Quantitative Easing and Tapering Uncertainty: Evidence from Twitter*

Annette Meinusch and Peter Tillmann
Justus-Liebig-University Gießen, Germany

In this paper we analyze the extent to which people's changing beliefs about the timing of the exit from quantitative easing ("tapering") affect asset prices. To quantify beliefs of market participants, we use data from Twitter, the social media application. Our data set covers the entire Twitter volume on Federal Reserve tapering in 2013. Based on the time series of beliefs about an early or late tapering, we estimate a structural VAR-X model under appropriate sign restrictions on the impulse responses to identify a belief shock. The results show that shocks to tapering beliefs have non-negligible effects on interest rates and exchange rates. We also derive measures of monetary policy uncertainty and disagreement of beliefs, respectively, and estimate their impact. The paper is one of the first to use social media data for analyzing monetary policy and also adds to the rapidly growing literature on macroeconomic uncertainty shocks.

JEL Codes: E32, E44, E52.

1. Introduction

After the 2008 economic crisis, the U.S. Federal Reserve adopted a series of unconventional monetary policy measures in order to

*We thank seminar participants at the Verein für Socialpolitik (Münster 2015), the European Economic Association (Mannheim 2015), the Money, Macro and Finance Group (Cardiff 2015), the 9th International Conference on Computational and Financial Econometrics (London 2015), the Deutsche Bundesbank Big Data Group, the CFCM conference (Nottingham 2016), and the Universities of Duesseldorf and Giessen for important feedback. We are grateful to Michael McMahon and Matthias Neuenkirch for helpful comments. Lucas Hafemann and Florian Viereck provided invaluable research assistance. Author e-mails: annette.meinusch@wirtschaft.uni-giessen.de and peter.tillmann@wirtschaft.uni-giessen.de.

enhance credit conditions and support economic recovery. Large-scale asset purchases known as quantitative easing (QE) led to a tripling of the Federal Reserve's balance sheet.[1] When Federal Reserve Chairman Ben Bernanke, while testifying before the U.S. Congress, first mentioned the possibility of reducing asset purchases on May 22, 2013, markets were wrong-footed. Bernanke's remarks triggered fears of a premature end of asset purchases and an earlier than expected increase in the federal funds rate. Markets coined the term "tapering" to describe the reduction of asset purchases by the Federal Reserve and the eventual end of QE.

Market jitters following the May 22, 2013 testimony led to a sharp increase in long-term interest rates in the United States, a period of high volatility on asset markets, and a dramatic appreciation of the U.S. dollar, particularly against emerging market currencies. Since a large part of these turbulences appeared exaggerated and panic-driven, observers referred to the "taper tantrum."

Federal Reserve Governor Jeremy Stein (2014) reflected on the revision of investors' expectations and the strong market movements in 2013 that gave rise to the "tantrum" notion:

> In early May 2013, long-term Treasury yields were in the neighborhood of 1.60 percent. Two months later, shortly after our June 2013 FOMC meeting, they were around 2.70 percent. Clearly, a significant chunk of the move came in response to comments made during this interval by Chairman Bernanke about the future of our asset purchase program. Perhaps it is not surprising that news about the future course of the asset purchase program would have a strong effect on markets. But here is the striking fact: According to the Survey of Primary Dealers conducted by the New York Fed, there was hardly any change over this period in the expectation of the median respondent as to the ultimate size of the program. Chairman Bernanke's comments . . . did not have any clear directional implications for the total amount of accommodation to be provided via asset purchases.[2]

---

[1] See D'Amico et al. (2012) and Rogers, Scotti, and Wright (2014) for recent surveys on the effectiveness of unconventional monetary policies.

[2] In April 2013, the pessimistic first quartile of institutions asked by the Federal Reserve Bank of New York's Survey of Primary Dealers showed that markets

In this paper, we provide an empirical analysis of the revision of expectations of market participants and its impact on asset prices that gave rise to the taper tantrum. First, we quantify the response of interest rates, exchange rates, and other asset prices to the belief shocks of market participants specific to tapering beliefs. Second, we measure monetary policy uncertainty as well as derive a measure of disagreement of market participants about future monetary policy and quantify their effects on financial variables. Third, we decompose the dynamics of asset prices in order to isolate the fraction of movements due to changes in tapering beliefs.

The primary difficulty of any study addressing sudden changes in beliefs and their consequences is that individual beliefs about the future course of monetary policy are not observable. Survey evidence is typically only available on a very low frequency, thus making an analysis of daily data impossible. An alternative would be to use beliefs extracted from futures prices or the yield curve. The disadvantage is that these market prices do not allow us to extract measures of disagreement of market participants.

In this paper, we offer a new approach to identify shocks to people's beliefs about monetary policy by using social media. We use data from Twitter.com, the popular social media application for short text messages ("tweets" of no more than 140 characters) since many market participants use their Twitter account to express and disseminate their views on the future stance of monetary policy. To the best of our knowledge, Twitter data has not been used to study monetary policy before.

The advantage of using Twitter data for research purposes is that (i) users not only receive information but can actively share information, (ii) tweets reflect personal views of market participants, (iii) tweets can be used to extract not only a consensus view on policy but also the degree of uncertainty and disagreement about policy, and (iv) Twitter users can respond immediately to news about policy such as Bernanke's testimony and also to other Twitter users'

expected the Federal Reserve to reduce its monthly purchases of assets worth 85 billion dollars at its December meeting. The events in May 2013 triggered a reassessment of expectations. In the July survey, market professionals were expecting purchases of only 65 billion dollars at the September FOMC meeting and only 50 billion dollars after the December meeting.

contributions. Our data set allows us to track the evolution of market beliefs about monetary policy up to the second.

We use the entire Twitter volume containing the words "Fed" and "taper," which amounts to almost 90,000 tweets for the period April to October 2013. From this, we identify tweets that express an explicit view about whether the reduction of bond purchases will occur soon or whether it will occur late. The resulting time series of beliefs of early or late tapering are then put into a vector autoregression (VAR-X) with daily data on interest rates and exchange rates and exogenous variables that control for FOMC meeting days and real economic activity.

With appropriate sign restrictions we are able to identify belief shocks and their dynamic effects. In addition, we use Twitter data to construct two indexes reflecting the uncertainty and disagreement of future Federal Reserve policy and estimate the impact of uncertainty shocks in our VAR model.

The results show that "tapering soon" belief shocks lead to a significant increase in long-term interest rates and a persistent appreciation of the U.S. dollar. A prototypical belief shock raises the share of all tweets considering an early tapering by 10 percentage points, and leads to a 3 basis point increase in long-term yields and a 0.2 percent appreciation of the dollar. These results are in line with the considerations of Krishnamurthy and Vissing-Jorgensen (2013). The data also allow us to study the effects of an increase in uncertainty and disagreement among Twitter users on asset prices. Thus, we can shed light on the points raised by Kashyap (2013), stressing the importance of disagreement about the course of tapering unconventional monetary policy.

Understanding market responses to exiting from QE and other unconventional monetary policies is important. Not only is the Federal Reserve about to gradually exit from unconventional monetary policy, but the Bank of England and, at some point in the future, the Bank of Japan and the European Central Bank are all on the brink of similar exits from unconventional monetary policies. Clearly communicating exit strategies from unconventional monetary policy to financial markets participants and the general public is essential for a smooth and frictionless return to normal. Analyzing data from social media is a useful way of cross-checking whether official communication was received by the markets as intended. In addition, it is important to quantify the impact of market beliefs on interest rates in light of forward guidance used by many central banks.

The remainder of the paper is organized into 7 sections. Section 2 briefly reviews the related literature. Section 3 introduces our data set on Twitter messages, which is used for the empirical analysis in section 4. The results are discussed in section 5. Section 6 is devoted to the analysis of extensions and the robustness of our results. Section 7 includes our concluding remarks and draws some policy implications.

2. The Literature on Federal Reserve Tapering and Uncertainty Shocks

This paper is related to various strands of the literature. There are several papers that focus on the impact tapering announcements have on asset prices. A very useful collection of facts related to the responses to tapering news is provided by Sahay et al. (2014).

Eichengreen and Gupta (2015) present the earliest systematic analysis of Federal Reserve tapering. They attribute the fluctuations in emerging economies in 2013 to Federal Reserve tapering and explain the magnitude of fluctuations in terms of initial macroeconomic conditions. It is shown that better macroeconomic fundamentals did not necessarily shield economies from the tapering fallout.

Aizenman, Binici, and Hutchison (2016) estimate a panel model with daily data for emerging economies and relate the response to tapering news to macroeconomic fundamentals. Similar to Eichengreen and Gupta (2015), they show that fundamentally stronger countries were more sensitive to tapering and argue that this is due to the massive capital inflows these countries received under the Federal Reserve's quantitative easing programs. Their paper uses dummies for FOMC meetings during 2013 as a proxy for tapering news.

Nechio (2014) provides descriptive evidence for the adjustment of emerging economies after Chairman Bernanke's May 22 testimony. She finds that the relative strength of emerging markets' responses reflect internal and external weaknesses specific to each market. Daily data on twenty-one emerging countries is used by Mishra et al. (2014). In contrast to Eichengreen and Gupta (2015) and Aizenman, Binici, and Hutchison (2016), their evidence supports the notion that countries with stronger macroeconomic fundamentals experienced a smaller depreciation of their currencies and smaller increases in borrowing costs. In this study, the market responses are measured in a two-day event window around an FOMC meeting or a publication day of FOMC minutes.

All of these papers proxy market expectations about Federal Reserve tapering by impulse dummies reflecting FOMC meetings and Chairman Bernanke's testimony, or by focusing on relatively narrow event windows. They do not measure market expectations directly. This is exactly where our paper adds to the literature. We extract information from Twitter messages to construct a high-frequency indicator of market beliefs. This indicator also reflects changes in policy perception between FOMC meetings and, in particular, mounting uncertainty before FOMC meetings, which cannot be appropriately proxied by meeting dummies.

Closest to this paper is the work by Dahlhaus and Vasishtha (2014) and Matheson and Stavrev (2014). The latter authors estimate a bivariate VAR model for U.S. stock prices and long-term bond yields. Sign restrictions are used to identify a fundamental-based news shock leading to an increase in both variables and a monetary shock implying an opposite response of stock prices and yields. The authors show that in the taper tantrum episode monetary shocks were important initially, while news shocks became important towards the end of 2013. Our research, however, measures market expectations from social media and avoids restricting the asset price response. Dahlhaus and Vasishtha (2014) identify a "policy normalization shock" using sign restrictions as one that raises federal funds futures but leaves current rates unchanged. They show that this shock has a significant impact on the common component of capital flows to emerging economies.

More generally, our paper adds to the growing body of literature concerned with the macroeconomic consequences of uncertainty shocks. In recent years, researchers develop indicators of uncertainty and analyze the data using VAR models. The first researcher to use this methodology was Bloom (2009). He presents a structural model of macroeconomic uncertainty affecting second moments and estimates a VAR model that replicates the theoretical findings. Baker, Bloom, and Davis (2016) focus on uncertainty about future economic policy. They construct an uncertainty index by referring to newspaper articles about uncertainty and show that this index has predictive power for several macroeconomic variables. On a business level, Bachmann, Elstner, and Sims (2013) use German survey data in a VAR model. They find that a heightened degree of uncertainty for businesses correlates to higher unemployment, lower investment, and higher refinancing costs.

The only paper focusing on monetary developments so far is Istrefi and Piloiu (2013). The authors use the Baker, Bloom, and Davis (2016) index of policy uncertainty for the United States, the United Kingdom, Germany, and the euro area and show that within a structural VAR model uncertainty raises long-term inflation expectations. In contrast to most of these contributions, our measure of policy uncertainty based on Twitter information directly addresses specific uncertainty about the future course of monetary policy.

Lately, there has been a growing interest in the use of social media (Twitter, Google, Facebook) as a data source for economic analyses. Among others, Choi and Varian (2009, 2012) use Google Trends data to forecast near-term values of economic indicators such as initial claims for unemployment. In the context of financial markets Da, Engelberg, and Gao (2011) derive a measure of investor attention based on Google search data. Vlastakis and Markellos (2012) find a link between Google keyword searches and stock trading volume and stock return volatility. Dergiades, Milas, and Panagiotidis (2015) have shown that social media provides significant short-run information for the Greek and Irish government bond yield differential. Acemoglu, Hassan, and Tahoun (2014) predict protests of Egypt's Arab Spring by a Twitter-based measure of general discontent about the government in power. Our study extends this field of research and is the first to analyze monetary policy based on Twitter messages.[3]

3. Tapering Beliefs on Twitter

We extract market participants' beliefs and their uncertainty about the future course of monetary policy from Twitter messages. Twitter

[3]Tillmann (2015) studies the transmission of tapering-related belief shocks to emerging market economies.

is becoming more and more popular among financial professionals. It allows them to comment on policy and market events and to distribute their view to either their followers or an even wider audience in real time.

For the purpose of this study, we obtained the entire Twitter traffic between April 15 and October 30, 2013 containing the words "Fed" and "taper" from Gnip.com, a provider of social media data. Since the debate was focused around the "taper" buzzword, we are confident that we did not miss important tweets when using these keywords. The data set includes nearly 90,000 tweets from about 27,000 users located in 135 countries and the exact time they were sent. This is a unique data set to study market views during the tapering tantrum episode. Panel A in figure 1 plots the daily evolution of Twitter traffic over time.[4]

It can be seen that the number of tweets increases around Bernanke's testimony and around each FOMC meeting. The use of Twitter peaks at the September 17/18 FOMC meeting, when the Federal Reserve finally decided to continue its QE policy and not begin tapering. The sample period covers the entire taper tantrum episode and is sufficiently long to perform a VAR analysis. Further, the data set comprises each tweet's text message of at most 140 characters as well as the name, the location, and the number of followers of the Twitter user.[5]

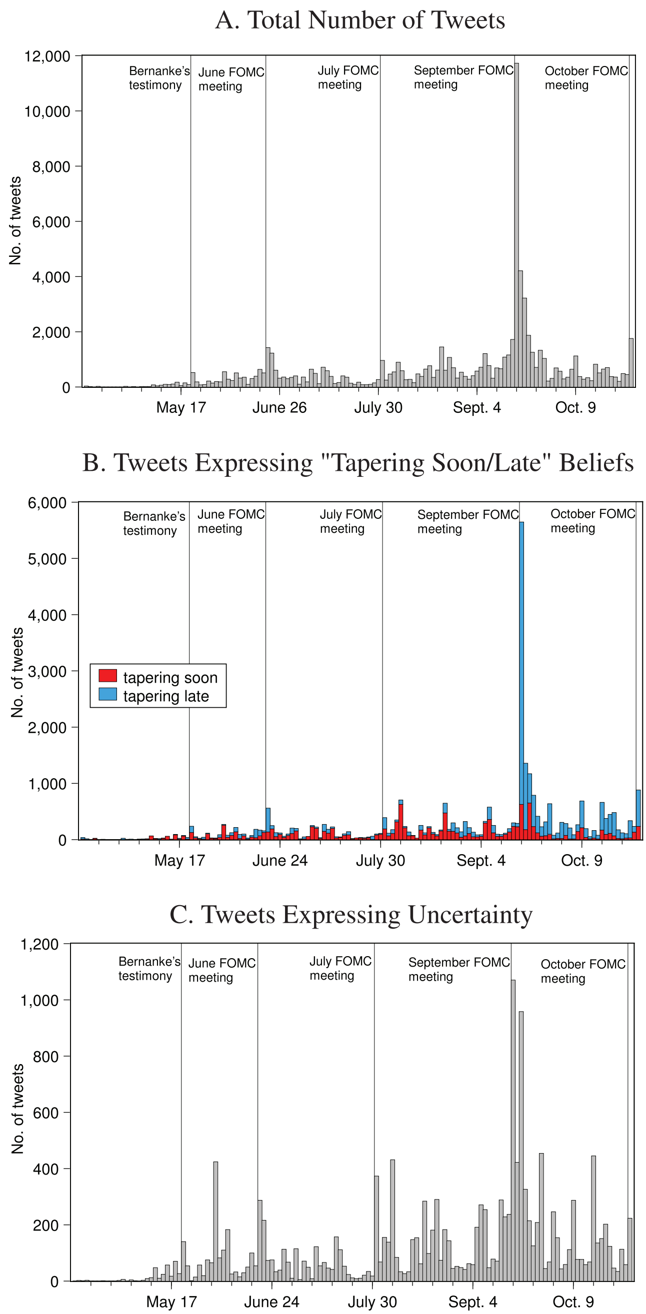
3.1 Beliefs about the Timing of Tapering

The tweets are separated into those expressing the belief of an early tapering, probably in the summer of 2013 or at the September 2013

[4]Retweets, which account for around 26 percent of all tweets, are included in this figure. For the purpose of this paper we interpret retweeted messages as an endorsement of the initial message's relevance and include it in our measure of beliefs. In the robustness section we present results from a model that excludes retweets.

[5]The online appendix (available at https://petertillmann.wordpress.com/research/) contains a plot of the (log) number of users on the ordinate versus the ranked (log) number of tweets on the abscissa. The resulting graph resembles a Zipf-like distribution, indicating that a small number of active users frequently share their opinions about the Federal Reserve's future policy stance and that a large number of users generate tweets about Federal Reserve tapering rather infrequently.

Figure 1. Data Series



Notes: Sample period: April 15–October 30, 2013. The vertical lines indicate FOMC meetings and Chairman Bernanke's May 22 testimony.

FOMC meeting, and those expressing the belief of tapering occurring later. A two-step procedure is used to interpret the content of tweets and allocate the tweets to $Tweets_t^{soon}$ and $Tweets_t^{late}$. In a first step, tweets are filtered according to a list of predefined keywords. In a second step, all remaining tweets are, if appropriate, manually allocated to one of the two categories. The appendix goes into more detail about this procedure. As a result, we are left with roughly 32,000 tweets that contain explicit views about future policy. Tweets that could not be assigned to one of those two categories mostly comment on market movements, point to the upcoming FOMC meeting, or formulate unspecific policy views (i.e., "to taper or not to taper"). Finally, the tweets are aggregated into daily series of beliefs.

As an example, consider the following tweet written on May 20, 2013:

> "Job market gains could lead Fed to taper QE3 early,"

which reflects the view of an early tapering and is allocated to $Tweets_t^{soon}$. Likewise, consider this tweet written on July 31, 2013:

> "The case for a September Fed taper just got a whole lot stronger."

This tweet is also counted as reflecting the view that the Federal Reserve will taper early. The following tweets, in contrast, suggest that Twitter users believe in a later tapering decision. On May 21, 2013, a tweet stated the following:

> "Fed's Bullard says doesn't see a good case for taper unless inflation rises ..."

and on September 18, 2013 the following was retweeted:

> "RT @DailyFXTeam: Economist Nouriel Roubini tweets that based on weak macro data, the Fed shouldn't taper today."

Panel B in figure 1 depicts the identified belief series. We clearly see sizable fluctuations in beliefs and the increased volatility before and after FOMC meetings. Our data contain a total of 15,422 tweets that refer to early tapering and 16,997 tweets that are associated with tapering late. Furthermore, the majority of tweets initially expressed the belief of an early tapering, which then changes in September 2013 in favor of a late tapering. Interestingly, both series, $Tweets_t^{soon}$ and $Tweets_t^{late}$, peak around the September 18 meeting of the FOMC, with a total of 652 "soon" tweets on September 20 and 3,472 "late" tweets on September 18. Further, they are positively correlated with a correlation coefficient of 0.5.

In the regressions below, we include our belief proxies, $Tweets_t^i$, as a fraction of the total amount of timing-related tweets on a particular day:

$$Beliefs_t^i = 100 \times \frac{Tweets_t^i}{Tweets_t^{soon} + Tweets_t^{late}} \ ,$$

with $i = \{soon, late\}$.

3.2 Uncertainty and Disagreement about Tapering

Beliefs of an early or a late tapering could fluctuate within the same day, indicating that there is substantial heterogeneity in market participants' beliefs about future policy. This possibility encourages us to construct two additional indicators that reflect the uncertainty and the disagreement between market commentators about future policy. The uncertainty measure, $Tweets_t^{uncertainty}$, is constructed by counting specific words reflecting uncertainty as in Loughran and McDonald (2011). These authors construct a comprehensive word list to describe uncertainty in text data, which is calibrated to financial applications. Details about the construction are also given in the appendix. Panel C in figure 1 plots the uncertainty indicator. Like the other belief series, uncertainty also seems to be sensitive to official Federal Reserve communication. In particular, uncertainty increases following each FOMC meeting in the sample period.

The indicator $Beliefs_t^{uncertainty}$ is expressed as a ratio of the number of uncertainty-related tweets, $Tweets_t^{uncertainty}$, and the total amount of all tweets on a particular day, i.e.,

$$Beliefs_t^{uncertainty} = 100 \times \frac{Tweets_t^{uncertainty}}{TotalTweets_t} \ .$$

A fourth indicator measures market participants' diverging views about the short-term path of monetary policy. This measure of disagreement, $Beliefs_t^{disagreement}$, is based on soon and late belief series and is defined as

$$Beliefs_t^{disagreement} = 1 - \sqrt{\left(\frac{Tweets_t^{soon}}{TotalTweets_t} - \frac{Tweets_t^{late}}{TotalTweets_t}\right)^2}.$$

It reaches its maximum value of 1 for cases in which the fraction of beliefs corresponding to early tapering is equal to the fraction of beliefs referring to later tapering. If one opinion concerning future monetary policy dominates the other, both fractions diverge and the disagreement index declines. In the following we will use fluctuations in tapering beliefs in a vector autoregressive model to identify unexpected shocks to tapering expectations, policy uncertainty, and investor disagreement.

4. The Model

We use a set of vector autoregressive models with an exogenous variable (VAR-X) to analyze the consequences of shocks to people's beliefs. A combination of sign restrictions is used to identify structural belief shocks.

4.1 The VAR-X Model

Our structural VAR-X model is assumed to have the standard form

$$B(L)Y_t = C + D^{FOMC} + \Theta X_t + \varepsilon_t, \quad \text{with } E[\varepsilon_t \varepsilon_t'] = \Sigma_\varepsilon,$$

where $B(L) \equiv B_0 - B_1 L - B_2 L^2 - ... - B_p L^p$ is a p^{th}-order matrix polynomial in the lag operator L. Further, Y_t is a k-dimensional time series of endogenous variables, X_t is an exogenous variable, and ε_t represents a serially uncorrelated prediction error with Σ_ε as its variance-covariance matrix. The variance-covariance matrix of the structural innovation is normalized to $E(\varepsilon_t \varepsilon_t') \equiv \Sigma_\varepsilon = I_k$.

Since figure 1 shows that our belief series peak on days of FOMC meetings, one could argue that the information content of Twitter beliefs stems from the fact that they simply reflect the official

Federal Reserve communication. To control for this information, we include five dummy variables, D^{FOMC}, for the FOMC meetings in our sample period. For our application it is also important to control for macroeconomic data releases that would affect monetary policy expectations and, as a consequence, asset prices. We include the Aruoba-Diebold-Scotti (2009) daily business conditions index (hereafter the ADS index) as an exogenous variable in order to accomplish this.

A reduced-form representation for this system of equations is

$$A(L)Y_t = \Omega + \Psi X_t + u_t, \quad \text{with } E[u_t u_t'] = \Sigma_u,$$

where $\Omega = B^{-1}(C + D^{FOMC})$ and $\Psi = B^{-1}\Theta$, $A(L) \equiv I - A_1 L - A_2 L^2 - ... - A_p L^p$ reflects the matrix polynomial, and u_t constitutes a white-noise process with variance-covariance matrix Σ_u. Further, the structural shocks are linked to the reduced-form shocks by $u_t = B^{-1}\varepsilon_t$.

The model-specific vector of endogenous variables is

$$Y_t = \left(Beliefs_t^i, \quad TotalTweets_t, \quad Rate_t, \quad FX_t\right)'$$

for $i = (soon, \ late)$ and

$$Y_t = \left(Beliefs_t^j, \quad Rate_t, \quad VIX_t, \quad FX_t\right)'$$

for $j = (uncertainty, \ disagreement)$, where $Rate_t$ is the ten-year constant maturity yield, $TotalTweets_t$ is the log number of daily tweets, and FX_t is the log USD–EUR exchange rate. The VIX index of implied stock market volatility, which is denoted by VIX_t, is needed for identification as discussed below.[6] In order to account for the overall Twitter activity on a particular day, which is a measure of the attention monetary policy receives, we include the total number of tapering tweets. We fit the VAR model to the data by including ten lags of the endogenous variables. All weekends and holidays for which no financial data is available are excluded. The sample period consists of 138 daily observations and covers April 15, 2013 to October 30, 2013 and hence is sufficiently long for reliably estimating a VAR.

[6] All data are taken from the FRED (Federal Reserve Economic Data) database of the Federal Reserve Bank of St. Louis.

4.2 Identification

The identification of belief shocks is crucial for this analysis. As the contemporaneous interaction among all variables at a daily frequency prevents us from using a triangular identification scheme, sign restrictions (Uhlig 2005) provide a useful alternative to identify a structural shock in this VAR analysis. In a sign-restriction approach, identification is achieved by imposing ex post restrictions on the signs of the response of the endogenous variables to a structural shock, e.g., our belief shock. We believe that using sign restrictions creates a VAR best suited to analyze the mutual interaction between market beliefs about policy, asset prices, and volatility indicators even though most of the literature on uncertainty shocks relies on triangular identification schemes instead (such as Bloom 2009 and Baker, Bloom, and Davis 2016, reviewed in section 2).

In order to identify economically meaningful structural shocks, ε_t, we need to find a matrix B_0^{-1} such that the structural innovations are linked to the reduced-form shocks by $u_t = B_0^{-1}\varepsilon_t$, and $\Sigma_u = B_0^{-1}\Sigma_\varepsilon B_0^{-1'} = B_0^{-1}B_0^{-1'}$ with $\Sigma_\varepsilon = I_k$ holds. We proceed in the following way: We estimate our model by OLS, which provides us the reduced-form coefficients $A(L)$ and the covariance matrix Σ_u. Since it is $P_c^{-1} = chol(\Sigma_u)$ so that $\Sigma_u = P_c^{-1}P_c^{-1'}$ and $\Sigma_u = P_c^{-1}\tilde{S}\tilde{S}'P_c^{-1} = B_0^{-1}B_0^{-1'}$ with $B_0^{-1} = P_c^{-1}\tilde{S}$, we randomly draw a matrix $\tilde{S}$ from a space of orthonormal matrices.

Further, we calculate impulse response functions for the restricted periods as $D(L) = A(L)^{-1}B_0^{-1}$ and check whether they satisfy the postulated sign restrictions. We discard those response functions that fail to meet the restrictions while a new orthonormal matrix and new impulse responses are drawn. This procedure is continued until 500 accepted impulse response functions are stored, for which we then compute impulse response functions for all desired periods.

The impact restrictions we use to identify a belief shock are summarized in table 1. We estimate several VAR-X models, one for each alternative series of beliefs and our measures of uncertainty and disagreement. The belief shocks are identified by imposing positive responses of $Beliefs_t^i$, $Beliefs_t^j$, and responses of interest rates and the VIX index. A shock to "tapering soon" beliefs in model I is identified as one that raises the respective belief series and leads to

Table 1. Sign Restrictions to Identify a Belief Shock

	$\boldsymbol{Beliefs_t^i}$	$\boldsymbol{TotalTweets_t}$	$\boldsymbol{Rate_t}$	$\boldsymbol{FX_t}$
Model I: Soon	+		+	
Model II: Late	+		−	
	$\boldsymbol{Beliefs_t^j}$	$\boldsymbol{Rate_t}$	$\boldsymbol{VIX_t}$	$\boldsymbol{FX_t}$
Model III: Uncertainty	+		+	
Model IV: Disagreement	+		+	

higher bond yields. These restrictions are imposed for three periods. The "tapering late" shock raises "tapering late" beliefs and lowers bond yields. We do not restrict the responses for the exchange rate but expect a belief shock to lead to a depreciation of foreign currencies against the U.S. dollar. Since our measure of Twitter beliefs does not include obvious comments on market movements but only firm views on the timing of tapering, we are confident that we can exclude problems of reverse causality.

Since we do not know how shocks to uncertainty and disagreement affect the long-term interest rate, we abstain from restricting those responses in our models III and IV. We assume, however, that both are associated with an increase in market volatility. Hence, we include and restrict the VIX index in our model which is often interpreted as a proxy for fluctuations in risk aversion. As the belief series for uncertainty and disagreement are calculated by using the total amount of tweets on a particular day, we avoid including $TotalTweets_t$ in these two specifications. Although we do not derive the restrictions from a particular asset pricing model, it seems plausible that any increase in policy uncertainty or disagreement among investors is associated with a higher implied volatility. The restrictions for models III and IV are also imposed for three days.

5. Results

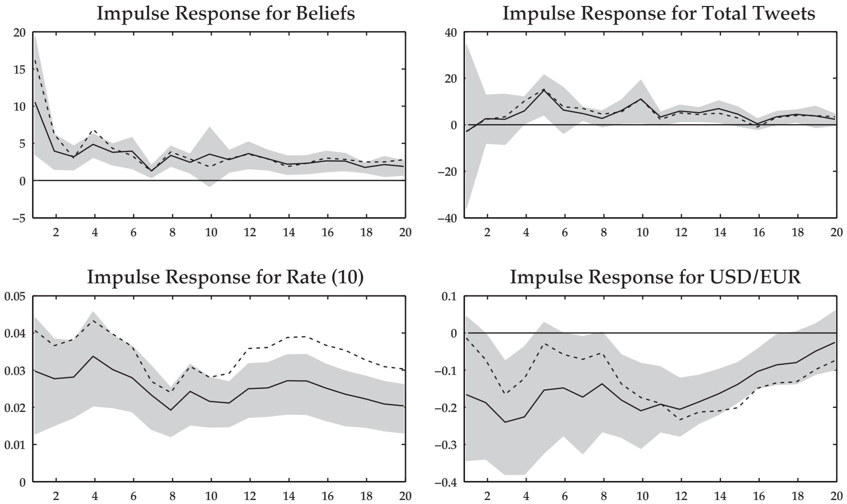
In this section we present the impulse responses following a shock to tapering beliefs, uncertainty, and disagreement about tapering, as identified by the set of sign restrictions. We show the median responses of the variables to a belief shock of one standard deviation for a horizon of twenty days after the shock together with the 16th and 84th percentiles of all accepted impulse responses. Additionally, we show the median-target impulse response (Fry and Pagan 2011), which corresponds to a single structural model that yields an impulse response closest to the median response. We also discuss the forecast error variance decomposition and the historical decomposition of the endogenous variables.

5.1 Shocks to Tapering Beliefs

The responses to a "tapering soon" belief shock are depicted in figure 2. It can be seen that a shift in beliefs towards an early tapering, demonstrated by a 10 percent increase of Twitter users foreseeing an early tapering, leads to a substantial tightening of monetary conditions. Long-term interest rates persistently increase by about 3 basis points. Since some days are characterized by swings in beliefs that are much stronger than the 10 percent depicted here, our model shows that belief shocks could lead to large increases in bond yields. Importantly, our three-day restriction on the direction of the response of bond yields seems to impose a fairly weak constraint on adjustment dynamics. We also find the exchange rate to appear sensitive to tapering beliefs. The dollar appreciates by 0.2 percent following a shift in beliefs. The immediate and persistent response of the dollar to tapering beliefs is in line with the arguments discussed by Krishnamurthy and Vissing-Jorgensen (2013). The total number of tweets, which is a measure of the overall attention of market participants to tapering, adjusts only moderately after a belief shock.

A key feature of our Twitter data set is that we can use the heterogeneity of users' beliefs to distinguish between "tapering soon" and "tapering late" belief shocks. In model II we therefore use the fraction of users expressing a late-tapering belief. Note that by construction, $Tweets_t^{soon}$ and $Tweets_t^{late}$ are not perfectly negatively correlated. Hence, we can estimate the VAR model for late-tapering beliefs in order to compare the strength of the responses to $Tweets_t^{soon}$ and $Tweets_t^{late}$. We obtain inverse results

Figure 2. Impulse Responses to "Tapering Soon" Belief Shock

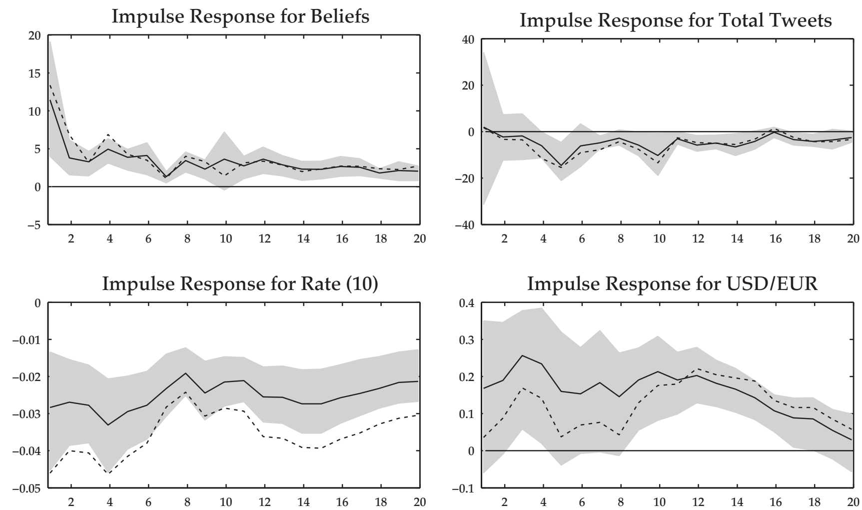


Notes: The solid line is the median impulse response function. The dotted line is the Fry-Pagan (2011) median-target impulse response. The confidence band reflects the 16th and 84th percentiles of all accepted draws.

for responses to a "tapering late" belief shock: the long-term interest rate decreases and the dollar depreciates; see figure 3. Broadly speaking, the responses to "tapering soon" or "tapering late" belief shocks appear symmetric. In both models we appear to underestimate the response of long-term interest rates relative to the Fry-Pagan impulse response.

Figure 4 presents a historical decomposition of interest rates and exchange rates. The contribution of the "tapering soon" belief shock to bond yields, which is depicted by the bars, is particularly pronounced after the June and before the September 2013 FOMC meeting. Likewise, for the exchange rate, the identified shock has large explanatory power before the September meeting. Table 2 decomposes the error variance of the one-, ten-, and twenty-step-ahead forecasts of all variables into the components accounted for by the identified "tapering soon" and "tapering late" belief shocks. For a twenty-day horizon, between 15 percent and 20 percent of the variances of asset prices can be attributed to each shock. Thus, both

Figure 3. Impulse Responses to "Tapering Late" Belief Shock



Notes: The solid line is the median impulse response function. The dotted line is the Fry-Pagan (2011) median-target impulse response. The confidence band reflects the 16th and 84th percentiles of all accepted draws.

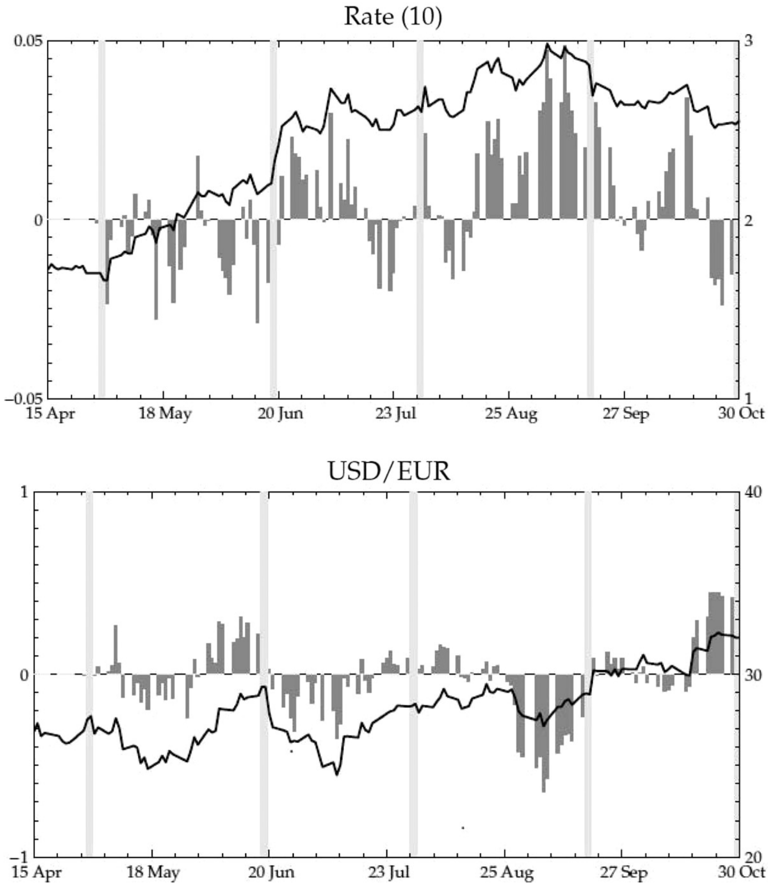
decompositions support the notion of belief shocks as important drivers of bond yields and exchange rates.

5.2 Shocks to Uncertainty and Disagreement

While the previous subsection studied shifts in the share of Twitter users believing in an early or late tapering, we now analyze the effect of uncertainty and disagreement. Figures 5 and 6 show the results for a shock to the uncertainty and the disagreement indexes, respectively, described in section 3. For these two specifications, no restrictions are imposed on the long-term interest rate and the exchange rate in order to let the data speak freely. However, we restrict the VIX index to respond positively after a belief shock.

We find a persistent response of the VIX index that goes beyond the three-day restriction imposed for identification purposes. Interestingly, an uncertainty shock seems to have no effect on the

Figure 4. Historical Decomposition for "Tapering Soon" Belief Shock



Notes: Contribution of "tapering soon" belief shock (left scale) and observable asset prices (right scale). The shaded areas indicate FOMC meetings.

long-term interest rate.[7] The dollar depreciates by 0.2 percent after an increase in monetary policy uncertainty. A shock to market participants' disagreement about the timing of the tapering decision

[7]Bekaert, Hoerova, and Lo Duca (2013) also use a VAR model and find a negative and persistent effect of uncertainty shocks on the real interest rate.

Table 2. Forecast Error Variance Decomposition

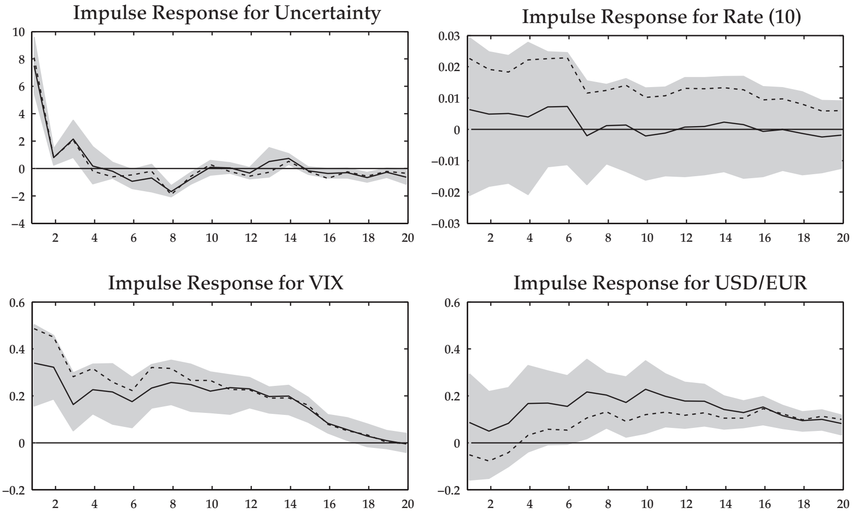
	Impact of Belief Shock (in % of Total Variation)					
	Tapering Soon			Tapering Late		
	At Horizon			At Horizon		
Variable	**1 Day**	**10 Day**	**20 Day**	**1 Day**	**10 Day**	**20 Day**
$Beliefs_t$	23.79	16.77	16.87	18.91	16.25	16.25
$TotalTweets_t$	15.42	14.84	14.81	15.19	14.71	14.65
$Rate_t$	13.02	19.58	19.68	14.36	19.63	19.51
FX_t	14.66	16.98	17.50	13.88	20.10	20.03
	Impact of Belief Shock (in % of Total Variation)					
	Uncertainty			**Disagreement**		
	At Horizon			**At Horizon**		
Variable	**1 Day**	**10 Day**	**20 Day**	**1 Day**	**10 Day**	**20 Day**
$Beliefs_t$	36.54	32.59	32.69	13.81	13.82	14.03
$Rate_t$	12.82	13.40	13.63	17.62	17.36	17.21
VIX_t	12.97	13.52	13.29	17.14	17.13	16.80
FX_t	13.57	14.73	14.69	20.92	19.49	19.32

leads to an increase in interest rates and an appreciation of the dollar. In terms of the signs and the persistence of the responses, these reactions are similar to those after a "tapering soon" shock. Our findings are in line with Kashyap's (2013) view that disagreement is an important explanatory factor for the taper tantrum.

Given that the estimated effects of uncertainty and disagreement diverge with regard to the exchange rate response, we calculate the correlation between the uncertainty and disagreement series. We obtain an unconditional correlation coefficient of 0.03, indicating the absence of a systematic correlation between those two indexes. Thus, we can conclude that shifts in investors' uncertainty or disagreement are separate and quantitatively important factors in the dynamics of interest rates and exchange rates.

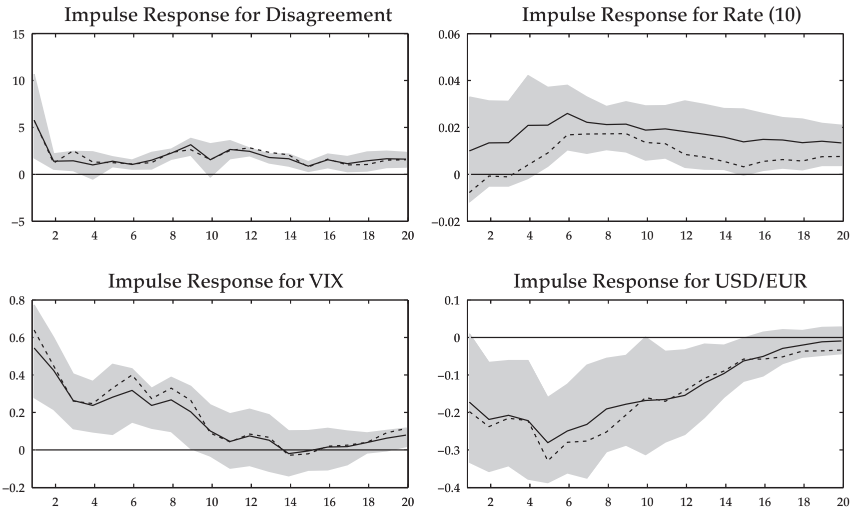
Finally, we report the contribution of uncertainty and disagreement shocks to the forecast error variance of the variables in the

Figure 5. Impulse Responses to Uncertainty Shock



Notes: The solid line is the median impulse response function. The dotted line is the Fry-Pagan (2011) median-target impulse response. The confidence band reflects the 16th and 84th percentiles of all accepted draws.

Figure 6. Impulse Responses to Disagreement Shock



Notes: The solid line is the median impulse response function. The dotted line is the Fry-Pagan (2011) median-target impulse response. The confidence band reflects the 16th and 84th percentiles of all accepted draws.

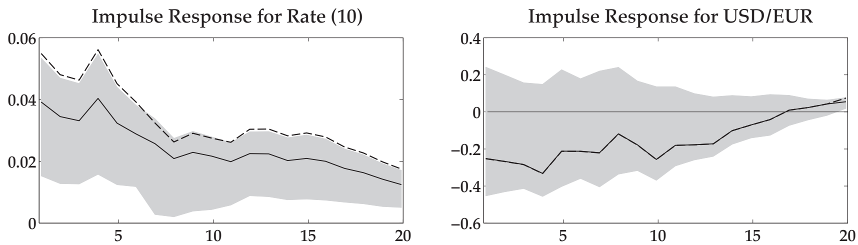
VAR model; see table 2. There is a substantial impact on all included series based on the median estimate of the responses. We find that for all horizons considered, a share between 12 percent and 20 percent of the variation in all variables (for the belief series more than 30 percent) is due to a shock in either variable. This again underlines the quantitative importance of shifts in market beliefs for interest rates and exchange rates.

6. Robustness

In this section, we analyze the robustness of our findings with regard to the role of the information contained in Twitter messages. The first robustness check investigates the role our belief series play in the identification and estimation of shocks. It could be argued that the effects of belief shocks on asset prices entirely stem from the restrictions imposed on bond yields. If the restrictions on the belief series were not binding, the information contained in Twitter would be irrelevant for the result. To shed light on this concern, we estimate a model without Twitter data, that is, we estimate a model that contains the long-term interest rate and the log exchange rate only. The "soon" shock is identified by imposing a positive response of yields. Figure 7 presents the resulting impulse response functions. While both median responses remain unchanged, the exchange rate response is no longer significant. Hence, the information contained in Twitter matters and is indispensable for estimating and identifying powerful shocks.

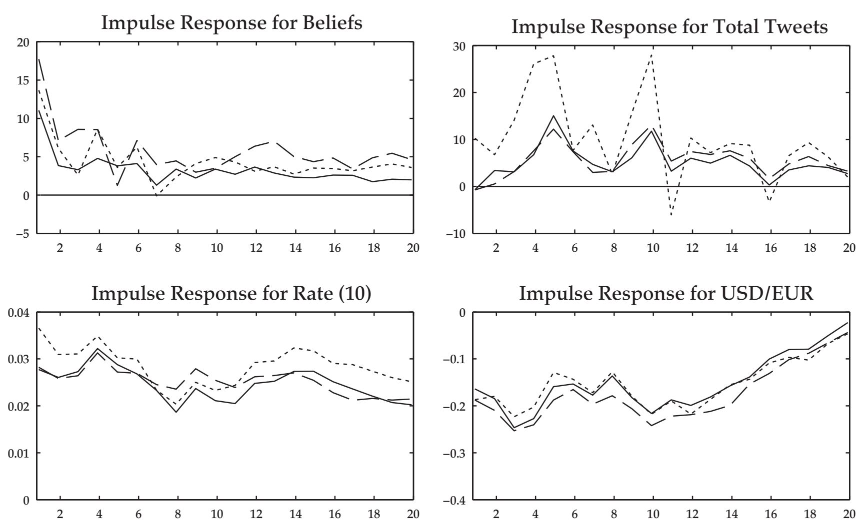
Next, we assess the results of model I by modifying the belief indicators. First, we exclude all retweets from the data set, i.e., forwarded Twitter messages. This is important because it could be argued that retweets do not contain original information and, as a consequence, should not matter for asset prices. We exclude retweets both from $Beliefs_t^i$ and from the total number of tweets. Second, we weight each tweet by the user's number of followers. The relevance of a given tweet should depend on whether the tweet is read by only a few followers or by thousands of subscribers. Again, this is done for $Beliefs_t^i$ and the total number of tweets. Figure 8 shows the impulse responses for both robustness checks. We also report the baseline results for better comparison.

Figure 7. Impulse Responses to "Tapering Soon" Belief Shock: Model Without Twitter Data



Notes: The solid line is the median impulse response function. The dotted line is the Fry-Pagan (2011) median-target impulse response. The confidence band reflects the 16th and 84th percentiles of all accepted draws.

Figure 8. Impulse Responses to "Tapering Soon" Belief Shock: Alternative Twitter Series



Notes: The solid line is the baseline model discussed before. The model that excludes retweets is represented by the dotted line. The dashed line is the result for a model in which tweets are weighted by the number of followers.

It can be seen that our benchmark results for a "sooner" shock (solid line) found in section 5.1 are still robust even when excluding retweeted messages (dotted line). Both specifications of user beliefs deliver almost the same results. This also implies that beliefs expressed through Twitter messages move markets but that the cascade of retweeted messages is of minor importance for financial markets. The results for Twitter series weighted by the number of followers (dashed line) also yield results that hardly differ from the baseline findings. Hence, we can conclude that our results are robust with respect to the way the Twitter data is included.

In addition to these two modifications, we estimate three other models, whose results are not reported in detail. All findings are available in the separate online appendix. In the first model, we include day-of-the-week dummies to account for the fact that new information that is relevant for asset prices is not equally distributed across all weekdays. For example, monetary policy decisions by the European Central Bank and other central banks are typically made on Wednesday or Thursday. FOMC decisions are already captured in the baseline model by appropriate meeting dummies. Likewise, on Monday markets respond to information received over the weekend. This modification leaves all results unchanged.

The second model assesses the identification restrictions. We use a Cholesky ordering of the variables instead of sign restrictions. Tapering beliefs are ordered first such that asset prices can respond contemporaneously to Twitter information. While the results are less significant, their overall magnitude is unchanged.

In a third variation we replace the ten-year bond yields with federal funds rate futures. These are measured by the historical continuous contract data for the CME CBOT thirty-day federal funds futures obtained from Quandl.com. We find that a "tapering soon" belief shock reduces futures prices. This is consistent with an increase in the expected federal funds rate. The response of the exchange rate remains unchanged. A corresponding historical decomposition shows that "soon" beliefs drive futures prices down before the June meeting and around the September FOMC meeting.

7. Conclusions

This paper provides an empirical analysis of the taper tantrum episode of U.S. monetary policy, in which the adjustment of expectations of a normalization of policy caused global market jitters. The analysis is based on a unique data set consisting of 90,000 Twitter messages on Federal Reserve tapering that we use to build series of investors' beliefs about an early or late tapering. A series of VAR estimates showed that shocks to market beliefs derived from Twitter messages have strong and persistent effects on bond yields and exchange rates. The paper is the first study on monetary policy using social media data.

The implications of the findings are threefold. First, our results show that beliefs about exiting QE have contractionary effects on asset prices. This is additional evidence that announcing QE had the intended expansionary effects in the first place.

Second, we show that market sentiment reflected in individual text messages matters for asset prices. Many papers use market prices such as federal funds futures or the yield curve to model expectations of future policy. However, market prices do not allow the researcher to extract information on the uncertainty of the policy outlook or the disagreement among market participants. Twitter data, which we used to show that beliefs of an early or a late tapering could change on the same day, allow such an analysis. Given the ubiquity of social media data and the ability to deal with a large data volume, the usage of this kind of data is an interesting field for future studies in monetary policy.

Third, the study sheds light on the importance of explicitly communicating an exit from unconventional monetary policy measures and offers some quantitative evidence to policymakers. Since many central banks such as the European Central Bank or the Bank of Japan are still heavily engaged in asset purchases and other unconventional policy measures, the challenges of preparing markets for the exit from those policies are yet to come. In this sense the taper tantrum episode of U.S. policy provides valuable lessons that may allow other central banks to avoid exceptional market volatility.

Moreover, our study stands out from many others that analyze the influence of social media on financial markets, because of the uniqueness of our data set. To our knowledge, most of the existing literature relies on, at best, the overall volume only. In comparison with that, our data set comprises each tweet's content, the exact timing the tweet was sent, and the name and location of the Twitter user. Hence, we are able to exploit the data in several dimensions. Given the speed at which news and information spread, it would be interesting to analyze the high-frequency impact of tapering beliefs during and around the FOMC meeting days. This is one task for further research.

Appendix. Construction of Beliefs

Here we describe our procedure of constructing our series of market beliefs, $Tweets^{soon}$ and $Tweets^{late}$, from our set of 87,621 Twitter messages that had been prefiltered out of the entire Twitter traffic by the words “taper” and “Fed.”

We prepare our data set by discarding a small number of tweets written in a language other than English. Then we take into account the fact that tweet data is given in UTC time while all other series, especially asset prices, are based on New York time. Hence, for an adequate estimation of our model, harmonization of the timing is required. Since UTC time is four hours ahead of New York time, we subtract four hours from UTC time to standardize it to New York time. As a consequence, tweets that were posted between 12:00 a.m. and 3:39 a.m. are now assigned to the previous day.

Further, we use a two-step approach to separate beliefs of early tapering from those of late tapering. In a first step, we employ dictionary methods that allow to filter tweets according to a list of predefined keywords. Table 3 and table 4 show the selected keywords for the categories “late” and “soon,” respectively.

It can be seen that both categories are separated into a list of keywords before and after September 18, 2013. This differentiation is necessary because some keywords imply tapering beliefs that depend on the date the corresponding tweet was sent, i.e., a tweet that includes the keyword “December” posted in May corresponds to expectations of a late tapering, while another tweet also referring to “December” but posted in October indicates an early tapering. Keywords that have this property are written in italics. We choose September 18, 2013 as our critical date because of the significant shift in tapering expectations that occurred after the September FOMC meeting, shown by figure 1.

For cases in which the tweets contain negations or keywords from both categories, our dictionary method is not able to allocate tweets to one of the two specified categories. Nevertheless, those tweets are identified by the algorithm that allows us, in a second step, to check and assign them manually.

Table 3. Predefined Keywords for Category "Late"

Late (until September 18, 2013)		**Late (from September 19, 2013)**	
3rd		*1st*	
2014		2014	
backed away		backed away	
bluff	incl. bluffing	bluff	incl. bluffing
dampen		dampen	
debt ceiling		debt ceiling	
deferred		deferred	
delay	incl. delayed	delay	incl. delayed
December			
dove	incl. dovish	dove	incl. dovish
Dudley		Dudley	
ease fears		ease fears	
end of this year		end of this year	
February		February	
		first	
increase		increase	
		January	
later in 2013		later in 2013	
less	incl. less likely	less	incl. less likely
November			
October			
shutdown		shutdown	
six months		six months	
third			
too soon		too soon	
until		until	
weak	incl. weakness	weak	incl. weakness
will take		will take	

Concerning keywords that are used to identify a series reflecting uncertainty, the reader is referred to Loughran and McDonald (2011). Basically, this procedure is similar to the procedure that is outlined above, except there is no need to create categories before and after September 18 for the specified tweets. For the entire sample it is sufficient to utilize a constant list of keywords.

Table 4. Predefined Keywords for Category "Soon"

Sooner (until September 18, 2013)		**Sooner (from September 19, 2013)**	
		4th	
		2013	
August			
begin		begin	
can taper		can taper	
confidence		confidence	
could taper		could taper	
drop		drop	
early		early	
end eas	incl. end easing	end eas	incl. end easing
expects to taper		expects to taper	
exit qe		exit qe	
fall		fall	
faster		faster	
		fourth	
fuel		fuel	
fell		fell	
Fisher		Fisher	
good news		good news	
hawk		hawk	
increasing expectations		increasing expectations	
in next		in next	
July			
June			
Lacker		Lacker	
likely		likely	
low unemployment		low unemployment	
lower unemployment		lower unemployment	
may begin		may begin	
may soon		may soon	
may taper		may taper	
midyear		midyear	
next few		next gew	
next meeting		next meeting	
		November	
		Nov	
now taper		now taper	
ought to taper		ought to taper	
		Octaper	
		October	
Plosser		Plosser	
pressure		pressure	
quicker		quicker	

(continued)

Table 4. (Continued)

Sooner (until September 18, 2013)		Sooner (from September 19, 2013)	
ready		ready	
reduce		reduce	
refine	incl. refining	refine	incl. refining
rumor		rumor	
septaper			
September			
set to taper		set to taper	
should taper		should taper	
slow down		slow down	
soon	incl. sooner	soon	incl. sooner
soonish		soonish	
still		still	
summer			
talk ongoing		talk ongoing	
taper hint		taper hint	
taper sooner		taper sooner	
taper talk		taper talk	
this summer			
unemployment drops		unemployment drops	
unemployment falls		unemployment falls	
unemployment fell		unemployment fell	
urge	incl. urged	urge	incl. urged
will taper off		will taper off	
will taper QE		will taper QE	
within months		within months	
would taper		would taper	
		December	

Although the dictionary approach to the content analysis of tweets is not immune to mistakes, we believe that in the aggregate the resulting belief series are representative for the true beliefs of Twitter users. For every wrongly coded "soon taper" belief, there might be a wrongly coded "late taper" belief. Aggregated over the day, these errors will potentially offset.

References

Acemoglu, D., T. A. Hassan, and A. Tahoun. 2014. "The Power of the Street: Evidence from Egypt's Arab Spring." NBER Working Paper No. 20665.

Aizenman, J., M. Binici, and M. H. Hutchison. 2016. "The Transmission of Federal Reserve Tapering News to Emerging Financial Markets." *International Journal of Central Banking* 12 (2, June): 317–56.

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

Bachmann, R., S. Elstner, and E. Sims. 2013. "Uncertainty and Economic Activity: Evidence from Business Survey Data." *American Economic Journal: Macroeconomics* 5 (2): 217–49.

Baker, S. R., N. Bloom, and S. J. Davis. 2016. "Measuring Economic Policy Uncertainty." Unpublished, Stanford University.

Bekaert, G., M. Hoerova, and M. Lo Duca. 2013. "Risk, Uncertainty and Monetary Policy." *Journal of Monetary Economics* 60 (7): 771–98.

Bloom, N. 2009. "The Impact of Uncertainty Shocks." *Econometrica* 77 (3): 623–85.

Choi, H., and H. Varian. 2009. "Predicting Initial Claims for Unemployment Benefits." Technical Report, Google.com.

———. 2012. "Predicting the Present with Google Trends." *Economic Record* 88 (s1): 2–9.

Da, Z., J. Engelberg, and P. Gao. 2011. "In Search of Attention." *Journal of Finance* 66 (5): 1461–99.

D'Amico, S., W. English, D. Lopez-Salido, and E. Nelson. 2012. "The Federal Reserve's Large-Scale Asset Purchase Programs: Rationale and Effects." Finance and Economics Discussion Series Paper No. 2012-85, Board of Governors of the Federal Reserve System.

Dahlhaus, T., and G. Vasishtha. 2014. "The Impact of U.S. Monetary Policy Normalization on Capital Flows to Emerging-Market Economies." Working Paper No. 2014-53, Bank of Canada.

Dergiades, T., C. Milas, and T. Panagiotidis. 2015. "Tweets, Google Trends, and Sovereign Spreads in the GIIPS." *Oxford Economic Papers* 67 (2): 406–32.

Eichengreen, B., and P. Gupta. 2015. "Tapering Talk: The Impact of Expectations of Reduced Federal Reserve Security Purchases on Emerging Markets." *Emerging Markets Review* 25: 1–15.

Fry, R., and A. Pagan. 2011. "Sign Restrictions in Structural Vector Autoregressions: A Critical Review." *Journal of Economic Literature* 49 (4): 938–60.

Istrefi, K., and A. Piloiu. 2013. "Economic Policy Uncertainty and Inflation Expectations." Unpublished, Goethe University Frankfurt.

Kashyap, A. 2013. "Commentary: The Ins and Outs of LSAPs." Proceedings of the 2013 Economic Policy Symposium on "Global Dimensions of Unconventional Monetary Policy," 113–24. Federal Reserve Bank of Kansas City.

Krishnamurthy, A., and A. Vissing-Jorgensen. 2013. "The Ins and Outs of LSAPs." In *Global Dimensions of Unconventional Monetary Policy*, 57–111. Proceedings of the 2013 Economic Policy Symposium sponsored by the Federal Reserve Bank of Kansas City, held in Jackson Hole, Wyoming, August 21–23.

Loughran, T., and B. McDonald. 2011. "When Is a Liability Not a Liability? Textual Analysis, Dictionaries and 10-Ks." *Journal of Finance* 66 (1): 35–65.

Matheson, T., and E. Stavrev. 2014. "News and Monetary Shocks at a High Frequency: A Simple Approach." *Economics Letters* 125 (2): 282–86.

Mishra, P., K. Moriyama, P. N'Diaye, and L. Nguyen. 2014. "Impact of Fed Tapering Announcements on Emerging Markets." IMF Working Paper No. 14/109.

Nechio, F. 2014. "Fed Tapering News and Emerging Markets." Economic Letter No. 2014-06, Federal Reserve Bank of San Francisco.

Rogers, J. H., C. Scotti, and J. H. Wright. 2014. "Evaluating Asset-Market Effects of Unconventional Monetary Policy: A Multi-country Review." *Economic Policy* 29 (80): 749–99.

Sahay, R., V. Arora, T. Arvanitis, H. Faruquee, P. N'Diaye, and T. Mancini-Griffoli. 2014. "Emerging Market Volatility: Lessons from the Taper Tantrum." IMF Staff Discussion Note No. 14/09.

Stein, J. C. 2014. "Challenges for Monetary Policy Communication." Speech at the Money Marketeers of New York University, New York City, May 6, 2014.

Tillmann, P. 2015. "Tapering Talk on Twitter and the Transmission to Emerging Economies." Unpublished (October).

Uhlig, H. 2005. "What Are the Effects of Monetary Policy on Output? Results from an Agnostic Identification Procedure." *Journal of Monetary Economics* 52 (2): 381–419.

Vlastakis, N., and R. N. Markellos. 2012. "Information Demand and Stock Market Volatility." *Journal of Banking and Finance* 36 (6): 1808–21.