Global Research Unit
Working Paper #2016-014

Tapering Talk on Twitter and the Transmission to Emerging Economies

Peter Tillmann, Justus-Liebig-University Giessen

Tapering Talk on Twitter and the Transmission to Emerging Economies*

Peter Tillmann†

Abstract

When in 2013 the Federal Reserve started to discuss unwinding its asset purchases and exiting unconventional monetary policy ("tapering talk"), markets adjusted expectations and asset prices dropped sharply, in particular in emerging markets. In this paper we quantify the effect of the tapering talk on emerging financial markets. We use the entire stream of tapering-related messages sent on Twitter.com, the social media network, to build a series of market participants' beliefs of an early tapering. This series is then included in a VAR system, in which a tapering belief shock is identified using sign restrictions. We find that the tapering shock has significant effects on emerging financial market and explains almost the entire dynamics of bond prices, stock prices, exchange rates and CDS-spreads during the "taper tantrum". The results remain robust if we exclude retweets and control for major policy events.

Keywords: Tapering, unconventional monetary policy, emerging markets, quantitative easing, Twitter

JEL classification: E32, E44, E52

*I thank Mahvash S. Qureshi and Annette Meinusch for extremely helpful discussions as well as participants of the ADBI Workshop on "Global Uncertainty, Macroeconomic Shocks and Growth in Asia" (Tokyo, 2015) for important feedback. Nhung Hong Dang, Lucas Hafemann and Florian Viereck provided excellent research assistance.

†Email: peter.tillmann@wirtschaft.uni-giessen.de

1 Introduction

Since 2008 the Federal Reserve (Fed) conducted unconventional monetary policies known as Quantitative Easing (QE). One element of QE was asset purchases from the private sector. In his regular testimony before the U.S. senate on May 22 2013, Fed chairman Ben Bernanke mentioned preparations for the eventual exit from QE. Unwinding QE and returning to conventional interest rate policy is known as tapering, while observers refer to the Fed's discussion of tapering as "tapering talk". The result of Bernanke's testimony was a period of heightened uncertainty about future monetary policy, high volatility and pressure on emerging market currencies as markets had not anticipated an exit from unconventional policies. Since the market responses appeared overly nervous, observers coined the phrase "tapering tantrum".[1]

Speculation about tapering gained momentum after the May 22 testimony and was intensified at the June FOMC meeting, on which the Fed continued the taper talk. Markets adjusted expectations and expected a reduction in monthly asset purchases for the September meeting. In September, however, tapering was postponed and eventually took off in 2014. At the time of writing, in April 2015, observers fear a "super taper tantrum" or a "tapering 2.0", i.e. a renewed period of high volatility, as the Fed prepares for a "lift-off" of the Federal Funds rate from the zero lower bound, the first interest rate increase in nearly a decade.[2]

This paper empirically quantifies the responses of emerging economies to the taper talk.[3] For that purpose we employ a vector autoregression model and identify a tapering shock using sign restrictions. The central innovation of this paper is to use data from social media in order to identify the tapering shock. We obtain all messages sent over Twitter.com on "tapering" and from that select the tweets which contain a view on the timing of the Fed's tapering decision. A dictionary approach is used to disentangle these tweets into those foreseeing an early tapering and those arguing in favor of a late tapering. The series of beliefs of an early tapering is used to estimate a vector autoregression (VAR) model that also includes U.S. long-term bond yields and a measure of financial conditions in emerging market economies, that is, bond and stock market indices, exchange rates and CDS spreads. We employ

[1] Interestingly, Bernanke never used the word "taper" in his congressional testimony. He just said "in the next few meetings, we could take a step down in our pace of purchases". This quote has been taken from the Wall Street Journal blog, see http://blogs.wsj.com/economics/2013/06/07/why-the-fed-hates-the-word-tapering/.

[2] See Financial Times (2015), The Economist (2015).

[3] There is very little research on exiting from unconventional policies. A notable exception is Krishnamurty and Vissing-Jorgensen (2013).

sign restrictions to identify a tapering shock as a shock that raises U.S. yields and the Twitter beliefs of an early tapering at the same time. It is shown that such a shock has a strong and significant impact on emerging markets' bond prices, stock prices and exchange rates against the U.S. dollar. A historical decomposition of the estimated VAR model shows that this model is able to replicate a large fractions of the dynamics of financial variables during the tapering period. These results should inform other central banks such as the European Central Bank of the Bank of Japan, which will eventually also have to communicate their exit from unconventional monetary policy.

The main contribution of the paper is to quantify the effects of an adjustment of investors' expectations on emerging economies using social media data. This novel data set has one important advantage over alternative data sources: it allows us to find a proxy for the degree of disagreement of market participants. We disentangle the Twitter messages according to the views on the timing of the taper. As a result we obtain two series, one highlighting an early taper and one foreseeing a late taper, that coexist. Thus, we can study the disagreement of market participants and are not forced to use the aggregate market view as, for example, incorporated in futures data.

Twitter data has not been used before in a macroeconomic context. In a companion paper, Meinusch and Tillmann (2015) work with the same data set, but model only the domestic effects of tapering beliefs on U.S. variables.

The remainder of the paper is organized as follows: Section two surveys the literature. Section three presents the data, while section four presents the model. The results and a battery of robustness checks are discussed in section five. Section six concludes.

2 The "taper talk" in the literature

Several papers address the spillovers of unconventional monetary policy in the U.S. to advanced and emerging economies, e.g. Bauer and Neely (2014), Lim et al. (2014), Bowman et al. (2014), Glick and Leduc (2012) and Neely (2015). Since we are interested in the spillovers from exiting QE, these papers offer some guidance as regards the likely (absolute) magnitude of the responses. A few papers aim at modelling the tapering episode empirically.[4] The problem faced by these studies is how to model shifts in market expectations. Eichengreen and Gupta (2015) and Aizenman et al. (2016) either interpret all fluctuations in emerging economies in

[4]See Sahay et al. (2014) and Rai and Suchanek (2014) for event studies on tapering.

2013 as resulting from the taper talk or employ dummy variables to measure the effect of FOMC decisions, respectively. The contribution of both studies is that they are able to link the strength of tapering-driven fluctuations to the fundamental weaknesses in the economies. While Eichengreen and Gupta (2015) and Aizenman et al. (2016) do not find that stronger macroeconomic fundamentals make emerging economies less vulnerable to tapering-related spillovers, the paper by Ahmed et al. (2015) finds that emerging countries with better fundamentals suffered less. After controlling for fundamentals, they find that financial conditions deteriorated more in countries that had earlier received larger capital inflows. Ahmed et al. (2015) also use the change in financial conditions between April and August 2013 as the dependent variable.

The drawback of these studies, however, is that measuring beliefs in this way might not be able to disentangle tapering beliefs from other driving forces. In fact, in 2013 emerging market were also affected by the war in Syria, the ongoing European debt crisis and the budgetary conflict in U.S. politics, to name just a few other potential sources of financial stress. Mishra et al. (2014) use an event-study approach and measure market reactions in a two-day window around FOMC meeting and days on which FOMC minutes were published.

Matheson and Stavrev (2014) and Dahlhaus and Vasishta (2014) employ VAR models with sign restrictions to identify structural shocks. Matheson and Stavrev (2014) identify a monetary shock that is restricted to lower asset prices and increase interest rates. Dahlhaus and Vasishta (2014) identify a "policy normalization shock", which raises expectations of future policy measured by Federal Funds Futures but have no effect on the current policy rate.

In this paper, we use data from Twitter.com to obtain a measure of attention to Fed policy and policy transmission to emerging economies. In a companion paper, Meinusch and Tillmann (2015), we use the same data set. However, this paper is different from Meinusch and Tillmann (2015) in one important aspect ways: we address the spillovers to emerging countries and the channels of these spillovers, while Meinusch and Tillmann (2015) focus on the domestic effects of tapering beliefs only.

3 Tapering on Twitter

We use data from Twitter.com, the social network application to send text messages 140 characters in lengths, to construct a measure of investors' beliefs about the tapering of unconventional policies. Twitter usage is very common among financial professionals and journalists to engage in a public debate and to follow the views of

others. In our empirical analysis we use this series of Twitter messages as a measure of financial professionals' attention to a shift in the Fed's policy stance from which we construct series of policy beliefs. We obtain 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 analyses. The data set includes 87.621 tweets from 27.276 users located in 136 countries and the exact time they were sent.[5] We remove public holidays and weekends in the U.S. from these series because there is no financial data available on these days. We end with a belief series of 139 days in length.
From the overall Twitter traffic we select those tweets that express a view on the timing of the tapering. A dictionary approach is used to identify a series expressing the belief of an early tapering and a late tapering, respectively. The list of keywords and the construction of these series are reported in the appendix. A dictionary approach is advantageous here because, by construction, all messages share a common context. There is no danger that a specific keyword has different meanings in alternative contexts. All tweets are short (140 characters as a maximum) and highly focused on the Fed's tapering decision. The belief series are now included in a VAR model.

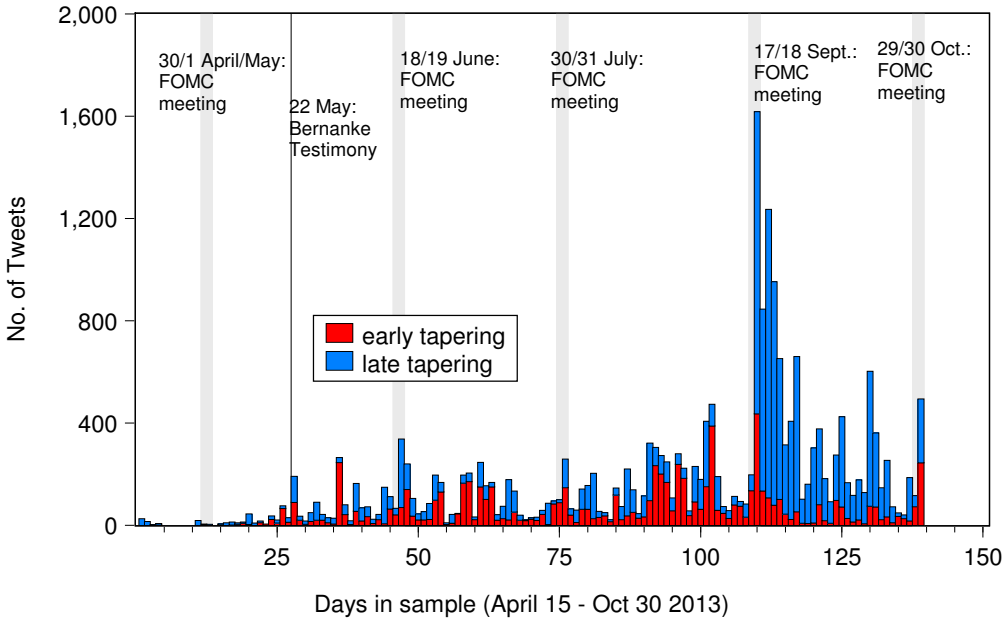

Figure 1: Tweets expressing beliefs on soon/late tapering

[5]Retweets 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 them in our measure of investors' beliefs. Below we show that the results are robust to excluding retweets from our belief series.

For the purpose of this paper, the tweets are aggregated into a series at daily frequency. Figure (1) presents the series of "early tapering" and "late tapering" beliefs. It can be seen that the number of tweets increases around Bernanke's testimony on May 22 and around each FOMC meeting depicted in the graph. The data volume explodes at the September 2013 FOMC meeting, when the Fed decided to postpone tapering until 2014. As expected, following the September FOMC meeting the majority of tweets foresee a late tapering decision.

4 The model

A vector autoregressive model is able to capture the dynamics of the international transmission of shocks. The reduced-form representation is

$$Y_t = A_0 + A(L)Y_t + u_t, \quad \text{with } E[u_t u_t'] = \Sigma_u$$

where $A(L)$ reflects the matrix polynomial in the lag operator of order p and u_t constitutes a white noise process with variance-covariance matrix Σ_u. We also add a constant to the model. The model is estimated on the following vector of endogenous variables at a daily frequency

$$Y_t = \left(\Delta Tweets_t^{soon}, \Delta R_t^{US}, \Delta EM_t\right)'$$

where Δ denotes first differences. The VAR model is kept small and includes three variables: (1) The series of Twitter messages on an early tapering relative to all tweets, $Tweets_t^{soon}$, (2) the yield on 10-year U.S. government bonds, R_t^{US} and (3) a variable reflecting financial conditions in emerging market economies, EM_t. This third variable is one of the following series: the EMBI+ emerging market bond index, the MSCI stock price index for emerging markets, the exchange rate of one of the "fragile five" economies, i.e. Indonesia, India, South Africa, Turkey and Brazil, or the 12-months CDS spreads of Indonesia, Turkey and South Africa.[6] We estimate the VAR system with 10 lags of the endogenous variables.

Figure (2) plots the evolution of the EMBI bond price and the MSCI stock price indices. The shaded areas reflect the testimony on May 22, the June and the September FOMC meeting. Both, stocks and bonds loose about 12% of their value during the first half of the sample with the maximum loss occurring after the June FOMC meeting when the Fed strengthened its tapering intention. After that, both prices recovered, and this recovery gained momentum around the September meeting. Fig-

[6] Unfortunately, CDS spreads were not available for India and Brazil.

ure (3) depicts the foreign exchange rates of the "fragile five". The currencies loose between 20% and 25% of their value during the tapering talk. The CDS spreads, see figure (4) increase by about 100 basis points during the episode of the taper talk.

As in other VAR models, estimating the structural shock that is driving the dynamics of the system is crucial. A "tapering talk" shock is identified by imposing restrictions on the signs of the responses following the work of Uhlig (2005). In particular, such a shock is assumed to raise the share of users on Twitter foreseeing an early tapering as reflected by an increase in $Tweets_t^{soon}$ and raises U.S. bond yields. The response of emerging markets is left unrestricted. The restrictions are imposed for two days and are summarized in following table:

Table 1: Sign restrictions to identify a tapering belief shock

$\Delta Tweets_t^{soon}$	ΔR_t^{US}	ΔEM_t
$+$ [for t=0,1]	$+$ [for t=0,1]	unrestricted

Sign resections appear superior to a recursive ordering of the variables because we avoid the need to prespecify an ordering of the variables' responses over time. Below we will also present results for an alternative specification in which R_t^{US} remains unrestricted.

5 Results

In this section the results of the empirical analysis are presented and the robustness of the findings is analyzed.

Impulse responses. The main results are presented in a series of impulse response functions in Figures (5) to (14). All figures depict the adjustment following a tapering shock one standard deviation in size (red line) surrounded by a confidence band which reflects the 16th and 84th percentiles of all draws. To account for the Fry-Pagan (2011) critique, we also report the mean-target impulse response (black line), i.e. the one response out of the large number of drawn responses that is closest to the median response.

In all results the number of tapering tweets rises strongly after a shock (not shown). Furthermore, following a tapering shock the long-term interest rate increases significantly in all specifications. The change in U.S. bond yields rises by three to four basis points. Together with the response of tweets, the numbers suggest that an

increase in tweets by 100%, which is observed quite often in the sample, raises rates by 20 basis points.
Most importantly, financial conditions in emerging economies are very sensitive to tapering shocks. The EMBI index of bond prices drops by 0.4%. Thus, spillovers from U.S. monetary policy affects the refinancing conditions of sovereigns in emerging countries. Likewise, stock prices in emerging economies, see Figure (6) drop by almost 0.7% after the shock. When we turn to the responses of local currencies' exchange rate against the USD, we see, as a general tendency, a depreciation of the local currency. While the Indonesian rupiah, Figure (7), depreciates by 0.4%, the Indian rupee, Figure (8), looses 0.8% of its value following the taper talk. A similar and also significant fall is observed for the Turkish lira, the South African rand and the Brazilian real, see Figures (9), (10), and (11), respectively.[7] CDS spreads, i.e. a measure of the country's risk of default, increase by about four to seven basis points.

Historical decomposition. Based on the VAR estimates the identified shock, we could isolate the contribution of the tapering shock to each of the endogenous variables over time. In Figures (15), (16) and (17)we plot these contributions, which are denoted in the units of the original variable. Each figure also shows three shaded dates, which correspond to the dates of the May 22 testimony, the June FOMC meeting and the September FOMC meeting.
The model is able to explain a large part of the evolution of the EMBI and the MSCI indices, respectively. The green areas, which reflect the contribution of the tapering shock, captures a large fraction of the observables. For the exchange rates series, the explanatory power is somewhat lower, which is not surprising given general the difficulty to explain exchange rates by macroeconomic fundamentals. Nevertheless, the model is again able to replicate a large share of the exchange rate dynamics. Furthermore, tapering talk shocks explain a large part of the CDS spreads of Indonesia, Turkey and South Africa. Overall, it can be seen that the taper talk has been responsible for most of the fluctuations in emerging financial markets in 2013.

Lag order. In Figures (18) and (19) we change the lag order of the VAR system from 10 to four or eight lags, respectively. The sensitivity of emerging markets' bond prices to tapering shocks does not change noticeably.

Controlling for policy events. Our series of Twitter messages peaks at the May 22 Testimony of chairman Bernanke, the June FOMC meeting where it was made

[7]These responses are closely in line with the findings of Aizenman et al. (2016).

clear that the path of asset purchases will be moderated before the end of 2013 and the September FOMC meeting at which the Fed surprised markets by maintaining the pace of asset purchases. On these dates, the series of tweets clearly reflect official Fed communication. It is interesting, though, to control for these events in order to analyze whether the Twitter series contains information beyond what could be captured by policy dummies. For that purpose we construct three series of dummies which are one at one of the policy events mentioned before and zero otherwise. These dummies then enter the VAR model as deterministic variables. The result is presented in Figure (20). It can be seen that all findings of the baseline specification remain unchanged. Hence, the information contained in the series of tapering tweets is not restricted to the policy events in the sample period.

The role of retweets. An intensified tapering talk moves financial markets. The overall volume of tweets, however, includes a large number of forwarded messages ("retweets"). One could argue that every time a message is retweeted, there is no information gained on financial markets. In order to quantify the role of original information, i.e. original tweets, and retweets, we net out the retweets from our series of Twitter volume and estimate the baseline model including the EMBI index again. The results based on the number of original tweets only are presented in Figure (21). The drop in the EMBI bond index is as strong as in the benchmark case. It seems that financial professional put equal attention to pieces of original information and on flow of endorsed or retweeted information.

Restricting only the tapering beliefs. The previous findings showed that a shock leads to an increase in U.S. bond yields and to spillovers to emerging economies. These results were derived from a set of sign restrictions that included a restriction of the response of bond yields. In order to find our whether the responses of emerging economies stem primarily from the information included in $Tweets_t^{soon}$ or from the reaction of R_t^{US}, we now present results from a specification in which the latter remains unrestricted. This alternative set of restrictions is summarized in the following table:

Table 2: Alternative sign restrictions to identify a tapering belief shock

$Tweets_t^{soon}$	R_t^{US}	EM_t
$+$ [for t=0,1]	unrestricted	unrestricted

In addition, we also add the three policy dummies mentioned before to the deterministic part of the VAR model, on which the information coming from bond yields is particularly important. Thus, we stack the cards against us and estimate a specification in which we estimate the information content of the belief series in the inter-meeting period. The results are shown in figures (22), (23) and (24). The response of the U.S. yield is no longer significant. Interestingly, however, the effect on emerging economies is still present. Emerging markets' bond prices, as reflected by the EMBI index, drop by 0.3%. The MSCI index falls strongly and the Indian rupee loses approximately 0.7% of its value. Hence, the spillovers to emerging economies remain even if we exclude important policy days and relax the constraint on U.S. yields.

Taken together, the results corroborate the view that the expected policy shift of the Fed as reflected in an increase in "early tapering" beliefs had strong spillovers to emerging economies.

6 Conclusions

In this paper we empirically quantified the sensitivity of emerging markets to the tapering talk in the U.S. in 2013. A central innovation of this paper is to use a data set consisting of almost 90,000 Twitter messages to extract market beliefs of an early tapering decision. Twitter data has not been used before in the monetary policy literature but proved very helpful for modelling the beliefs. A VAR model in which shocks were identified using sign restrictions revealed a significant and strong effect of a tapering talk-shock to emerging economies, with bond prices and stock prices falling, exchanges rate depreciating against the dollar and CDS spreads widening.
While the Fed did already exit from unconventional monetary policies, many other central banks will follow eventually. This paper shows the spillover effects of a major shift in policy on emerging market economies to be large. The paper can also be seen as an estimate of how a misguided communication on the part of the Fed is transmitted to global asset markets. Other central banks will study the tantrum-episode in order to learn about how to communicate the exit from unconventional monetary policy.
This paper also highlighted the usefulness of social media data for macroeconomic analysis. When peoples' expectations play a role, i.e. for asset pricing and inflation projections, social media data is a particularly interesting field of future research.

References

[1] Ahmed, S., B. Coulibaly and A. Zlate (2015): "International financial spillovers to emerging market economies: How important are economic fundamentals", *International Finance Discussion Paper* No. 1135, Board of Governors of the Federal Reserve System.

[2] 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* June 2016, 317-356.

[3] Bauer, M. D. and C. J. Neely (2014): "International channels of the Fed's unconventional monetary policy", *Journal of International Money and Finance* 44, 24-46.

[4] Bowman, D. J. M. Londono and H. Sapriza (2014): "U.S. unconventional monetary policy and transmission to emerging market economies", *International Finance Discussion Paper* No. 1109, Board of Governors of the Federal Reserve System.

[5] 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.

[6] The Economist (2015): "Mismatch point", The Economist, March 21 2015, 11-12.

[7] 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.

[8] Financial Times (2015): "IMF warns of 'super taper tantrum' ahead", Financial Times, April 15 2015.

[9] Fry, R. and A. Pagan (2011): "Sign restrictions in structural vector autoregressions", *Journal of Economic Literature* 49, 938-960.

[10] Glick, R. and S. Leduc (2012): "Central bank announcements of asset purchases and the impact of the global financial and commodity markets", *Journal of International Money and Finance* 31, 2078-2101.

[11] Krishnamurty, A. and A. Vissing-Jorgensen (2013): "The Ins and Outs of LSAPs", Proceedings of the 2013 Economic Policy Symposium on "Global Dimensions of Unconventional Monetary Policy", Federal Reserve Bank of Kansas City.

[12] Lim, J. J., S. Mohapatra and M. Stocker (2014): "Tinker, taper, QE, bye? The effect of quantitative easing on financial flows to developing countries", *Policy Research Working Paper* No. 6820, The World Bank.

[13] Matheson, T. and E. Stavrev (2014): "News and monetary shocks at a high frequency: a simple approach", *Economics Letters* 125, 282-286.

[14] Meinusch, A. and P. Tillmann (2015): "Unconventional monetary policy and tapering uncertainty: evidence from Twitter", *unpublished*, University of Giessen.

[15] 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, International Monetary Fund.

[16] Neely, C. J. (2015): "Unconventional monetary policy had large international effects", *Journal of Banking and Finance* 52, 101-111.

[17] Rai, V. and L. Suchanek (2014): "The effect of the Federal Reserve's tapering announcement on emerging markets", *Working Paper* No. 2014-50, Bank of Canada.

[18] 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* SDN/14/09, International Monetary Fund.

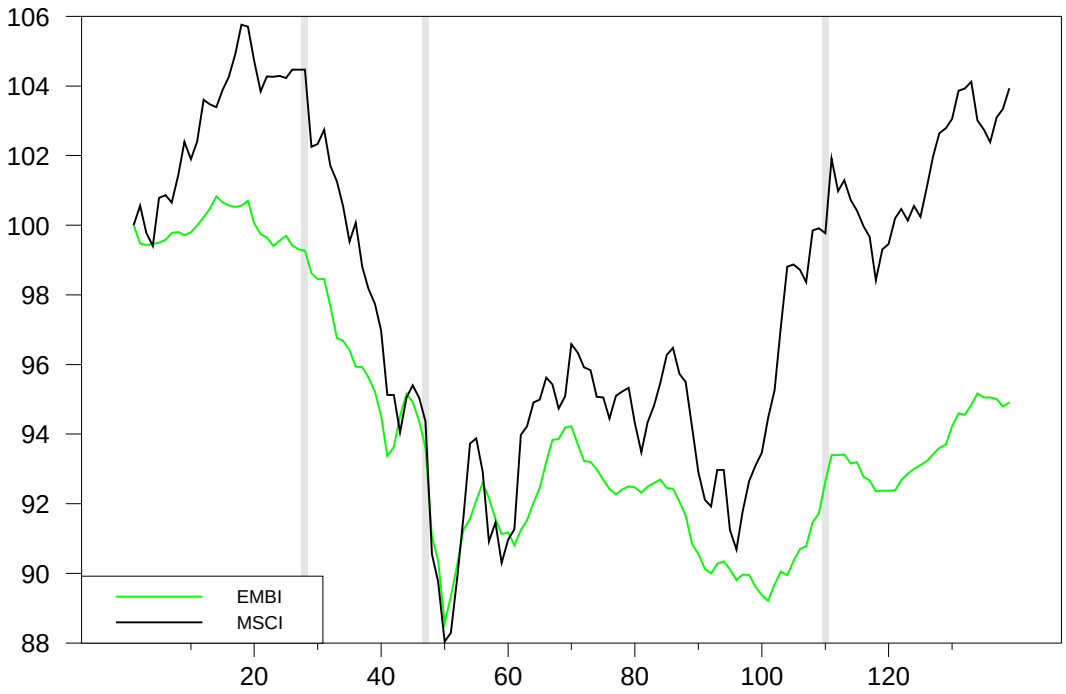

Figure 2: EMBI and MSCI series (April 15 2013 = 100)

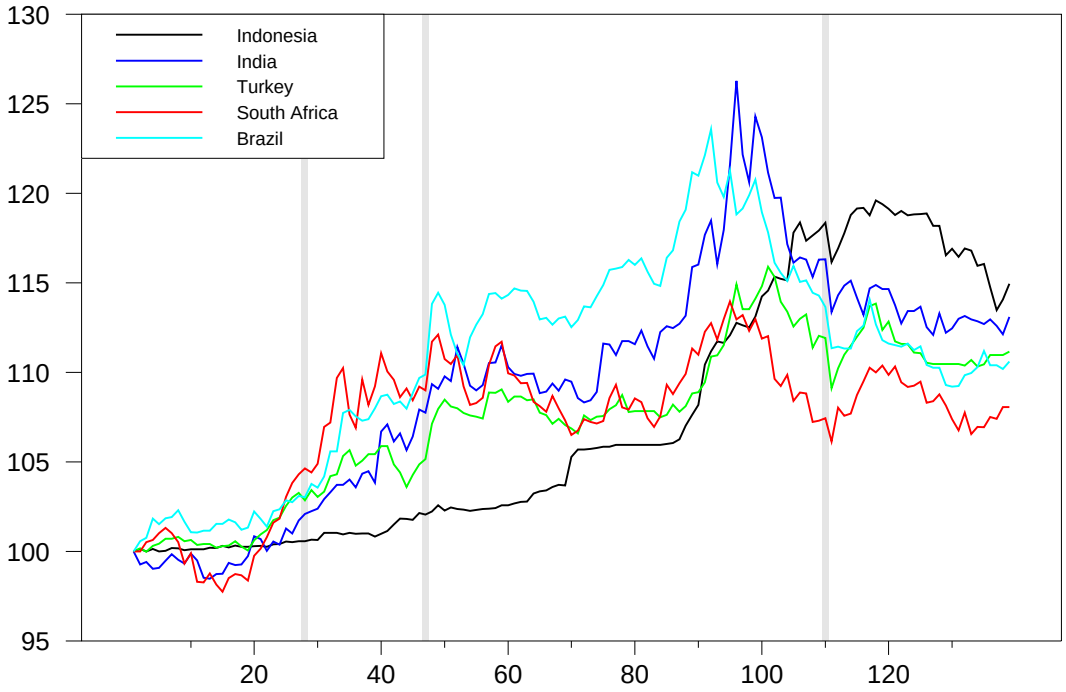

Figure 3: Exchange rates of the "fragile five" against the USD (April 15 2013 = 100)

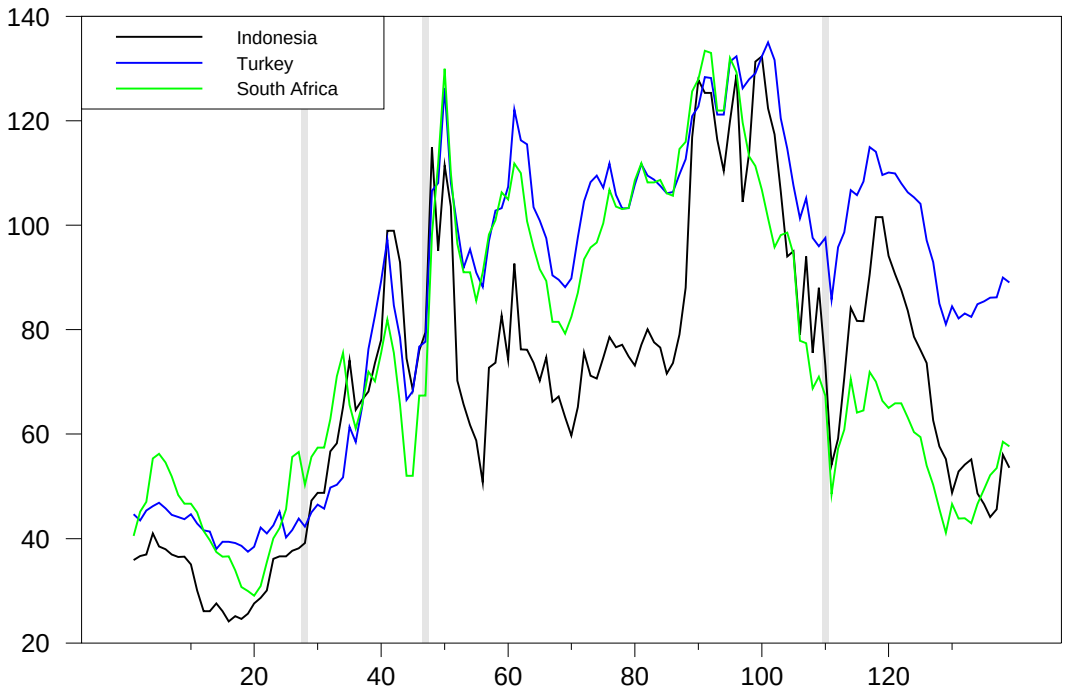

Figure 4: 12 months CDS spreads

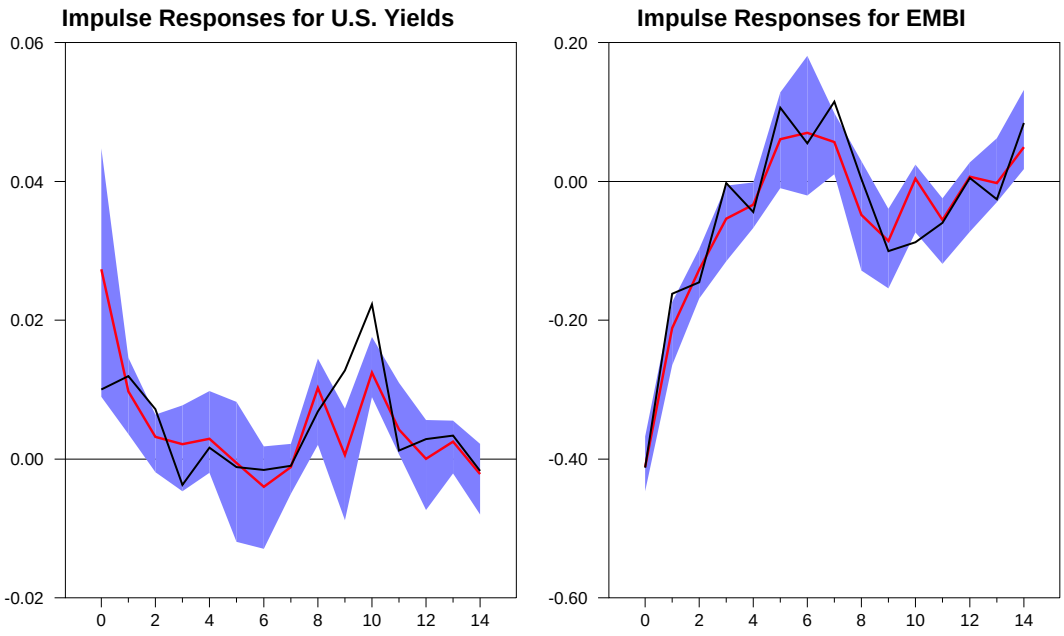

Figure 5: Impulse responses to tapering shock: EMBI+ index

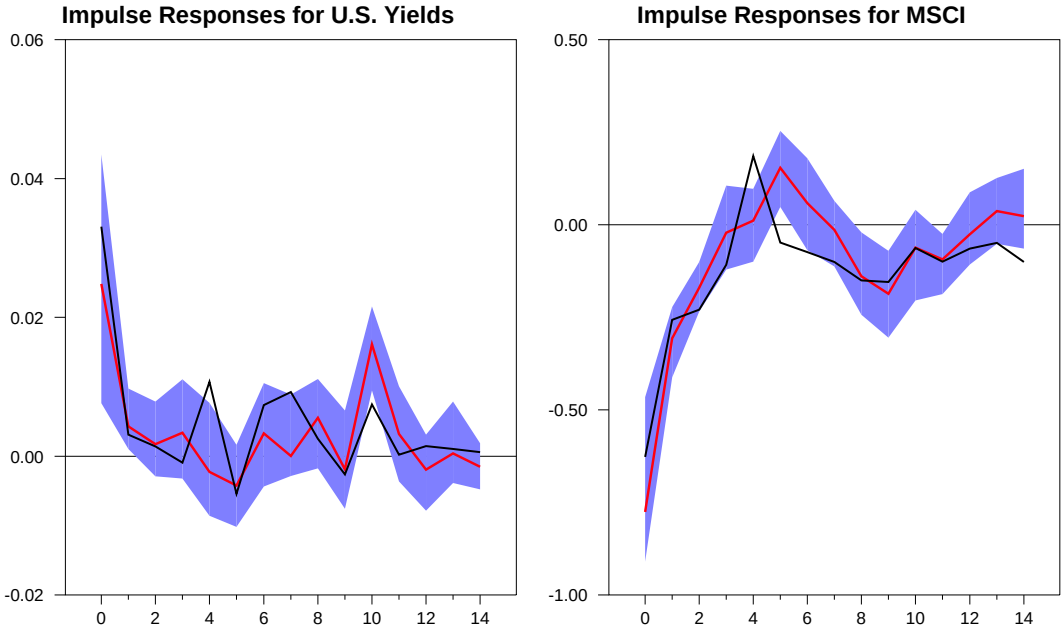

Figure 6: Impulse responses to tapering shock: MSCI index

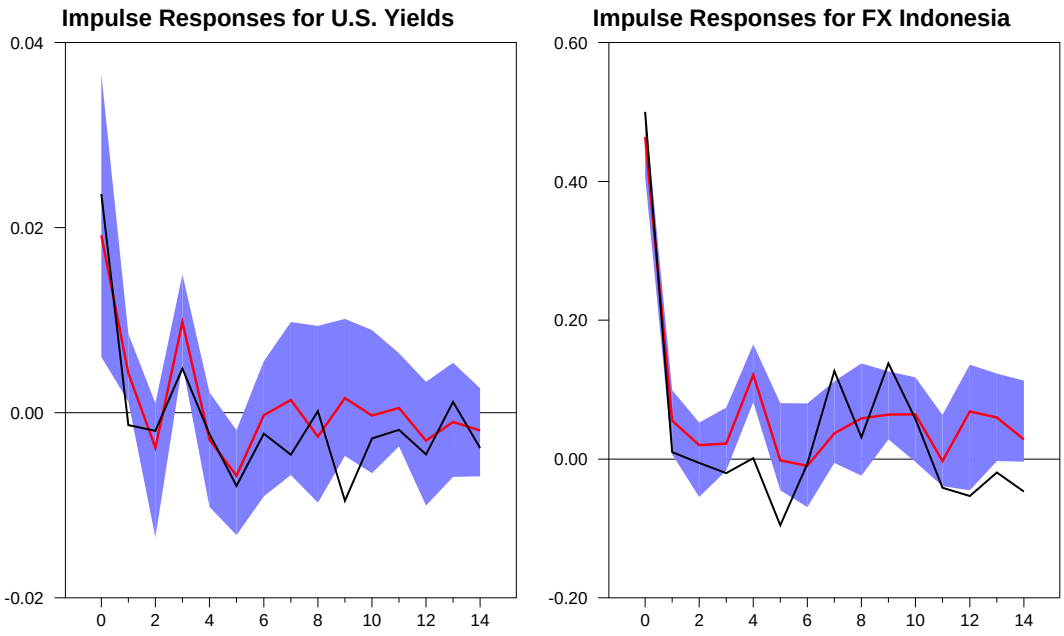

Figure 7: Impulse responses to tapering shock: exchange rate Indonesia

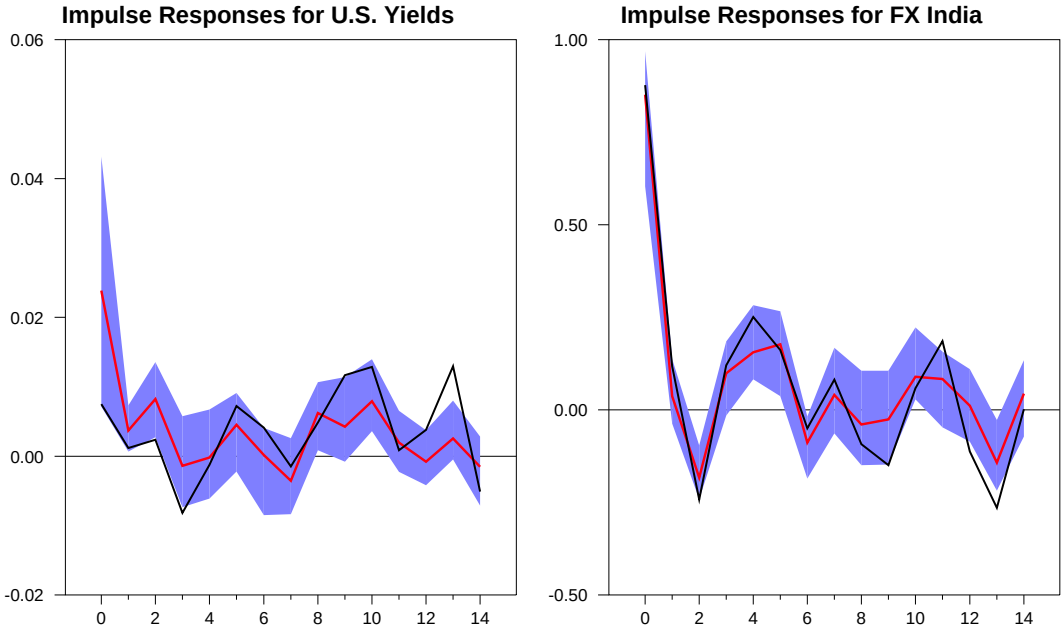

Figure 8: Impulse responses to tapering shock: exchange rate India

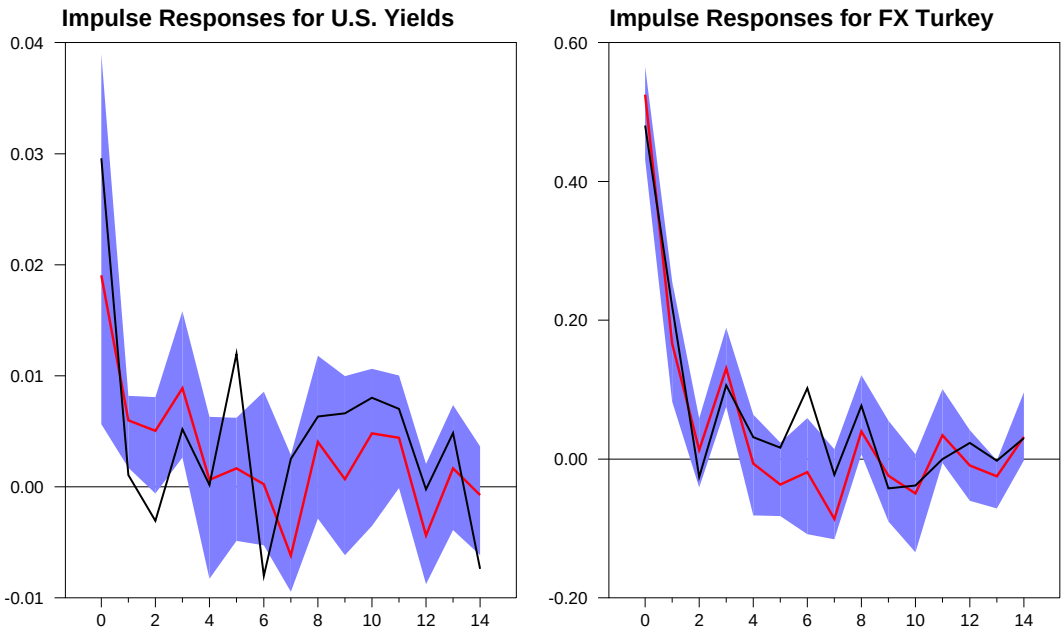

Figure 9: Impulse responses to tapering shock: exchange rate Turkey

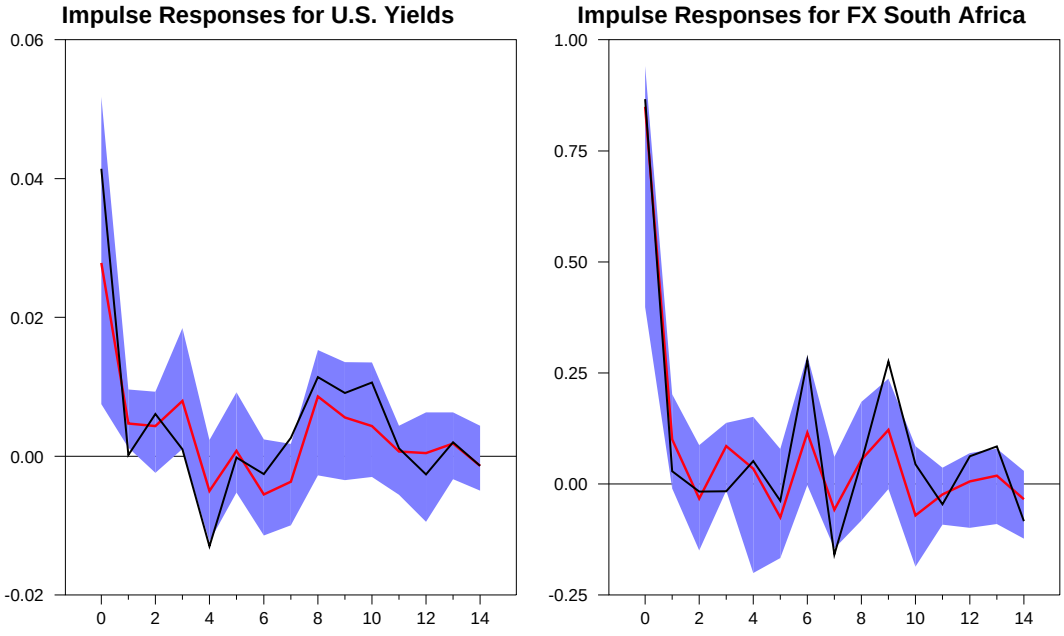

Figure 10: Impulse responses to tapering shock: exchange rate South Africa

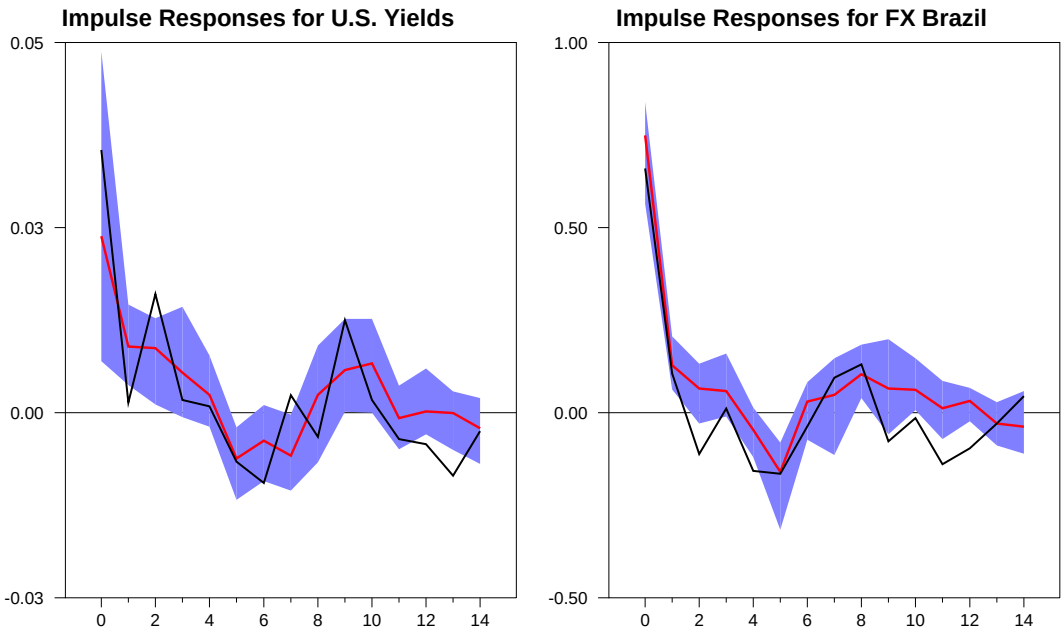

Figure 11: Impulse responses to tapering shock: exchange rate Brazil

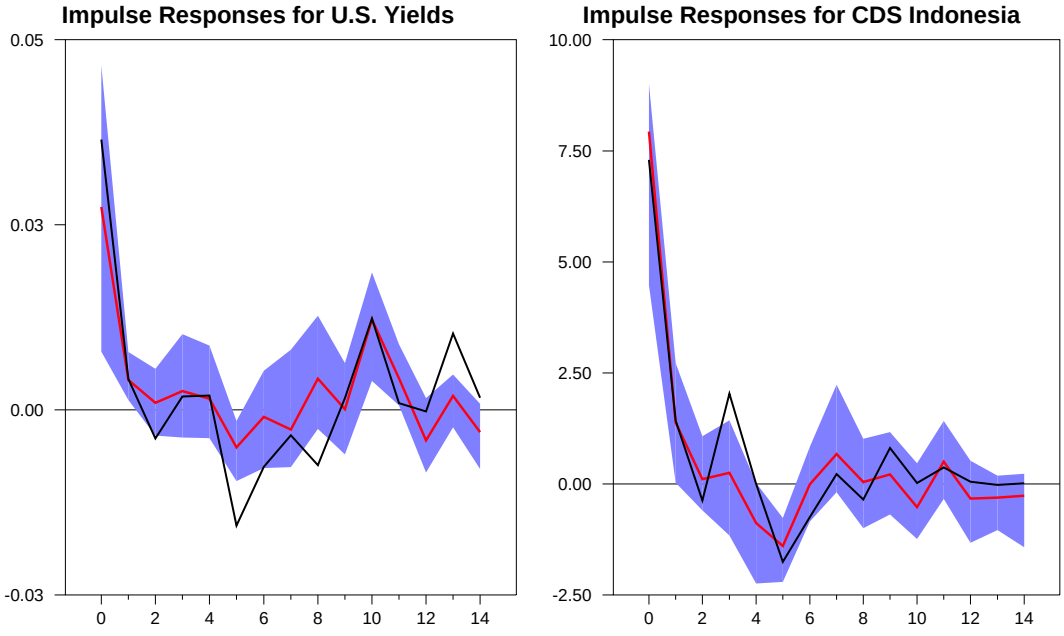

Figure 12: Impulse responses to tapering shock: CDS spread Indonesia

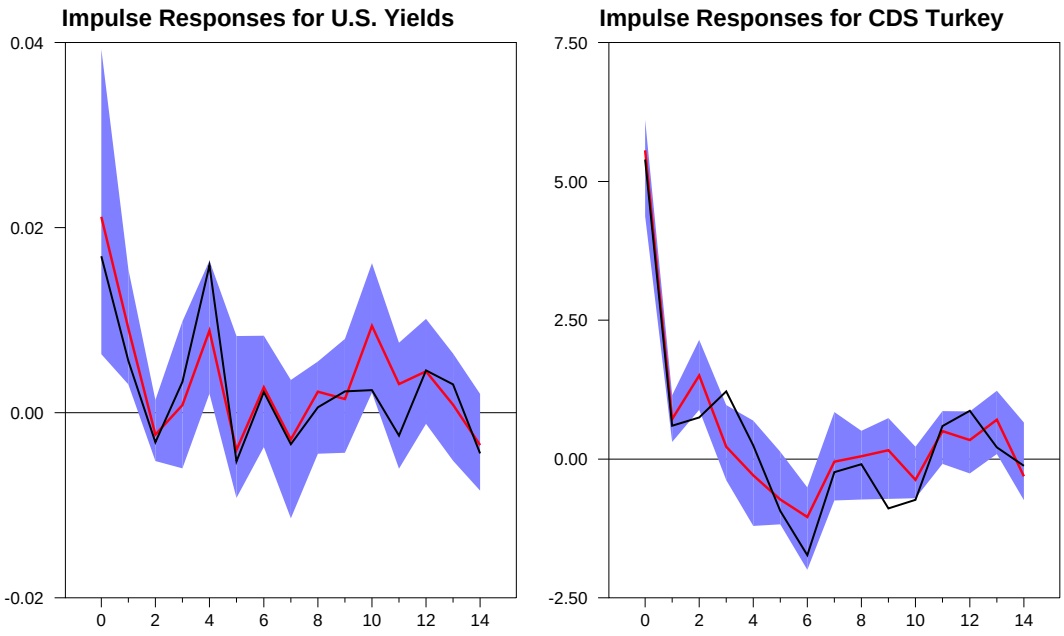

Figure 13: Impulse responses to tapering shock: CDS spread Turkey

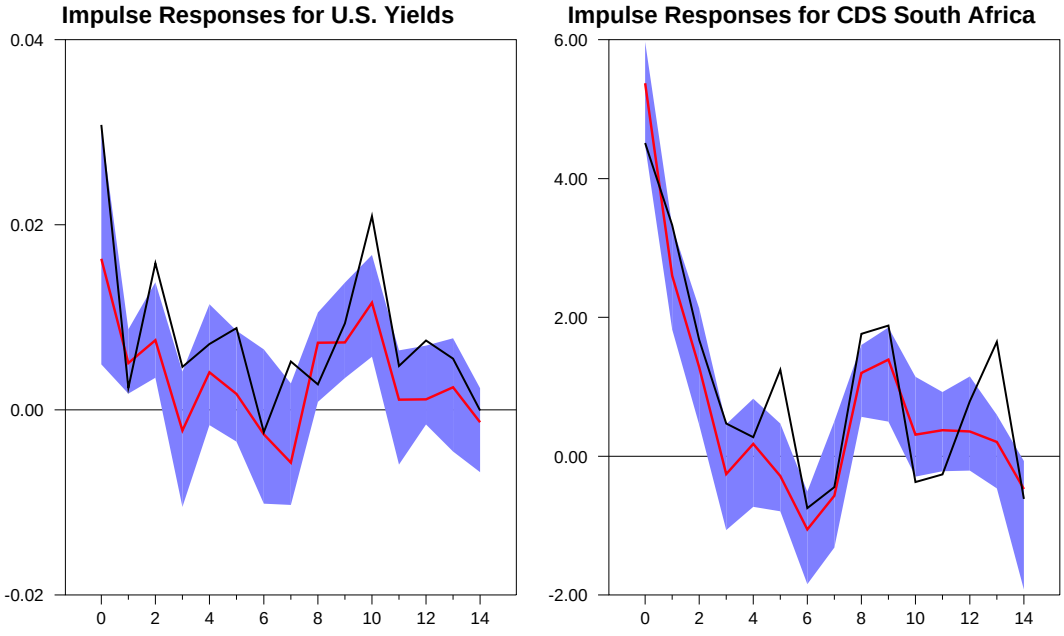

Figure 14: Impulse responses to tapering shock: CDS spread South Africa

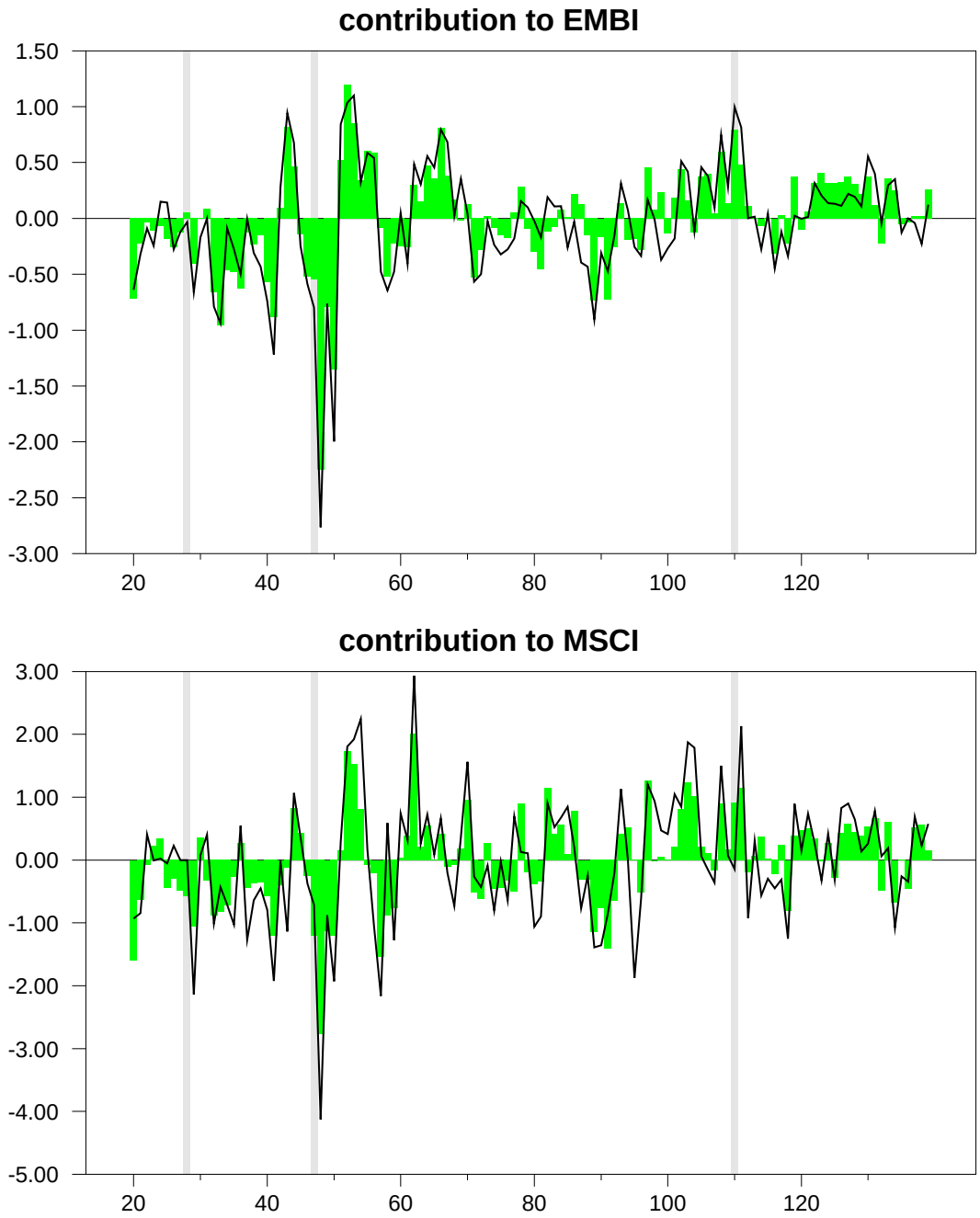

Figure 15: Fraction explained by tapering shock (green) and original series (black)

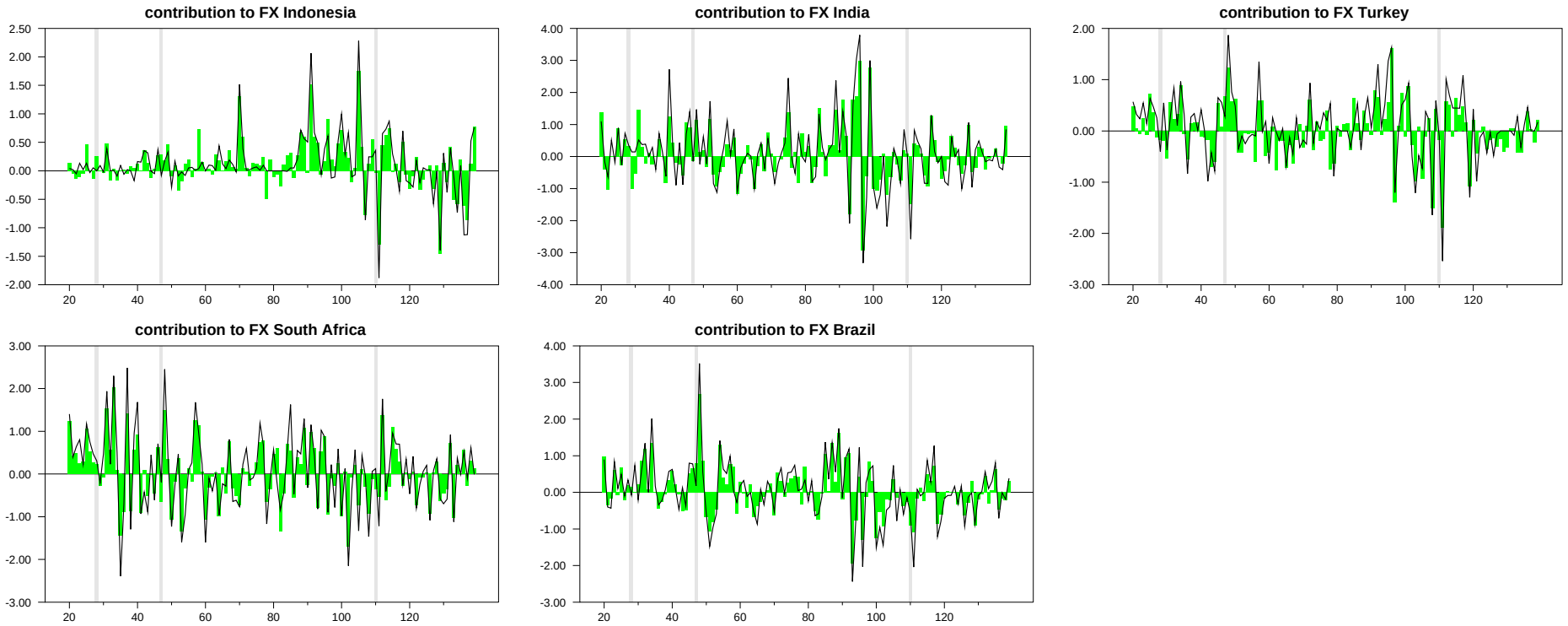

Figure 16: Fraction explained by tapering shock (green) and original series (black)

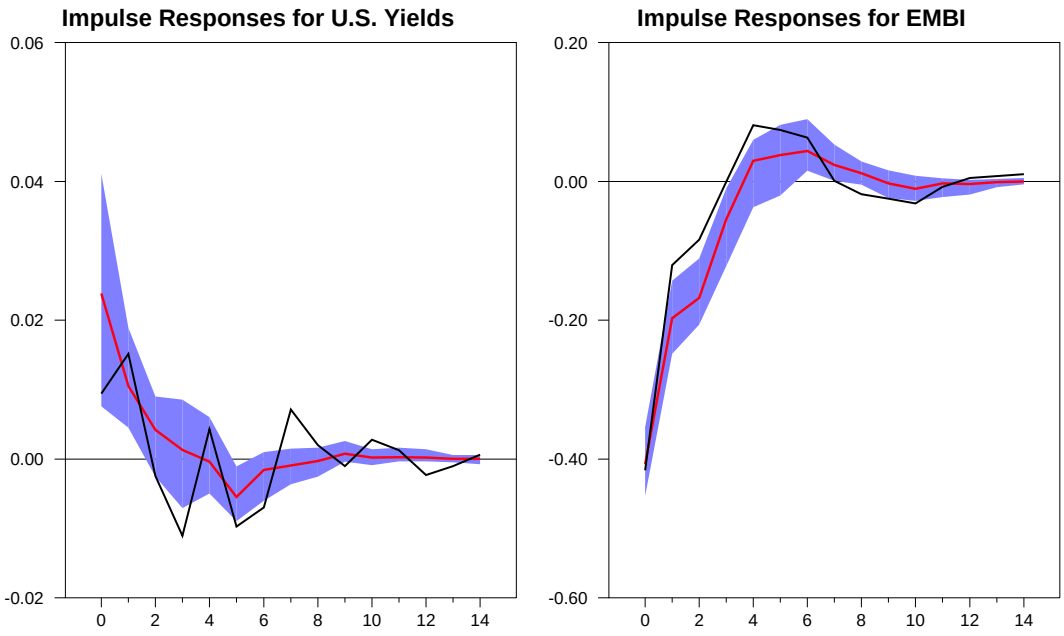

Figure 18: Impulse responses to tapering shock: EMBI+ index, four lags

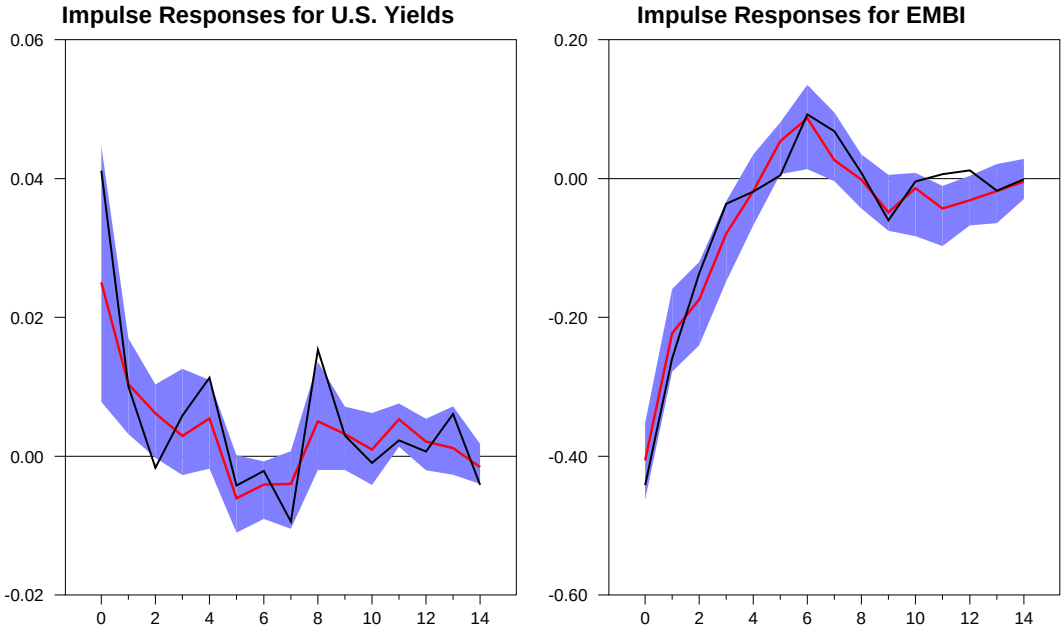

Figure 19: Impulse responses to tapering shock: EMBI+ index, eight lags

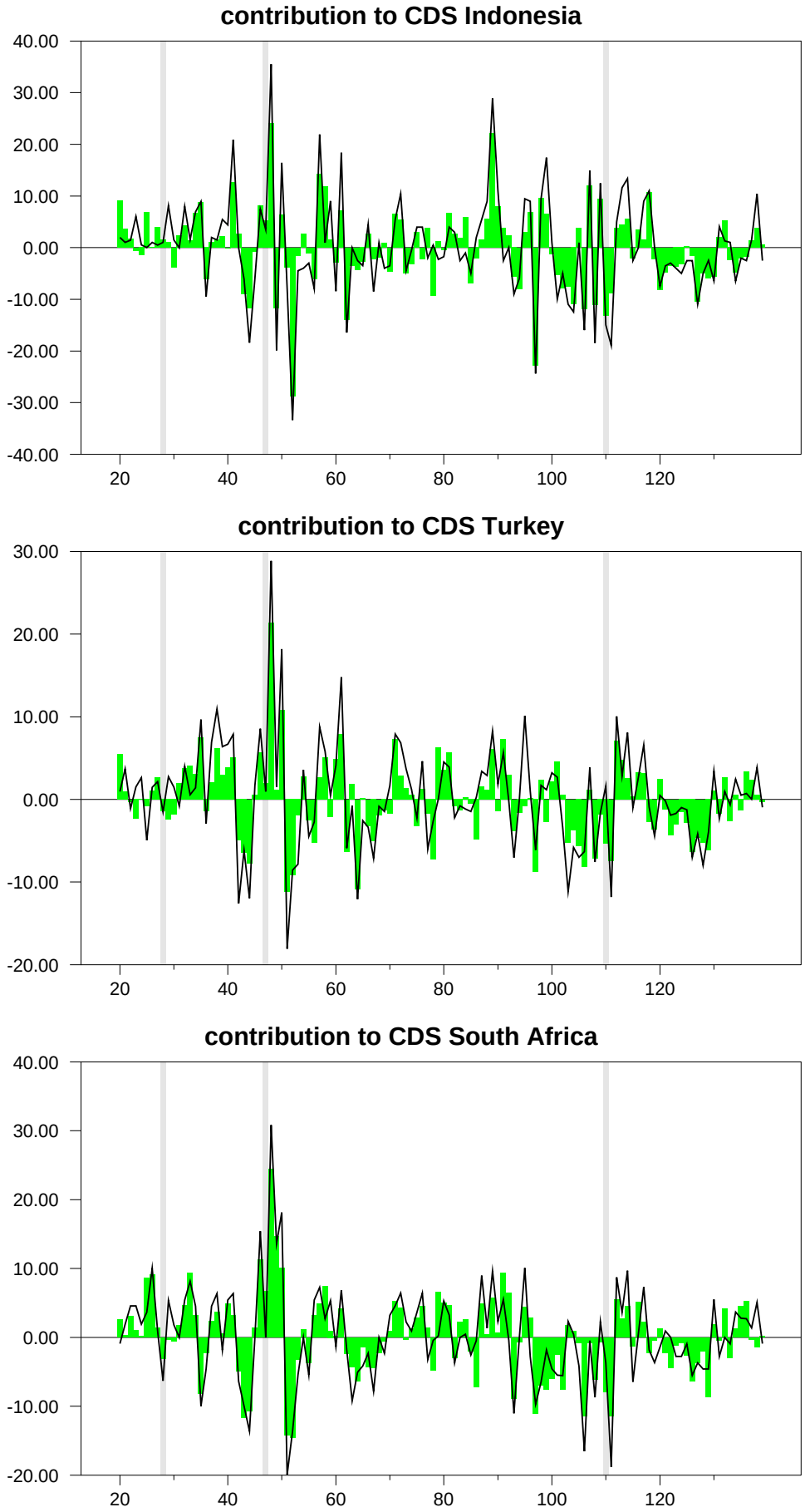

Figure 17: Fraction explained by taperinshock (green) and original series (black)

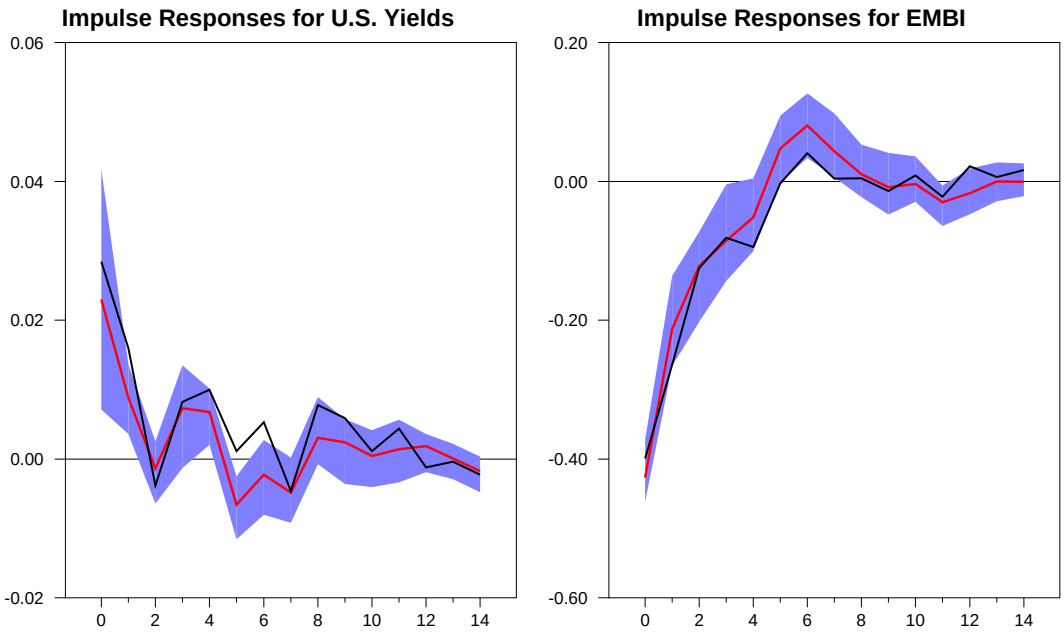

Figure 20: Impulse responses to tapering shock: EMBI+ index, Fed dummies

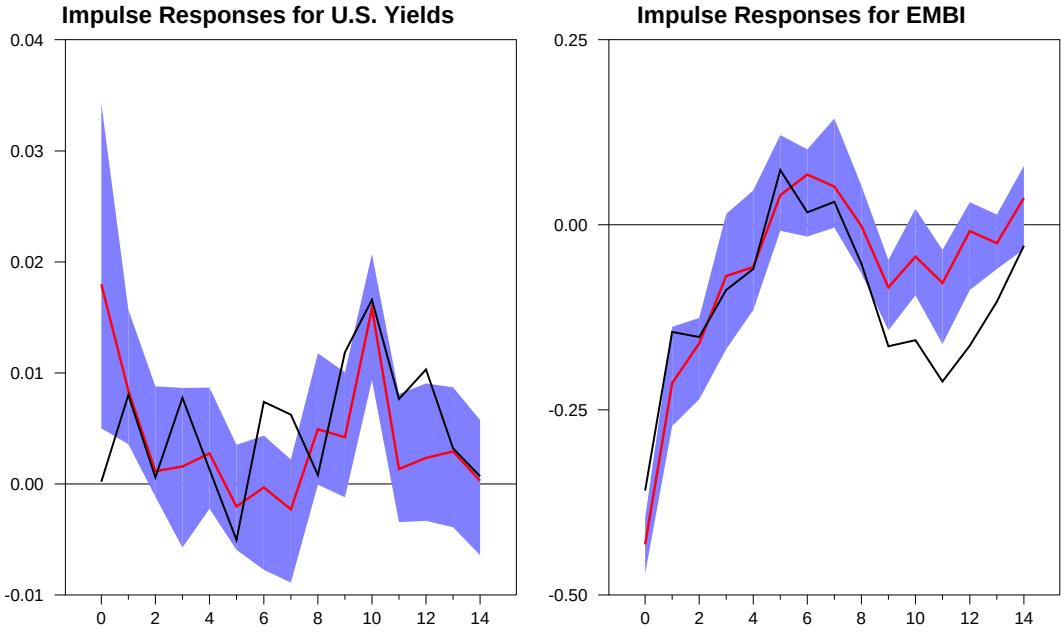

Figure 21: Impulse responses to tapering shock: EMBI+ index, excluding retweets

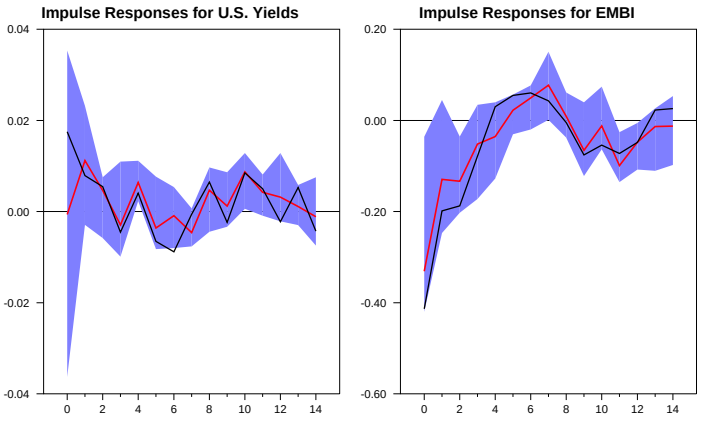

Figure 22: Impulse responses to tapering shock: EMBI+ index, alternative sign restrictions

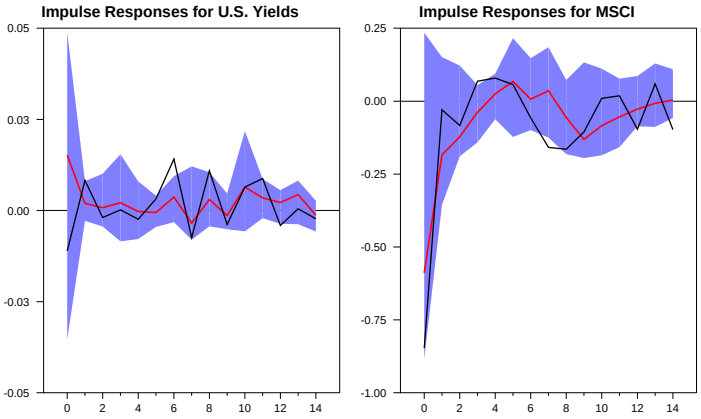

Figure 23: Impulse responses to tapering shock: MSCI index, alternative sign restrictions

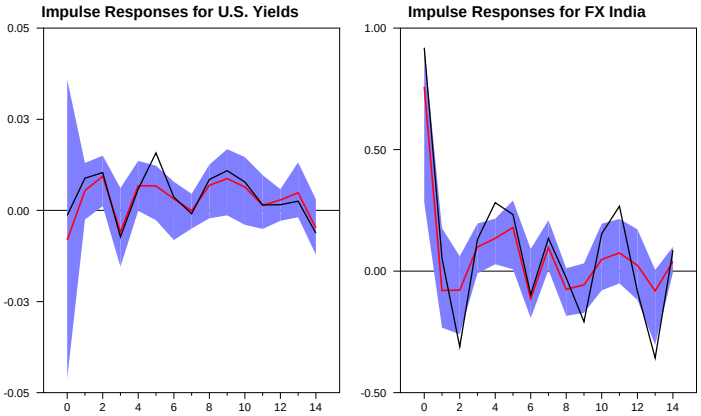

Figure 24: Impulse responses to tapering shock: exchange rate India, alternative sign restrictions

Appendix

Here we describe the steps of construction our series of beliefs. We proceed in three steps:

Step 1

We obtain all tweets sent between April 15 and October 30 2013 that contain the keywords "taper" and "Fed". We prepare our dataset 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 it is required to harmonize the timing. 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 am and 3:39 am are now assigned to the previous day.

Step 2

Based on this raw data the $Tweets^{soon}$ and $Tweets^{late}$series are constructed. First, a dictionary with specific keywords on an early or late tapering, respectively, is set up. Based on that, the tweets are allocated to one of these two categories or to a third category of tweets which contain no view on the timing of the tapering decisions. Tables (3) and (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 valid pre and post 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.

Step 3

In a third step, all tweets that include negations of keywords from both categories are allocated manually.

Table 3: Predefined keywords for category "late"

Late (until September 18, 2013)		**Late (from September 19,2013)**	
2014		2014	
backed away		backed away	
bluff	incl. bluffing	bluff	incl. bluffing
dampen		dampen	
delay	incl. delayed	delay	ink. delayed
December		*March*	
debt ceiling		debt ceiling	
doesn't soon		doesn't soon	
doesn't taper		doesn't taper	
dove	incl. dovish	dove	incl. dovish
Dudley		Dudley	
ease fears		ease fears	
end of the year		end of the year	
in 3rd		*in 1st*	
increase		increase	
isn't happening		isn't happening	
isn't soon		isn't soon	
later in 2013		later in 2013	
less	incl. less likely	less	incl. less likely
no exit		no exit	
no taper		no taper	
not enough		not enough	
not exit QE		not exit QE	
not fast		not fast	
not so fast		not so fast	
not soon		not soon	
not yet		not yet	
November		*February*	
October		*January*	
shutdown		shutdown	
six months		six months	
third		*first*	
this year		not this year	
too soon		too soon	
until		until	
weak	incl. weakness	weak	incl. weakness
will not		will not	
will not taper off		will not taper off	
will take		will take	

Table 4: Predefined keywords for category "soon"

Sooner (until September 18, 2013)		**Sooner (from September 19, 2013)**	
		2013	
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
		end of the year	
expects to taper		expects to taper	
exit qe		exit qe	
fall		fall	
faster		faster	
		fourth	
		4th	
fuel		fuel	
ready		ready	
fell		fell	
Fisher		Fisher	
good news		good news	
hawk	incl. hawkish	hawk	incl. hawkish
in next		in next	
increasing expectations		increasing expectations	
June		June	
July		July	
August		August	
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 few	
next meeting		newt meeting	
		November	
now taper		now taper	
ought to taper		ought to taper	
Plosser		Plosser	
pressure		pressure	
quicker		quicker	
reduce		reduce	
refine	incl. refining	refine	incl. refining
rumour		rumour	
septaper		septaper	
September		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		summer	
talk ongoing		talk ongoing	
taper hint		taper hint	
taper sooner		taper sooner	
taper talk		taper talk	
this summer		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	