AARHUS UNIVERSITY
BUSINESS AND SOCIAL SCIENCES
DEPARTMENT OF ECONOMICS AND BUSINESS

CREATES
Center for Research in Econometric Analysis of Time Series

On the Effects of Private Information on Volatility

Anne Opschoor, Michel van der Wel, Dick van Dijk
and Nick Taylor

CREATES Research Paper 2012-08

Department of Economics and Business
Aarhus University
Bartholins Allé 10
DK-8000 Aarhus C
Denmark

Email: oekonomi@au.dk
Tel: +45 8716 5515

On the Effects of Private Information on Volatility

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

This version: February 20, 2012†

Abstract

We study the impact of private information on volatility. 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 effect of public information on volatility. Using high-frequency 30-year U.S. Treasury bond futures data, we find that private information, measured by order flow, is statistically and economically significant for explaining volatility. Private information is more important than public information, with the effect of an order flow shock on volatility being 18% larger than the effect of the most influential macroeconomic announcement.

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

JEL: G14, E44.

---

*Corresponding author, e-mail address: vanderwel@ese.eur.nl. Postal address: Erasmus University Rotterdam, Erasmus School of Economics, P.O.Box 1738, 3000 DR, Rotterdam, The Netherlands

†Anne Opschoor is from Erasmus University Rotterdam and the Tinbergen Institute. Van der Wel is from Erasmus University Rotterdam, CREATES, Tinbergen Institute and ERIM. Van Dijk is from Erasmus University Rotterdam, Tinbergen Institute and ERIM. Taylor is from Cardiff Business School, Cardiff University, UK. We appreciate the comments of participants at the European Finance Association meetings (Stockholm, August 2011), the 7th Annual Central Bank Workshop on the Microstructure of Financial Markets (Stavanger, August 2011), the Marie Curie Initial Training Network workshop on High Frequency Research (Berlin, May 2011), the 4th Erasmus Liquidity Conference (Rotterdam, June 2011) and the 4th International Conference on Computational and Financial Econometrics (London, December 2010). We are grateful to Mark van Achter, Alain Chaboud, Yong Chen, Albert Menkveld, Asani Sarkar and Christian Wagner for their detailed comments. Michel van der Wel is grateful to Netherlands Organisation for Scientific Research (NWO) for a Veni grant; and acknowledges support from CREATES, funded by the Danish National Research Foundation. We are responsible for all errors.

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 are equally informed at the same time, the arrival of public information typically causes an immediate change in asset prices and volatility. Private information refers to news that is distributed asymmetrically among 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.[1]

The effect of public information on prices and volatility and the effect of private information on prices have been established for many financial markets.[2] 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 (Hasbrouck, 1991), foreign exchange (Evans and Lyons, 2008), and the Treasury market (Brandt and Kavajecz, 2004; Green, 2004; Pasquariello and Vega, 2007; Menkveld, Sarkar and Van der Wel, 2011).

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 futures, using high-frequency (5 minutes) intraday data for the period 2004-2009.

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

[2]This literature dates back to (at least) French and Roll (1986) and Cutler, Poterba and Summers (1989), and includes important contributions by Ederington and Lee (1993), Berry and Howe (1992), Fleming and Remolona (1999), Balduzzi, Elton and Green (2001), Boyd, Hu and Jagannathan (2005), Andersen, Bollerslev, Diebold and Vega (2003, 2007), Faust, Rogers, Wang and Wright (2007), Bartolini, Goldberg and Sacarny (2008), and Brenner, Pasquariello and Subrahmanyam (2009).

We conduct our analysis in a comprehensive modeling 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. We proxy private information by order flow, which is defined as the difference between the volume in buyer-initiated transactions and the volume in seller-initiated transactions.

Our results demonstrate that this proxy indeed affects volatility. Specifically, a higher 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 a one standard deviation shock in order flow on volatility is around 8.0 basis points at the daily level. Interestingly, we find that the effect of private information on volatility is 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 6.8 basis points, which is 15% lower than the daily 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 volatility.

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 among the most actively traded long-term interest rate contracts 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.

Two ideas need some further clarification. First, the use of order flow (also sometimes referred to as signed volume) to proxy private information. As mentioned earlier, order flow measures the difference between buyer-initiated and seller-initiated trade volume. Trades get initiated by traders wanting to trade immediately. For this immediacy the initiaters have to pay a premium to traders willing to facilitate the need for such immediacy. The idea is best illustrated in a limit order market setting, where trade initiators would submit orders meeting or crossing prices of previously submitted limit orders and the providers of immediacy are the traders that submitted these previous limit orders. The aforementioned literature on private information and prices finds that a reason for paying for such immediacy is trading due to (short-lived) informational advantages.[3] In this literature, order flow is found to be significant in explaining returns. This significance is interpreted as order flow transmitting information relevant to prices to the market. As this information has not been incorporated in prices earlier, order flow is seen as to reveal information that is held asymmetrically among market participants. A large order flow (in absolute sense) thus coincides with a period of high information asymmetry and thus a high level of private information.

A second idea worth clarifying further is how private information in the Treasury market can be thought of. 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 straightforwardly 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 such as order flow.

Apart from uncovering the effect of private information on volatility, 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 returns and volatility. Specifically, we propose a model specifying the dynamics of both returns and volatility in such a way that both equations

[3] Based on this idea, Sarkar and Schwartz (2009) develop a measure of 'market sidedness': a market is more one-sided (two-sided) if the correlation between the numbers of buyer- and seller-initiated trades decreases (increases). Their analysis shows that trading is more one-sided in markets with a large degree of information asymmetry, such as prior to merger news.

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.

We provide several robustness checks and extensions of our main result that private information affects volatility. First, we investigate the use of the bid-ask spread as an alternative proxy of private information. This does not change the main result that private information matters for volatility. Second, we document that the effect of private information depends on the state of the economy, in the sense that the effect of order flow on volatility is higher during recessions. Third, we find that the effect of private information on volatility depends on the heterogeneity of analysts' expectations concerning upcoming public information releases. Order flow influences volatility to a greater extent in times characterized by a high level of dispersion of beliefs.

The remainder of the paper is organized as follows. Section 2 discusses related literature. Section 3 describes the data and presents summary statistics. Section 4 develops the joint modeling framework for analyzing the effects of public and private information on returns and volatility and provides the main empirical results. Section 5 reports extensions of the main analysis, including the effect of the bid-ask spread on volatility, and allowing for the effects of public and private information to depend on the state of the economy or on the level of disagreement among analysts concerning upcoming public information releases. Section 6 concludes.

2 Related Literature

Only a few papers study issues related to private information and volatility. The paper closest to ours is Berger, Chaboud and Hjalmarsson (2009), who take a long term perspective and find volatility variation is directly related to information flow and the impact of that information flow on prices for the foreign exchange market. In the analysis they use order flow to measure information flow. We follow suit but label order flow as a measure for private information. We do so as in addition we consider public information to explicitly contrast the two and determine the relative contribution of each.[4] Berger, Chaboud and Hjalmarsson (2009) consider daily volatility, measured by the realized variance being the sum of squared returns over all 5-minute intervals in the day. We differ by focusing on the raw higher-frequency data at the 5-minute level, and jointly modeling return and volatility to ensure no return effects are attributed to volatility (and vice versa).

Related contributions identify private information effects on volatility only through 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 consider a link between 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, a number that includes the impact through order flow. Likewise, He, Lin, Wang and Wu (2009) start with a structural model for return changes due to public information shocks and information asymmetry and obtain a model-implied variance of price changes in the US Treasury market. In contrast to these two papers, in our work the term volatility implies a parameterization of the conditional volatility. Within this conditional context, other studies have considered alternate measures of private information flow. For instance, 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.

Several strands of literature touch related subjects. First, Chordia, Sarkar and Subrahmanyam (2005) analyze the comovement of returns, volatility, liquidity and order flow. One

[4] In footnote 18 Berger, Chaboud and Hjalmarsson (2009) point out that the largest residuals remaining after modeling information flow are due to important macro news announcements. This finding highlights the importance of incorporating macro news announcements in the model as we do. In fact, it may be the case that part of the effect Berger, Chaboud and Hjalmarsson (2009) attribute to information flow is actually due to macro news releases as literature shows order flow is more informative around such releases (Green, 2004).

of their main findings is that volatility shocks are informative for predicting shifts in liquidity, highlighting the interrelation between the two. We 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)). Second, our analysis of private information effects on volatility relates to the literature on the effects of buying- and selling-pressure in options markets (Bollen and Whaley, 2004; Garleanu, Pedersen and Poteshman, 2009). The focus in this literature is typically on implied volatility rather than option prices directly, as a way of controlling for price differences due to option contract specification (such as strike and maturity).

Obviously there is a relation between order flow and volume. Order flow is defined as the difference between buyer-initiated and seller-initiated trades, while volume is the sum of all trades. Literature on the relation between volume and volatility dates back to Clark (1973) and Tauchen and Pitts (1983), who develop the mixture of distribution hypothesis for the joint movement of return and volume based on an unobserved information arrival variable. Andersen (1996) modifies this framework to include both information asymmetry and liquidity needs, and shows it is useful for analyzing the economic factors behind volatility clustering. Further empirical evidence on the volume-volatility relation is provided by Karpoff (1987), who also surveys the literature. For our purpose it is important to realize that private information is measured not by the total amount of trading that takes place, but by how this volume can be attributed to either buying- or selling-pressure.

Finally, there is also a link between private information and liquidity. For example, Green (2004) finds that information asymmetry increases when liquidity is high while Brandt and Kavajecz (2004) find the opposite. In addition, Jiang and Lo (2011) argue that heterogeneous private information is followed by low trading volume, low total market depth and hidden depth. If we only include (unanticipated) order flow in our volatility specification, we discard the effect of liquidity (shocks) on volatility. Jiang and Lo (2011) control for trading volume to differentiate between private information and liquidity. We follow their approach and include volume as a control variable in our analysis.

3 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 an illustration for the effects of public and private information on Treasury futures returns and volatility, focusing on the announcements of Nonfarm Payroll Employment.

3.1 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 Time (ET). Electronic trading occurs through Globex from 6.30 p.m. to 5.00 p.m. from Sunday to 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.[5] We construct a single time series of transaction prices and volumes using the most nearby contract, which is the most intensely 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, obtained from Tickdata Inc., records the timestamp (in seconds), price and volume for

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

each transaction. We aggregate the data to 5-minute intervals.[6] For computing returns we use the last observed transaction price in each intraday interval. Overnight returns are excluded.

3.2 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. We use the first one for our main analysis, while the second proxy is used in the robustness section.

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 among participants.

We follow Pasquariello and Vega (2007) by using the unanticipated portion of aggregate order flow. Specifically, we clean order flow from autocorrelations that are (partly) due to microstructure imperfections causing 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 scale order flow by its empirical standard deviation. Related to the definition of order flow OF_t as given above, this measure is also referred to as the 'signed' trading volume. In a second step, we estimate an appropriate ARMA model to remove the 'expected' order flow.[7]

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

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

The residuals of this model represent the aggregate unanticipated order flow OF_t^* over each time interval t.[8] While our empirical analysis is based on unanticipated order flow, we refer to order flow and unanticipated order flow interchangeably.

Our second measure for private information is the bid-ask spread during a certain interval. 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) average buy and sell prices during a particular interval, see also Manaster and Mann (1996). We set negative spread estimates equal to zero and trim observations that exceed the 99.95th percentile of the empirical distribution to avoid having noisily measured spreads dominating our results.

The bid-ask spread proxy may be slightly more noisy, as beyond effects due to asymmetric (private) information it could also include other market microstructure effects such as costs for holding inventory and order processing (Biais, Glosten and Spatt, 2005). To examine to which extent this is the case we follow Huang and Stoll (1997) and decompose the bid-ask spread into these components.[9] In additional results, available from the authors upon request, we show that over a longer horizon a substantial part of the spread (20%) is attributable to asymmetric information.

Our data set does not include the 'sign' variable $q_{t,j}$, which indicates whether a trade is initiated by the buying or selling party. Hasbrouck (2004b) proposes to estimate the sign following sequential trade models such as Roll (1984). In these models the transaction price series is decomposed into an unobserved efficient price and a deviation from this efficient price which is equal to the (half-)spread times the trade sign. The trade sign is thus also an unobserved series that can be filtered from the data. This filtering can be done through Monte Carlo techniques (as Hasbrouck (2004b) proposes), or by state space estimation (Van der Wel, Menkveld and Sarkar (2009)). To apply this method on our data set consisting of on average 15,562 daily transactions over 6 years, we utilize the latter approach.[10]

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

[9]Ideally we would do such a decomposition to obtain a decomposed spread series at the 5-minute frequency, which is the same frequency as our aggregated futures data. Unfortunately a 5-minute level spread decomposition requires much finer higher-frequency data with a good amount of transactions, which is not always available. For this reason we do not pursue this approach, and keep order flow as our main proxy and take the bid-ask spread (not decomposed) as alternative.

[10]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

Table 1: Summary statistics

This table shows the mean, standard deviation, skewness, and kurtosis, as well as the minimum and maximum values, over the full 2004-2009 sample period for the variables in our analysis. We show return (in %, denoted with R_t) which is defined as 100 times the difference between the logarithm of the last observed transaction price in each intraday interval, the total number of trades in each interval (N_t), order flow (OF_t) with unanticipated order flow (OF_t^*), the total trading volume of the trades in each interval (Qnt_t), and the spread (Spr_t) computed as the difference between the volume weighted buy price and the average sell price during each intraday interval. All numbers are based on 157,500 aggregated 5-minute observations, the sample goes from January 2, 2004 through December 31, 2009.

Variable		Mean	Std.dev.	Skewness	Kurtosis	Min	Max
Return (in %)	R_t	0.00	0.058	0.599	118.0	-2.21	3.37
Number of transactions	N_t	144	157	3.527	32.3	0	3,924
Order flow	OF_t	-1	1,249	-0.039	11.7	-16,754	16,351
Unanticipated order flow	OF_t^*	0	1,249	-0.041	11.6	-16,750	16,385
Trading volume	Qnt_t	2,335	2,578	3.362	24.1	0	63,924
Spread (in $)	Spr_t	0.0186	0.0170	1.660	15.7	0	0.3250

Table 1 shows summary statistics of the data. All numbers are based on aggregated 5-minute data for the full sample period 2004-2009. The data set covers 1,420 full days with observations from 8:00 a.m. to 5:00 p.m. and 69 'half days' with observations from 8:00 until 1:00 p.m. The latter occurs when the market is partly closed in the afternoon, for example when the preceding day was Thanksgiving. Mean futures returns are small or zero, as is expected for 5-minute observation frequency. The volatility is 5.8 bps at 5-minute level, which roughly translates into 53 basic points (bps) daily volatility. The table also shows the high activity of the 30-year futures market as there are on average 144 trades in each 5-minute interval. In addition, the difference in the standard deviations of order flow and unanticipated order flow is rather small, showing that the AR(1) coefficient is close to zero or, in other words, the autocorrelation in order flow is rather limited.

3.3 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. These data are obtained from Econoday, and are derived from information that is published

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 order flow and the the bid-ask spread based on the tick-test.

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 International and Thomson Financial. To obtain the analysts' forecasts the aforementioned companies hold a survey among 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 (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 the analysts forecasts, and σ_k is the sample standard deviation of their difference.

Table 2 lists the 24 macroeconomic variables in our data set.[11] 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. Out of the 1,489 days that are covered by our dataset, 1,000 days contain at least one announcement. 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 (partly) closed on the day of the announcement. On average, for each (monthly) macro variable we are able to use 70 of the 72 announcements. For the later part of the sample, since June 2007, we also have the minimum and maximum analysts' forecasts available. We will use this information to calculate a dispersion measure that we use in one of the extensions in Section 5. Table 2 shows for each variable the number of minimum and maximum forecasts that are available.

[11] We do not include FOMC announcements of the federal funds target rate since the surprises for this variable are mostly equal to zero during our sample period.

Table 2: 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[a]
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	8.30 a.m.	72	69	31
5. Retail sales	8.30 a.m.	72	70	31
6. Industrial production	9.15 a.m.	72	69	31
7. Capacity utilization	9.15 a.m.	72	69	31
8. Personal income	8.30 a.m.	72	71	33
9. Consumer credit	3.00 p.m.	72	63	31
Consumption				
10. Personal consumption exp.	8.30 a.m.	72	71	33
11. New home sales	10.00 a.m.	72	69	32
Investment				
12. Durable goods orders	8.30 a.m.	72	70	32
13. Construction spending	10.00 a.m.	72	72	31
14. Factory orders	10.00 a.m.	72	72	31
15. Business inventories	10.00 a.m.[b]	72	69	30
Government Purchases				
16. Government budget deficit	2.00 p.m.[c]	72	65	29
Net Exports				
17. Trade balance	8.30 a.m.	72	71	31
Prices				
18. Producer price index	8.30 a.m.	72	69	31
19. Consumer price index	8.30 a.m.	72	71	32
Forward-looking				
20. Consumer confidence index	10.00 a.m.	72	72	33
21. NAPM index	10.00 a.m.[d]	72	72	33
22. Housing starts	8.30 a.m.	72	71	30
23. Index of leading indicators	10.00 a.m.	73	68	32
Weekly				
24. Initial unemployment claims	8.30 a.m.	312	304	135

[a] Dispersions available only after June 2007. [b] Earlier part of sample often at 8:30.
[c] Announcement at 1:30 on October 14, 2005; at 1:00 on October 14, 2008 and at 3:30 on October 16, 2009. [d] Latest part of sample at 9:45 (since January 2007).

3.4 Public and private information effects: A first impression

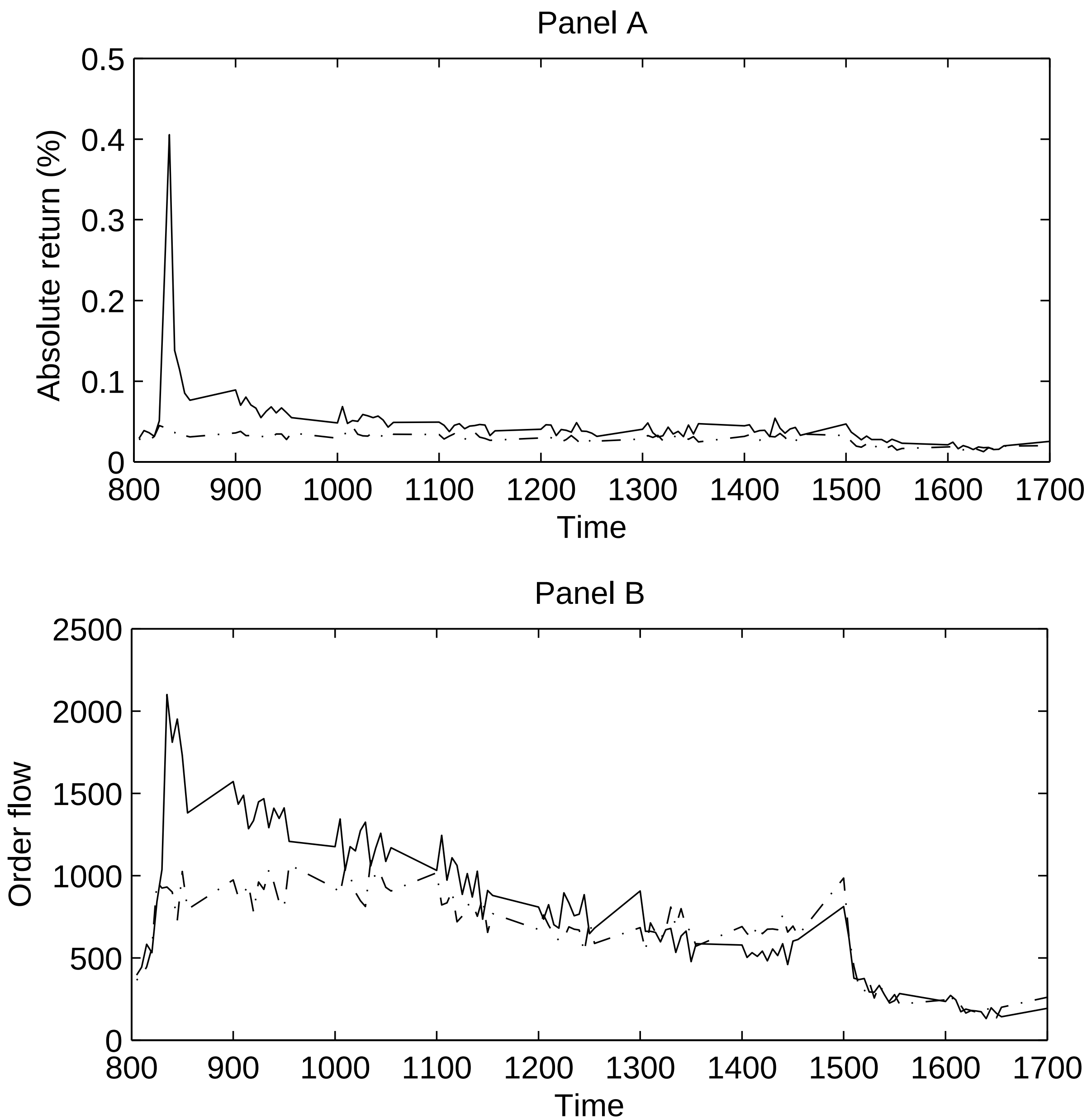
Figure 1 gives a first impression of the effect of public and private information on volatility. The chart considers patterns in the volatility and order flow during all (69) days with scheduled Nonfarm Payroll Employment announcements at 8:30 during our sample period. For this purpose, we compute the average of these variables across these announcement days (where we use the absolute value of order flow) as well as corresponding averages across 69 randomly selected trading days without any macro announcement. Figure 1 depicts the average volatility and order flow during these announcement and non-announcement days. The graphs provide an important insight. Volatility spikes during the interval after the announcement time, reflecting the initial reaction of the market to the announcement. A spike immediately after the announcement is also seen for unanticipated order flow, though less pronounced. After the announcement, volatility does not immediately revert to its pre-announcement level. Similarly, unanticipated order flow also remains at a higher level for some time after the announcement. It is this pattern in volatility and order flow that encourages our research question whether and how private information affects volatility.

4 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 and provide our main results. 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 these can be estimated jointly. We develop our general modeling framework in different stages. First, we consider the model specification for 5-minute intraday returns. Second, we augment this with a specification for (conditional) volatility.

Figure 1: Order flow and volatility during announcement and non-announcement days

This figure shows time series of the volatility 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 69 trading days where no announcements were made (dotted line), and 69 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.



4.1 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 unexpected order flow as the proxy for private information, such that the model reads

$$R_t = \beta_0 + \beta_a D_a + \sum_{i=1}^{I} \beta_i R_{t-i} + \sum_{k=1}^{K} \sum_{j=0}^{J} \beta_{kj} S_{k,t-j} + \beta_{OF} OF_t^* + \varepsilon_t, \tag{3}$$

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 D_a is a dummy variable that equals one on announcement days. 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 heteroscedasticity (and any remaining serial correlation not captured by the autoregressive terms) by using Newey-West standard errors.

Table 3 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 order flow. 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 $J = 0$. 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 Treasury bond futures returns. The magnitude and significance of the coefficients are not affected by the inclusion of the private information variable. 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

Table 3: Effect of public and private information on returns

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

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

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: 30-year 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	-0.306^{***}	-0.290^{***}
5. Retail sales	-0.105^{***}	-0.097^{***}
6. Industrial production	0.003	0.002
7. Capacity utilization	-0.016	-0.010
8. Personal income	-0.023	-0.019
9. Consumer credit	-0.003	-0.002
Consumption		
10. Personal consumption expenditures	-0.020	-0.017
11. New home sales	-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	0.014
Government Purchases		
16. Government budget deficit	0.006	0.008
Net Exports		
17. Trade balance	-0.026^{**}	-0.020^{**}
Prices		
18. Producer price index	-0.054^{**}	-0.041^{**}
19. Consumer price index	-0.036	-0.037
Forward Looking		
20. Consumer confidence index	-0.063^{***}	-0.054^{***}
21. NAPM index	-0.020	-0.022
22. Housing starts	-0.015	-0.020
23. Index of leading indicators	-0.047^{***}	-0.039^{***}
24. Initial unemployment claims	0.041^{***}	0.036^{***}
R^2	0.019	0.212
Nr. Observations	157,500	157,500

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 procyclical variables such as Nonfarm Payroll Employment, Retail Sales and New Home Sales, indeed affect bond prices negatively, while countercyclical 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 affects 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 30 basis points in the futures returns.

The estimation results in the rightmost column of Table 3 demonstrate that unanticipated order flow significantly influences returns, which is consistent with prior literature (Evans and Lyons, 2002; Green, 2004). Including unanticipated order flow increases the regression R^2 from 1.9% to 21.2%. A decomposition of the explained return variance assigns 91% to (orthogonalized) order flow and 9% to public information variables. Hence order flow accounts for roughly 20% of the price variance. This number is in line with the literature, as Evans and Lyons (2008, Table 6) report a corresponding number of 22%. 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 positive shock of order flow increases the return by two and a half basis points. Although this number is lower than for example the corresponding value of Nonfarm Payroll Employment, we 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 reduction in seller-initiated trades. In both cases, this signals positive private information, which increases 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 1.9% (when order flow is not included). 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 tiny fraction of the total sample size. To highlight the importance

Table 4: Effect of public information on returns during announcements

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

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

where R_t is the 5-minute log return of the 30-year U.S. Treasury bond futures from period $t-1$ to t and $S_{k,t}$ is the standardized surprise of announcement type k, $k = 1, \ldots, 24$. We include only the pairs $(R_t, S_{k,t})$ when an announcement of fundamental k was made at time t. The superscripts ***, ** and * indicate significance at the 1%, 5% and 10% level respectively. The sample goes from January 2, 2004 through December 31, 2009.

Dependent variable: 30-year futures returns			
	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	-0.297^{***}	0.268	69
5. Retail sales	-0.110^{***}	0.223	70
6. Industrial production	-0.011	0.009	68
7. Capacity utilization	-0.014	0.022	69
8. Personal income	-0.019	0.021	71
9. Consumer credit	-0.003	0.003	63
Consumption			
10. Personal consumption expenditures	-0.019	0.022	71
11. New home sales	-0.054^{***}	0.237	69
Investment			
12. Durable goods orders	-0.073^{***}	0.135	70
13. Construction spending	-0.020	0.013	72
14. Factory orders	-0.017	0.019	72
15. Business inventories	0.014	0.011	69
Government Purchases			
16. Government budget deficit	0.007	0.011	65
Net Exports			
17. Trade balance	-0.029^{**}	0.057	71
Prices			
18. Producer price index	-0.071^{***}	0.132	69
19. Consumer price index	-0.031	0.017	71
Forward Looking			
20. Consumer confidence index	-0.067^{***}	0.285	72
21. NAPM index	-0.016	0.014	72
22. Housing starts	-0.009	0.003	71
23. Index of leading indicators	-0.046^{***}	0.271	68
24. Initial unemployment claims	0.039^{***}	0.075	304

of macro news during the announcement periods, we estimate the simplified model

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

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 4. 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 20% for the Consumer Confidence Index, Nonfarm Payroll Employment, the Conference Board Index of Leading Indicators, New Home Sales, and Retail Sales.

4.2 Public and private information also matter for volatility

ABDV (2003, 2007) focus on modeling the effects of macro announcements on returns. 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 capturing the intraday pattern of volatility, that is,

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

where $|\hat{\varepsilon}_t|$ is the absolute value of the residual of equation (3), and the sine and cosine terms aim to capture the intraday volatility pattern. As we consider 5-minute returns during the period from 8:00 a.m. to 5:00 p.m., we set N equal to 108 (9 hours $\times$ 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 can produce negative fitted values. This happens when, e.g., 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.[12] 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. We adopt the basic idea of the Engle and Rangel (2008) Spline-GARCH model, and decompose the conditional volatility in 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 contains the intra-day volatility pattern (by means of a Fourier series), dummy variables 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

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

given by

$$
\begin{aligned}
R_t &= \beta_{0,m} + \beta_{a,m} D_a + \sum_{i=1}^{I} \beta_i R_{t-i} + \sum_{k=1}^{K} \sum_{j=0}^{J} \beta_{kj} S_{k,t-j} + \beta_{OF,m} OF_t^* + \varepsilon_t, \\
\varepsilon_t &= \sqrt{g_t \tau_t} u_t, \\
g_t &= (1 - \alpha - \beta_{GA}) + \alpha \left(\frac{\varepsilon_{t-1}^2}{\tau_{t-1}} \right) + \beta_{GA} g_{t-1}, \\
\tau_t &= \exp \Bigg(\beta_{0,v} + \beta_{a,v} D_a + \sum_{q=1}^{Q} \left(\delta_q \cos(\frac{q 2\pi t}{N}) + \phi_q \sin(\frac{q 2\pi t}{N}) \right) + \\
&\qquad \sum_{k=1}^{K} \sum_{j'=0}^{J'} \beta_{kj'} |S_{k,t-j}| + \beta_{OF,v} |OF_t^*| + \beta_{Qnt} Qnt_t^* \Bigg)
\end{aligned}
\tag{6}
$$

where $u_t \sim N(0,1)$, the conditional variance $E[\varepsilon_t^2 | \mathcal{I}_{t-1}] = g_t \tau_t$ 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. The exponential specification of τ_t obviously avoids the problem whereby the conditional variance can become negative by construction. Finally, Qnt_t^* denotes unanticipated volume, which is added as a control variable to account for the link between private information and liquidity as discussed in Section 2.[13]

Both De Goeij and Marquering (2006) and Brenner, Pasquariello and Subrahmanyam (2009) model volatility effects of macroeconomic announcements through GARCH models. We differ in set-up not only by in addition considering private information, but also by use of the Spline-GARCH model. An important difference of our set-up compared to the augmented-GARCH specifications of these two papers is that the latter model implies by definition a persistent effect of surprises on volatility. This is the result of, in effect, introducing a dummy variable that accounts for the announcements in the GARCH equation. In this case, a high impact of macro announcements through the GARCH coefficients (that is, the usual α and β from, e.g., a GARCH(1,1) specification) always models a persistent effect of public information on volatility.

The Spline-GARCH model allows controlling for the length of the effect of the surprises by simply adding lags of surprises in the specification of τ_t, which does not influence g_t.

[13]Similar to the use of unanticipated order flow, we include the unanticipated part of volume, Qnt_t^*. Specifically, we apply an ARMA(2,1) filter on the logarithm of volume after controlling for the intraday effects of volume via a Fourier series.

Hence we are able to distinguish between the effect of private information on volatility and the effect of macroeconomic surprises on volatility.

Table 5 provides estimation results for the GARCH model in (6) based on the complete sample period from January 2004 until December 2009. We set $Q = 5$ and $J^{'} = 1$ in the Spline-GARCH model and estimate all parameters in the model (6) simultaneously, by (quasi) maximum likelihood. Panel A reports results of three different specifications. The first specification does not include private information. The other two specifications add private information to the model and controls for liquidity effects by including the volume of trades. The order flow impact of the return is significant, though somewhat lower than when only the return equation is estimated: compared to Table 3 the coefficient decreases from 0.025 to 0.022. Interestingly, Panel A of the table indicates that (unanticipated) order flow has a significant positive effect on volatility, after controlling for public information and liquidity effects. The positive coefficient on volatility suggests that the learning process of the agents to clear the market increases uncertainty about the bond returns. The market aggregates private information which in turn affects the returns and uncertainty on the futures market. In addition, the table shows the importance of considering volume as additional explanatory variable for volatility. The coefficient of order flow in the volatility equation does reduce, from 0.43 to 0.26 when volume is included, but remains statistically significant. Thus, liquidity effects only partly explain our result that private information, measured by order flow, matters for volatility.[14]

Panel B of Table 5 reports the estimated coefficients for the surprises in macro announcements corresponding with the third specification of Panel A.[15] The results of the conditional mean equation indicate that 13 out of the 24 announcement surprises significantly affect subsequent returns, where in general the same reasoning holds as described in the discussion about Table 3, i.e. procyclical variables cause a negative bond return whereas countercyclical variables positively influence bond returns.[16] In addition, the impact of public information

[14] The table also confirms the positive relation between volatility and volume. This result is consistent with the literature on the volume-volatility relation that is mentioned in Section 2. In fact, specification (3) of Table 5 shows that both volume and order flow are important for explaining volatility.

[15] The estimated public information coefficients corresponding to the other two specifications are of similar order.

[16] As a separate analysis, we include dummy variables to capture the effect that the mere presence of an announcement boosts volatility, apart from the size of the associated surprise, see ABDV (2003). Results suggest that indeed the presence of an announcement matters for volatility. In line with previous literature,

Table 5: Public and private information effects

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

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

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, D_a is a dummy variable that is 1 on announcement days and 0 else, $S_{k,t}$ is the standardized news announcement for $k = 1, \ldots, 24$, OF_t^* denote the unanticipated order flow and Qnt_t^* represents the unanticipated volume of trades from period t to $t+1$. Further, $I = 3$, $J = 0$, $J' = 1$ and $Q = 5$. We show only the effects of public and private information variables on returns and volatility. In Panel A we report parameter estimates corresponding with private news for three various specifications, whereas panel B provides the public news coefficients of the last specification. In Panel B the first column denotes the contemporaneous effect while the second column denotes the one-period lagged effect of public information on volatility. 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: Private information and trading volume			
Dependent variable: 30-year futures returns			
	(1)	(2)	(3)
Conditional mean equation			
Order flow		0.022***	0.022***
Macro announce.	yes	yes	yes
Conditional volatility equation			
Order flow		0.431***	0.256***
Trading volume			0.327***
Macro announce.	yes	yes	yes
log-likelihood	257,562	284,311	289,425
Nr. Observations	157,500	157,500	157,500

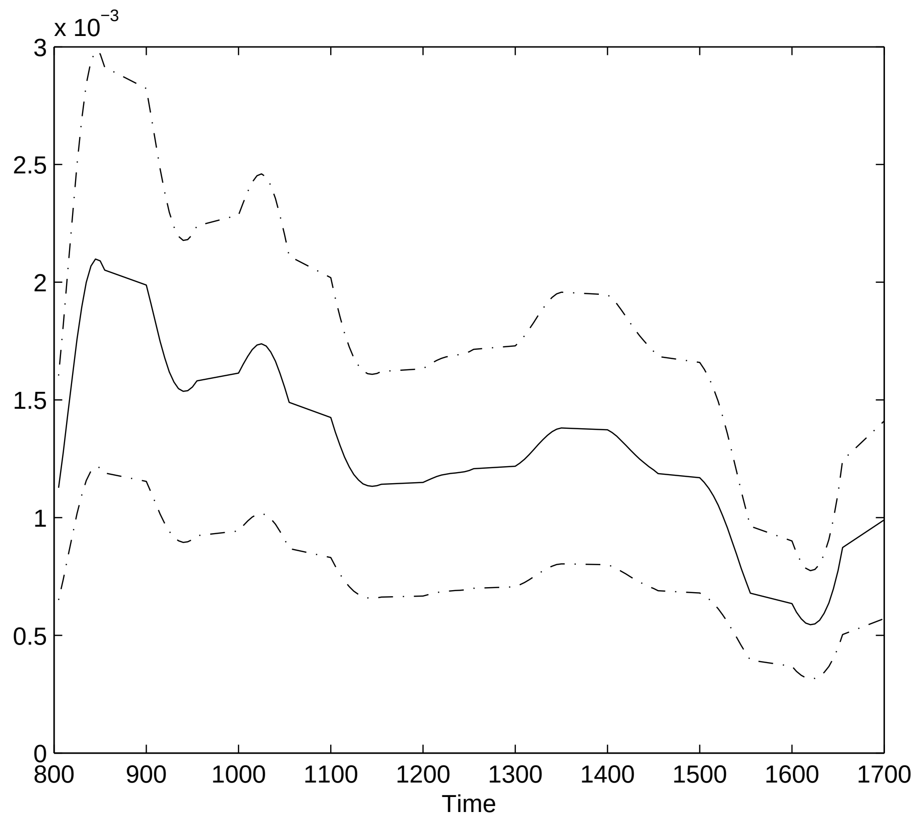
(continued from previous page)

Panel B: Public information			
	Cond Mean	Cond Volatility	
1. GDP advance	−0.125**	1.137***	0.291
2. GDP preliminary	−0.033	0.687***	−0.124
3. GDP final	0.003	0.374*	0.006
Real Activity			
4. Nonfarm payroll employment	−0.228***	2.264***	0.717***
5. Retail sales	−0.091***	0.736***	0.451***
6. Industrial production	−0.009	0.285	0.059
7. Capacity utilization	−0.026	0.096	0.192
8. Personal income	−0.023	0.279	−0.024
9. Consumer credit	−0.002	−0.102	−0.161
Consumption			
10. Personal consumption expenditures	−0.025*	0.342*	0.015
11. New home sales	−0.061***	0.552***	0.293**
Investment			
12. Durable goods orders	−0.053**	0.800***	0.002
13. Construction spending	−0.016	0.910***	0.405***
14. Factory orders	−0.015	0.650***	0.389***
15. Business inventories	−0.003	0.070	−0.026
Government Purchases			
16. Government budget deficit	0.001	0.030	−0.005
Net Exports			
17. Trade balance	−0.024***	0.008	0.027
Prices			
18. Producer price index	−0.044**	0.743***	0.218*
19. Consumer price index	−0.076***	0.922***	0.033
Forward Looking			
20. Consumer confidence index	−0.062***	0.369***	0.195*
21. NAPM index	−0.014	0.534***	0.098
22. Housing starts	−0.024*	0.417***	−0.039
23. Index of leading indicators	−0.021***	0.002	0.144
24. Initial unemployment claims	0.035***	0.213***	−0.017

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 studies, see for example ABDV (2007). We find that 16 out of the 24 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: 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.



we consider the specification with surprises rather than dummies for our main analysis. As another additional analysis, we introduce dummy variables to interact with surprises of low and high size, allowing for a different effect of low surprises on volatility. This effect is not present in our data. For both analyses our main results on private information and volatility are qualitatively similar.

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. The remaining estimated parameters of the model are the constants, dummy variables and the GARCH parameters. Both $\beta_{0,m}$ and $\beta_{a,m}$ are almost zero and not significantly different from zero. In contrast, the corresponding values in the volatility equation ($\beta_{0,v}$ and $\beta_{a,v}$) are statistically significant. Moreover, $\beta_{a,v}$ is positive, indicating that on announcement days the overall level of volatility is higher compared to non-announcement days. Finally the GARCH parameters indicate high persistence in volatility, since β_{GA} and α are equal to 0.96 and 0.03 respectively.

Table 6: Economic significance of public and private information on volatility

This table provides economic significance of the volatility part of the Spline-GARCH model:

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

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, D_a is a dummy variable that is 1 on announcement days and 0 else, $S_{k,t}$ is the standardized news announcement for $k = 1, \ldots, 24$, OF_t^* denotes unanticipated order flow and Qnt_t^* represents unanticipated trading volume from period t to $t+1$. Further, $I = 3$, $J = 0$, $J' = 1$ and $Q = 5$. We measure the economic significance by taking the partial derivative of the conditional volatility ($\sqrt{h_t}$) with respect to the variable S_{kt} and OF_t^*. In case of order flow, we scale up the effect to a daily level. The numbers should be interpreted as the partial effect on volatility of a one standard deviation shock to the public or private variable. The sample goes from January 2, 2004 through December 31, 2009.

Economic significance in basis points		
Public information	GDP advance	3.4
	Nonfarm payroll employment	6.8
	Consumer price index	2.8
Private information	Order flow	8.0

Table 6 shows the economic significance of the estimated coefficients of the Spline-GARCH model. To assess the economic significance of a certain variable in this framework we analyze the partial derivative of the volatility specification with respect to that variable. Multiplying this partial derivative with the sample standard deviation of the variable under review provides the economic significance. To get a feeling for the economic significance of public information we consider the three news announcements that have the strongest effect on volatility (GDP advance, Nonfarm Payroll Employment and the CPI). To deal with the difference that order flow is available at each of the 5-minute intervals but announcements only in one 5-minute interval each month we consider the economic significance at the daily level. A striking result of the analysis is that the economic significance of private information is higher than that of public information. In fact, the daily impact of order flow is 18% larger than the impact of any type of announcement. A one standard deviation change in unanticipated order flow implies on average an increase of 8.0 basis points of the volatility, while a similar change in the most influential surprise, Nonfarm Payroll Employment, increases volatility by 6.8 basis points. The economic significance of the CPI and GDP advance announcements is considerably lower and equals on average roughly 35% of the economic significance of order flow.

Our result of the statistically and economically significant impact of order flow on volatility confirms the finding of Berger, Chaboud and Hjalmarsson (2009). However, one of the key differences between the two types of analyses is that we control for the effect of order flow on the first moment (returns), whereas Berger, Chaboud and Hjalmarsson (2009) solely consider (realized) volatility. Ignoring the first moment has the risk that return effects are attributed to the second moment. In fact, running the Spline-GARCH model from (6) but setting all mean effects to 0 (thus taking $R_t = \varepsilon_t$) would overestimate the economic significance of order flow by 6 basis points. An additional key difference is the inclusion of public information in the form of macro announcements when analyzing the impact of order flow on volatility. The effect of this incorporation is twofold: it identifies the impact of order flow while controlling for the effect of public information, and allows to compare the relative contribution of public and private information.

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

5 Robustness checks and extensions

We consider several modifications and extensions of the Spline-GARCH model from the previous section to examine the robustness of our result that private information matters for volatility. First, we use the bid-ask spread as an alternative proxy for private information. Second, we relate private information effects on volatility to the state of the economy. Finally, we consider whether the dispersion of beliefs among analysts influences the importance of private information for both returns and volatility.

5.1 Bid-ask spread as an alternative proxy for private information

We use order flow as the main measure of private information. In section 3.2 we discuss that the bid-ask spread could be an alternative proxy, though it may be noisier. To examine the relation between spread and volatility, we simply include it as additional explanatory variable in equation (6) to obtain as joint public and private news component τ_t:

$$\begin{aligned}\tau_t \; &= \exp\Big(\beta_{0,v} + \textstyle\sum_{v=1}^{K} \beta_v D_{a,v} + \sum_{q=1}^{Q}\left(\delta_q \cos(\frac{q2\pi t}{N}) + \phi_q \sin(\frac{q2\pi t}{N})\right) + \\ &\textstyle\sum_{k=1}^{K}\sum_{j'=0}^{J'} \beta_{kj'}|S_{k,t-j}| + \beta_{OF,v}|OF_t^*| + \beta_{Qnt} Qnt_t^* + \beta_{Spr} Spr_t\Big)\,,\end{aligned} \tag{7}$$

where Spr_t denotes the spread from period t to $t+1$, and the specifications for R_t, ε_t and g_t remain as they are in equation (6).[17]

Table 7 provide results associated with using the bid-ask spread as an alternative proxy for private information, where we only provide the estimates of the private information effects to conserve space. We run two variations of this regression. First, in column (1) we report results for the specification where only the spread is included as private variable. Interestingly, the spread is also significant in explaining volatility. The second column confirms this finding when we control for liquidity effects by means of the volume of trades. As a final variation, we include both order flow and spread while again controlling for liquidity effects. Column (3) shows that in this case both the spread and order flow are significant for explaining volatility movements. All in all, these results show that our main conclusion on the relationship between private information and volatility is robust to considering an

[17] In contrast to the volume of trades and order flow, we include the spread instead of the unanticipated spread into our analysis as the spread does not contain any ARMA structure.

Table 7: Private information and the bid-ask spread

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

$$\begin{aligned}
R_t &= \beta_{0,m} + \beta_{a,m} D_a + \textstyle\sum_{i=1}^{I} \beta_i R_{t-i} + \sum_{k=1}^{K} \sum_{j=0}^{J} \beta_{kj} S_{k,t-j} + \beta_{OF,m} OF_t^* + \varepsilon_t, \\
\varepsilon_t &= \sqrt{\tau_t g_t} u_t, \\
g_t &= (1 - \alpha - \beta_{GA}) + \alpha \left(\frac{\varepsilon_{t-1}^2}{\tau_{t-1}} \right) + \beta_{GA} g_{t-1}, \\
\tau_t &= \exp \Big(\beta_{0,v} + \beta_{a,v} D_a + \textstyle\sum_{q=1}^{Q} \left(\delta_q \cos(\frac{q 2\pi t}{N}) + \phi_q \sin(\frac{q 2\pi t}{N}) \right) + \\
&\quad \textstyle\sum_{k=1}^{K} \sum_{j'=0}^{J'} \beta_{kj'} |S_{k,t-j}| + \beta_{OF,v} |OF_t^*| + \beta_{Qnt} Qnt_t^* + \beta_{Spr} Spr_t \Big),
\end{aligned}$$

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, D_a is a dummy variable that is 1 on announcement days and 0 else, $S_{k,t}$ is the standardized news announcement for $k = 1, \ldots, 24$, OF_t^* and Qnt_t^* denote the unanticipated order flow and unanticipated trading volume respectively and Spr_t represents the bid-ask spread from period t to $t+1$. Further, $I = 3$, $J = 0$, $J' = 1$ and $Q = 5$. We show only the effects of private information variables on returns and volatility. The superscripts ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively. The sample goes from January 2, 2004 through December 31, 2009.

Dependent variable: 30-year futures returns			
	(1)	(2)	(3)
Conditional mean equation			
Order flow	0.020***	0.020***	0.021***
Macro announce.	yes	yes	yes
Conditional volatility equation			
Bid-ask spread	0.098***	0.120***	0.112***
Order flow			0.236***
Trading volume		0.411***	0.338***
Macro announce.	yes	yes	yes
log-likelihood	278,964	289,425	290,694
Nr. Observations	157,500	157,500	157,500

alternative proxy.

5.2 Conditioning on the state of the economy

We investigate how the effects of public and private information on both returns and volatility depend on the state of the economy. Previous studies, including Beber and Brandt (2010) and ABDV (2007), find that the impact of macro announcements depends on whether the economy is in expansion or recession. Similar to the existing literature, we use the NBER index to define periods of recessions and expansions. Our data set contains an expansion period from 2004 until December 2007, followed by a recession until June 2009. The remaining months of 2009 are again defined as an expansion period. We introduce a dummy variable D_e that takes the value one during expansion periods. 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:

$$
\begin{aligned}
R_t &= \beta_{0,m} + \beta_{a,m} D_a + \beta_{e,m} D_e + \sum_{i=1}^{I} \beta_i R_{t-i} + \sum_{k=1}^{K}\sum_{j=0}^{J} \beta_{e,kj} D_e S_{k,t-j} + \\
&\quad \sum_{k=1}^{K}\sum_{j=0}^{J} \beta_{r,kj} D_r S_{k,t-j} + \beta_{e,OF} D_e OF_t^* + \beta_{r,OF} D_r OF_t^* + \varepsilon_t, \\
\varepsilon_t &= \sqrt{g_t \tau_t}\, u_t, \\
g_t &= (1 - \alpha - \beta_{GA}) + \alpha \left(\frac{\varepsilon_{t-1}^2}{\tau_{t-1}} \right) + \beta_{GA} g_{t-1}, \\
\tau_t &= \exp\Bigg(\beta_{0,v} + \beta_{a,v} D_a + \beta_{e,v} D_e + \sum_{q=1}^{Q} \left(\delta_q \cos(\frac{q 2\pi t}{N}) + \phi_q \sin(\frac{q 2\pi t}{N}) \right) + \\
&\quad \beta_{e,OF} D_e |OF_t^*| + \beta_{r,OF} D_r |OF_t^*| + \beta_{e,Qnt} D_e Qnt_t^* + \beta_{r,Qnt} D_r Qnt_t^* \\
&\quad \sum_{k=1}^{K}\sum_{j'=0}^{J'} \beta_{e,kj'} D_e |S_{k,t-j}| + \sum_{k=1}^{K}\sum_{j'=0}^{J'} \beta_{r,kj'} D_r |S_{k,t-j}| \Bigg),
\end{aligned}
\tag{8}
$$

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 8 reports coefficient estimates of the adapted Spline-GARCH model, together with Wald tests for the null hypothesis that the return and volatility impact of the public and private information variables do not depend on the state of the economy. Panel A shows

Table 8: Recessions and expansions

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

$$
\begin{aligned}
R_t &= \beta_{0,m} + \beta_{a,m} D_a + \beta_{e,m} D_e + \textstyle\sum_{i=1}^{I} \beta_i R_{t-i} + \sum_{k=1}^{K}\sum_{j=0}^{J} \beta_{e,kj} D_e S_{k,t-j} + \\
&\quad \textstyle\sum_{k=1}^{K}\sum_{j=0}^{J} \beta_{r,kj} D_r S_{k,t-j} + \beta_{e,OF} D_e OF_t^* + \beta_{r,OF} D_r OF_t^* + \varepsilon_t, \\
\varepsilon_t &= \sqrt{g_t \tau_t}\, u_t, \\
g_t &= (1-\alpha-\beta_{GA}) + \alpha\left(\frac{\varepsilon_{t-1}^2}{\tau_{t-1}}\right) + \beta_{GA} g_{t-1}, \\
\tau_t &= \exp\Big(\beta_{0,v} + \beta_{a,v} D_a + \beta_{e,v} D_e + \textstyle\sum_{q=1}^{Q}\left(\delta_q \cos(\frac{q2\pi t}{N}) + \phi_q \sin(\frac{q2\pi t}{N})\right) + \\
&\quad \beta_{e,|OF|} D_e |OF_t^*| + \beta_{r,|OF|} D_r |OF_t^*| + \beta_{e,Qnt} D_e Qnt_t^* + \beta_{r,Qnt} D_r Qnt_t^* \\
&\quad \textstyle\sum_{k=1}^{K}\sum_{j'=0}^{J'} \beta_{e,kj'} D_e |S_{k,t-j}| + \sum_{k=1}^{K}\sum_{j'=0}^{J'} \beta_{r,kj'} D_r |S_{k,t-j}|\Big),
\end{aligned}
$$

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, D_a is a dummy variable that is 1 on announcement days and 0 else, $S_{k,t}$ is the standardized news announcement for $k = 1, \ldots, 24$, OF_t^* and Qnt_t^* denotes the unanticipated order flow and the unanticipated trading volume from time t to $t+1$. The dummy variables D_e and D_r denote an expansion and recession period, based on NBER data. Further, $I = 3$, $J = 0$, $J' = 1$ and $Q = 5$. Panel A contain estimated coefficients regarding the return part of the model whereas panel B contains the volatility part. We show only the contemporaneous effect and discard the coefficients corresponding with the first lag. We perform a Wald test to test on equal impact during expansions and recessions. 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			
Dependent variable: 30-year futures returns			
	Expansion	Recession	Wald test
Private information			
Order flow	0.019^{***}	0.061^{***}	$3,888^{***}$
Public information			
1. GDP advance	−0.064	0.049	1.81
2. GDP preliminary	−0.038	−0.031	0.01
3. GDP final	0.004	0.008	0.02
Real Activity			
4. Nonfarm payroll employment	-0.187^{***}	−0.002	2.00
5. Retail sales	-0.082^{***}	-0.062^{***}	0.26
6 . Industrial production	−0.017	0.043^{***}	7.73^{***}
7. Capacity utilization	-0.036^{*}	-0.043^{**}	0.06
8. Personal income	-0.040^{*}	-0.049^{*}	0.06
9. Consumer credit	−0.004	0.013	1.45
Consumption			
10. Personal consumption expenditures	−0.023	-0.024^{**}	0.01
11. New home sales	-0.059^{***}	-0.083^{**}	0.31
Investment			
12. Durable goods orders	-0.048^{**}	−0.030	0.20
13. Construction spending	−0.021	−0.006	0.24
14. Factory orders	−0.017	−0.021	0.01
15. Business inventories	−0.002	−0.007	0.11
Government Purchases			
16. Government budget deficit	−0.004	0.010^{**}	2.41
Net Exports			
17. Trade balance	-0.029^{***}	−0.023	0.06
Prices			
18. Producer price index	-0.046^{*}	−0.005	1.29
19. Consumer price index	-0.077^{**}	-0.030^{*}	1.40
Forward Looking			
20. Consumer confidence index	-0.076^{***}	-0.036^{**}	3.05
21. NAPM index	-0.027^{*}	0.024	3.38^{*}
22. Housing starts	−0.024	0.023	2.12
23. Index of leading indicators	-0.018^{**}	-0.038^{***}	2.32
24. Initial unemployment claims	0.036^{***}	0.034^{***}	0.02

(continued from previous page)

Panel B: Conditional Volatility			
	Expansion	Recession	Wald test
Private information			
Order flow	0.213***	0.403***	137***
Control variable			
Trading Volume	0.274***	0.754***	2,150***
Public information			
1. GDP advance	1.257***	0.417	3.15*
2. GDP preliminary	0.942**	0.331	1.40
3. GDP final	0.519**	−0.690	4.61**
Real Activity			
4. Nonfarm payroll employment	2.257***	0.900***	15.52***
5. Retail sales	0.892***	0.232	6.75***
6. Industrial production	0.616**	−0.567	8.89***
7. Capacity utilization	−0.172	0.437*	1.81
8. Personal income	0.664**	0.678*	0.00
9. Consumer credit	−0.162	−0.445	1.36
Consumption			
10. Personal consumption expenditures	0.339	−1.971**	7.17***
11. New home sales	0.565***	−0.002	1.11
Investment			
12. Durable goods orders	0.920***	−0.042	12.26***
13. Construction spending	1.017***	0.027	15.43***
14. Factory orders	0.991***	0.255*	7.26***
15. Business inventories	0.151	−0.084	1.07
Government Purchases			
16. Government budget deficit	−0.205	−0.417	0.14
Net Exports			
17. Trade balance	−0.087	0.136	0.95
Prices			
18. Producer price index	0.947***	−0.184	19.18***
19. Consumer price index	1.215***	−0.116	31.43***
Forward Looking			
20. Consumer confidence index	0.532***	−0.033	6.36**
21. NAPM index	0.578***	0.145	2.90*
22. Housing starts	0.497***	0.298	0.34
23. Index of leading indicators	0.067	−0.317	2.14
24. Initial unemployment claims	0.308***	0.088	3.44*
log-likelihood	294,692		
Nr. observations	157,500		

results for the return equation, and suggests that private information has a statistically significant different effect on returns in the different states of the economy. The impact of order flow on returns is more than three times higher during recessions. Also reported in the table is the effect of public information on returns conditional on the state of the economy. Surprises generally have a larger effect in the expansion regime compared to the recession regime, which is in line with the aforementioned literature. For most macroeconomic announcements the differences are small however, hence the Wald test does not reject the null hypothesis.

Panel B of the table shows that also the effect of private information on volatility is significantly different between expansions and recessions. Again, order flow is more important for volatility during recessions. This difference is also economically significant as the coefficient increases from 0.21 in expansions to 0.40 in recessions. Regarding public information, panel B suggests that the effect on volatility is larger in expansions for most announcements, except for Capacity utilization and Personal income. Based on the Wald test, most of these differences are statistically significant. Finally, the table highlights again the importance of including the volume of trades in our specification. The liquidity effect is much stronger during contractions than expansions. Ignoring this variable may overestimate the impact of order flow on volatility at each state of the economy.

In sum, private information matters for volatility, irrespective of the state of the economy. The effect of order flow on returns and volatility is much stronger during recessions. In addition, public information has a larger influence on volatility during expansions. Particularly in times of recession, it is important to consider private information when modeling volatility.

5.3 Conditioning on the dispersion of beliefs

This paper claims that order flow is important for explaining volatility. As noted earlier, order flow conveys private information about the agent's optimization problem at the micro level. When public information arrives to the market, order flow captures the aggregation of this private information. The value of this information depends on the degree of information heterogeneity among market participants. When the market consensus is low, order flow is more informative. This can be beneficial for traders who observe order flow. As a final additional analysis we investigate the impact of private information on both returns and volatility for different levels of 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. Specifically, 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 is constant throughout the quarter. The monthly proxy for information heterogeneity is defined as a sum of monthly (scaled) dispersions across announcements,

$$SRA_{P,t} = \sum_{k=1}^{K} \frac{RA_{k,t} - \hat{\mu}(RA_{k,t})}{\hat{\sigma}(RA_{k,t})}, \tag{9}$$

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. Due to the limited sample period, the number of observations for this analysis drops from 157,500 to 68,904.

Table 9 reports estimation results for the Spline-GARCH model conditioning the effects

Table 9: Order flow and information heterogeneity

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

$$
\begin{aligned}
R_t &= \beta_{0,r} + \beta_{a,r} D_a + \beta_{m,r} D_m + \beta_{h,r} D_h + \textstyle\sum_{i=1}^{I} \beta_i R_{t-i} + \sum_{k=1}^{K} \sum_{j=0}^{J} \beta_{kj} S_{k,t-j} + \\
&\quad \gamma_{OF,l} OF_t^* D_l + \gamma_{OF,m} OF_t^* D_m + \gamma_{OF,h} OF_t^* D_h + \varepsilon_t, \\
\varepsilon_t &= \sqrt{\tau_t g_t} u_t, \\
g_t &= (1 - \alpha - \beta_{GA}) + \alpha \left(\frac{\varepsilon_{t-1}^2}{\tau_{t-1}} \right) + \beta_{GA} g_{t-1}, \\
\tau_t &= \exp \Big(\beta_{0,v} + \beta_{a,v} D_a + \beta_{m,v} D_m + \beta_{h,v} D_h + \textstyle\sum_{q=1}^{Q} \left(\delta_q \cos(\frac{q 2\pi t}{108}) + \phi_q \sin(\frac{q 2\pi t}{108}) \right) + \\
&\quad \textstyle\sum_{k=1}^{K} \sum_{j'=0}^{J'} \beta_{kj'} |S_{k,t-j}| + \gamma_{|OF|,l} |OF_t^*| D_l + \gamma_{|OF|,m} |OF_t^*| D_m + \gamma_{|OF|,h} |OF_t^*| D_h + \\
&\quad \gamma_{Qnt,l} Qnt_t^* D_l + \gamma_{Qnt,m} Qnt_t^* D_m + \gamma_{Qnt,h} Qnt_t^* D_h \Big),
\end{aligned}
$$

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, D_a is a dummy variable that is 1 on announcement days and 0 else, $S_{k,t}$ is the standardized news announcement for $k = 1, \ldots, 24$, OF_t^* and Qnt_t^* denote the unanticipated order flow and the unanticipated volume of trades from period t to $t+1$. We introduce three additional dummy variables, D_l, D_m, and D_h which represent a low, medium or high level of information heterogeneity respectively. Results in the table correspond with $I = 3$, $I' = 12$, $J = 0$, $J' = 1$ 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 and the volume of trades. The superscripts ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively. The sample goes from June 2, 2007 through December 31, 2009.

Dependent variable: 30-year futures returns		
	Cond Mean	Cond Vol
Private information		
Order flow (low)	0.023***	0.159***
Order flow (medium)	0.028***	0.230***
Order flow (high)	0.131***	0.373***
Control variable		
Trading volume (low)		0.464***
Trading volume (medium)		0.553***
Trading volume (high)		0.766***
Macro announce.	yes	yes
log-likelihood	114,368	
Nr. observations	68,904	

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 5. 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 Pasquariello 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 – a result that is consistent with the notion that a sizeable proportion of macro news is transmitted via order flow (see Evans and Lyons 2008, for evidence of this in the DM/$ exchange rate market).

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 daily conditional volatility is equal to 11.7 basis points. As a comparison, the corresponding values in times of medium and low dispersion are equal to 7.2 and 5.0 respectively. Hence order flow is at least two times more important when the dispersion of beliefs is high compared to when it is low.

6 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, and use order flow as proxy for private information.

Our main finding is that private information significantly affects volatility, with this type of information more important than public information in this respect. Indeed, the effect of a shock to order flow on volatility is larger than the effect of a surprise of the same magnitude in the most influential macroeconomic news announcement. We extend our main finding by demonstrating that our finding is robust to consider the bid-ask spread as an alternative proxy of private information. In addition, we show that the effect of private information depends on the state of the economy in the sense that the effect of order flow on volatility is larger during bad states of the economy. Third, we include the dispersion of beliefs about macroeconomic announcements among analysts into account, with the effect of private information on volatility being more than two times larger when there is high dispersion compared to low dispersion. Our results imply that risk managers, portfolio managers and regulators should take into account private information variables as a determinant of volatility.

References

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.

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.

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.

Bartolini, L., L. Goldberg and A. Sacarny (2008), How economic news moves markets, *Current issues in economics and Finance* **14**. Federal Reserve Bank of New York.

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.

Berger, D., A. Chaboud and E. Hjalmarsson (2009), What drives volatility persistence in the foreign exchange market?, *Journal of Financial Economics* **94**, 192–213.

Berry, T.D. and K.M. Howe (1992), Public information arrival, *Journal of Finance* **49**, 1331–1346.

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.

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.

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.

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.

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.

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.

Chordia, T., A. Sarkar and A. Subrahmanyam (2005), An empirical analysis of stock and bond market liquidity, *Review of Financial Studies* **18**, 85–129.

Clark, P.K. (1973), A subordinated stochastic process model with finite variance for speculative prices, *Econometrica* **41**, 135–155.

Cutler, D.M., J.M. Poterba and L.H. Summers (1989), What moves stock prices?, *Journal of Portfolio Management* **15**, 4–12.

De Goeij, P. and W. Marquering (2006), Macroeconomic announcements and asymmetric volatility in bond returns, *Journal of Banking and Finance* **30**, 2659–2680.

Ederington, L. and J. Lee (1993), How markets process information: news releases and volatility, *Journal of Finance* **45**, 1161–1191.

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.

Evans, M. and R. Lyons (2002), Order flow and exchange rate dynamics, *Journal of Political Economy* **110**, 170–180.

Evans, M. and R. Lyons (2008), How is macro news transmitted to exchange rates?, *Journal of Financial Economics* **88**, 26–50.

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.

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.

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.

Garleanu, N., L.H. Pedersen and A. Poteshman (2009), Demand-based option pricing, *Review of Financial Studies* **22**, 4259–4299.

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.

Green, T.C. (2004), Economic news and the impact of trading on bond prices, *Journal of Finance* **59**, 1201–1233.

Hasbrouck, J. (1991), Measuring the information content of stock trades, *Journal of Finance* **46**, 179–207.

Hasbrouck, J. (2004a), *Empirical Market Microstructure: The Institutions, Economics, and Econometrics of Securities Trading*, Oxford: Oxford University Press.

Hasbrouck, J. (2004b), Liquidity in the futures pit: Inferring market dynamics from incomplete data, *Journal of Financial and Quantitative Analysis* **39**, 305–326.

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.

Huang, R.D. and H.D. Stoll (1997), The components of the bid-ask spread: A general approach, *Review of Financial Studies* **10**, 995–1034.

Jiang, G.J. and I. Lo (2011), Private information flow and price discovery in the U.S. Treasury market, Working Paper.

Karpoff, J.M. (1987), The relation between price changes and trading volume: a survey, *Journal of Financial and Quantitative Analysis* **22**, 109–126.

Kyle, A.S. (1985), Continuous auctions and insider trading, *Econometrica* **53**, 1315–1335.

Lee, C. and M. Ready (1991), Inferring the trade direction from intraday data, *Journal of Finance* **46**, 733–746.

Manaster, S. and S.C. Mann (1996), Life in the pits: Competitive market making and inventory control, *Review of Financial Studies* **9**, 953–975.

Menkveld, A.J., A. Sarkar and M. Van der Wel (2011), Customer flow, intermediaries, and discovery of the equilibrium riskfree rate, *Journal of Financial and Quantitative Analysis (forthcoming)* .

Pasquariello, P. and C. Vega (2007), Informed and strategic order flow in the bond market, *Review of Financial Studies* **20**, 1975–2019.

Roll, R. (1984), A simple implicit measure of the effective bid-ask spread in an efficient market, *Journal of Finance* **39**, 1127–1139.

Sarkar, A. and R.A. Schwartz (2009), Market sidedness: insights into notives for trade initiation, *Journal of Finance* **64**, 375423.

Tauchen, G.E. and M. Pitts (1983), The price variability-volume relationship on speculative markets, *Econometrica* **51**, 485–505.

Van der Wel, M., A.J. Menkveld and A. Sarkar (2009), Are market makers uninformed and passive?, Working Paper.

2011-45:	Peter Christoffersen and Hugues Langlois: The Joint Dynamics of Equity Market Factors
2011-46:	Peter Christoffersen, Kris Jacobs and Bo Young Chang: Forecasting with Option Implied Information
2011-47:	Kim Christensen and Mark Podolskij: Asymptotic theory of range-based multipower variation
2011-48:	Christian M. Dahl, Daniel le Maire and Jakob R. Munch: Wage Dispersion and Decentralization of Wage Bargaining
2011-49:	Torben G. Andersen, Oleg Bondarenko and Maria T. Gonzalez-Perez: Coherent Model-Free Implied Volatility: A Corridor Fix for High-Frequency VIX
2011-50:	Torben G. Andersen and Oleg Bondarenko: VPIN and the Flash Crash
2011-51:	Tim Bollerslev, Daniela Osterrieder, Natalia Sizova and George Tauchen: Risk and Return: Long-Run Relationships, Fractional Cointegration, and Return Predictability
2011-52:	Lars Stentoft: What we can learn from pricing 139,879 Individual Stock Options
2011-53:	Kim Christensen, Mark Podolskij and Mathias Vetter: On covariation estimation for multivariate continuous Itô semimartingales with noise in non-synchronous observation schemes
2012-01:	Matei Demetrescu and Robinson Kruse: The Power of Unit Root Tests Against Nonlinear Local Alternatives
2012-02:	Matias D. Cattaneo, Michael Jansson and Whitney K. Newey: Alternative Asymptotics and the Partially Linear Model with Many Regressors
2012-03:	Matt P. Dziubinski: Conditionally-Uniform Feasible Grid Search Algorithm
2012-04:	Jeroen V.K. Rombouts, Lars Stentoft and Francesco Violante: The Value of Multivariate Model Sophistication: An Application to pricing Dow Jones Industrial Average options
2012-05:	Anders Bredahl Kock: On the Oracle Property of the Adaptive LASSO in Stationary and Nonstationary Autoregressions
2012-06:	Christian Bach and Matt P. Dziubinski: Commodity derivatives pricing with inventory effects
2012-07:	Cristina Amado and Timo Teräsvirta: Modelling Changes in the Unconditional Variance of Long Stock Return Series
2012-08:	Anne Opschoor, Michel van der Wel, Dick van Dijk and Nick Taylor: On the Effects of Private Information on Volatility