday returns. Second, we augment this with a second model equation for (conditional) volatility. We also discuss the resulting empirical findings in this section.

3.1 Public and private information matter for returns

We adopt the approach of Andersen, Bollerslev, Diebold and Vega (2003, 2007), ABDV hereafter. We assume a linear model specification for the 5-minute returns on the 30-year Treasury bond futures, including I autoregressive terms and J lags of the announcement surprises of each of the K macroeconomic fundamentals. In addition, we include L lags of K' private information variables, such that the model reads

$$R_t = \beta_0 + \beta_a D_a + \sum_{i=1}^{I} \beta_i R_{t-i} + \sum_{k=1}^{K} \sum_{j=0}^{J} \beta_{kj} S_{k,t-j} + \sum_{k'=1}^{K'} \sum_{l=0}^{L} \gamma_{k'l} P_{k',t-l} + \epsilon_t, \qquad (3)$$

for $t = 1, \ldots, T$ where R_t is the 5-minute return from time $t-1$ to time t, $S_{k,t}$ is the standardized surprise of announcement k at time t as defined in (2) and $P_{k't}$ denotes the k'-th private information variable. Finally, D_a is a dummy variable that equals one on announcement days. We follow Pasquariello and Vega (2007) by including unanticipated order flow in the above specification.[10] The spread is not used for the conditional mean equation, since it can only takes positive values. The disturbance term ϵ_t in (3) is likely to be heteroscedastic. We subsequently specify a separate model equation for the conditional volatility of ϵ_t, which allows us to examine the effects of public and private information on the Treasury bond futures volatility explicitly. First, however, we document the effects on returns by estimating (3) only, accounting for the heteroskedasticity (and any remaining serial correlation not captured by the autoregressive terms) by using Newey-West standard errors.

Table 2 reports results based on estimating the model in (3) using the full sample period from January 1, 2004 to December 31, 2009. We estimate the model both with and without the private information variables. We choose the number of lags of the explanatory variables by means of the Akaike and Schwarz information criteria. Both criteria suggest $I = 3$ and

---

[10]An AR(1) filter was sufficient to capture the autocorrelation in order flow.

Table 2: Effect of public and private information on returns

This table reports the estimation results of the following regression model:

$$R_t = \beta_0 + \beta_a D_a + \sum_{i=1}^{I} \beta_i R_{t-i} + \sum_{k=1}^{K} \beta_k S_{k,t} + \beta_{OF} OF_t^* + \epsilon_t,$$

where R_t is the 5-minute log return of 30-year U.S. Treasury bond futures from time period $t-1$ to t, $S_{k,t}$ is the standardized surprise of announcement type k, $k = 1, \ldots, 24$ and OF_t^* denotes the unanticipated order flow. We include a dummy variable that equals one on announcement days. The first column presents results of the regression where only public information is included. The second column provide results of the regression that contains both public and private information. The superscripts $***,**$ and $*$ indicate significance at the 1%, 5% and 10% level respectively, where the significance is assessed using Newey-West standard errors. The sample goes from January 2, 2004 through December 31, 2009.

Dependent variable:	30Y futures returns	
	(1)	(2)
Private Information		
Order flow		0.025***
Public Information		
1. GDP advance	−0.069	−0.056
2. GDP preliminary	−0.036	−0.030
3. GDP final	0.008	0.002
Real Activity		
4. Nonfarm payroll employment	−0.302***	−0.286***
5. Retail sales	−0.104***	−0.097***
6. Industrial production	0.045	0.040
7. Capacity utilization	−0.058***	−0.043**
8. Personal income	−0.024	−0.020
9. Consumer credit	−0.003	−0.004
Consumption		
10. Personal consumption expenditures	−0.020	−0.015
11. New home sales Investment	−0.046***	−0.040***
Investment		
12. Durable goods orders	−0.070***	−0.053***
13. Construction spending	−0.021	−0.019
14. Factory orders	−0.017	−0.022
15. Business inventories	0.017	0.013
Government Purchases		
16. Government budget deficit	0.007	0.008
Net Exports		
17. Trade balance	−0.027**	−0.020**
Prices		
18. Producer price index	−0.054**	−0.041**
19. Consumer price index	−0.042	−0.041*
Forward Looking		
20. Consumer confidence index	−0.063***	−0.054***
21. NAPM index	−0.020	−0.022
22. Housing starts	−0.015	−0.023
23. Index of leading indicators	−0.045***	−0.039***
24. Initial unemployment claims	0.041***	0.035***
R^2	0.019	0.215
Nr. Observations	159,646	159,646

$J = 0$, as well as $J' = 0$ when the private information variables are included. Note that, while the macroeconomic surprises are only included contemporaneously with the return during the first 5-minutes following the announcements, they do affect subsequent returns through the autoregressive terms.

In line with earlier research (e.g., ABDV (2007), Table 5A) we find that many of the fundamentals have a significant effect on bond futures returns. The magnitude and significance of the coefficients is not affected by the inclusion of the private information variables. We explain the reaction of the bond market to macroeconomic news mainly in terms of revisions of inflationary expectations, which is in accordance with commentaries in the financial press. In line with the view of the Phillips curve, inflation should be positively correlated with economic activity. Higher inflation leads to higher interest rates, hence the returns on treasuries decline. The estimation results support this interpretation. We find that pro-cyclical variables such as Nonfarm Payroll Employment, Retail Sales and New Home Sales Investment, indeed affect bond prices negatively, while counter-cyclical variables such as the Initial Unemployment Claims have a positive impact on bond prices. Further, expected inflation also is a key focal point for fixed-income investors since news associated with the Producer Price Index (PPI) significantly affect bond prices. Comparing the magnitudes of the coefficients across surprises, we observe that new public information in Nonfarm Payroll Employment announcements are most important in economic terms. A one standard deviation surprise in this variable implies a change of almost 30 basis points in the futures returns.

The estimation results in the rightmost column of Table 2 demonstrate that unanticipated order flow significantly influences returns, which is consistent with prior literature (see, e.g., Evans and Lyons (2002) and Green (2004)). Including unanticipated order flow increases the regression R^2 from 0.019 to 0.215. The coefficient of unanticipated order flow is equal to 0.025 and significantly different from zero. Hence, on a 5-minute basis, a one standard deviation shock of order flow increases the return by 2.5 basis points. Although this number is lower than for example the corresponding value of Nonfarm Payroll Employment, one should take into account that this macroeconomic announcement is released only once per month, while order flow is available at each time interval. Note that a positive shock to order flow can be due either to an increase in buyer-initiated trades or a reduc-

tion in seller-initiated trades. In both cases, this signals positive private information, which increased bond returns.

Considering the importance of macroeconomic announcements in terms of explanatory power for the 5-minute returns, the R^2 of the specification (3) is rather small at 0.019. This is somewhat misleading though, since these announcements occur relatively rarely, in the sense that the number of observations for which a surprise actually occurs is only a very small fraction of the total sample size. To highlight the importance of macro news during the announcement periods, we estimate the simplified model

$$R_t = \alpha_k + \beta_k S_{k,t} + \epsilon_t, \tag{4}$$

using only those observations when an announcement of variable k was made at time t. The estimation results for the 24 different announcements are shown in Table 3. Note that the magnitude of the least squares coefficient estimates as well as their standard errors are rather similar to those obtained with the general model specification in (3). The R^2 values are considerably higher for most announcements, exceeding 0.2 for the Consumer Confidence Index, Nonfarm Payroll Employment, the Conference Board Index of Leading Indicators, New Home Sales, and Retail Sales.

3.2   Public and private information also matter for volatility

ABDV (2003, 2007) focus on modeling the effects of macro announcements on returns. In order to account for heteroskedasticity, the return equation as given in (3) is estimated by means of a two-step Feasible Weighted Least Squares (FWLS) procedure, where the weights are inversely related to estimates of the volatility of the returns. It is well-documented that the main determinants of volatility of high-frequency returns (apart from its mere persistence) are macroeconomic news announcements and a pronounced deterministic pattern related to the trading activity during different parts of the day, see Bollerslev, Cai and Song (2000), among many others. For this reason, in the 'ABDV approach' a linear specification for the time-varying volatility of the intra-day unexpected returns ε_t in (3) is proposed that includes the standardized surprises, autoregressive terms and a flexible Fourier series

16

Table 3: Effect of public information on returns during announcements

This table reports the estimation results of the following regression model:

$$R_t = \alpha_k + \beta_k S_{k,t} + \epsilon_t,$$

where R_t is the 5-minute log return of the 30-year U.S. Treasury bond futures from period $t-1$ to t. We include only the pairs $(R_t, S_{k,t})$ when an announcement of fundamental k was made at time t. The superscripts $***, **$ and $*$ donate significance at the 1%, 5% and 10% level respectively. The sample goes from January 2, 2004 through December 31, 2009.

	β_k	R^2	Nr. Obs
1. GDP advance	-0.068	0.094	24
2. GDP preliminary	-0.038^*	0.082	24
3. GDP final	0.008	0.005	24
Real Activity			
4. Nonfarm payroll employment	-0.294^{***}	0.268	71
5. Retail sales	-0.110^{***}	0.225	71
6. Industrial production	0.000	0.000	70
7. Capacity utilization	-0.016	0.029	70
8. Personal income	-0.018	0.020	72
9. Consumer credit	-0.003	0.003	65
Consumption			
10. Personal consumption expenditures	-0.020	0.023	72
11. New home sales Investment	-0.054^{***}	0.235	70
Investment			
12. Durable goods orders	-0.072^{***}	0.136	72
13. Construction spending	-0.020	0.013	72
14. Factory orders	-0.017	0.019	72
15. Business inventories	0.014	0.011	70
Government Purchases			
16. Government budget deficit	0.007	0.011	65
Net Exports			
17. Trade balance	-0.029^{**}	0.059	72
Prices			
18. Producer price index	-0.071^{***}	0.132	69
19. Consumer price index	-0.035	0.021	72
Forward Looking			
20. Consumer confidence index	-0.067^{***}	0.285	72
21. NAPM index	-0.016	0.014	72
22. Housing starts	-0.008	0.003	72
23. Index of leading indicators	-0.045^{***}	0.260	71
24. Initial unemployment claims	0.039^{***}	0.074	309

capturing the intraday pattern of volatility, that is,

$$|\hat{\epsilon}_t| = \beta_0 + \sum_{i=1}^{I'} \beta_i |\hat{\epsilon}_{t-i}| + \sum_{k=1}^{K} \sum_{j'=0}^{J'} \beta_{kj'} |S_{k,t-j}| + \sum_{q=1}^{Q} \left(\delta_q \cos(\frac{q2\pi t}{N}) + \phi_q \sin(\frac{q2\pi t}{N}) \right) + u_t, \quad (5)$$

where $|\hat{\epsilon}_t|$ is the absolute value of the residual of equation (3), $S_{k,t}$ denotes again the public information variables and the sine and cosine terms the calender effect. This effect consists of a intraday pattern that capture the high-frequency rhythm of deviations of intraday volatility from its daily average. ABDV (2003) model this pattern by a Fourier flexible form with trigonometric terms that follows a periodicity of one day. In case of 5-minute returns, we set N equal to 108 (9 hours × 12 five-minute intervals). The specification in (5) is estimated by Ordinary Least Squares (OLS) and the inverse of the fitted values of the dependent variable are used as weights to perform a FWLS estimation of the return equation in (3).

Although the above specification is flexible and easy to estimate, a numerical problem may occur: the conditional volatility equation may produce negative fitted values. In particular, this may happen when the Fourier series attains its minimum value and a macro announcement does not occur during this interval. Such a negative value implies negative weights for the FWLS step of the return equation, which invalidates the approach.[11] Note that negative fitted values need not necessarily occur to create problems. Even if the fitted value of (5) remains positive but becomes very small, the corresponding observations receive a disproportionally large weight in the FWLS estimation of the return specification (3), which may not be desirable. A second issue is that the ABDV two-step approach neglects parameter uncertainty of the conditional mean equation when estimating the volatility equation. As we are particularly interested in the effects of public and private information on volatility, it is important to take this uncertainty into account.

We overcome both econometric issues discussed above by using a GARCH-type approach, see also De Goeij and Marquering (2006) and Brenner, Pasquariello and Subrahmanyam (2009). We deviate from these studies by adopting the basic idea of the Spline-GARCH model developed by Engle and Rangel (2008) to decompose the conditional volatility in

---

[11]Estimating (5) for the 5-minute treasury futures returns using the full sample period, the fitted value becomes negative for 70 observations.

two (multiplicative) parts. One of these components is a standard GARCH term (but normalized to have unconditional mean equal to one) to capture high-frequency movements in volatility. In Engle and Rangel (2008), the other component captures low-frequency movements in volatility, which is achieved by using a quadratic spline function of time t. For the purpose of our analysis, we specify the second component such that it describes the intra-day volatility pattern (by means of a Fourier series) and the effects of public and private information. Combined with the return equation from the ABDV approach, i.e. (3), the complete model specification is given by

$$R_t = \beta_{0,m} + \beta_{a,m}D_a + \sum_{i=1}^{I}\beta_i R_{t-i} + \sum_{k=1}^{K}\sum_{j=0}^{J}\beta_{kj}S_{k,t-j} + \sum_{k'=1}^{K'}\sum_{l=0}^{L}\gamma_{k'l}P_{k',t-l} + \epsilon_t,$$

$$\epsilon_t = \sqrt{g_t\tau_t}u_t,$$

$$g_t = (1 - \alpha - \beta_{GA}) + \alpha\left(\frac{\epsilon_{t-1}^2}{\tau_{t-1}}\right) + \beta_{GA}g_{t-1},$$

$$\tau_t = \exp\left(\beta_{0,v} + \beta_{a,v}D_a + \sum_{q=1}^{Q}\left(\delta_q\cos(\frac{q2\pi t}{N}) + \phi_q\sin(\frac{q2\pi t}{N})\right) + \right.$$

$$\left. \sum_{k=1}^{K}\sum_{j'=0}^{J'}\beta_{kj'}|S_{k,t-j}| + \sum_{k'=1}^{K'}\sum_{l'=0}^{L'}\gamma_{k'l'}|P_{k',t-l'|}\right), \tag{6}$$

where $u_t \sim N(0,1)$, the conditional variance $E[\epsilon_t^2|\mathcal{I}_{t-1}] = g_t\tau_t$ with $\mathcal{I}_t$ the information set at time t, g_t represents the unit-GARCH term and τ_t denotes the joint public and private news component. We include both unanticipated order flow and the bid-ask spread as private news proxies. The exponential specification of τ_t obviously avoids that the conditional variance can become negative by construction.

An important difference between the specification of the Spline-GARCH model and the augmented-GARCH model of de Goeij and Marquering (2006) is that the latter model implies by definition a permanent effect of surprises on volatility. This is the result of introducing a dummy variable that accounts for the announcements in the GARCH equation. A high value of β_{GA} implies then the permanent effect of public information on volatility. The Spline-GARCH model however allows us to control for the length of the effect of the surprises; we simply add lags of surprises in the specification of τ_t, which does not permanently influence g_t. Hence we are able to distinguish between the effect of private information on

19

volatility and the effect of macroeconomic surprises on volatility.

We estimate all parameters in the model (6) simultaneously, by (quasi) maximum likelihood. Given the large number of parameters, this can be a tough non-linear maximization problem. For this reason we first estimate (3) and (5) by OLS, and use the resulting coefficient estimatesas meaningful starting values for the maximization of the likelihood.[12] In addition, we reduce the computational burden by estimating the conditional likelihood: first we estimate the volatility parameters while holding the return parameters fixed, and vice versa. In a final step we optimize the joint likelihood.

Table 4 provides estimation results for the GARCH model in (6) based on the complete sample period from January 2004 until December 2009. Panel A reports results of four different specifications. The first specification does not include private information. The other three specifications add private information variables to the model. We set $Q = 5$ and $J' = 1$ in the Spline-GARCH model for all specifications.[13] The results again indicate both (unanticipated) order flow and the bid-ask spread have significantly positive effects on the volatility. The coefficients for the two proxies for private information have approximately the same magnitude when corrected for the difference in standard deviation of order flow and the bid-ask spread. Both positive coefficients suggest that the learning process of the agents to clear the market increases uncertainty about the bond returns. A higher spread means lower liquidity which in turn implies a higher volatility. In summary, the market aggregates private information, which in turn affects the returns and uncertainty on the futures market.

Panel B of Table 4 reports the estimated coefficients for the surprises in macro announcements. We only report the public information coefficients for the most general specification that includes both order flow and the bid-ask spread. In the other three specifications the surprise coefficients are comparable. The results indicate that 15 out of the 24 announcement surprises significantly affect subsequent returns, where in general the same reasoning holds as described in the discussion about Table 2, i.e. procyclical variables cause a negative

---

[12]In more detail, we transform the estimated coefficients of (5) by taking the exponent as a result of the specification of the volatility part of the spline GARCH model.

[13]The Akaike and Schwarz criteria propose $Q = 5$ and $J' = 0$ in the 'ABDV approach', which is used for meaningful starting values. However, a bottom-up procedure, i.e. including a lag of each public news variable until the F-test of joint significance of all estimated coefficients cannot be rejected, allows for one lag. To be sure that we do not miss any information of lagged public news surprises, we set J' equal to one in the 'ABDV approach'.

Table 4: Public and private information effects

This table reports the estimation results of the Spline-GARCH model:

$$
\begin{aligned}
R_t &= \beta_{0,m} + \beta_{a,m}D_a + \sum_{i=1}^{I} \beta_i R_{t-i} + \sum_{k=1}^{K} \sum_{j=0}^{J} \beta_{kj} S_{k,t-j} + \gamma_1 OF_t^* + \epsilon_t, \\
\epsilon_t &= \sqrt{h_t} u_t, \\
h_t &= \tau_t g_t, \\
g_t &= (1 - \alpha - \beta_{GA}) + \alpha \left(\frac{\epsilon_{t-1}^2}{\tau_{t-1}} \right) + \beta_{GA} g_{t-1}, \\
\tau_t &= \exp \left(\beta_{0,v} + \beta_{a,v}D_a + \sum_{q=1}^{Q} \left(\delta_q \cos(\tfrac{q2\pi t}{N}) + \phi_q \sin(\tfrac{q2\pi t}{N}) \right) + \right. \\
&\quad \left. \sum_{k=1}^{K} \sum_{j'=0}^{J'} \beta_{kj'} |S_{k,t-j}| + \gamma_1 |OF_t^*| + \gamma_2 |P_t| \right),
\end{aligned}
$$

where R_t is the 5-minute log return of the 30-year U.S. Treasury bond futures from time period $t-1$ to t, $S_{k,t}$ is the standardized news announcement for $k = 1, \dots, 24$ and P_t and OF_t^* denote the bid-ask spread and the unanticipated order flow from period t to $t+1$. Further, $I = 3$, $J = 0$, $J' = 1$ and $Q = 5$. In Panel A we report parameter estimates corresponding with private news for four various specifications, whereas panel B provides the public news coefficients of the last specification. The superscripts $***$, $**$, and $*$ donate significance at the 1%, 5% and 10% level, respectively. The sample goes from January 2, 2004 through December 31, 2009.

Panel A: Private information				
Dependent variable:		30Y futures returns		
	(1)	(2)	(3)	(4)
Conditional mean equation				
Order flow		0.022***	0.020***	0.021***
Macro announce.	yes	yes	yes	yes
Conditional volatility equation				
Order flow		0.551***		0.497***
Bid-ask spread			18.63***	15.54***
Macro announce.	yes	yes	yes	yes
log-likelihood	247,112	273,370	270,429	275,776
Nr. Observations	153,132	153,132	153,132	153,132

21

Panel B: Public information			
	Cond Mean part	Cond Volatility part	
Macroeconomic announcements			1st lag
1. GDP advance	−0.151**	2.350***	0.517
2. GDP preliminary	−0.061	2.191***	−0.104
3. GDP final	−0.010	0.895***	0.101
Real Activity			
4. Nonfarm payroll employment	−0.349***	4.441***	1.303***
5. Retail sales	−0.091***	1.528***	0.862***
6. Industrial production	−0.006	1.078**	−0.209
7. Capacity utilization	−0.044**	−0.223	0.471
8. Personal income	−0.028*	0.457	−0.027
9. Consumer credit	−0.002	−0.214	−0.325
Consumption			
10. Personal consumption expenditures	−0.044**	0.972***	0.102
11. New home sales Investment	−0.075***	1.143***	0.417**
Investment			
12. Durable goods orders	−0.055*	1.660***	0.032
13. Construction spending	−0.018	1.946***	0.851***
14. Factory orders	−0.007	1.457**	0.632**
15. Business inventories	−0.005	0.290	−0.122
Government Purchases			
16. Government budget deficit	−0.002	0.032	−0.076
Net Exports			
17. Trade balance	−0.027***	0.102	0.262
Prices			
18. Producer price index	−0.053**	1.576***	0.389**
19. Consumer price index	−0.123***	1.883***	0.014
Forward Looking			
20. Consumer confidence index	−0.068***	1.091***	0.300
21. NAPM index	−0.014	0.977***	0.198
22. Housing starts	−0.031**	1.067***	−0.091
23. Index of leading indicators	−0.023***	0.112	0.122
24. Initial unemployment claims	0.038***	0.578***	−0.025

Figure 2: Intraday volatility effects
This figure depicts the intraday effects that capture the high-frequency pattern of deviations of intraday volatility from its daily average, as estimated in the Spline-GARCH model defined in (6). In addition, the figure depicts 95% confidence bounds.
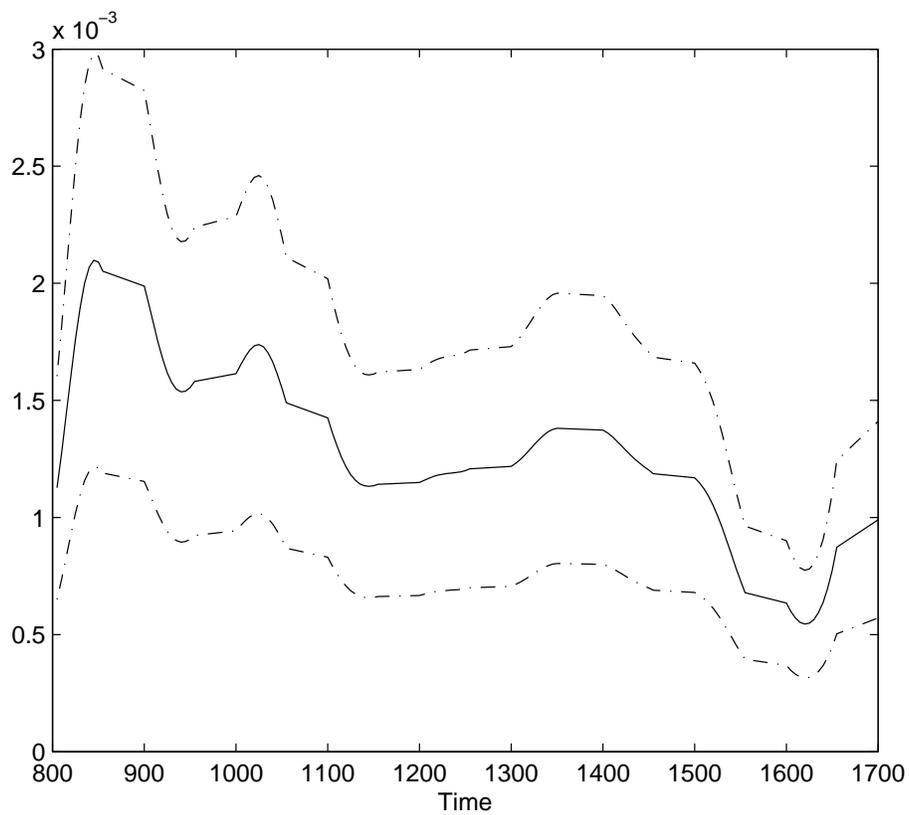
bond return whereas countercyclical variables positively influence bond returns. In addition, the impact of public information is short-lived, since the return equation requires no lag of the macro variables (guided by a Likelihood-Ratio test). Thus the price discovery process is quick. This is consistent with earlier found results, see for example ABDV (2007).

We find that almost all macro surprises affect volatility in the first five minutes after the announcement, whereas some variables also have a significant impact on volatility during the next five minutes. The impact of public information on volatility is less short-lived compared to the impact on returns, although a likelihood-ratio test indicates only one lag of the macro announcement variables is required.

Figure 2 depicts the Fourier series that captures the intraday deviations from the daily volatility. Typically, when floor trading starts, volatility is relatively high, whereas during lunch and after closing time (15:00) volatility is relatively low. Incorporating this deterministic pattern contributes substantially to explaining the intraday behavior of volatility.

Table 5 shows the economic significance of the estimated coefficients of the Spline-GARCH model. We take the three most influential news announcements on volatility, GDP Advance, Nonfarm Payroll Employment and the CPI, and compute the partial derivative of the volatility with respect to these announcements. In addition, we include Initial Unemployment Claims. Recall that news announcements occur rarely, while order flow and the bid-ask spread are available at each 5-minute interval. We add up the influence of private variables on the volatility over a whole day to compare both public and private variables in an economically meaningful way. Specifically, we add up all values of the square root of τ_t over all intraday intervals including both private and public variables and subtract the sum of the volatilities without including the public variables.

The most striking result is the daily impact of the private information proxies. Unanticipated order flow and the bid-ask spread increase on average the daily volatility by 80 and 75 basis points respectively. This number is robust against special announcement releases, as we average it out over all 1,509 trading days. This is quite large compared to the effect of public information variables. Nonfarm Payroll Employment releases are the most influential news announcements: A one standard deviation change in the surprise implies on average an increase of 17 basis points of the volatility. The economic significance of the GDP advance and the CPI index is of the same order, whereas a release of the Initial Unemployment

Table 5: Economic significance of public and private information on volatility
This table provides economic significance of the volatility part of the Spline-GARCH model:

$$R_t = \beta_{0,m} + \beta_{a,m}D_a + \sum_{i=1}^{I}\beta_i R_{t-i} + \sum_{k=1}^{K}\sum_{j=0}^{J}\beta_{kj}S_{k,t-j} + \gamma_1 OF_t^* + \epsilon_t,$$
$$\epsilon_t = \sqrt{h_t}u_t,$$
$$h_t = \tau_t g_t,$$
$$g_t = (1 - \alpha - \beta_{GA}) + \alpha\left(\frac{\epsilon_{t-1}^2}{\tau_{t-1}}\right) + \beta_{GA}g_{t-1},$$
$$\tau_t = \exp\left(\beta_{0,v} + \beta_{a,v}D_a + \sum_{q=1}^{Q}\left(\delta_q \cos(\frac{q2\pi t}{N}) + \phi_q \sin(\frac{q2\pi t}{N})\right) + \right.$$
$$\left. \sum_{k=1}^{K}\sum_{j'=0}^{J'}\beta_{kj'}|S_{k,t-j}| + \gamma_1|OF_t^*| + \gamma_2|P_t|\right),$$

where R_t is the 5-minute log return of the 30-year U.S. Treasury bond futures from time period $t-1$ to t, $S_{k,t}$ is the standardized news announcement for $k = 1, \ldots, 24$ and P_t and OF_t^* denote the bid-ask spread and the unanticipated order flow from period t to $t+1$. Further, $I = 3$, $J = 0$, $J' = 1$ and $Q = 5$. In case of the public information variables, the economic significance is equal to the partial derivative of the conditional volatility ($\sqrt{h_t}$) with respect to the variable S_{kt}. We measure the economic impact of the private information variables on volatility by first adding up all contributions of these variables to the conditional volatilities over all intraday intervals on each day. In addition we take the mean over all sample days. We round all values to integers. The sample goes from January 2, 2004 through December 31, 2009.

Economic significance in basis points		
		Volatility
Public information	GDP advance	9
	Nonfarm payroll employment	17
	Consumer price index	8
	Initial unemployment claims	2
Private information	Order flow	80
	Bid-ask spread	75

Claims changes volatility by two basis points on average. Hence in comparison with public information, the effect of private information is striking.

In summary, the results suggest that private information variables are statistically and economically significant explanatory variables for volatility. Higher absolute values of order flow and lower market liquidity signal the presence of a higher level of private information, which in turn increases the uncertainty on the Treasury market.

4   Extensions

We analyze two extensions of the Spline-GARCH model of the previous section to examine the robustness of our results on the effects of private information on volatility. First, we explore whether this effect depends on the state of the economy. In particular, we extend the model to allow for different effects during expansion and recession periods. Second, we consider whether the dispersion of beliefs among analysts influences the importance of private information for both returns and volatility.

4.1   Conditioning on the state of the economy

We investigate the effect of public information on returns and volatility conditional on the state of the economy, following previous studies including Beber and Brandt (2010) and ABDV (2007). Compared to the existing literature on this issue, we adopt a novel approach to identify a recession or an expansion by using the Aruoba-Diebold-Scotti (ADS) business conditions index, see Aruoba, Diebold and Scotti (2009). In contrast with the NBER index, which is an ex-post measure, the ADS index is designed to track real business conditions at high frequency. Since this may describe the state of the economy more adequately, we use the ADS index to identify recessions and contractions. Underlying the ADS index are economic indicators that mix high- and low-frequency information and stock and flow data.[14] The average value of the ADS index is zero. Relatively larger positive (negative) values indicate relatively better-than-average (worse-than-average) conditions. Based upon the ADS index we construct expansion and recession regime indicators. Specifically, we define the dummy

---

[14]In particular, the (seasonally adjusted) economic indicators are: weekly Initial Jobless Claims; monthly Payroll Employment, Industrial Production, Personal Income less transfer payments, Manufacturing and Trade Sales; and quarterly real GDP.

variable D_E to be equal to one when the ADS index is positive (Expansion) and zero otherwise. A recession indicator is then obtained as $D_R = 1 - D_E$. Using these dummy variables for the state of the economy, the Spline-GARCH model is modified as follows:

$$R_t = \beta_{0,m} + \beta_{a,m}D_a + \beta_{E,m}D_E + \sum_{i=1}^{I}\beta_i R_{t-i} + \sum_{k=1}^{K}\sum_{j=0}^{J}\beta_{E,kj}D_E S_{k,t-j} +$$

$$\sum_{k=1}^{K}\sum_{j=0}^{J}\beta_{R,kj}D_R S_{k,t-j} + \beta_{E,OF}D_E OF_t^* + \beta_{R,OF}D_R OF_t^* + \epsilon_t,$$

$$\epsilon_t = \sqrt{g_t \tau_t}u_t,$$

$$g_t = (1 - \alpha - \beta_{GA}) + \alpha\left(\frac{\epsilon_{t-1}^2}{\tau_{t-1}}\right) + \beta_{GA}g_{t-1},$$

$$\tau_t = \exp\left(\beta_{0,v} + \beta_{a,v}D_a + \beta_{E,v}D_E + \sum_{q=1}^{Q}\left(\delta_q \cos(\frac{q2\pi t}{N}) + \phi_q \sin(\frac{q2\pi t}{N})\right) +\right.$$

$$\beta_{E,OF}D_E|OF_t^*| + \beta_{R,OF}D_R|OF_t^*| + \beta_{E,P}D_E|P_t| + \beta_{R,P}D_R|P_t| +$$

$$\left.\sum_{k=1}^{K}\sum_{j'=0}^{J'}\beta_{E,kj'}D_E|S_{k,t-j}| + \sum_{k=1}^{K}\sum_{j'=0}^{J'}\beta_{R,kj'}D_R|S_{k,t-j}|\right), \qquad (7)$$

where again $u_t \sim N(0,1)$, τ_t denotes the public and private news component, and g_t represents the unit-GARCH term.

Table 6 reports coefficient estimates of the adapted Spline-GARCH model, together with Wald tests for the null hypothesis that $\beta_{E,kj}$ and $\beta_{R,kj}$ or $\beta_{E,kj'}$ and $\beta_{R,kj'}$ are equal. Panel A suggests that private information has a statistically significant different effect on returns in the different states of the economy. Economically this difference is negligible, however, with the coefficients of order flow being equal to 0.021 and 0.022 in expansions and recessions, respectively. The effect of public information on returns conditional on the state of the economy is less clear. Only three variables are significantly different in the recession regime compared to the expansion regime: Factory Orders, Consumer Price Index and the Producer Price Index. In addition, several variables are significant in the recession period while they are not in the expansion period. This is in contrast with ABDV (2007) and Beber and Brandt (2010), who find that surprises generally have a larger effect in the expansion regime compared to the recession regime. This difference might be caused by differences in the sample period, as their data set covers the period 1992-2003, while our

Table 6: Recessions and expansions

This table reports the estimation results of the Spline-GARCH model:

$$
\begin{aligned}
R_t &= \beta_{0,m} + \beta_{a,m}D_a + \beta_{E,m}D_E + \sum_{i=1}^{I}\beta_i R_{t-i} + \sum_{k=1}^{K}\sum_{j=0}^{J}\beta_{E,kj}D_E S_{k,t-j} + \\
&\quad \sum_{k=1}^{K}\sum_{j=0}^{J}\beta_{R,kj}D_R S_{k,t-j} + \beta_{E,OF}D_E OF_t^* + \beta_{R,OF}D_R OF_t^* + \epsilon_t, \\
\epsilon_t &= \sqrt{h_t}u_t, \\
h_t &= \tau_t g_t, \\
g_t &= (1 - \alpha - \beta_{GA}) + \alpha\left(\frac{\epsilon_{t-1}^2}{\tau_{t-1}}\right) + \beta_{GA}g_{t-1}, \\
\tau_t &= \exp\Big(\beta_{0,v} + \beta_{a,v}D_a + \beta_{E,v}D_E + \sum_{q=1}^{Q}\left(\delta_q \cos(\tfrac{q2\pi t}{N}) + \phi_q \sin(\tfrac{q2\pi t}{N})\right) + \\
&\quad \beta_{E,OF}D_E|OF_t^*| + \beta_{R,OF}D_R|OF_t^*| + \beta_{E,P}D_E|P_t| + \beta_{R,P}D_R|P_t| + \\
&\quad \sum_{k=1}^{K}\sum_{j'=0}^{J'}\beta_{E,kj'}D_E|S_{k,t-j}| + \sum_{k=1}^{K}\sum_{j'=0}^{J'}\beta_{R,kj'}D_R|S_{k,t-j}|\Big),
\end{aligned}
$$

where R_t is the 5-minute log return of the 30-year U.S. Treasury bond futures from time period $t-1$ to t, $S_{k,t}$ is the standardized news announcement for $k = 1, \ldots, 24$, OF_t^* denotes the unanticipated order flow and P_t the bid-ask spread from period t to $t+1$. The dummy variables D_E and D_R denotes the expansion and he recession regimes, which are constructed with use of the ADS-index. Further, $I = 3$, $J = 0$, $J' = 1$ and $Q = 5$. Panel A contain estimated coefficients regarding the return part of the model. We use the Wald-test to test on statistical difference of the coefficients during expansions and recessions. Panel B contains the volatility part. We show only the contemporaneous effect and discard the coefficients corresponding with the first lag. The superscripts $***,**$ and $*$ indicate significance at the 1%, 5% and 10% level respectively. The sample goes from January 2, 2004 through December 31, 2009.

Panel A: Conditional Return			
	Expansion	Recession	Wald test
Private information			
Order flow	0.021***	0.022***	7.58***
Public information			
1. GDP advance	0.095	−0.179***	2.40
2. GDP preliminary	−0.117	0.021	1.37
3. GDP final	−0.021	0.026	1.16
Real Activity			
4. Nonfarm payroll employment	−0.194**	−0.350***	1.15
5. Retail sales	−0.117***	−0.056***	1.73
6 . Industrial production	−0.007	0.040	1.52
7. Capacity utilization	−0.047*	−0.045**	0.01
8. Personal income	−0.024	−0.013	0.13
9. Consumer credit	−0.008*	0.001	1.67
Consumption			
10. Personal consumption expenditures	−0.016*	−0.058**	2.39
11. New home sales Investment	−0.083***	−0.072***	0.19
Investment			
12. Durable goods orders	−0.054	−0.053**	0.00
13. Construction spending	−0.042	−0.015	0.38
14. Factory orders	0.216	−0.022	14.26***
15. Business inventories	0.003	−0.016	1.43
Government Purchases			
16. Government budget deficit	0.004	−0.013	0.22
Net Exports			
17. Trade balance	−0.009	−0.031**	1.90
Prices			
18. Producer price index	0.024	−0.100***	8.10***
19. Consumer price index	−0.048	−0.165***	3.63*
Forward Looking			
20. Consumer confidence index	−0.057***	−0.059***	0.00
21. NAPM index	−0.017	−0.005	0.37
22. Housing starts	−0.022	−0.043**	0.42
23. Index of leading indicators	−0.022**	−0.023**	0.00
24. Initial unemployment claims	0.033***	0.042***	0.64

28

Panel B: Conditional Volatility			
	Expansion	Recession	Wald test
Private information			
Order flow	0.513***	0.515***	0.04
Bid-ask spread	10.34***	16.19***	94.41***
Public information			
1. GDP advance	4.040***	1.442***	4.51**
2. GDP preliminary	3.237***	1.280**	2.43**
3. GDP final	1.135**	0.448	1.08
Real Activity			
4. Nonfarm payroll employment	4.233***	4.551***	0.30
5. Retail sales	1.807***	1.035***	3.10*
6 . Industrial production	0.501	0.610	0.02
7. Capacity utilization	0.706	−0.417	1.85
8. Personal income	1.048*	0.239	1.49
9. Consumer credit	−0.708*	−0.087	2.66
Consumption			
10. Personal consumption expenditures	−0.631	1.109***	11.44***
11. New home sales Investment	1.174***	1.207***	0.01
Investment			
12. Durable goods orders	2.236***	1.326***	4.09**
13. Construction spending	2.065***	1.941***	0.07
14. Factory orders	4.615***	0.934***	38.68***
15. Business inventories	−0.026	0.557**	1.81
Government Purchases			
16. Government budget deficit	0.080	0.039	0.01
Net Exports			
17. Trade balance	0.154	0.035	0.11
Prices			
18. Producer price index	1.464***	1.387***	0.03
19. Consumer price index	1.937***	1.673***	0.37
Forward Looking			
20. Consumer confidence index	1.254***	0.996***	0.37
21. NAPM index	0.558	1.116***	1.65
22. Housing starts	1.017**	1.034***	0.00
23. Index of leading indicators	0.209	0.125	0.04
24. Initial unemployment claims	0.644***	0.544***	0.21

sample ranges from 2004 until 2009.

Panel B shows that also the effect of private information on volatility is statistically significant different in expansions and recessions. This holds in particular for the bid-ask spread, since the difference is huge from an economic perspective: the coefficient increases from 10.34 in expansions to 16.19 during contractions. Regarding public information, panel B suggests that the effect on volatility is larger in expansions for most announcements, except for Personal Consumption Expenditures. Based on the Wald test, the effect of public information on volatility is significantly higher during expansions for five variables: GDP Advance, GDP Preliminary, Retail Sales, Durable Goods Orders, and Factory Orders.

In sum, private information matters for volatility, irrespective of the state of the economy. The bid-ask spread has more impact during recessions. In addition, public information has a larger influence on volatility during expansions.

4.2 Conditioning on the dispersion of beliefs

We investigate the impact of private information on both returns and volatility for different levels of the dispersion of beliefs among analysts. This extends Pasquariello and Vega (2007), who study this issue for returns only.

We measure the dispersion of beliefs using the range of analysts' forecasts for the 24 listed macro announcements. This is slightly different from Pasquariello and Vega (2007), who take the standard deviation of analysts' forecasts to estimate the dispersion. We follow their methodology for incorporating the dispersion of beliefs into the model. For this purpose, we first convert the weekly and quarterly dispersions to a monthly frequency. For the weekly announcements of Initial Unemployment Claims this conversion is done by simply averaging the range across four weeks. For the three quarterly announcements in our data set, GDP Advance, Preliminary, and Final, we assume that the dispersion of beliefs in the first month of the quarter is constant throughout the quarter. The monthly proxy for information heterogeneity is defined as a sum of monthly (scaled) dispersions across announcements,

$$SRA_{P,t} = \sum_{k=1}^{K} \frac{RA_{k,t} - \hat{\mu}(RA_{k,t})}{\hat{\sigma}(RA_{k,t})}, \tag{8}$$

where $RA_{k,t}$ is the highest minus the lowest professional forecast of announcement k at time t

and $\hat{\mu}(RA_{k,t})$ and $\hat{\sigma}(RA_{k,t})$ are its sample mean and standard deviation, respectively. Given the monthly dispersion estimates, we divide the empirical distribution function of $SRA_{P,t}$ in a low $(0 - 30\%)$, medium $(30 - 70\%)$ and high $(70 - 100\%)$ dispersion regime. Then we interact order flow with three dummy variables, which are constructed on these three regimes. In principle, it would also be possible to interact the standardized surprises with these dummy variables to allow for different effects of public information for different levels of dispersion. However, the range of the analysts' forecasts used to construct the dispersion measure is available only from June 2007 onwards. This implies that the subsample that can be used in this analysis contains only a limited number of announcements (30 for the monthly variables and 12 for the quarterly GDP figures), making it difficult to obtain reliable results.

Table 7 reports estimation results for the Spline-GARCH model conditioning the effects of private information on the level of the dispersion of analysts' forecasts. We do not report the estimates for the macro announcements, which are similar to those obtained for the complete sample period as shown in Table 4. The informativeness of order flow is considerably higher for months with highest dispersion in analyst forecasts, both in the return and volatility equation. We find a monotonic increase in the coefficients, indicating that the higher the dispersion, the more informative order flow is for return and volatility. This result confirms the findings of Paquariello and Vega (2007), but also provides new insight in explaining the conditional volatility. Order flow has a considerably stronger impact on uncertainty of Treasury futures in times of high dispersion of beliefs among traders about macro fundamentals.

To shed light on the economic significance of this result, the partial effect of a one standard deviation shock in order flow in months with high dispersion on the conditional volatility is equal to 3.35 basis points in a 5-minute interval.[15] As a comparison, the corresponding values in times of low and medium dispersion are equal to 0.85 and 1.18 respectively. Hence order flow is almost four times more important when dispersion of beliefs is high compared to when it is low.

---

[15]We compute this by taking the partial effect of each public/private variable on the conditional volatility.

Table 7: Order flow and information heterogeneity

This table reports the estimation results of the following Spline-GARCH model:

$$
\begin{aligned}
R_t &= \beta_{0,r} + \beta_{a,r}D_a + \beta_{m,r}D_m + \beta_{h,r}D_h + \sum_{i=1}^{I}\beta_i R_{t-i} + \sum_{k=1}^{K}\sum_{j=0}^{J}\beta_{kj}S_{k,t-j} + \\
&\quad \gamma_1 OF_t^* D_l + \gamma_2 OF_t^* D_m + \gamma_3 OF_t^* D_h + \epsilon_t, \\
\epsilon_t &= \sqrt{h_t}u_t, \\
h_t &= \tau_t g_t, \\
g_t &= (1-\alpha-\beta_{GA}) + \alpha\left(\frac{\epsilon_{t-1}^2}{\tau_{t-1}}\right) + \beta_{GA}g_{t-1}, \\
\tau_t &= \exp\left(\beta_{0,v} + \beta_{a,v}D_a + \beta_{m,v}D_m + \beta_{h,v}D_h + \sum_{q=1}^{Q}\left(\delta_q\cos(\tfrac{q2\pi t}{108}) + \phi_q\sin(\tfrac{q2\pi t}{108})\right) + \right. \\
&\quad \left. \sum_{k=1}^{K}\sum_{j'=0}^{J'}\beta_{kj'}|S_{k,t-j}| + \gamma_1|OF_t^*|D_l + \gamma_2|OF_t^*|D_m + \gamma_3|OF_t^*|D_h + \gamma_4 P_t\right),
\end{aligned}
$$

where R_t is the 5-minute log return of the 30-year U.S. Treasury bond futures from time period $t-1$ to t, $S_{k,t}$ is the standardized news announcement for $k = 1, \ldots, 24$, OF_t^* denotes the unanticipated order flow and P_t the bid-ask spread from period t to $t+1$. We introduce three dummy variables, D_l, D_m, and D_h which represent a low, medium or high level of information heterogeneity. Results below are for $I = 3$, $I' = 12$, $J = 0$, $J' = 1$ and $Q = 5$. Since the focus is on the interaction between order flow and the dummy variables, we only report the coefficients corresponding with private information. The superscripts $***$, $**$, and $*$ donate significance at the 1%, 5% and 10% level, respectively. The sample goes from June 2, 2007 through December 31, 2009.

	Private information	
Dependent variable	30Y futures returns	
	Cond Mean	Cond Vol
Order flow (low)	0.023***	0.335***
Order flow (medium)	0.029***	0.468***
Order flow (high)	0.128***	1.329***
Bid-ask spread		15.65***
Macro announce.	yes	yes
log-likelihood	111,161	
Nr. Observations	68,683	

5 Conclusion

We study the impact of private information on volatility in financial markets. We design a unified framework, inspired by the Spline-GARCH model of Engle and Rangel (2008), to study the relationship between public and private information and prices and volatility simultaneously. We apply the model to 5-minute returns and volatility for the 30-year US Treasury futures over the period 2004-2009. We use surprises in 24 key macroeconomic variables to capture public information, while we construct measures of order flow and the bid-ask spread as proxies for private information.

Our main finding is that private information significantly affects volatility. Private information is much more important than public information in this respect. The effect of a shock to order flow on volatility is more than four times larger than the effect of a surprise of the same magnitude in the most influential macroeconomic news announcement. Moreover, we document an interaction between public and private information effects on volatility. The impact of order flow on volatility depends on the dispersion of beliefs about macroeconomic announcements among analysts, with the effect being almost four times larger when there is high dispersion compared to low dispersion. Finally, we find that the effect of private information depends on the state of the economy, in the sense that the effect of order flow and the bid-ask spread on volatility is larger during contractions. Our results imply that risk managers, portfolio managers and regulators should take into account private information variables as a determinant of volatility.

References

[1] Andersen, T.G. (1996), "Return Volatility and Trading Volume: An Information Flow Interpretation of Stochastic Volatility", *Journal of Finance*, 51, 169-204.

[2] Andersen, T.G., T. Bollerslev, F.X. Diebold, and C. Vega (2003), "Micro Effects of Macro Announcements: Real-Time Price Discovery in Foreign Exchange", *American Economic Review*, 93, 38-62.

[3] Andersen, T.G., T. Bollerslev, F.X. Diebold, and C. Vega (2007), "Real-Time Price Discovery in Global Stock, Bond and Foreign Exchange Markets", *Journal of International Economics*, 73, 251-277.

[4] Aruoba, S.B., F.X. Diebold, and C. Scotti (2009), "Real-Time Measurement of Business Conditions", *Journal of Business and Economic Statistics*, 27, 417-427.

[5] Balduzzi, P., E.J. Elton, and T.C. Green (2001), "Economic News and Bond Prices: Evidence from the U.S. Treasury Market", *Journal of Financial and Quantitative Analysis*, 36, 523-543.

[6] Bartolini, L., L. Goldberg, and A. Sacarny (2008), "How Economic News Moves Markets", Federal Reserve Bank of New York *Current Issues in Economics and Finance*, 14.

[7] Beber, A., and M.W. Brandt (2010), "When It Cannot Get Better or Worse: The Asymmetric Impact of Good and Bad News on Bond Returns in Expansions and Recessions", *Review of Finance*, 14, 119-155.

[8] Berry, T.D., and K.M. Howe (1994), "Public Information Arrival", *Journal of Finance*, 49, 1331-1346.

[9] Biais, B., L.R. Glosten, and C. Spatt (2005), "Market Microstructure: A Survey of Microfoundations, Empirical Results, and Policy Implications", *Journal of Financial Markets*, 8, 217-264.

[10] Bollen, N.P.B., and R.E. Whaley (2004), "Does Net Buying Pressure Affect the Shape of Implied Volatility Functions?", *Journal of Finance*, 59, 711-753.

[11] Bollerslev, T., J. Cai, and F.M. Song (2000), "Intraday Periodicity, Long Memory Volatility, and Macroeconomic Announcement Effects in the US Treasury Bond Market", *Journal of Empirical Finance*, 7, 37-55.

[12] Boyd, J.H., J. Hu, and R. Jagannathan (2005), "The Stock Market's Reaction to Unemployment News: Why Bad News Is Usually Good for Stocks", *Journal of Finance*, 60, 649-672.

[13] Brandt, M.W., and K.A. Kavajecz (2004), "Price Discovery in the U.S. Treasury Market: The Impact of Orderflow and Liquidity on the Yield Curve", *Journal of Finance*, 59, 2623-2654.

[14] Brenner, M., P. Pasquariello, and M. Subrahmanyam (2009), "On the Volatility and Comovement of U.S. Financial Markets around Macroeconomic News Announcements", *Journal of Financial and Quantitative Analysis*, 44, 1265-1289.

[15] Chordia, T., A. Sarkar, and A. Subrahmanyam (2005), "An Empirical Analysis of Stock and Bond Market Liquidity", *Review of Financial Studies*, 18, 85-129.

[16] Cutler, D.M., J.M. Poterba, and L.H. Summers (1989), "What Moves Stock Prices?", *Journal of Portfolio Management*, 15, 4-12.

[17] De Goeij, P. and W. Marquering (2006), "Macroeconomic Announcements and Asymmetric Volatility in Bond Returns", *Journal of Banking and Finance*, 30, 2659-2680.

[18] Ederington, L., and J. Lee (1993), "How Markets Process Information: News Releases and Volatility", *Journal of Finance*, 45, 1161-1191.

[19] Engle, R.F., and J. Rangel (2008), "The Spline-GARCH Model for Low-Frequency Volatility and Its Global Macroeconomic Causes", *Review of Financial Studies*, 21, 1188-1222.

[20] Evans, M., and R. Lyons (2002), "Order Flow and Exchange Rate Dynamics", *Journal of Political Economy*, 110, 170-180.

[21] Evans, M., and R. Lyons (2008), "How is Macro News Transmitted to Exchange Rates?", *Journal of Financial Economics*, 88, 26-50.

[22] Faust, J., J.H. Rogers, S.-Y.B. Wang, and J.H. Wright (2007), "The High-Frequency Response of Exchange Rates and Interest Rates to Macroeconomic Announcements", *Journal of Monetary Economics*, 54, 1051-1068.

[23] Fleming, M.J., and E.M. Remolona (1999), "Price Formation and Liquidity in the U.S. Treasury Market: The Response to Public Information", *Journal of Finance*, 54, 1901-1915.

[24] French, K.R., and R. Roll (1986), "Stock Return Variances: The Arrival of Information and the Reaction of Traders", *Journal of Financial Economics*, 17, 5-26.

[25] Garleanu, N., L.H. Pedersen, and A. Poteshman (2009), "Demand-Based Option Pricing", *Review of Financial Studies*, 22, 4259-4299.

[26] Garman, M. (1976), "Market Microstructure", *Journal of Financial Economics*, 3, 257-275.

[27] Glosten, L.R., and P. Milgrom (1985), "Bid, Ask and Transaction Prices in a Specialist Market with Heterogeneously Informed Agents", *Journal of Financial Economics*, 14, 71-100.

[28] Green, T.C. (2004), "Economic News and the Impact of Trading on Bond Prices", *Journal of Finance*, 59, 1201-1233.

[29] Hasbrouck, J. (1991), "Measuring the Information Content of Stock Trades", *Journal of Finance*, 46, 179-207.

[30] Hasbrouck, J. (2004a), "Empirical Market Microstructure: The Institutions, Economics, and Econometrics of Securities Trading", Oxford University Press, USA.

[31] Hasbrouck, J. (2004b), "Liquidity in the Futures Pit: Inferring Market Dynamics from Incomplete Data", *Journal of Financial and Quantitative Analysis*, 39, 305-326.

[32] He, Y., H. Lin, J. Wang, and C. Wu (2009), "Price Discovery in the Round-the-Clock U.S. Treasury Market", *Journal of Financial Intermediation*, 18, 464-490.

[33] Jiang, G.J., and I. Lo (2011), "Private Information Flow and Price Discovery in the U.S. Treasury Market", Bank of Canada Working Paper, 2011-5.

[34] Kyle, A.S. (1985), "Continuous Auctions and Insider Trading", *Econometrica*, 53, 1315-1335.

[35] Lee, C., and M. Ready (1991), "Inferring the Trade Direction from Intraday Data", *Journal of Finance*, 46, 733-746.

[36] Manaster, S., and S.C. Mann (1996), "Life in the Pits: Competitive Market Making and Inventory Control", *Review of Financial Studies*, 9, 953-975.

[37] Pasquariello, P., and C. Vega (2007), "Informed and Strategic Order Flow in the Bond Market", *Review of Financial Studies*, 20, 1975-2019.

[38] Van der Wel, M., A.J. Menkveld, and A. Sarkar (2009), "Are Market Makers Uninformed and Passive? Signing Trades in the Absence of Quotes", Working paper.