## Working Papers / Documents de travail

# Quantitative Easing and the Term Premium as a Monetary Policy Instrument

Etienne Vaccaro-Grange

# Quantitative Easing and the Term Premium as a Monetary Policy Instrument

Etienne Vaccaro-Grange*

Aix-Marseille Univ., CNRS, EHESS, Ecole Centrale, AMSE, Marseille, France

November, 2019

## Abstract

The transmission of Quantitative Easing to aggregate macroeconomic variables through the yield curve is disentangled in two yield channels: the term premium channel, captured by a term premium series, and the signalling channel, that corresponds to the interest rate expectations counterpart. Both yield components are extracted from a term structure model and plugged into a Structural VAR with Euro Area macroeconomic variables in which shocks are identified using sign restrictions. With this set-up, I show how the central bank can use the term premium as a single monetary policy instrument to foster output and prices. However, I also show that there has been a cost channel in the transmission of QE to inflation between 2015 and 2017. This cost channel provides a new explanation as to why inflation has been so muted during this period, despite the easing monetary environment. Finally, a policy rule for the term premium is estimated.

***JEL classification:*** C32, E43, E44, E52
***Keywords:*** Quantitative Easing, Shadow-Rate Term Structure Model, BVAR, sign restrictions

# 1    Introduction

It was the German Economist Richard A. Werner who first talked about Quantitative Easing (QE) in a famous article published in 1995. First applied in Japan to fight strong deflationary pressures, the idea has since flourished and has been implemented in the U.S., U.K., Sweden, and Switzerland in the aftermath of the Great Financial Crisis. With its key policy rates reaching the Effective Lower Bound (ELB) in January 2015,[1] the ECB also decided to take the plunge with the introduction of a Public Sector Purchase Programme (PSPP).[2]

A considerable amount of literature has emerged in the past ten years to try to understand the different transmission channels of QE. Among them, Krishnamurthy & Vissing-Jorgensen (2011) provides a clear theoretical characterization. This paper contributes to the literature by capturing the unexplored term premium channel of QE to the economy, and by showing how the bond term premium can be used as a monetary policy instrument. In particular, I explain how the counterpart signalling channel of QE can be neutralized and why it can be beneficial. A focus is made on the yield curve channels of the PSPP programme, as this constitutes the primary transmission mechanism to interest rates.

The signalling channel of QE was initially brought to light by Eggertson & Woodford (2003) and Bernanke, Reinhart & Sack (2004), later emphasized by Bauer & Rudebusch (2016). It acts on the interest rate expectations component of long-term yields and relies on an implicit commitment by the central banker not to raise policy rates for a long period after starting QE. Because the central bank would incur heavy financial losses on the purchased bonds if interest rates go up, the *costly* measure has only been implemented at the ELB until now: that is, when there was no more room for conventional monetary policy. As opposed, the term premium channel of QE propagates through the second component of long-term yields, in rejection of the *Pure Expectation hypothesis*. Interest rate expectations can only explain part of long-term yields, the rest arising from time-varying and maturity-specific term premia. The term premium channel of QE works through the *portfolio balance mechanism*, assuming that investors have certain preferred-habitat demands and different appetence for risk (Vayanos & Vila (2009), Hamilton & Wu (2010)). Consequently, large central bank interventions on the secondary market force investors to re-balance the composition of their portfolio towards shorter maturity sovereign bonds (duration risk removal effect) or maturity equivalent corporate bonds. This leads to global drops in interest rates as asset prices surge.

A source of debate among economists is the relative importance of the two yield curve channels of QE on inflation and real GDP- a debate which has never been properly assessed. On one hand, VAR studies are unable to capture the term premium of bonds and have therefore used workarounds, such as the use of the term spread (Kapetanios, et al. (2012), Baumeister & Benati (2013)), or interest rate futures as a proxy for expectations (Weale & Wieladek (2016)). That is inconsistent with a clear identification of the two yield curve channels of QE since both embed a term premium and an expectation component. On the other hand, New-Keynesian models

---

[1] At the time of the QE announcement on January $22^{nd}$, 2015, the deposit facility rate was at -0.2%, the main refinancing operation rate at 0.05% and the marginal lending facility rate at 0.3%.

[2] In addition to credit-easing measures (lending operations such as (T)LTROs), the ECB implemented an asset purchase programme named Security Markets Program (SMP) (May 2010 - June 2014). Aimed to "address the malfunctioning of securities markets and restore an appropriate monetary policy transmission mechanism" according to the ECB's press release, the programme was coupled with sterilizing operations to ensure a constant central bank's balance sheet. For this reason, the programme was not Quantitative Monetary Easing strictly speaking. Still, the ECB has also been buying Covered bonds (since July 2009) and Asset-Backed Securities (since September 2014). Genuine QE was announced in January 2015 and launched in March 2015 with the PSPP and lasted until December 2018. It originally consisted in a monthly net purchase on the secondary markets of around €60 billion of public securities with targeted residual maturity of 1 to 30 years. 90% of which had been issued by Euro Area central governments and 10% by supranational organizations and multilateral developments banks located in the Euro Area. The amount was later raised to €80 billion and later downsized to €30 billion and €15 billion. A Corporate Sector Purchase Programme (CSPP) targeting corporate bonds had also been introduced in June 2016. See Claeys, Leandro & Mandra (2015) for a detailed presentation of the APP programme.

fail to build sizeable endogenous bond premia, due to their linear framework. In no case in the macroeconomic literature, have authors been able to capture bond premia originating from interest rate risk and thus isolating the pure term premium transmission channel of QE. Quite the contrary, the macro-finance literature has shined in building term structure models able to design term premia since the introduction of affine models (Duffie & Kan (1996)). They, however, usually do not link the term premium to the macroeconomy. In this paper, I reconcile the two strands of the literature by combining the use of a cutting-edge term structure model with an advanced VAR setting to analyze the term premium transmission channel of QE to aggregate macroeconomic variables. I therefore identify the *pure* term premium and signalling channels of QE to output and inflation.

In particular, I first extract the 10-year Euro Area nominal term premium and terminal rate from a shadow rate term structure model based on Wu & Xia (2016), with a factor rotation as in Lemke & Vladu (2017). To make sure that the path of the expected short-term rate matches the perception of market participants, I anchor interest rate expectations with data from surveys, following Geiger & Schupp (2018). I then plug the two yield components into a monthly Bayesian Structural VAR with the price level, real GDP and a series of market-consistent PSPP purchase that I construct using the Euro Area Monetary Policy Database of Altavilla et al. (2019). Next, I identify a pure QE signalling and QE term premium shock thanks to zero and sign restrictions, using the algorithm of Arias, Rubio-Ramirez & Waggoner (2018). In addition, to sharpen the identification of the QE term premium shock I employ Narrative sign restrictions, based on Antolín-Díaz & Rubio-Ramírez (2018) and developed beyond the original framework. I label them *developed* Narrative sign restrictions. To the best of my knowledge, this paper is the first to combine zero restrictions, traditional sign restrictions and Narrative sign restrictions, thus providing the sharpest possible identification scheme while beating a technical challenge.

The results are twofold. First, I find that the term premium channel of QE had a positive impact on real GDP with a lag of three quarters, while the signalling channel played little role. I also show how the central bank can use the term premium as a single monetary policy instrument, by orthogonalizing the signalling and term premium channels of QE. Second, I show that both yield channels of QE contributed negatively to consumer prices in 2015-2017 and positively afterwards. Using a DCC-GARCH model and time-varying coefficient regressions, I find evidence that this result can be interpreted as a temporary cost-channel in the transmission mechanism of QE to consumer prices: the reduction in firms' cost of borrowing generated by a decrease in the term premium, pushed both PPI and CPI inflation down. That is, assuming that firms have to borrow in advance to finance capital and labor, a drop in the marginal cost of production encouraged producers to lower prices so as to gain in market share (Barth & Ramey (2002)). This cost channel of working capital provides a new explanation as to why inflation has been so muted during this period, despite the easing monetary environment.

Finally, I estimate a policy rule for the term premium to serve as a governing rule for QE. Especially, I measure how the term premium has responded to deviation of inflation from its target and output from its natural level. I find that during the QE period, the term premium has responded more than one to one to output gap and a little less than one to one to the inflation gap, with statistically significant coefficients. This implies that the central bank has actively used the term premium to stabilize the macroeconomic aggregates.

This paper relates to financial studies on the impact of QE on asset prices in the Euro Area: Altavilla, Carboni & Motto (2015), Andrade et al. (2016), and Georgiadis & Gräb (2016) for event studies, Blattner & Joyce (2016), Koijen et al. (2016), De Santis (2016), and Dedola et al. (2017) for time-series analysis, and Eser et al. (2019) for a yield curve modeling approach. In this paper I analyze the impact of all PSPP-related announcements on the term premium, terminal rate and 10-year Euro Area yield at a daily frequency. As in the literature, I find that QE has globally flattened the yield curve, mostly via a compression of the term premium. This study

is also linked to empirical research on the impact of QE on macroeconomic aggregates: Garcia Pascual & Wieladek (2016), Gambetti & Musso (2017) for a time-series approach, and Andrade et al. (2016), Mouabbi & Sahuc (2018), and Darracq Pariès & Papadopoulou (2019) for the theory. Similarly, I find that the overall macroeconomic impact of QE was positive, with the exception of the highlighted temporary cost-channel. Finally, this paper is related to Ireland (2015), that builds an affine term structure model to analyze the impact of a risk shock on macroeconomic aggregates in the U.S. However, if his risk variable originates from bond risk premia, no link is made with QE in his study. I am the first to isolate the pure term premium channel of QE to the economy, to the best of my knowledge.

The rest of the paper is organized as follows: **Section 2** introduces the term structure models used to extract a term premium and terminal rate, and **Section 3** analyzes the yield curve channels of QE thanks to the Bayesian SVAR. Then, **Section 4** sheds light on the cost-channel of monetary policy and **Section 5** estimates a policy rule for the term premium. **Section 6** concludes.

# 2 Extracting a term premium

In the Eurozone, QE has been implemented at the effective lower bound, when there was no more room for conventional monetary policy. In January 2015, the Deposit Facility Rate (DFR) was indeed at -0.20% and there was little scope left for any further rate cut.[3] Consequently, any term structure model employed to recover the yield curve should account for the non-symmetric probability distribution in the pricing of the short-term rate near the ELB. To do so, I introduce an arbitrage-free discrete-time Shadow Rate Term Structure Model (SRTSM) in the spirit of Black (1995), in which the actual short rate is the maximum of the *shadow rate* and a lower bound. As in no-arbitrage Gaussian Affine Term Structure Models (GATSMs) that serve as a benchmark in the term structure literature, I will be able to dissociate the term premium from the expected future short-term rate. However, as multifactor SRTSMs are non-linear and computationally cumbersome, I employ Wu & Xia (2016) linear approximation, that renders the model more tractable while providing an excellent approximation to the data.

## 2.1 Gaussian Affine Term Structure framework

GATSMs find roots in Duffie & Kan (1996), but the model presented here is similar to Ang & Piazzesi (2003) and Duffee (2012). In this framework, the *one*-period nominal interest rate $r_t$ is an affine function of a state vector $X_t$:

$$r_t = \delta_0 + \delta_1' X_t \tag{1}$$

where $\delta_0 \in \mathbb{R}$ and $\delta_1 \in \mathbb{R}^N$

The transition equation that governs the state vector under the physical probability measure $\mathbb{P}$ and the risk-neutral probability measure $\mathbb{Q}$ is:[4]

$$X_t = \mu^{\mathbb{P}} + \Theta^{\mathbb{P}} X_{t-1} + \Sigma_\varepsilon \varepsilon_t^{\mathbb{P}} \tag{2}$$

$$X_t = \mu^{\mathbb{Q}} + \Theta^{\mathbb{Q}} X_{t-1} + \Sigma_\varepsilon \varepsilon_t^{\mathbb{Q}} \tag{3}$$

where $X_t \in \mathbb{R}^N$, $\mu^{\mathbb{P},\mathbb{Q}} \in \mathbb{R}^N$, $\Theta^{\mathbb{P},\mathbb{Q}} \in \mathcal{M}_{N \times N}(\mathbb{R})$, $\Sigma_\varepsilon \in \mathcal{M}_{N \times N}(\mathbb{R})$ and $\varepsilon_t^{\mathbb{P},\mathbb{Q}} \overset{iid}{\sim} \mathcal{MVN}(0, I_N)$

Besides, parameters linking the risk-neutral and physical dynamics are related as follows:

$$\mu^{\mathbb{Q}} = \mu^{\mathbb{P}} - \lambda_0 \tag{4}$$

---

[3]Still, the DFR was lowered to -0.30% in December 2015, -0.40% in March 2016 and -0.50% in September 2019.
[4]See Geiger (2011), Section 3.5.2 for more details.

$$\Theta^{\mathbb{Q}} = \Theta^{\mathbb{P}} - \lambda_1 \tag{5}$$

where $\lambda_0 \in \mathbb{R}^N$ and $\lambda_1 \in \mathcal{M}_{N \times N}(\mathbb{R})$

Furthermore, assuming an essentially affine form of the *log* of the nominal stochastic discount factor (Duffee (2012)), leads the time-varying nominal market price of risk $\Lambda_t$ to be an affine function of the state vector:

$$\Lambda_t = \Sigma_\varepsilon^{-1}(\lambda_0 + \lambda_1 X_t) \tag{6}$$

where $\Lambda_t \in \mathbb{R}^N$

Then, zero-coupon bond prices are exponential affine functions of the state vector:

$$P_{n,t} = e^{A_n + B_n' X_t} \tag{7}$$

So are nominal zero-coupon bond yields:

$$y_{n,t} = -\frac{1}{n}(A_n + B_n' X_t) \tag{8}$$

where coefficients $A_n \in \mathbb{R}$ and $B_n \in \mathbb{R}^N$ are such that:

$$A_n = A_{n-1} + B_{n-1}'\mu^{\mathbb{Q}} - \delta_0 + \frac{1}{2}B_{n-1}'\Sigma_\varepsilon \Sigma_\varepsilon' B_{n-1} \tag{9}$$

$$B_n' = B_{n-1}'\Theta^{\mathbb{Q}} - \delta_1' \tag{10}$$

Complete derivation is presented in **Appendix A.1**

## 2.2 Extension to the Shadow Rate Term Structure Model

In GATSMs, yields can freely become negative around the ELB. This is inconsistent with economic theory,[5] as well as with the historical distribution of observed yields. Therefore, I now introduce a shadow rate model to prevent fitted yields from becoming *too* negative. As in Black (1995), I define the short-term nominal interest rate $r_t$ as the maximum between the nominal *shadow rate* $s_t$ and a lower bound $\underline{r}_t$:

$$r_t = max(s_t, \underline{r}_t) \tag{11}$$

Hence, the *shadow rate* $s_t$ contains more information than the actual nominal short-term rate $r_t$ when the former is below the lower bound $\underline{r}_t$, while the latter remains economically meaningful.

To extend the GATSM of **Section 2.1** to a SRTSM, let us replace the nominal short-term interest rate $r_t$ of Equation (1) by the *shadow rate* $s_t$. Then, we are interested in finding the form of the $n$-period nominal zero-coupon bond $y_{n,t}$ as a function of $X_t$. However, Equation (11) renders the bond pricing formula non-linear and an exact analytical solution of the model is not available beyond one factor in the state vector. Consequently, I employ Wu & Xia (2016) discrete-time approximation.

In their framework, the form of the $n$-period nominal zero-coupon bond yield $y_{n,t}$ under the risk-neutral probability measure $\mathbb{Q}$ is:

$$y_{n,t} = \frac{1}{n}\sum_{j=0}^{n-1} f_{j,j+1,t}$$

---

[5]Agents can always hold cash at a zero interest rate.

5

where

$$f_{n,n+1,t} \approx \underline{r}_t + \sigma_{s,n}^{\mathbb{Q}} \times g\left(\frac{a_n + b_n'X_t - \underline{r}_t}{\sigma_{s,n}^{\mathbb{Q}}}\right) \tag{12}$$

is the forward rate for a loan starting at date $t+n$ and maturing at date $t+n+1$,

with $g(x) = x\Phi(x) + \phi(x)$ where $\Phi(x)$ is the CDF of $x$ and $\phi(x)$ the PDF. In addition,
$a_n = \delta_0 + \delta_1'\left(\sum_{j=0}^{n-1}\Theta^{\mathbb{Q}j}\right)\mu^{\mathbb{Q}} - \frac{1}{2}\delta_1'\left(\sum_{j=0}^{n-1}\Theta^{\mathbb{Q}j}\right)\Sigma_\varepsilon\Sigma_\varepsilon'\left(\sum_{j=0}^{n-1}\Theta^{\mathbb{Q}j}\right)'\delta_1 \in \mathbb{R}$
$b_n' = \delta_1'\Theta^{\mathbb{Q}n} \in \mathbb{R}^N$
$\sigma_{s,n}^{\mathbb{Q}} = [\sum_{q=0}^{n-1}\delta_1'(\Theta^{\mathbb{Q}})^q\Sigma_\varepsilon\Sigma_\varepsilon'(\Theta^{\mathbb{Q}})^{q'}\delta_1]^{1/2} \in \mathbb{R}$

Complete derivation is presented in **Appendix A.2**

## 2.3  Transformation of the factors

To ease economic interpretation, I conduct an affine transformation of the generic latent factors $X_t$, following Lemke & Vladu (2017). Consequently, the new latent factors $\mathcal{P}_t$ will resemble principal components and I will be able to label them *level*, *slope*, and *curvature* as in Nelson-Siegel language. The same rotation is employed in the GATSM and the SRTSM. Note that this leaves model-implied yields, term premia and the *shadow rate* unaffected.

First, define a full rank loading matrix $W \in \mathcal{M}_{N \times J}(\mathbb{R})$ with $N$ the dimension of the state vector $X_t$ from Equation (2) and $J$ the number of selected yields. $W$ contains the weights that enable to obtain linear combinations of the observed yields $Y_t^\circ$ to map the observable principal components $\mathcal{P}_t^\circ$, such that:

$$\mathcal{P}_t^\circ = WY_t^\circ \tag{13}$$

Then, the affine transformation can be re-written:

$$\mathcal{P}_t = P_0 + P_1 X_t \tag{14}$$

where $P_0 \in \mathbb{R}^N$ and $P_1 \in \mathcal{M}_{N \times N}(\mathbb{R})$

Both under the physical $\mathbb{P}$ and risk-neutral measures $\mathbb{Q}$, the rotated latent factors $\mathcal{P}_t$ follows a VAR(1) process as in Equations (2) and (3):

$$\mathcal{P}_t = \mu_{\mathcal{P}}^{\mathbb{P}} + \Theta_{\mathcal{P}}^{\mathbb{P}}\mathcal{P}_{t-1} + \Sigma_{\varepsilon,\mathcal{P}}\varepsilon_t^{\mathbb{P}} \tag{15}$$

$$\mathcal{P}_t = \mu_{\mathcal{P}}^{\mathbb{Q}} + \Theta_{\mathcal{P}}^{\mathbb{Q}}\mathcal{P}_{t-1} + \Sigma_{\varepsilon,\mathcal{P}}\varepsilon_t^{\mathbb{Q}} \tag{16}$$

where $\mu_{\mathcal{P}}^{\mathbb{P},\mathbb{Q}} \in \mathbb{R}^N$, $\Theta_{\mathcal{P}}^{\mathbb{P},\mathbb{Q}} \in \mathcal{M}_{N \times N}(\mathbb{R})$ and $\Sigma_{\varepsilon,\mathcal{P}} \in \mathcal{M}_{N \times N}(\mathbb{R})$

Complete derivation is presented in **Appendix A.3**

## 2.4  Parametrization

For the parametrization of the GATSM and the SRTSM, I largely follow Lemke & Vladu (2017) once again, who calibrate a SRTSM for the Euro Area. Their normalization is based on Joslin, Singleton & Zhu (2011), which leads to a uniquely identified set of parameters.

In this setting, the state vector $X_t$ from Equation (2) contains three factors. It is indeed acknowledged in the literature that three factors are sufficient to account for most cross-sectional

variations of yields. Moreover, $\mu^{\mathbb{Q}} = (\mu_1^{\mathbb{Q}}, 0, 0)'$ and $\Theta^{\mathbb{Q}}$ from Equation (3) is in ordered real Jordan form, while the variance-covariance matrix $\Sigma_\varepsilon$ is supposed to be lower triangular. In addition, parameters from Equation (1) are such that $\delta_0$ is the vector null and $\delta_1$ a unit vector. On the other hand, parameters $\mu^{\mathbb{P}}$ and $\Theta^{\mathbb{P}}$ from the transition Equation (2) that governs the state vector under the physical probability measure are left unrestricted. The market price of risk parameters $\lambda_0$ and $\lambda_1$ from Equation (6) are also left unrestricted. Note that imposing a restriction on these parameters will impact estimates of term premia (Bauer (2017)). Indeed, setting $\lambda_0$ and $\lambda_1$ to zero leads the risk-neutral and physical measures to coincide. This parametrization falls in the case of the *strong form* of the *Expectation hypothesis*, i.e. investors require no term premium (or no compensation for interest rate risk) to hold long-term bonds. In the so-called Vasicek (1977) model, $\lambda_0 \neq 0$ and $\lambda_1 = 0_N$, so that the term premium is non trivial but time invariant and the model falls in the *weak form* of the *Expectation hypothesis*. On the other hand, my non-trivial parametrization leads to a non-zero time-varying market price of risk and term premium, thus departs from the *Expectation hypothesis*.

In this framework, the measurement equation giving the relation between the model-implied zero-coupon bond yields $Y_t$ and the observed zero-coupon bond yields $Y_t^\circ$ is:

$$Y_t^\circ = Y_t + \Sigma_\varsigma \varsigma_t \tag{17}$$

where $Y_t^\circ = (y_{1,t}^\circ, ..., y_{n,t}^\circ,)'$, $Y_t = (y_{1,t}, ..., y_{n,t})'$ and $\varsigma_t = (\varsigma_{1,t}, ..., \varsigma_{n,t})'$ with $\varsigma_t \sim \mathcal{MVN}(0, I_J)$ the vector of observation errors[6] and $\Sigma_\varsigma \in \mathcal{M}_{J \times J}(\mathbb{R})$ with $\Sigma_\varsigma = diag(\sigma_\varsigma, ..., \sigma_\varsigma)$.

To consistently recover the yield curve, I focus on $J = 12$ yields with maturity three- and six-month, one-, two-, three-, five-, seven-, ten-, fifteen-, twenty-, twenty-five-, and thirty-year, which corresponds to $n \in \{1, 2, 4, 8, 12, 20, 28, 40, 60, 80, 100, 120\}$ as my *one*-period (short-term) interest rate has a maturity of three months. In this study, it is important to price long-term yields as QE focuses on long maturity bonds. Additionally, data employed are monthly average of the AAA Euro Area implied yield curve from September 2004 to December 2018.[7]

Besides, a crucial choice is the one of the lower bound $\underline{r}_t$. Basically, the lower bound represents the minimum achievable level of interest rates. Passed it, agents would prefer holding cash to preserve their savings and the level of interest rates would be insignificant. Unobserved by definition, the "true" level of the lower bound has recently been a topic of attention. If some define a process for its law of motion (Wu & Xia (2017), Monfort et al. (2017)), most remain in the deterministic framework for tractability and better robustness (Wu & Xia (2016), Kortela (2016) and Lemke & Vladu (2017)). This is also what I do, studying two cases:

<u>Case A</u> (constant deterministic lower bound):

$$\underline{r}_t = \underline{r} = min(0, Y_{1:T}^\circ, r_{t+1}^{surveys}) \approx -0.857 \tag{18}$$

where $Y_t^\circ$ is the vector of observed yields at time $t$ and $r_{t+1}^{surveys}$ is the one-period ahead expected short-term interest rate from surveys, as described below. This equation shows that the ELB is actually negative. Indeed, investors currently agree to pay a little premium to be able to invest in scarce credit-risk free government bonds.[8] The spread between zero and the minimum of the negative yields reflects this premium.

---

[6]Standard in the literature, this means that the variance of observed errors is supposed to be the same across all maturities, which could of course be questioned.

[7]Data are available at: http://www.ecb.europa.eu/stats/money/yc/html/index.en.html

[8]The 3-month German government bond yield has been negative since July 2014.

<u>Case B</u> (time-varying deterministic lower bound):

$$\underline{r}_t = min(\underline{r}_{t-1}, 0, Y_t^{\circ}, r_{t+1}^{surveys}) \tag{19}$$

where $\underline{r}_0 = 0$. Equation (19) implies that the lower bound is a decreasing function that turns negative as soon as any observed yield or survey does.

In this framework, my state-space model is composed of a linear transition Equation (2) and a measurement Equation (17). The latter is linear in the GATSM (**Section 2.1**) and non-linear in the SRTSM (**Section 2.2**). This is because of the nature of $y_{n,t}$, that comes from linear Equation (8) in the GATSM and non-linear Equation (12) in the SRTSM. Hence, I estimate the model by maximum likelihood, based on the Kalman filter for the GATSM (details in **Appendix A.4.1**) and on the extended Kalman filter[9] for the SRTSM (details in **Appendix A.4.2**).

A comparison of the term structure models fit is presented in **Table 7 - Appendix A.5** to save space. Clearly, the SRTSM beats the benchmark Gaussian model in terms of fitting (RMSE is about 3.5 bps lower for case (B)). Besides, SRTSM (B) with the time-varying deterministic lower bound performs better than SRTSM (A) with a constant lower bound, both in terms of log-likelihood and RMSE. The choice of the time-varying lower bound is therefore backed by a better empirical fit. It is the one used to obtain latent factors, term premia and terminal rates estimates described below.

## 2.5 Expression of the term premium

I follow the literature in defining the term premium on an $n$-period nominal zero-coupon bond as the difference between the $n$-period zero-coupon yield and the expected short-term rate under the physical probability measure $\mathbb{P}$:[10]

$$\kappa_{n,t} = y_{n,t} - e_{n,t} \tag{20}$$

where

$$e_{n,t} = \frac{1}{n} \sum_{j=0}^{n-1} E_t^{\mathbb{P}}(r_{t+j})$$

That is, the term premium is the difference in expected return between a buy and hold strategy for a $n$-period zero-coupon bond and an instantaneous rollover strategy at the risk-free rate. It is a reward for interest rate risk (that is, uncertainty about future interest rates). As data employed are AAA Euro Area implied yields, the term premium should here not embed any compensation for credit or liquidity risk. Purposely, this choice is justified by the will to capture the *pure* term premium channel of QE conceived by the preferred-habitat and duration-risk removal theories. Thus, the channel captured here will be clean from any effect stemming from a lower default risk or a better market liquidity.

In addition, the terminal rate is defined as the 10-year ahead expected short-term interest rate:

$$exp_t = E_t^{\mathbb{P}}(r_{t+10Y}) \tag{21}$$

---

[9]Other nonlinear filtering techniques include the unscented Kalman filter and the particle filter. Yet, Christensen & Rudebusch (2015) finds little difference in the results between the extended and the unscented Kalman filters. Plus, the particle filter is computationally cumbersome, while it would provide little benefit (see Krippner (2015), Section 4.2.1 for a discussion).

[10]Other definitions of the term premium can be found in Cochrane & Piazzesi (2008).

This rate represents market participants view at time $t$ of the level of the 3-month interest rate in 10 years. It therefore embeds anticipations driven both by monetary policy and the state of the economy.

To make sure that the path of the expected short-term rate $e_{n,t}$ matches the perception of market participants, I anchor interest rate expectations with data from surveys[11] as in Geiger & Schupp (2018):

$$r_{t+j}^{surveys} = E_t^{\mathbb{P}}(r_{t+j}) + \Sigma_\varsigma \varsigma_t \tag{22}$$

with $\varsigma_t \sim \mathcal{MVN}(0, I_J)$ a vector of observation errors as in Equation (17).

Finally, note that assuming $E_t^{\mathbb{P}}(r_{t+j}) \approx f_{j,j+1,t}^\circ$ with $f_{j,j+1,t}^\circ$ the observed one-period forward rate, implies that $\kappa_{n,t} = 0 \ \forall \ n$.[12] That is, proxying the terminal rate with forwards is inconsistent with having a non-zero term premium. This is simply because forwards are not pure expectations but also embed a term premium. Thus, such approximation would not allow to isolate the term premium channel of QE and one understands the necessity of a term structure model to obtain estimates of a time-varying term premium.

## 2.6 Analysis of the term premium

**Figure 1** presents 10-year term premia from the GATSM and the SRTSM, with and without anchoring short-term interest rate expectations. On the other hand, **Figure 18 - Appendix A.5** shows results regarding the terminal rate, i.e the 10-year ahead expected short-term rate.

The first striking characteristics of term premia and terminal rates on those figures is that GATSM and SRTSM estimates are similar before 2012 but diverge after that date, which corresponds to the start of the low interest rate period in the Eurozone (see **Figure 16 - Appendix A.5**). Indeed, failure to account for the lower bound in pricing the term structure at short maturities can lead to under-estimate the expected path of the short-term rate and thus over-estimate term premia. Consistent with Carriero, Mouabbi & Vangelista (2015) for the UK yield curve, I find that term premia from the GATSM appear to be biased and that a SRTSM is necessary to obtain lower bound-consistent estimates of the yield decomposition.

Moreover, anchoring short-term interest rate expectations seem to matter little, both in the GATSM and SRTSM, with the exception of the 2012-2015 period for the SRTSM. Two reasons for that: first, the unavailability of survey data at longer maturities makes it difficult to fully anchor the path of the expected short-term rate. It therefore differs little from the model-based one. Second, the negative contribution of surveys in the estimates of the terminal rate from the SRTSM between 2012-2015 can be explained by the fact that market participants probably expected further policy rate cuts. Indeed, that cannot be accounted for in the model as the lower bound is purely deterministic and agents are not forward looking. This sheds light on the usefulness of anchoring interest rate expectations with survey data when using a deterministic lower bound.

Other than that, the counter-cyclicality of the term premium and pro-cyclicality of the terminal rate is stringent: the former rises by about 2% during the Great Financial Crisis (GFC), while the latter decreases by the same amount. This reflects two things: firstly, market participants must have expected a policy rate cut during the crisis. Secondly, investors usually do not like to take risk in 'bad' times, so the supply for risk becomes higher and its price goes down (its return goes up) (Bekaert, Engstrom & Ermolov (2016)). It is so because during recessions a higher macroeconomic uncertainty and volatility drives the compensation for risk that investors require to postpone consumption (Cochrane & Piazzesi (2005)). However, this pattern is not observable during the

---

[11] I use surveys of the 3-month Euribor, three and twelve months ahead, that I adjust for the spread with the AAA Euro Area implied yield curve.

[12] The sum of forward premia is equal to the yield premium.

eurozone sovereign debt crisis of 2012-2013 because data employed are AAA rated bonds. That is, my term premium is a credit-risk free term premium and a pure measure of interest rate uncertainty. It is therefore expected that it did not hike like lower grade periphery countries' credit risk premia.[13] In addition, the low level of term premia may have been driven by unconventional monetary policy measures implemented by the ECB in this period (LTROs in particular).

Other than that, the current low level of the term premium since 2014 can be attributed to the QE period that started in January 2015, but might have been expected since Draghi's Jackson Hole speech in August 2014. In contrary, the fact that the term premium was small around 2007 can be attributed to the *conundrum* period, during which yields remained low while expected short-term rates were rising due to the global saving glut in 2004-2006 (Rudebusch, Swanson & Wu (2006)). Nevertheless, the role of a lower inflation risk-premium is not to be ruled-out during this period.

Finally, the increase in the term premium in May-June 2015 corresponds to a short-lived deterioration in market liquidity (Cohen, Hördahl & Xia (2018)), while the increase in November-December 2016 coincides with the lower than expected extension of the APP programme.

Tuning to the terminal rate on **Figure 18 - Appendix A.5**, it has been continuously decreasing since 2012 with the introduction of unconventional monetary policy measures and negative policy rate environment. It is now stuck near the ELB, indicating that market participants do not see the AAA short-rate at a higher level than today in ten years.

Overall, my estimate of the 10-year term premium from the SRTSM with anchored interest rate expectations is largely in line with the ones of Wu & Xia (2017), Lemke & Vladu (2017), and Cohen, Hördahl & Xia (2018).[14]

Figure 1: Term premia
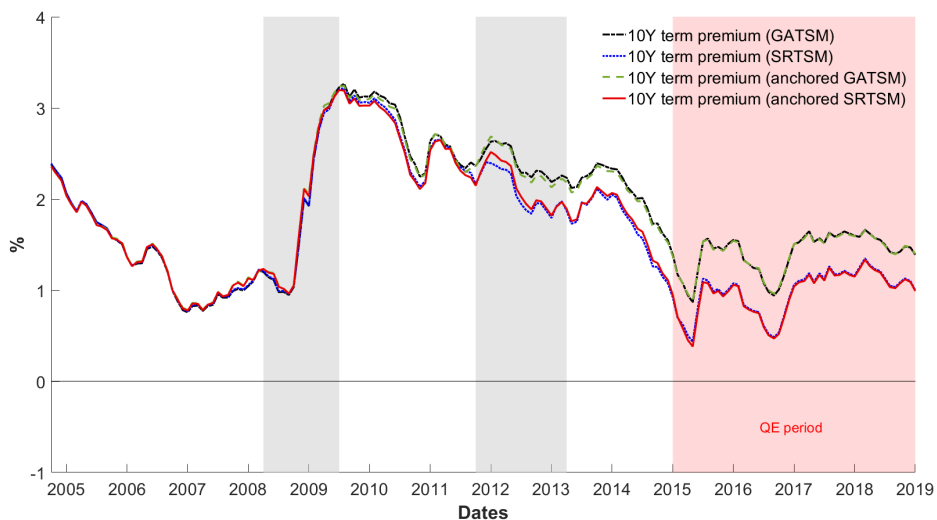


Term premia between 2004:M9 and 2018:M12
Solid grey areas correspond to recession periods as defined by the Center for Economic Policy Research (CEPR)
Solid red area corresponds to the QE period

Besides, **Figure 17 - Appendix A.5** shows the *level*, *slope*, and *curvature* after rotation of the latent factors. It is noteworthy that the yield curve has been upward slopping throughout the

---

[13]Portugal, Spain, Italy and Greece especially.

[14]Dynamics are very similar even though levels slightly differ as their estimates are based on the OIS Swap curve.

sample, as one can judge by the negative sign of its slope.[15] Plus, the level of the yield curve has been decreasing since 2009 and stabilizes around zero at the end of the sample. This is consistent with the unconventional monetary policies pushing long-term interest rate expectations and term premia down during that period.

# 3 The yield curve channels of QE

## 3.1 Discussion

In the context of the term structure literature, the are two yield curve channels of QE: the term premium channel and the signalling channel. The term premium channel of QE works through the so-called *portfolio-balance* mechanism, as market participants re-balance the composition of their portfolio after selling assets to the central bank on the secondary market. This channel would be neutral in a frictionless economy with perfect assets substitutability, as financial market participants would be indifferent as to swap a government bond for another asset with the central bank. In the real world, investors have preferred-habitat demands and different appetence for risk (Vayanos & Vila (2009), Hamilton & Wu (2010)): demand and supply of bonds are maturity and risk-bearing specific. In that framework, a government bond purchase by the central bank creates a shortage of long-term safe bonds on the market, that leads to an increase in these bonds prices and a decrease in the term premium component of the associated yields. The purchase of long-term bonds also removes duration-risk in the hands of investors and thereby reduces the long-end of the sovereign yield curve relative to the short-end. Indeed, investors selling bonds to the central bank adjust the composition of their portfolio, replacing long maturity bonds sold by shorter-ones, or invest in maturity-equivalent corporate bonds (and other riskier assets). The domino effect of portfolio adjustments finally leads to a decrease in interest rates on loans to firms and households, initiating a surge in consumption and investment in the economy. This credit channel is coupled with a wealth effect driven by financial market profits as asset prices surge and the monetary base expands.

On the other hand, the signalling channel of QE works through market expectations of the future policy rate, via the so called *Expectation hypothesis*. That is, market participants build their anticipations about the path of the short-term interest rate over a long period, thus driving the expectation component of long-term rates (Eggertson & Woodford (2003)). Such forward looking anticipations are induced from current and future monetary policy as well as the state of the economy.

From the central banker point of view, it has become very relevant to try to shape these interest rate expectations, so as to reduce long-term yields in the binding ELB environment for the policy rates. Undeniably, it is also a cheap way to help anchor inflation expectations and encourage investment. Indeed, in the standard New-Keynesian framework of Woodford (2003), the linearized IS curve iterated forward only involves expectations of future real short-term interest rate and inflation. Lowering the path of the expected real short-term rate therefore increases output and inflation through the IS and Phillips curve. In practice, lowering long-term yields through the signalling channel also reduces benchmark rates used to price corporate loans. This boosts credit, therefore corporate investment (as well as household consumption to a certain extent).

Thus, to shape interest rate expectations a central banker has to options: either simply credibly communicating about his intentions, or conditioning the implementation of a *costly* measure to a last resort. That is, when policy rates are seen as stuck at the ELB for a long time and when there is no more room for conventional monetary policy. The former is called forward guidance,

---

[15]The slope is defined as the difference between the short-term minus the long-term rate here.

while the latter is simply the signalling channel of QE. Especially, this signalling channel relies on an implicit commitment by the central banker not to raise policy rates for a long period, to the extent that the central bank would incur heavy financial losses on the large amount of bonds it has bought, would interest rates go up (and prices drop). In this view, the central bank may want to keep the policy rate lower than what the Taylor rule would recommend, for its own interest. Therefore, the higher the perceived coefficient related to financial losses in the central bank's objective function, the more powerful the signalling channel. However, I argue that the central bank could easily freeze this channel to orthogonalize QE with the expected short-term rate component of long-term yields. I also argue that it may be desirable for efficiency purposes.

Indeed, the central bank could freeze the signalling channel of QE simply by not conditioning its implementation to policy rates being at the ELB. For example, the central bank could clearly state that financial losses due to interest rates hike are fully acceptable, manageable and that it is ready to conduct QE whatever the level of the policy rates. Hence, market participants would not infer from the introduction of a QE programme that it is a way to artificially lower short-term interest rates more that it is possible. Furthermore, the central bank could even allow itself to conduct Quantitative Tightening (QT) in order to reduce its balance sheet exposure to interest rate risk. That is, the same way the central bank purchased a large amount of assets through QE, it could sell those off through a large program of Quantitative (monetary) Tightening. Thus, market participants would not infer that the central banker is implicitly committing not to raise policy rates for a long time and QE would not impact the expectation component of long-term yields anymore. The two yield curve channels of QE would therefore be orthogonal.

Above all, freezing the signalling channel of QE so that the policy measure would only be transmitted through the term premium channel would be beneficial to the central bank for three reasons. First, it would create a new instrument at disposal for the central bank, by granting it the possibility to control the term premium. Outside of the context of QE, it has actually been shown that there are welfare gains to having the central bank respond to the term premium, for example by including it in an augmented Taylor rule (Carlstrom, Fuerst & Paustian (2017)).[16]  Second, isolating the term premium channel of QE would strengthen the impact of the monetary policy measure. Indeed, the relative importance of the transmission channels of QE is unclear and so is the sum of their impacts on yields. It is not certain that the two channels do not counteract each other. In fact, I will show in the next section that there is at least a situation in which relying on the two channels provides diminished effects, as changes in interest rate expectations phagocyte the impact of the term premium channel on prices and real GDP. Besides, with the signalling channel frozen, the central bank could directly target a certain level of the term premium and adjust the purchases accordingly, instead of choosing an amount of purchase in euro and guess the impact it will have on yields. Last, it would also ease the management of interest rate expectations, by strengthening the effect of forward guidance relative to indirect market inference from a QE programme.

Still, why would shaping long-term yields by acting on the term premium be a desirable monetary policy? As mentioned earlier, textbook models clearly show that the output gap only depends on interest rate expectations and not the term premium. The straight answer is that these linearized toy models are unfortunately not able to feature any term premium and remain in the rigid framework of the *Pure Expectation hypothesis*. If no structural relationship between bond term premia and real GDP has been found so far (Rudebusch & Swanson (2012)), Economists tend to agree that lower term premia is expansionary for economic activity (Bernanke (2006)). Indeed,

---

[16]As this would require a rich model, I deliberately do not dig that dimension further and leave that for future research.

long-term sovereign yields (and so term premia by extension) serve as benchmark to price most corporate and household loans, and therefore have a direct impact on lending and investment. As a consequence, not only expectations of the policy rate matter, (which partly explains why forward guidance has much smaller effects in reality than in these theoretical models). Hence, there is undeniably a need for models deviating from the *Pure Expectation hypothesis* and that is why the literature has increased its focus on this topic after GFC. It is now common to have models featuring both a short-term and a long-term bond, even though it is difficult to build a New-Keynesian DSGE model featuring a sizeable and time-varying endogenous term premium. Even in models approximated at the third-order as in Rudebusch & Swanson (2012),[17] term premia produced are too low in level and volatility compared to empirical findings. Thus, tricks are often used, such as labeling transaction costs or market frictions as such (Chen, Cúrdia & Ferrero (2012), Kiley (2014), Carlstrom, Fuerst & Paustian (2017)). Macro-Finance DSGE models perform much better in terms of fitting the data, but either do not allow the term premium to freely feedback to the macroeconomy or impose restrictions that limit the feedback effect. A more structural type of models featuring a term premium relies on consumption-based asset pricing à la Creal & Wu (2019), in which the term premium is derived from the conditional covariance between the bond and the stochastic discount factor (SDF). In this framework, the SDF is derived from Euler equation and it is equal to the expected ratio of marginal utilities of consumption at two periods. It characterizes future consumption relative to current consumption and the term premium is the compensation investors require to hold an asset that does not deliver wealth when it is the most valuable to the agent (Campbell (2000)). Yet, once again no link is made with interest rate risk and its impact on spending. Thus, a last strand of the literature focuses on reduced-form models that are more flexible. Nevertheless, studies mostly concern the predictive power of the term premium for economic activity and they disagree. On the one hand, Hamilton & Kim (2002), Favero, Kaminska & Söderström (2005), and Wright (2006) find that lower term premia in level predict a slow growth, while Ang, Piazzesi & Wei (2006) and Dewachter, Lania & Lyrio (2014) find no predictive power in the term premium. On the other hand, Rudebusch, Sack & Swanson (2007) and Jardet, Monfort & Pegoraro (2013) find that a decline in the term premium is followed by a faster GDP growth. In this paper, I put an end to the debate by providing clear evidence that a negative term premium shock stimulates real GDP in the context of QE.

Thus, if I argue that lowering the term premium stimulates aggregate demand through QE, I state that it may also be beneficial in some circumstances to raise it through QT. Indeed, let us think of a situation where the bank lending transmission channel of monetary policy is impaired, due to shrunk bank margins in a flat yield curve environment. In addition, suppose the short-end of the curve is at the ELB. Here, it could be desirable for the central bank to repair the credit channel by restoring bank margins through a steepening of the yield curve. However, as short-term rates could not be lowered further, the only way would be to raise medium to long-term yields through QT. Another situation could be the case of a central bank implementing yield curve control with a clear target for the 10-year yield as a monetary policy, such as the Bank of Japan since September 2016. In that spirit, a recent debate has emerged about the benefits for the central bank not only to target the short-term natural rate of interest but also its long-term counterpart or even the whole natural yield curve, in order to fill the output gap (Dufrénot, Rhouzlane & Vaccaro-Grange (2019), Jordà & Taylor (2019)). Then, it may be desirable for the central bank to only adjust long-term yields so as to address a long-term yield gap, while keeping a short-term interest rate

---

[17]Other examples include Rudebusch, Sack & Swanson (2007), Hördahl, Tristani & Vestin (2008), and Van Binsbergen et al. (2012). These models need to be solved at the third-order through the so-called perturbation method and can therefore not be estimated. Plus, term premia produced are far too low in terms of size and volatility, unless one makes controversy assumptions such as a very high curvature for the utility function, a high degree of habit persistence or that stochastic shocks are large and highly persistent too. A second-order approximation would provide a constant non-zero term premium, while the first-order produces no term premium.

gap put at zero.

## 3.2 Bayesian Structural VAR approach

In what follows, I set up a Bayesian Structural VAR including the two yield channels of QE, output, and prices, to analyse the effect of each channel on the macroeconomy. To do so, I orthogonolize the signalling and term premium channels of QE using zero restrictions, as the central bank would do if it dissociates QE from interest rate expectations. In addition, the use of sign restrictions and *developed* Narrative sign restrictions provides a structural identification of the shocks driving the economy. This framework enables me to quantify the impact of QE through Impulse Response Functions, Historical Decompositions and Conditional Forecasts.

Let us first consider the general case of a vector of size $N$, $\Upsilon_t$, containing the endogenous variables of interest and $X_t$, a vector of size $M$ containing some exogenous variables. Its representation in a VAR form is:

$$\Phi(L)\Upsilon_t = DX_t + \epsilon_t \tag{23}$$

where $\Phi(L) = I_N - \Phi_1 L - ... - \Phi_p L^p$ is the matrix of lag polynomials of order $p$, $D \in \mathcal{M}_{N \times M}(\mathbb{R})$ and $\epsilon_t \in \mathbb{R}^N$ with $\epsilon_t \sim \mathcal{MVN}(0, \Sigma_\epsilon)$ is the vector of residuals. Beside, let us assume that all the roots of $det[\Phi(L)]$ lie outside the unit disk, so that $\Phi(L)$ is invertible and $\Upsilon_t$ is stationary.

Structural representation of Equation (23) is:

$$B^{-1}\Upsilon_t = \sum_{k=1}^{p} Z_k \Upsilon_{t-k} + DX_t + w_t \tag{24}$$

where $B \in \mathcal{M}_{N \times N}(\mathbb{R})$ and $Z_k \in \mathcal{M}_{N \times N}(\mathbb{R})$ are the matrices of structural parameters, and $w_t \in \mathbb{R}^N$ is the vector of structural shocks with $w_t \sim \mathcal{MVN}(0, I_N)$.

In the baseline specification of this paper, $\Upsilon_t$ contains the *log*-real GDP $y_t$, the consumer price level $p_t$, a market-consistent measure of the announced amount of bonds to be purchased through the APP programme $b_t$, the 10-year term premium $\kappa_t$, and the (10-year) terminal rate $exp_t = E_t^{\mathbb{P}}(r_{t+10Y})$ extracted from the SRTSM (Equation (20) and (21)):

$$\Upsilon_t = (y_t, p_t, b_t, \kappa_t, exp_t)' \tag{25}$$

In addition, $X_t$ is a vector of *ones*, so that the VAR only includes a constant.

An obvious criticism of constant parameter SVARs is that they do not allow for a change in the underlying structural relationships, that may arise through time. Yet, time-varying parameter SVARs identified with sign restrictions have recently been criticized for their inconsistency with Bayesian inference (Bognanni (2018)). Plus, no algorithm has been developed for the use of Narrative sign restrictions in time-varying parameter SVARs so far. For these reasons, I stick to the use of a constant parameter SVAR and restrict the analysis to the post GFC period, to limit the possibility of a structural change in the parameters.

## 3.3 Identification strategy

To identify the structural shocks $w^i$, I employ a combination of zero restrictions and traditional sign restrictions on the impulse response functions as well as *developed* Narrative sign restrictions on the structural shocks and on the historical decompositions. In particular, I identify a QE term premium shock, a QE signalling shock, a Risk shock, a Supply shock, and a Demand shock. Even

though I am not directly interested in the Risk, Supply, and Demand shocks, it is acknowledged that adding plausible restrictions to identify other shocks helps pin down the ones of interest (Canova & Paustian (2011), Kilian & Murphy (2012)).

The first identification scheme I study is the following:

Table 1: Identification scheme I

|  | QE premium | QE signalling | Risk | Supply | Demand |
|---|---|---|---|---|---|
| $\kappa$ | - with N | 0 | - | $\star$ | - |
| $exp$ | 0 | - | $\star$ | $\star$ | $\star$ |
| $b$ | + | + | 0 | $\star$ | $\star$ |
| $p$ | 0 | 0 | $\star$ | - | + |
| y | 0 | 0 | $\star$ | + | + |

$\star$ Means no restriction is imposed
$N$ stands for (developed) Narrative sign restrictions

In this setting, a QE term premium shock is a shock that lowers the term premium in the context of an announcement of a purchase of bonds by the central bank. That shock leaves the terminal rate unchanged to capture the fact that this pure QE term premium shock should not affect interest rate expectations if the central bank has orthogonalized the two yield curve channels of QE. In addition, the impact of the QE term premium shock on the macro variables is set to zero on impact, to capture both its qualitative and quantitative impact on the economy. The usual reasonable assumption that macroeconomic variables can only respond with a lag to a monetary policy shock is made. Furthermore, a pure QE signalling shock is a shock that decreases the terminal rate in the context of an announcement of a purchase of bonds by the central bank. It should also not impact the term premium and macroeconomic variables are only allowed to respond with a lag. Besides, a (negative) Risk shock is a shock that moves the term premium (i.e, the reward for bearing risk) down, where that movement is not due to a QE programme. The identification of this shock is crucial to separate movements of the term premium that are due to QE and those due to market movements (flight-to-safety, risk, etc.). Then, a standard Supply shock is identified and its impact on the policy variables is left unrestricted. Indeed, the reaction of the central bank to that type of shock depends on its loss function regarding the real GDP-inflation trade-off. Finally, a Demand shock is expansionary for real GDP, prices and it also leads to a reduction in the term premium as risk shrinks[18] (Hördahl, Remonola & Valente (2015)).[19] Details of the identification with traditional sign restrictions and zero restrictions are provided in **Appendix B.1.1** and **B.1.2**.

However, from this set of restrictions one could wonder whether strict identification is achieved between the Risk shock and the two QE shocks. The answer is yes: the probability to "draw a zero" when the restriction is positive or negative is exactly zero. Indeed, over an infinite number of possibilities within the set of real numbers $\mathbb{R}$, there is no chance to draw a specific value. In this setting, shocks are therefore set identified. Nevertheless, I couple the traditional sign restrictions with exclusive Narrative sign restrictions to sharpen the identification of the QE term premium shock. The Narrative sign restrictions approach was developed by Antolín-Díaz & Rubio-Ramírez

---

[18]Note that remaining agnostic on the response of the term premium only shortens the positive response of real GDP.

[19]The authors show using events-studies that a positive announcement on the state of the economy leads to a reduction of the risk premium.

(2018) and has since become very popular. This cutting-edge identification strategy helps pin down structural shocks using prior Narrative information. In this paper, I develop the Narrative sign restrictions approach beyond the framework given by the authors and I label them *developed* Narrative sign restrictions. Then, I use these type of restriction to identify the QE term premium shock by setting restrictions at key announcement dates of the PSPP by the ECB. In particular, I focus on major events and study their impact on the term premium using daily data (see **Figure 2**).

I impose a restriction on the sign of the QE term premium shock and on the amplitude of its contribution to the Historical Decomposition (HD) of the term premium at key dates. In Antolín-Díaz & Rubio-Ramírez language, these are respectively referred to as **Class I** and **Class II** restrictions. Still, I innovate by also adding restrictions on the relative amplitude of the QE term premium shock with respect to other shocks and with respect to itself during the rest of the sample. I also impose a restriction on the sign of the contribution of the QE term premium shock to the HD of the term premiumn. Useful to sharpen the identification, I name these types of restrictions *developed* Narrative sign restrictions. A Summary of the restrictions employed is detailed in **Table 3**. Mathematical expressions of the *developed* Narrative sign restrictions are available in **Appendix B.1.3**. For a treatment of the traditional sign restrictions approach on the IRFs, see for example Chapter 13 of Kilian & Lütkepohl (2016).

Let us now focus on the key dates related to the PSPP presented on **Table 2**, for which imposing *developed* Narrative sign restrictions may be relevant. Note that I ignore announcements related to the CSPP programme, that could have impacted the sovereign yield curve via spillover effects. This decision is justified by the fact that I want to focus on the direct effects of the PSPP only. In order to find out which of these announcements related to the QE programme had a substantial impact on the term premium of government bonds, I plot the latter on **Figure 2** at a daily frequency during the months of the announcements. Similar plots for the terminal rate and the 10-year yield are presented on **Figure 19** and **Figure 20 - Appendix A.5**.

Table 2: PSPP-related events

| Dates | Event |
|---|---|
| 22/08/2014 | Draghi Jackson Hole Speech |
| 22/01/2015 | Announcement of the Expanded APP |
| 09/03/2015 | Start and release of implementation details of APP |
| 03/09/2015 | Increase in PSPP issue share limit |
| 10/03/2016 | PSPP amount raised and CSPP announced |
| 08/12/2016 | PSPP recalibration |
| 19/01/2017 | Extension of the PSPP below the DFR |
| 26/10/2017 | PSPP recalibration |
| 14/06/2018 | PSPP recalibration |

APP stands for Asset Purchase Programme
PSPP stands for Public Sector Purchase Programme
CSPP stands for Corporate Sector Purchase Programme
DFR stands for Deposit Facility rate
Source : ECB Press Release

Figure 2: Daily 10Y term premium during months of key announcements
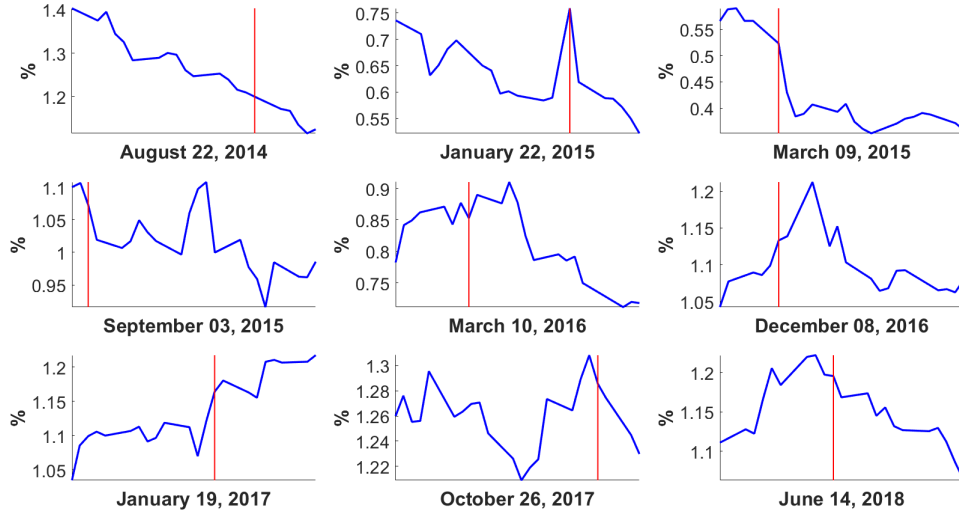


10Y term premium in the month of the major QE-related announcements
Red vertical bar corresponds to announcement

Strikingly, January $22^{nd}$ and March $9^{th}$, 2015 were major events that moved the term premium significantly. On the former date, the APP programme was introduced and on the latter it was implemented (with a release of important details). Besides, March $10^{th}$, 2016 also seems to be a relevant date that contributed to an upward movement of the term premium. On that date, market went reverse as participants were disappointed by the policy rate cut coupled with the PSPP recalibration (monthly APP purchases increased to €80bn and introduction of the CSPP).[20] As a consequence, the term premium went up, which could be assimilated to a negative QE shock. However, it is not clear that this event was the main market mover of the month. So it is not for the other dates studied. Especially, some researchers (for example Middeldorp (2015) and De Santis (2016)), believe that QE was very much anticipated by market participants so much so that the actual announcement date could be placed in August 2014, following President Draghi's Jackson Hole speech (Garcia Pascual & Wieladek (2016)). However, once again it is here not clear that this event was an obvious market mover for the term premium, even though it may have contributed to its global downward trend around that period. Therefore, I will only use January $22^{nd}$ and March $9^{th}$, 2015 as dates to impose *developed* Narrative sign restrictions. Unfortunately for my identifying scheme, other news came out during the months related to the QE announcements. For example, January $22^{nd}$, 2015 also corresponds to an announcement related to the LTRO programme. One could then wonder whether I identify a pure QE shock and not a more global unconventional monetary policy shock. However, the fact that the QE shock is conditioned on an announced amount of bonds to be purchased to go up thanks to a traditional sign restriction, guarantees the identification of a pure QE term premium shock. Overall, the choice of January $22^{nd}$ and March $9^{th}$, 2015 is in line with Andrade et al. (2016) and Dedola et al. (2017),[21] that find a statistically significant impact on asset prices at these dates.

---

[20]http://www.theguardian.com/business/live/2016/mar/10/ecb-stimulus-measures-mario-draghi-negative-rates-qe-business-live

[21]See Appendix B, table 1 in particular.

Eventually, the *developed* Narrative sign restrictions imposed are the following:

Table 3: Developed Narrative sign restrictions

| Dates | Restrictions |
|---|---|
| 22/01/2015 | QE shock is positive, its amplitude is higher than other shocks at this date and the highest QE shock of the sample. Its contribution to the HD of the term premium is negative and higher than the contribution of the other shocks. |
| 09/03/2015 | QE shock is positive and its contribution to the HD of the term premium is negative. |

## 3.4   Data and estimation

I estimate the SVAR model with 12 lags and a constant on monthly data ranging from January 2009 to December 2018. The idea is to focus on the post-GFC period, which also corresponds to the introduction of unconventional monetary policies. Economic activity is the *log*-real GDP[22] interpolated at a monthly frequency using a *Denton* approach and the OECD Composite leading indicator (CLI) as an index.[23] Price index is the Harmonized Index of Consumer Prices (HICP).[24] Finally, I build a market-consistent series of amount of bonds announced to be purchased monthly through the APP programme on **Figure 3**. That is, the official amount of bonds purchased is adjusted to be placed at the date of announcement (versus implementation) and the size is corrected with respect to market reactions to the press release to capture surprises. Especially, I take a look at the impact of the press releases[25] on the 10-year German yield (Bund) using the database from Altavilla et al. (2019) and adjust the size of the APP to be purchased accordingly.[26] Thus, I ensure that dates such as March $9^{th}$, 2015 actually correspond to an increase in the series as markets reacted positively to the announcement (the Bund went down), even though the size of the programme did not changed in reality. On the opposite, market participants were disappointed by the amplitude of the extension of the APP programme on December $8^{th}$, 2016. It therefore now corresponds to a decrease in the perceived APP purchase.

---

[22]Data are available at: http://fred.stlouisfed.org/series/CLVMEURSCAB1GQEA19
[23]Data are available at: http://data.oecd.org/leadind/composite-leading-indicator-cli.htm
[24]Data are available at: http://sdw.ecb.europa.eu
[25]Press realeses can be found here http://www.ecb.europa.eu/press/pr/date/2019/html/index.en.html
[26]The adjusted size is negatively proportional to the change in the 10Y German yield.

Figure 3: APP purchase
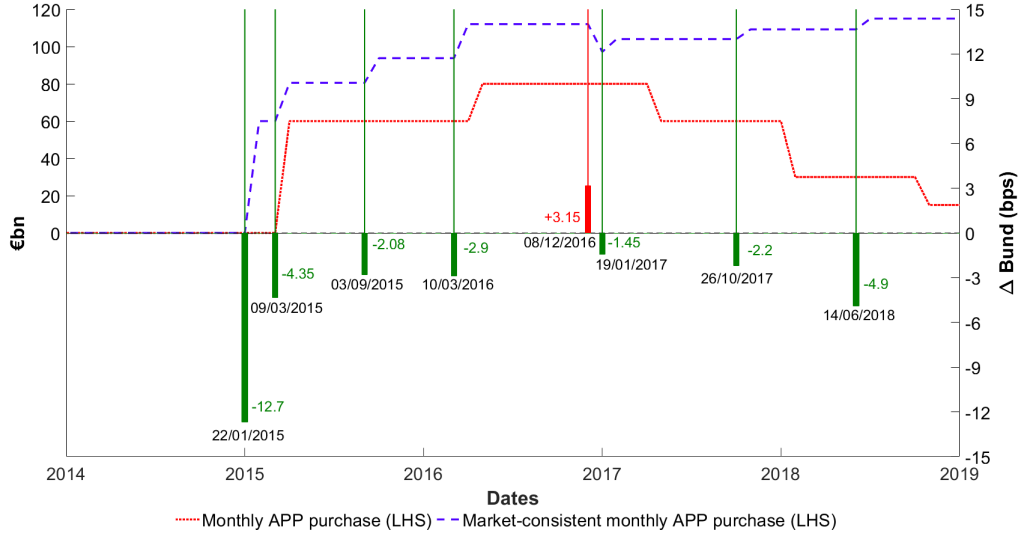


Vertical bars corresponds to the announcements
Numbers correspond to the change in bps in the Bund 45 min around the press conference

Moreover, the SVAR is estimated using Bayesian methods with standard natural conjugate (Normal-Wishart) priors following the most efficient algorithm currently available: the one of Arias, Rubio-Ramirez & Waggoner (2018). In this algorithm, coefficients are drawn from a conjugate uniform-normal-inverse-Wishart posterior distribution over the orthogonal reduced-form parametrization and then transformed into a structural parametrization. This transformation induces a normal-generalized-normal posterior distribution over the structural parameterization. It requires the use of an important sampler to re-weight the successful draws in the presence of zero restrictions. Similarly, the use of Narrative sign restriction truncates the support of the likelihood function and therefore requires a re-weighting of the likelihood. To the best of my knowledge, this paper is the first to combine zero restrictions, traditional sign restrictions and (*developed*) Narrative sign restrictions, thus providing the sharpest possible identification scheme while beating a technical challenge. Details of the Bayesian estimation are provided in **Appendix B.2**.

## 3.5 The term premium as a monetary policy instrument

To analyze the impact of the structural shocks on the endogenous variables following identification scheme I, I compute the Impulse Response Functions (IRFs) on **Figure 4**, using the fact that:

$$\Upsilon_{t+h} = \sum_{q=0}^{\infty} C_q L^q \epsilon_{t+h}^i = \sum_{q=0}^{\infty} C_q \epsilon_{t+h-q}^i = \sum_{q=0}^{\infty} C_q b_i w_{t+h-q}^i \tag{26}$$

where $\Upsilon_t = \sum_{q=0}^{\infty} C_q L^q \epsilon_t = [\Phi(L)]^{-1}\epsilon_t$ is the Wold decomposition, $b_i$ is the $i^{th}$ column of $B$, and $w_t$ is the vector of structural shocks from Equation (24).

Figure 4: Point-wise median Impulse Response Functions - Identification Scheme I
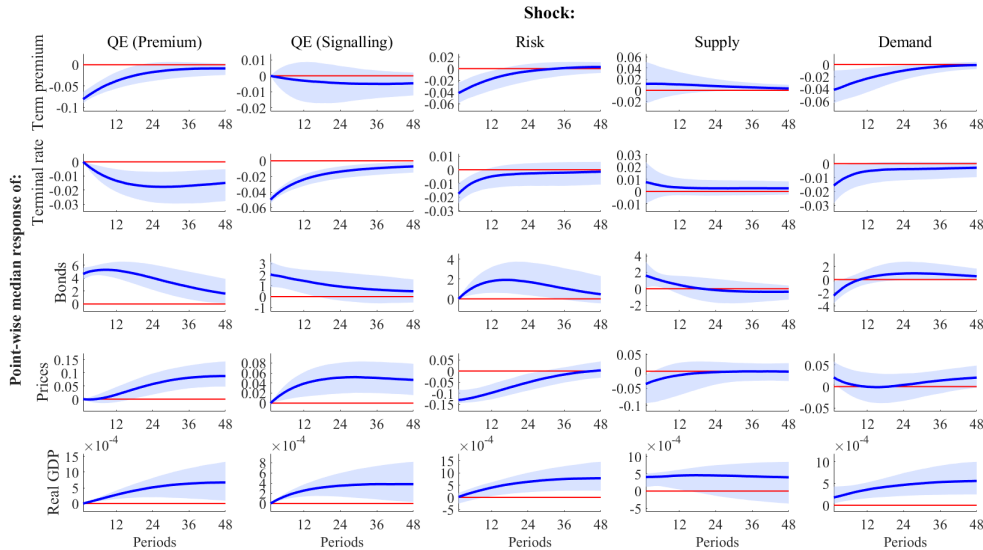


In particular, shocks are calibrated to be of one standard deviation (structural shocks are presented on **Figure 23 - Appendix B.4**). I present the point-wise median and the 68% confidence interval response of each variables to all shocks. Nonetheless, I acknowledge the drawbacks of basing an analysis on such a reductive statistics as highlighted by Kilian & Inoue (2013). Therefore, I also compute the Fry & Pagan (2011) median target response and the confidence bands quantile targets on **Figure 24 - Appendix B.4**. Overall, this robustness check shows that there is very little difference in the statistics response displayed.[27] Especially, a decrease in the 10-year term premium of about 0.075%, coupled with an announcement of a € 5 billion of monthly purchase, gradually raises prices to a peak of about 0.1 and increases *log*-real GDP by about $7.5 \times 10^{-4}$ after three years. For comparison, this means that assuming linearity, a 1% decrease in the term premium, requiring the announcement of a € 66 billion monthly purchase, would lead the price level to jump by about 1.3 point and real GDP by about 13.3 point. Besides, it is worth noting that the increase in macro variables is quite gradual, potentially reflecting a certain high degree of price rigidity and a slow transmission mechanism through the credit channel. Interestingly, that improvement in the economy is combined with a slight decrease in the terminal rate, indicating that there might remain a trace of the signalling channel present in the data despite the orthogonalization of the shocks.

Besides, a QE signalling shock that decreases the terminal rate by 0.05%, raises prices slightly less and real GDP slightly more than the QE term premium shock. Equivalently, a 1% decrease in the terminal rate, coupled with an announcement of a € 40 billion monthly purchase, leads to a hike of 1 in the price level and about 20 in real GDP. There is however no clear impact on the term premium, as expected by the orthogonalization.

Otherwise, a negative Risk shock decreases both the term premium and the terminal rate as confidence increases. More interestingly, it leads to an adverse effect on prices (that go down) and real GDP (that go up). I conjecture that the extra confidence mostly concerns firms rather that household consumers. That is, a decrease in the long-term yield components not driven by monetary policy is a gain in confidence by firms that decide to increase investment. On the contrary, households' confidence is not so much impacted by a long-term rate drop and don't change their

---

[27]Responses are based on 10,000 successful draws of which 6,117 are unique after re-weighting.

consumption habit. As a consequence, demand remains stable but production increases: real GDP goes up and prices go down.

Moreover, Demand and Supply shocks have a much more transitory impact on inflation: the effect dies out after a year. On the other hand, most of the increase in real GDP arises on impact and seems to be permanent. Both responses are of the same order of magnitude as the one standard deviation QE shocks, confirming the credibility of the previous results. However, this does not mean that Macro and QE shocks contributed to output and prices in the same proportion, as it will be shown on the historical decompositions. Finally, the Supply shock does not significantly move the term premium and terminal rate, while the Demand shock decreases both. Indeed, this improvement in the economy is accompanied with an increase in confidence and investment that pushes long-term rates down.

Finally, it is worth mentioning that the results are broadly robust to changes in the prior tightness, the number of lags, the inclusion or exclusion of the exogenous variables and the removal of Narrative sign restrictions.

I now study the case in which the central bank has not orthogonalized the term premium and signalling channels of QE in **Table 4**. Thus, the contemporaneous and endogenous response of market participants to each of the QE shocks is taken into account. In this setting, a QE term premium shock now leads market participants to revise their expectations about the terminal rate upward, inferring a future hawkish move from the central bank. Plus, a QE signalling shock now lowers the term premium, as risk shrinks due to a better expected economic outlook.

Table 4: Identification scheme II

|  | QE premium | QE signalling | Risk | Supply | Demand |
|---|---|---|---|---|---|
| $\kappa$ | - | - | - | $\star$ | - |
| $exp$ | + | - | $\star$ | $\star$ | $\star$ |
| $b$ | + | + | **0** | $\star$ | $\star$ |
| $p$ | + | + | $\star$ | - | + |
| $y$ | + | + | $\star$ | + | + |

$\star$ Means no restriction is imposed
$N$ stands for (developed) Narrative sign restrictions

Figure 5: Point-wise median Impulse Response Functions - Identification Scheme II
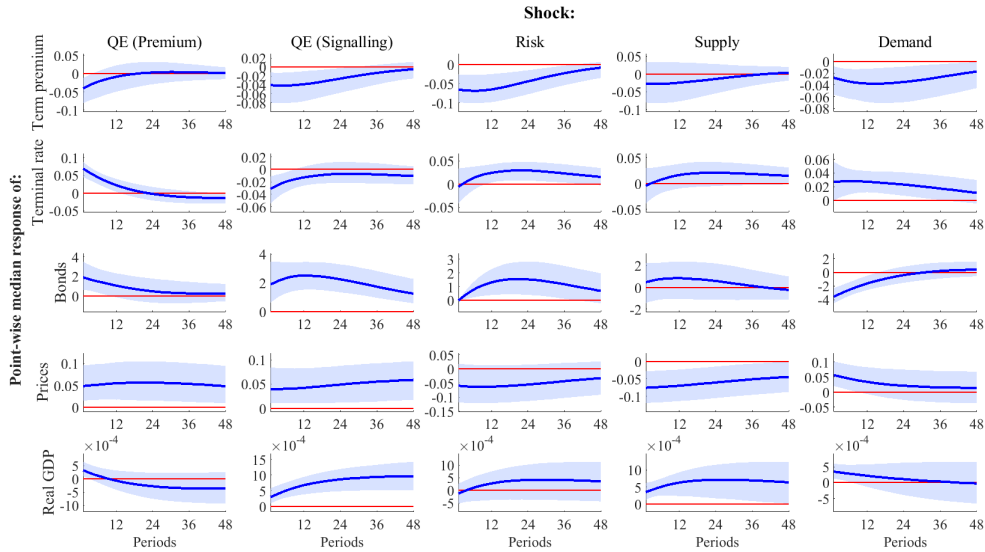


Identification scheme II shows that the QE term premium shock still fosters prices and real GDP, but the impact on real activity is now much more transitory. The positive effect dies out after a year as a result of higher interest rate expectations. Moreover, the peak impact on prices averages 0.05 points, whereas it reached 0.1 in the first identification scheme. On the other hand, the impact of the signalling channel is stronger in this second identification scheme, as the decrease in the terminal rate is accompanied by a reduction in the term premium. This raises real GDP more ($1x10^{-3}$ versus $4x10^{-4}$), while the response of prices remains broadly similar.

Overall, a first overview shows that these results confirm previous studies about the positive impact of QE on the Euro Area economy (Bundesbank Report (2016), Garcia Pascual & Wieladek (2016), Gambetti & Musso (2017), Mouabbi & Sahuc (2018), and ECB Economic Bulletin (2019)). Yet, the most important takeaway is that orthogonalizing the yield curve channels of QE is meaningful and can have a positive impact on the amplitude of the response of the macroeconomic variables. Indeed, Identification Scheme II shows that the signalling channel can phagocyte the effects of the term premium channel if those are not fully orthogonalized. In addition, the IRFs of prices and real GDP clearly show that the term premium can be used as a monetary policy instrument to foster the economy, to the extent that the central bank is capable to control it. Still, as soon as the monetary institution credibly states that there is no link to be inferred on the level of the policy rates from QE, the signalling channel will freeze and QE will enable the central bank to control the term premium uniquely.

Additionally, if controlling the term premium arises in the context of QE, that is of non-sterilized asset purchase, it might very well be the case that sterilized asset purchase such as the Operation Twist in the U.S or even the (never implemented) Outright Monetary Transaction (OMT) or Securities Market Program (SMP) in the Euro Area could be sufficient to give control of the term premium. Whether the institution should keep the size of its balance sheet constant or not following market interventions is up to other considerations such as the rate of growth of money, the level of inflation, and the liquidity of financial markets. Unfortunately, the framework of this paper does not allow to give a stance on the effects of non-sterilized asset purchases on the economy.

While the previous IRFs enable to observe the hypothetical propagation of shocks, historical decompositions provide a historical perspective to the analysis, by showing the contribution of all shocks in driving the variables over the sample period. Thus, I produce the median target historical decomposition of real GDP and prices on **Figure 6** and **7**. HDs of the term premium and terminal rate are presented on **Figure 25** and **26** in **Appendix B.4**. To do so, I use:

$$H_{j,w^i,t_s} = e_j' \sum_{q=0}^{t_s-1} C_q b_i w_{t_s-q}^i \tag{27}$$

where $H_{j,w^i,t_s}$ is the HD of variable $j$ for a shock on $w^i$ at time $t_s$ and $e_j$ is the $j^{th}$ column of $I_5$

Figure 6: Historical Decomposition of real GDP - Identification Scheme I
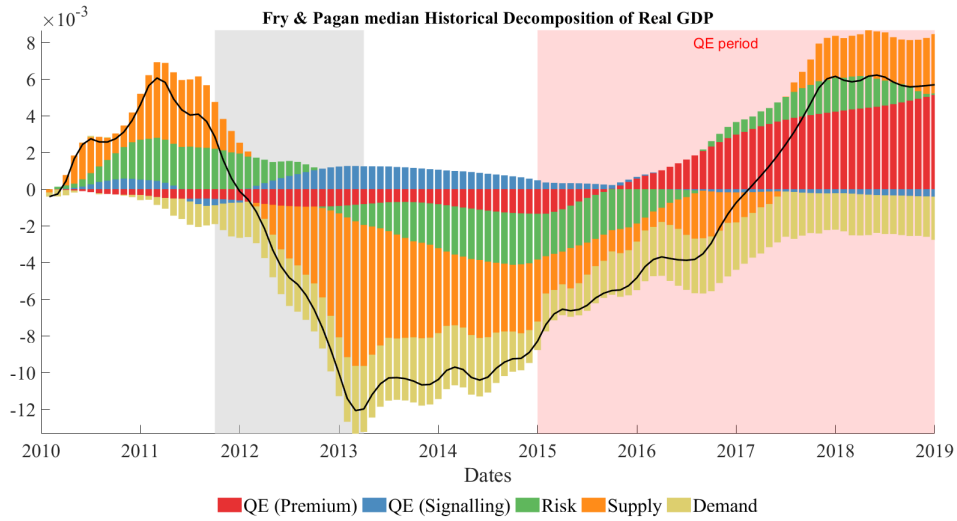


Historical decomposition is expressed in deviation from initial conditions
Solid grey area corresponds to a recession period as defined by the Center for Economic Policy Research (CEPR)
Solid red area corresponds to the QE period

Figure 7: Historical Decomposition of Prices - Identification Scheme I
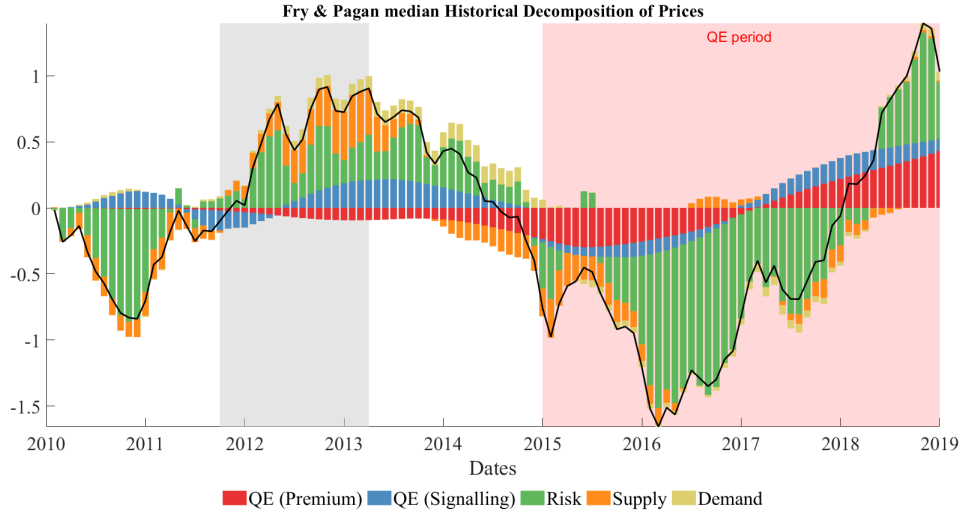


**Fry & Pagan median Historical Decomposition of Prices**

Legend: QE (Premium), QE (Signalling), Risk, Supply, Demand

Historical decomposition is expressed in deviation from initial conditions
Solid grey area corresponds to a recession period as defined by the Center for Economic Policy Research (CEPR)
Solid red area corresponds to the QE period

The first thing that comes to sight when taking a look at the HD of real GDP is the prominence of the role of the Supply and Demand shocks as drivers of real Economic activity. Especially, the addition of the contributions of these two shocks astonishingly dominates between mid 2011 and 2017, whereas the contributions of the QE shocks are comparatively negligible until mid 2016. From there on however, the positive contribution of the QE term premium shock becomes meaningful and gradually grows until the end of the sample. That is, about three quarters after the announcement of the QE programme and two quarters after its implementation, real GDP starts being positively driven by the reduction in the term premium. On the other hand, the signalling channel played no role at all after 2016. Nevertheless, during 2012-2015 real GDP has seen a slight positive contribution of the QE signalling channel and a negative contribution of the QE term premium channel. I refer to this as the *whatever it takes* period, in reference to the famous phrase pronounced by Mario Draghi in London on July $26^{th}$, 2012. That is, Draghi's commitment and threat of the OMT programme is somehow captured in my identification scheme and has had a positive effect on real activity via a reduction in interest rate expectations. As a contrary, deep turmoil on periphery sovereign debts added to tensions linked to a late QE programme weighted down on real GDP. The takeaway from this chart is that only the term premium channel of QE played a significant role in increasing real GDP. Also, this role came with a lag of three to six months, consistent with economic theory. Indeed, the domino effect of the portfolio balance mechanism from sovereign bonds to financing the real economy does need a few months to arise.

Turning to the decomposition of consumer prices, it is noteworthy that shocks related to the term premium played a big role over the sample, whether through the Risk or QE shocks. Alongside, the Supply and Demand shocks played little role in driving prices between 2009 and 2018, a period marked with a higher volatility of the price level (Stdev. of 0.63 for 2004-2008 and 0.96 for 2009-2018), much of it being driven by the Risk shock. In addition, as for real GDP, the positive contribution of the QE signalling shock from mid 2012 to mid 2014 is probably also due to the

*whatever it takes* period, in the sens that lower interest rate expectations must have raised inflation expectations and therefore prices. Besides, the amplitude of the term premium and signalling QE shocks surge from mid 2014, when QE started being anticipated by the markets (especially after Draghi speech at Jackson Hole in August 2014). However, the sum of the contribution of the QE shocks was negative between 2014 and 2017. It therefore means that during 2015Q3-2017, a the reduction in interest rate expectations and term premia had a positive effect on real GDP but a contractionary impact on consumer prices. From 2017 to 2019 yet, QE shocks' contributions to prices revert and follow the ones of real GDP, encouraging the rise in prices. Thus, it seems that the initial response of prices to QE is puzzling and in contradiction to Economic theory at first glance.

To test the robustness of this price puzzle, I next conduct a counterfactual analysis as it is often the case when evaluating the impact of QE (Kapetanios et al. (2012), Baumeister & Benati (2013) among others). That is, I calculate the values of real GDP and of the price level, had QE not been implemented through the term premium channel. However, I depart from the quoted studies as I compute shock-specific conditional forecasts in addition to the usual shock shut-down in the historical decompositions. Especially, I assume that the conditional forecasts of real GDP and prices are generated by the QE term premium shock only, so as to fully isolate the term premium channel. To do so, I rely on the method developed by Dieppe, Legrand & Van Roye (2016). In particular, I conduct an in-sample forecasting exercise from January 2015 to December 2018 to capture the entire period of the QE programme. Based on the SRTSM estimated at a daily frequency, I estimate that the amplitude of the impact of the announcement of the QE programme on January $22^{nd}$, 2015 on the term premium lasted until February $2^{nd}$, 2015. Past that date, the downward impact started reverting. More precisely, I find that the 10-year AAA Euro Area term premium fell by about 25 basis points between these dates (see **Figure 22 - Appendix A.5**). In addition, **Figure 21 - Appendix A.5** shows that most of the drop in the yield curves arose from the drop in the term premium; the expectation curve being stuck at the ELB. This is in line with **Figure 19** and **20** in **Appendix A.5**, that show the amplitude of the drop in the term premium, terminal rate, and 10-year yield following the QE-related events.

Ultimately, I plot the forecasts of the macroeconomic variables on **Figure 8** on the counterfactual that the term premium had not fallen by 25 basis points as a result of the QE term premium shock, using:

$$\Upsilon_{t+h} = \widetilde{\Upsilon}_{t+h} + \sum_{q=1}^{h} C_{h-q} B w_{t+q} \tag{28}$$

where $\Upsilon_{t+h}$ is the conditional forecast of $\Upsilon$ at time $t = t + h$ and

$$\widetilde{\Upsilon}_{t+h} = \sum_{q=1}^{p} Z_q^{(h)} \Upsilon_{t-q+1}$$

is the unconditional forecast at time $t = t+h$, with $Z_q^{(h)}$ products of $Z_k$ and $B$ that can be obtained recursively.

Details regarding the derivation of the posterior distribution to produce conditional forecasts generated by a specific shock are presented in **Appendix B.3**

Figure 8: Conditional Forecasts - Identification Scheme I

**Point-wise median Conditional Forecasts:**
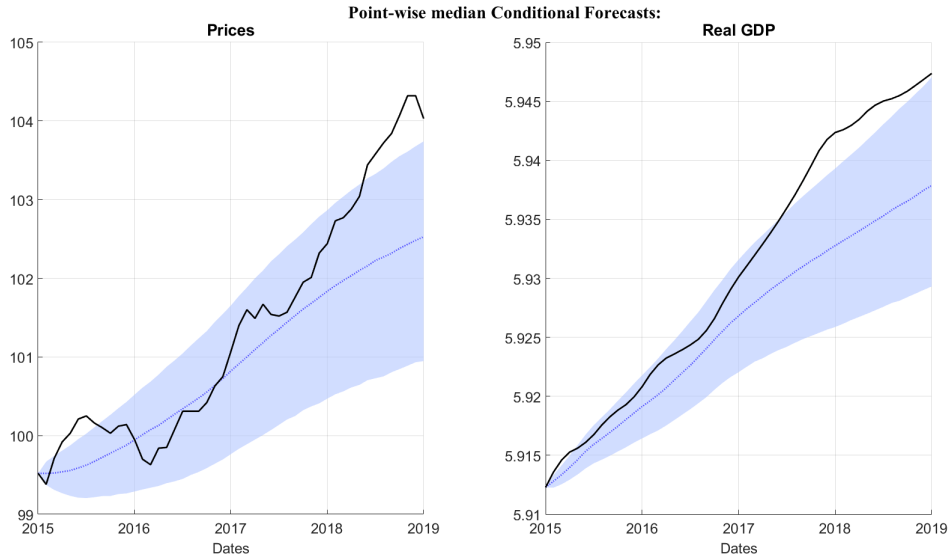


Solid black line shows the data
Doted-blue line is the point-wise median counterfactual with the confidence interval

The conditional forecasting exercise of **Figure 8** broadly confirms the analysis of the historical decompositions: the term premium channel of QE fostered real GDP (by about 1 point as of 2018M12), but weighted down on prices in 2016. However, note that the counterfactual forecast shows an increase in prices between 2015 and 2016, as opposed to the historical decomposition. This is potentially due to the difference in the draw analysed (median target v.s. point-wise median), but this may also indicate that a the source of the price puzzle only appeared with a lag. As a robustness check, counterfactual prices and real GDP originating from the HDs are presented on **Figure 28** and **Figure 27 - Appendix B.4**. Overall, they show that prices would have been slightly higher without the QE term premium shock until mid 2016 and lower afterwards, while real GDP would have been on average lower by 1 point over the period.

Finally, the initial adverse effect of the term premium channel on real GDP and prices is confirmed by the scatter plot of the macroeconomic variables conditional of the QE shocks, displayed on **Figure 9**. Indeed, this scatter plot clearly shows that there has been a trade-off between the two variables in 2014-2016 conditional on the term premium channel, as shown by the negative (-93.77) slope of the light blue regression line in the top right-hand corner plot. However, this trade-off disappears right before 2017 and the term premium channel of QE is then able to foster both prices and real GDP at the same time (the red regression line displays a big positive slope of 247.84). Moreover, the positive relation of the second part of the QE period actually counterbalances the early sample trade-off when the whole QE period is considered (black line). This confirms the hypothesis of a temporary price puzzle, rather than a well rooted effect. In addition, the data scatter plot shows that the unconditional correlation has been upward slopping over the period, indicating that the trade-off is only conditional on the QE term premium shock. This is consistent with the theoretical framework of a Phillips Curve, where a higher output comes at the cost of a higher inflation. Besides, the absence of any trade-off conditional on the signalling channel of QE corroborates the HD displaying a marginal impact of interest rate expectations on real GDP through the programme.

Figure 9: Scatter Plot - Identification Scheme I
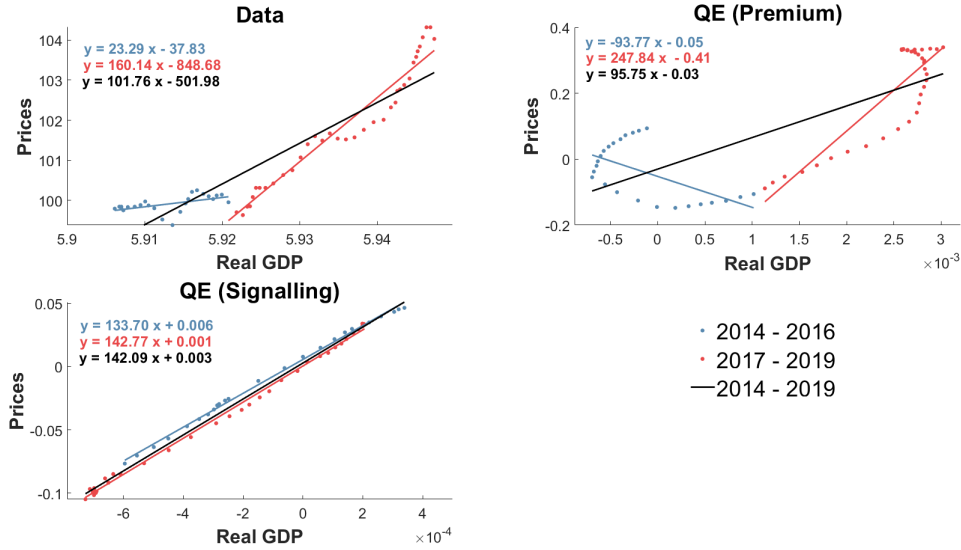


Scatter plot from 2014:M01 to 2018:M12

If the previous exercises confirm the presence of a temporary prize puzzle in the transmission of QE through the term premium channel, they however remain silent on its origin. Plus, the B-SVAR framework employed relies on a structural interpretation of the shocks hitting the economy and does not allow for any potential time-variation in its parameters. These issues are addressed in the next section, using reduced-form regressions with time-varying coefficients as well as a DCC-GARCH model.

# 4   Cost-channel of QE

There has been a vast literature trying to solve empirical price puzzles of monetary policy, starting with Sims (1992). I am however the first, to the best of my knowledge, to shed light on a price puzzle in the transmission mechanism of QE in the Euro Area. To highlight the transmission channel of the term premium to the consumer price level, I conduct three reduced-form regressions with time-varying parameters featuring the dynamics of the cost of capital for firms, producer price inflation, and consumer price inflation. I am therefore able to capture any potential time variation in the functioning of the term premium channel, while remaining agnostic vis-a-vis the shocks driving the economy.

First, the cost of borrowing dynamics is:

$$r_t^K = \omega_t r_{t-1}^K + \psi_t E_t[r_{t+10Y}] + \phi_t \kappa_t + \epsilon_{1,t} \tag{29}$$

where the corporate cost of borrowing for long-term loans $r_t^K$ is a function of its own lag as well as the (10-year) terminal rate $E_t[r_{t+10Y}]$ and the 10-year term premium $\kappa_t$. This framework relies on the assumption that corporate loan rates are mostly driven by the 10-year sovereign yield, which is a reasonable hypothesis as long as one believes that investment is funded by long-term loans rather than short-term. It enables to quantify the contribution of the yield components to the cost

27

of borrowing.

Second, Producer Price Inflation (PPI) dynamics is:

$$\pi_t^{PPI} = \mu_t \pi_{t-1}^{PPI} + \alpha_t \pi_t^{wages} + \beta_t r_{t-1}^K + \gamma_t \pi_t^{energy} + \zeta_t \pi_t^{import} + \epsilon_{1,t} \tag{30}$$

where the industrial PPI inflation (excluding construction) $\pi_t^{PPI}$ is a function of its own lag, wage inflation (based on compensation per employee) $\pi_t^{wages}$, the cost of borrowing for corporations a quarter before $r_{t-1}^K$, energy price inflation (based on HICP energy) $\pi_t^{energy}$, and imported inflation (excluding construction and energy) $\pi_t^{import}$. This setting will capture the importance of the term premium in driving PPI inflation.

Finally, Consumer Price Inflation (CPI) dynamics is:

$$\pi_t^{CPI} = \rho_t \pi_{t-1}^{CPI} + (1 - \rho_t) E_t[\pi_{t+1}^{CPI}] + \theta_t \pi_{t-1}^{PPI} + \eta_t r_{t-1}^K + \epsilon_{3,t} \tag{31}$$

where the CPI inflation (HICP) $\pi_t^{CPI}$ is driven by its own lag, the one quarter ahead consumer price inflation expectation from ECB Survey of Professional Forecasters $E_t[\pi_{t+1}^{CPI}]$,[28] by PPI inflation $\pi_{t-1}^{PPI}$, and the cost of borrowing for corporations a quarter before $r_{t-1}^K$. Hence, should a reduction of the term premium weight down on CPI inflation through a cost-channel, $\eta_t$ would be positive.

Besides, time-varying coefficients $(\omega_t, \psi_t, \phi_t, \mu_t, \alpha_t, \beta_t, \gamma_t, \zeta_t, \rho_t, \theta_t, \eta_t) \in \mathbb{R}$ are constrained random walks $(0 < \omega_t, \mu_t, \rho_t < 1$ and $\psi_t, \phi_t, \alpha_t, \gamma_t, \zeta_t, \theta_t > 0)$, while $\epsilon_{i,t} \overset{iid}{\sim} \mathcal{N}(0, \sigma_i^2)$ are the residuals. Equation (29), (30), and (31) are estimated via maximum a posteriori through the extended Kalman Filter from 2004Q4 to 2018Q4 using quarterly data. Three-year length rolling window OLS regressions estimates serve as prior to guide the filter. Estimates of time-varying coefficients $\beta$, $\eta$, and $\phi$ are presented on **Figure 10**. Of course, an obvious criticism is that one could see the three equations as linked and that they should be derived from a micro-founded model. I acknowledge this point but stress the fact that the exercise is deliberately left reduced-form so as to add to the structural analysis conducted in **Section 3.2**.

---

[28]Data are available at: http://sdw.ecb.europa.eu

Figure 10: Cost channel
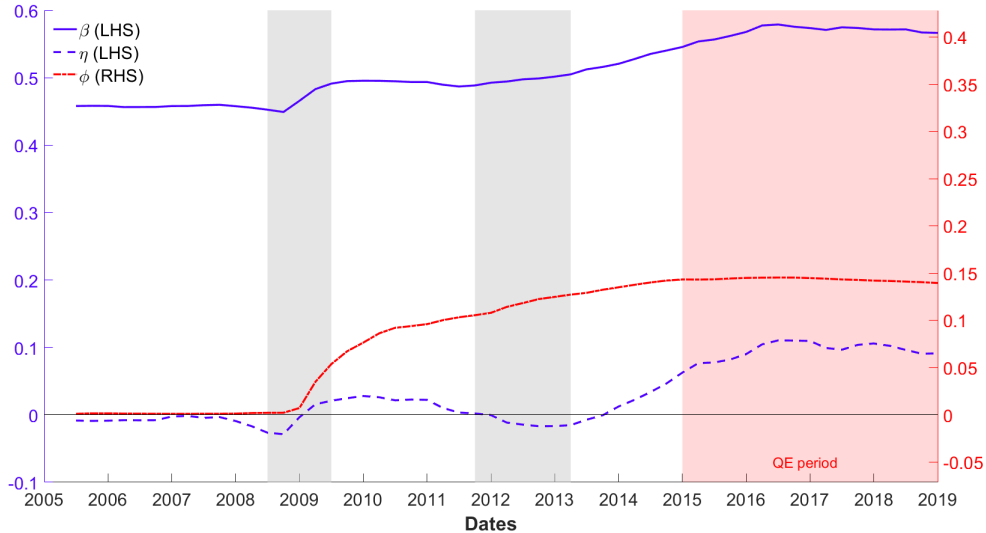


Sample period: 2004Q9-2018Q4
Solid grey area corresponds to a recession period as defined by the Center for Economic Policy Research (CEPR)
Solid red area corresponds to the QE period

Strikingly, the role of the term premium in the dynamics of the cost of borrowing for corporations, $\phi$, surged during GFC, jumping from zero to stabilize at around 0.15 in the QE period. This is in line with the unconventional monetary policies implemented in that period, where the central bank tried to act on firms financing conditions through long-term yields. Besides, this hypothesis is strengthen by the conditional covariance plotted on **Figure 29 - Appendix C.1**. Extracted from a Dynamic Conditional Correlation (DCC)-GARCH model based on Engle (2002) (details of the model are presented in **Appendix C**), the covariance between the term premium and the cost of borrowing hiked during GFC and since the start of the QE programme. It is positive between 2015 and 2019, indicating that the reduction in the cost of borrowing was also associated with a reduction in the term premium. This confirms the role of QE on firms financing.

More interestingly, if the role of the cost of capital in driving PPI inflation, $\beta$, plotted on **Figure 10** is positive throughout the sample and increased by about 25% since GFC, the contribution to CPI inflation, $\eta$, switches sign twice. This means that the decreasing cost of borrowing observed on **Figure 31 - Appendix D** since mid 2014, contributed to the drop in PPI inflation between 2012Q3 to 2016Q1. However, this contraction in the cost of borrowing alternatively pushed CPI inflation up and down. The relationship is indeed slightly positive in 2009-2012 and 2014-2018, implying that a lower cost of borrowing caused a lower CPI inflation. This is in contradiction with economic theory, which says that an interest rate cut should foster demand, therefore production and consumer prices. Instead, results provide evidence that the reduction in the cost of capital provoked by a decrease in the term premium has enabled producers to temporarily reduce selling prices. They would do so in order to gain in market share, while preserving their margins thanks to the diminution of the cost of production (PPI inflation here). This price puzzle is thus most probably solved through the cost channel developed by Christiano, Eichenbaum & Evans (1996), Barth & Ramey (2002), Ravenna & Walsh (2006) and Chowdhury, Hoffmann & Schabert (2006) among others. Especially, the cost channel of monetary policy focuses on the supply (cost-push) effects of a monetary policy shock. Not denying the demand-induced effect of an interest rate cut,

cost channel theory sheds light on the role of the cost of capital in the marginal cost of production determining firms' prices. That is, assuming that firms have to borrow in advance to finance capital and labor, a decrease in the cost of capital caused by a monetary policy easing lowers firms marginal cost of production. Then, some of those firms are able to lower selling prices in order to gain in market share: CPI goes down, implying a strong counter-cyclicality of price mark-ups.[29] After a certain period depending on price rigidity and the price-demand elasticity, demand goes up following the prices cut and CPI inflation starts picking up again.

In this paper, evidence from the B-SVAR of **Section 3.2** shows that the term premium channel of QE has contributed negatively to consumer prices between 2015 and 2017, after which the sign of the effect reverts. In addition, evidence from the regressions with time-varying coefficients show that the role of the cost of borrowing in the dynamics of CPI inflation has been positive since 2014. This validates the story of a cost-channel of working capital in the transmission of QE to the economy, that disappeared with a rising demand around 2017. Why the role of cost of capital in driving consumer prices switches sign throughout the sample can be explained by the net cash position of firms, working capital requirements, menu costs as well as the degree of competition in specific market segments. Put differently, it is not surprising that the sign and amplitude of the cost of borrowing-CPI inflation relationship, $\eta$, is time-varying, as the amplitude of the demand-induced versus the supply-side effects of monetary policy may not be constant. Eventually, the sign is determined by the dominance of one of these two effects. This finding provides a new explanation to why inflation has been so muted between 2015 and 2017, despite the easing monetary environment. If the negative contribution of energy inflation has often been highlighted (ECB Economic Bulletin (2019)), it is quite plausible that a cost channel effect has also contributed to slow prices growth down during that period.

Before closing this section, I introduce the concept of cost of capital pass-through to quantify the cost channel effect revealed here. In particular, I compute the ratio of the elasticity of CPI inflation with respect to the cost of borrowing, $\eta$, to the elasticity of PPI inflation with respect to the cost of borrowing, $\beta$. Thus, a positive pass-through implies the presence of the cost channel and the higher the pass-through, the stronger the cost channel. It therefore measures the ability of producers to reset prices when marginal cost of production goes down. It is presented on **Figure 11**. As of December 2018, the pass-through reached 0.16, implying that a one-percent drop in the cost of borrowing would imply a decrease in CPI inflation by 0.16 of the drop in PPI inflation. It started rising about a year before the introduction of the QE programme, when term premia were pushed down by anticipations. The fact that the pass-through is also positive between 2009 and 2012 may be attributed to the sharp decrease in yields initiated by policy rate cuts after GFC.

---

[29]The literature lacks recent study on the conditional cyclicality of price mark-ups in the Euro area. For a recent paper on the U.S. see Nekarda & Ramey (2019).

Figure 11: Cost of capital pass-through
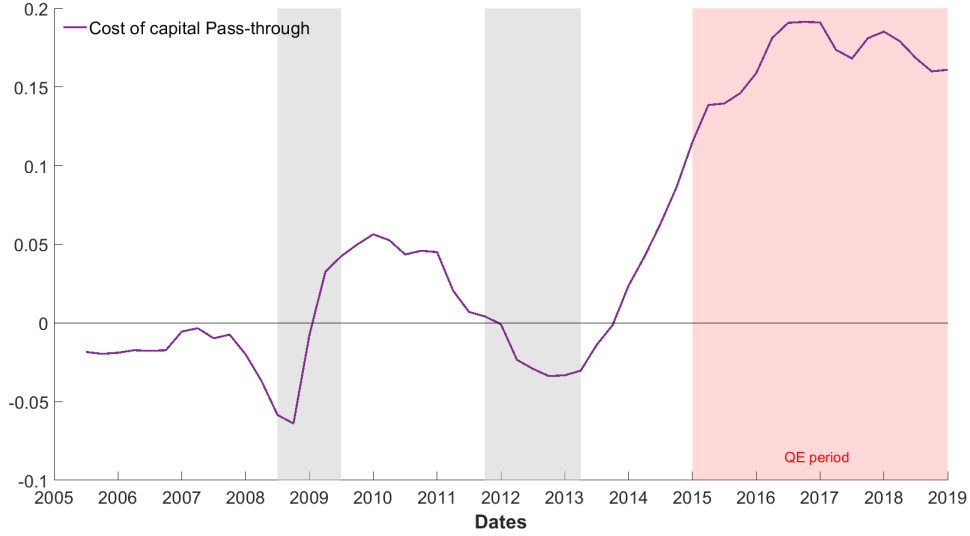
# 5 Policy rule for the term premium

Now that I have shown that the term premium can be used as a monetary policy instrument and that I have quantified its impact on prices and real GDP, I compute a policy rule for the term premium to serve as a base thought to design a governing rule for Quantitative Easing. In particular, to see if there is a parameter instability in the contingent relation, I estimate both a constant and a time-varying parameter version of the following rule:

$$\kappa_t = \rho_t \kappa_{t-1} + (1 - \rho_t)(\tau_{\pi,t} \pi_t^{gap} + \tau_{y,t} y_t^{gap} + \tau_{r,t} r_t) \tag{32}$$

where $\kappa_t$ is the 10-year term premium, $\pi_t^{gap}$ is annualized CPI inflation minus the inflation target (2%), $y_t^{gap}$ is the detrended $log$-real GDP growth using an HP-Filter, and $r_t$ is the 3-month Euro Area interest rate minus the one quarter ahead annualized inflation expectation from ECB SPF.[30] Besides, $0 < \rho_t < 1$ and $(\tau_{\pi,t}, \tau_{y,t}, \tau_{\kappa,t}) \in \mathbb{R}$ are random walks. Constant parameter version of Equation (32) is estimated via maximum likelihood, while the time-varying version is estimated via maximum a posteriori through the extented Kalman Filter using OLS estimates as priors to guide the filter.

---

[30]Interpolated at the monthly frequency using a Quadratic approach matching average.

Table 5: Constant parameter Policy rule for the term premium

| LLH | RMSE | $\rho$ | $\tau_\pi$ | $\tau_y$ | $\tau_r$ |
|------|-------|-----------|----------|----------|-----------|
| -92 | 0.11 | 0.982*** | 0.950* | -2.525* | -0.887** |
|      |       | (0.009) | (0.555) | (1.31) | (0.389) |

Sample period: 2009M01-2018M12
Standard errors are indicated in parentheses
LLH is the Log-Likelihood
RMSE is the Root Mean Squared Error
* Rejection of the null hypothesis of the coefficient equal to zero at 10%
** Rejection of the null hypothesis of the coefficient equal to zero at 5%
*** Rejection of the null hypothesis of the coefficient equal to zero at 1%

Figure 12: Time-varying parameter Policy rule for the term premium
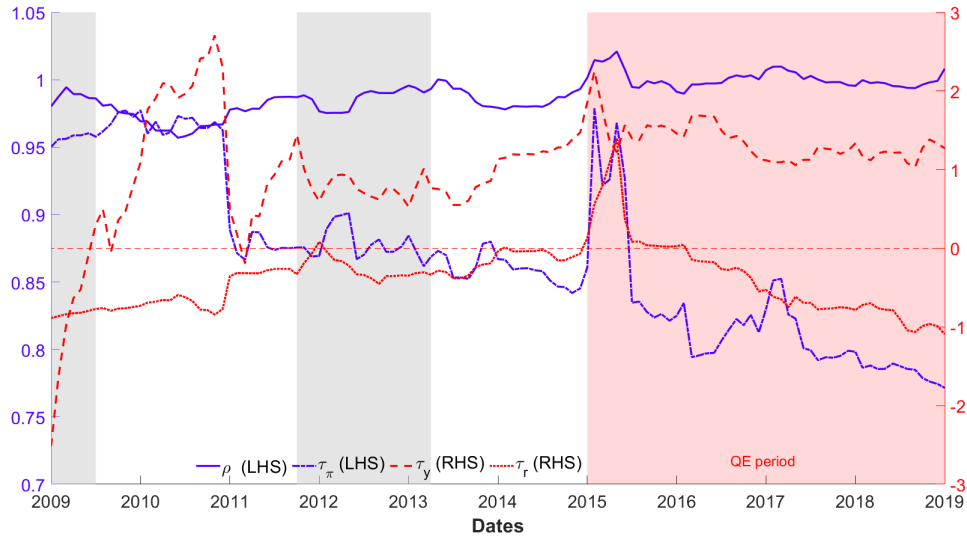


Sample: 2009:M1-2018:M12
LLH=59 - RMSE=0.11
Solid grey areas correspond to recession periods as defined by the Center for Economic Policy Research (CEPR)
Solid red area corresponds to the QE period

In the constant-parameter regression presented in **Table 5**, all coefficients are statistically significant at 10% and the fit is relatively good (RMSE=0.11). Besides, the sign of the coefficients $\tau_{\pi,t}, \tau_{y,t},$ and $\tau_{\kappa,t}$ is worth commenting. First, it seems that over the period considered, inflation feedback to the term premium was positive, as it would be in a conventional Taylor rule. Thus, the term premium responded positively to a rise in inflation but slightly lower than one-to-one. Yet, taking a look at the potential time-variation in that coefficient on **Figure 12** shows that the strength of the feedback decreased over the sample, with the exception of the QE shock episodes in January and March 2015. Second, the output gap coefficient is strongly negative, averaging -2.5 as displayed in **Table 5**. However, a quick look at **Figure 12** reveals that the negativity is actually driven by the non-stable first part of the sample (GFC). Aside from that, the feedback is positive and stabilizes at about 1.3 in December 2018, implying that the term premium responded more

than one-to-one to an increase in the output gap.[31] Finally, the negative sign of $\tau_r$ for most part of the sample means that the term premium responded adversely to monetary policy shocks, implying that market participants tended to revise their perception of risk, inferring an improvement in the economy from a policy rate hike.

Of course, the micro-foundation and determinacy of such a rule is unexplored here. This would require a rich model, that is beyond the scope of that paper. However, this purely reduced-form exercise sheds light on the relevance of a policy rule for the term premium and lays the ground for future research.

# 6    Conclusion

In this paper, I quantify the effects of the ECB PSPP programme on the Eurozone economy through the yield curve channels and show that the term premium can be used as a single monetary policy instrument. My motivation for this study lies in the fact that it has been found in the literature that QE mostly impacts the term premium component of long-term yields, relative to short-term interest rate expectations. However, this term premium channel of QE to the economy had never been properly isolated and studied. Plus, practitioner's view that a decline in the term premium is expansionary for economic activity is not backed by the existing literature, both theoretical and empirical, that struggles to capture the impact of a change in the term premium on macroeconomic aggregates. This paper contributes to fill this gap in the literature.

To capture the macroeconomic implications of the ECB QE programme, both through the signalling and term premium channels, I first extract a term premium and terminal rate series from a shadow rate term structure model with anchored interest rate expectations. I then plug the two yield components into a monthly Bayesian Structural VAR with the price level, real GDP, and a series of market-consistent PSPP purchase. I next identify a pure QE signalling and QE term premium shock thanks to zero restrictions, sign restrictions, and Narrative sign restrictions that I develop beyond the framework given by their authors. With this set-up, I show that the term premium channel of QE had a positive impact on real GDP with a lag of three quarters, while the signalling channel played little role. In addition, I show that both the signalling and term premium channels contributed negatively to consumer prices in 2015-2017 and positively afterwards. Indeed, I later expose, using a DCC-GARCH model and time-varying coefficient regressions, that there has been a cost channel in the transmission mechanism of QE to inflation. That is, the decrease in firms' cost of borrowing associated to lower term premia, allowed producers to cut selling prices so as to gain in market share. Assuming firms have to borrow in advance to finance capital and labor, this cost channel of QE provides a new explanation as to why inflation has been so muted between 2015 and 2017, despite the easing monetary environment.

Above all, in this paper I shed light on the benefits of using the term premium as a monetary policy instrument. In particular, I argue that in addition to giving the central bank a new policy instrument and to strengthen the impact of QE, it would also enable the monetary institution to shape the long-end of the yield curve. This may be desirable in the context of a yield curve control policy or simply to steepen the yield curve in order to restore bank margins. Thus, I explain how the central bank could directly control the term premium by neutralizing the signalling channel of QE. Briefly, the central bank would just need to dissociate interest rate expectations from asset purchase by credibly stating that it is ready to conduct QE- whatever the level of the policy rates. Hence, market participants will not infer from the introduction of a QE programme that the central bank will keep the policy rates low for a long period. The monetary institution will therefore be able to shape the long-end of the yield curve by acting on its term premium only. Besides, as it

---

[31]Of course, an obvious flaw in this exercise is the measure of output gap, simply inferred from filtering real GDP growth.

can shrink the term premium through QE, the central bank could also very well raise it through Quantitative Tightening (QT), if necessary to match the long-end of the curve with a target.

Finally, like a Taylor rule for the short-term interest rate, a policy rule for the term premium could be a major ingredient in the composition of a governing rule for asset purchase-and-sale. Here, I lay the groundwork to find one by estimating a constant as well as a time-varying coefficient version of such a rule. I am, however, well aware of the fact that my approach in this paper is purely empirical and that more structure from a model would be welcome to fully understand the interactions between the term premium and the macroeconomy. I leave that for future research.

# References

[1] Altavilla C., Brugniolini L., Gürkaynak R.S., Motto R. & Ragusa G., 2019, "Measuring euro area monetary policy", ECB Working Paper Series No. 2281

[2] Altavilla C., Carboni G. & Motto R., 2015, "Asset purchase programmes and financial markets: lessons from the Euro Area", ECB Working Paper Series No. 1864

[3] Andrade P., Breckenfelder J., De Fiore F., Karadi P. & Tristani O., 2016, "The ECB's asset purchase programme: an early assessment", ECB Working Paper Series No. 1956

[4] Ang A. & Piazzesi M., 2003, "A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables", Journal of Monetary Economics, 50(4), pp 745-787

[5] Ang A., Piazzesi M. & Wei M., 2006, "What Does the Yield Curve Tell Us about GDP Growth?", Journal of Econometrics, 131(1/2), pp 359-403

[6] Antolín-Díaz J. & Rubio-Ramírez J.F., 2018, "Narrative Sign Restrictions for SVARs", American Economic Review, 108(10), pp 2802-2829

[7] Arias J. E., Rubio-Ramirez J.F. & Waggoner D.F., 2018, "Inference based on structural vector autoregressions identified with sign and zero restrictions: theory and applications" Econometrica, 86(2), pp 685-720

[8] Barth III, M.J. & Ramey V.A., 2002, "The cost channel of monetary transmission", NBER chapters, NBER macroeconomics annual 2001(16), MIT Press, Cambridge MA, pp 199-256

[9] Bauer M.D., 2017, "Restrictions on Risk Prices in Dynamic Term Structure Models", Journal of Business & Economic Statistics, pp 1-16

[10] Bauer M.D. & Rudebusch G.D., 2016, "Monetary Policy Expectations at the Zero Lower Bound", Journal of Money, Credit and Banking, 48(7) pp 1439–1465

[11] Bauer M.D. & Rudebusch G.D., 2014, "The signaling channel for Federal Reserve bond purchases", International Journal of Central Banking, 10(3), pp 233-289

[12] Baumeister C. & Benati L., 2013, "Unconventional Monetary Policy and the Great Recession: Estimating the Macroeconomic Effects of a Spread Compression at the Zero Lower Bound", International Journal of Central Banking, June 2013

[13] Bekaert G., Engstrom E. & Ermolov A., 2016, "Macro risks and the term structure of interest rates", NBER Working Paper Series No. 22839

[14] Bernanke B.S., 2006, "Reflections on the Yield Curve and Monetary Policy", Speech presented at the Economic Club of New York, March 20

[15] Bernanke B.S., Reinhart V. & Sack B., 2004, "Monetary policy alternatives at the zero bound: an empirical assessment", Brookings Papers on Economic Activity, 2, pp 1-78

[16] van Binsbergen J.H., Fernández-Villaverde J., Koijen R.S.J. & Rubio-Ramírez J., 2012, "The term structure of interest rates in a DSGE model with recursive preferences", Journal of Monetary Economics, 59(7), pp 634-648

[17] Black F., 1995, "Interest Rates as Options", Journal of Finance, 50(5), pp 1371–1376

[18] Blattner T.S. & Joyce M.A.S., 2016, "Net debt supply shocks in the euro area and the implications for QE", ECB Working Paper No. 1957

[19] Bognanni M., 2018, "A Class of Time-Varying Parameter Structural VARs for Inference under Exact or Set Identification, mimeo

[20] Campbell J.Y., 2000, "Asset pricing at the millenium", The Journal of Finance, 55(4), pp 1515–1567

[21] Canova F., 2007, "Methods for Applied Macroeconomic Research", Princeton University Press

[22] Canova F. & Paustian M., 2011, "Business Cycle Measurement with Some Theory", Journal of Monetary Economics, 58(4), pp 345–361

[23] Carlstrom C.T., Fuerst T.S. & Paustian M., 2017, "Targeting long rates in a model with segmented markets", American Economic Journal: Macroeconomics, 9(1), pp 205-242

[24] Carriero A., Mouabbi S. & Vangelista E., 2015, "The UK term structure at the zero lower bound", Queen Mary university of London School of Economics and Finance Working Paper No. 755.

[25] Chen H., Cúrdia V. & Ferrero A., 2012, "The Macroeconomic effects of large-scalse asset purchase programmes", The Economic Journal, 122, pp 289-315

[26] Chowdhury I., Hoffmann, M. & Schabert A., 2006, "Inflation dynamics and the cost channel of monetary transmission", European Economic Review, 50(4), pp 995–1016

[27] Christensen J.H.E. & Rudebusch G.D., 2016, "Modeling Yields at the Zero Lower Bound: Are Shadow Rates the Solution?", Advances in Econometrics, 35, pp 75–125

[28] Christiano L.J., Eichenbaum M. & Evans C.L., 1996, "The effects of monetary policy shocks: evidence from the flow of funds", The Review of Economics and Statistics, 78(1), pp 16–34

[29] Claeys G., Leandro Á. & Mandra A., 2015, "European Central Bank Quantitative Easing: The Detailed Manual", Bruegel Policy Contribution, Issue 2015/02

[30] Cochrane J.H. & Piazzesi M., 2008, "Decomposing the Yield Curve", Manuscript, University of Chicago

[31] Cochrane J.H. & Piazzesi M., 2005, "Bond risk premia", American Economic Review, 95(1), pp 138–160

[32] Cohen B.H., Hördahl P. & Xia F.D., 2018, "Term premia: models and some stylised facts", BIS Quarterly Review, September 2018

[33] Creal D.D. & Wu J.C., 2016, "Bond risk premia in consumption-based models", NBER Working Paper No. 22183

[34] Darracq Pariès M. & Papadopoulou N., 2019, "On the credit and exchange rate channels of central bank asset purchases in a monetary union", ECB Working Paper Series No. 2259

[35] Dedola L., Georgiadis G., Gräb J. & Mehl A., 2017, "Does a Big Bazooka Matter? Central Bank Balance-Sheet Policies and Exchange Rates", ECB Working Paper Series No. 2197

[36] Deutsche Bundesbank Monthly Report, 2016, "The macroeconomic impact of quantitative easing in the Euro Area", Issue 2016/06, pp 29-53

[37] Dewachter H., Lania L. & Lyrio M., 2014, "Information in the Yield Curve: a Macro-Finance Approach", 29(1), Journal of Applied Econometrics, pp 42–64

[38] Dieppe A., Legrand R. & Van Roye B., 2016, "The BEAR Toolbox", ECB Working Paper Series N1934

[39] Duffee G.R., 2012, "Bond Pricing and the Macroeconomy", Handbook of the Economics of Finance, Volume 2, Part B, Chapter 13, pp 907-967

[40] Duffee G.R., 2002, "Term Premia and Interest Rate Forecasts in Affine Models", Journal of Finance, 57(1), 405–443

[41] Duffie D. & Kan R., 1996, "A yield-factor model of interest rates", Mathematical Finance, 6(4), pp 379-406

[42] Dufrénot G, Rhouzlane M. & Vaccaro-Grange E., 2019, "Potential growth and natural yield curve in Japan", AMSE Working Paper No. 12

[43] Durbin J. & Koopman S.J., 2012, "Time Series Analysis by State Space Methods", Oxford Statistical Science Series, Second Edition

[44] ECB Economic Bulletin, 2019, Issue 2

[45] Engle R.F. & Sheppard K., 2001, "Dynamic Conditional Correlation", NBER Working Papers No. 8554

[46] Eggertson G.B. & Woodford M., 2003, "The zero bound on interest rates and optimal monetary policy", Brookings papers on economic activity, 34, pp 139–235

[47] Eser F., Lemke W., Nyholm K., Radde S. & Vladu L., 2019, "Tracing the impact of the ECB's asset purchase programme on the yield curve", ECB Working Paper Series No. 2293

[48] Favero, C.A., Kaminska, I. & Söderström U., 2005, "The Predictive Power of the Yield Spread: Further Evidence and a Structural Interpretation", Unpublished manuscript, Università Bocconi

[49] Fry R. & Pagan A., 2011, "Sign Restrictions in Structural Vector Autoregressions: A Critical Review", Journal of Economic Literature, 49(4), pp 938-960

[50] Gambetti L. & Musso A., 2017, "The macroeconomic impact of the ECB's expanded asset purchase programme (APP)", ECB Working Paper No. 2075

[51] Garcia Pascual A. & Wieladek T., 2016, "The European Central Bank's QE: A New Hope", CESifo Working Paper 5946

[52] Geiger F., 2011, "The Yield Curve and Financial Risk Premia, Implication for Monetary Policy", Springer-Verlag Berlin Heidelberg, Lecture Notes in Economics and Mathematical Systems, 564

[53] Geiger F. & Schupp F., 2018, "With a little help from my friends: Survey-based derivation of euro area short rate expectations at the effective lower bound", Deutsche Bundesbank Discussion Paper No. 27

[54] Georgiadis G. & Gräb J., 2016, "Global financial market impact of the announcement of the ECB's asset purchase programme", Journal of Financial Stability, 26, pp 257–265

[55] Greenberg E., 2008, "Introduction to Bayesian Econometrics", Cambridge University Press

[56] Grewal, M. & Andrews A., 2008, "Kalman Filtering", Wiley and Sons, Third Edition

[57] Hamilton J.D., 1994, "Time Series Analysis", Princeton University Press

[58] Hamilton, J.D. & Kim D.H., 2002, "A Reexamination of the Predictability of Economic Activity Using the Yield Spread" Journal of Money, Credit, and Banking, 34(2), pp 340-360

[59] Hamilton J.D., & Wu C., 2010, "The Effectiveness of Alternative Monetary Policy Tools in a Zero Lower Bound Environment", Journal of Money, Credit and Banking, 44(s1), pp 3–46

[60] Hördahl P., Remonola E.M. & Valente G., 2015, "Expectations and Risk Premia at 8:30AM: Macroeconomic Announcements and the Yield Curve", BIS Working Paper No. 527

[61] Hördahl P., Tristani O. & Vestin D., 2008, "The Yield Curve and Macroeconomic Dynamics", The Economic Journal, 118(533), pp 1937-1970

[62] Ireland P.N., 2015, "Monetary policy, bond risk premia, and the economy", Jounal of Monetary Economics, 76, pp 124-140

[63] Jardet C., Monfort A. & Pegoraro F., 2013, "No-arbitrage Near-Cointegrated VAR(p) term structure models, term premia and GDP growth", Journal of Banking & Finance, 37(2), pp 389-402

[64] Jordà O. & Taylor A.M., 2019, "Riders on the Storm", Federal Reserve Bank of Kansas City Economic Policy Symposium "Challenges for Monetary Policy," Jackson Hole, Wyoming, August 22–24, 2019.

[65] Joslin, S., Singleton K. J. & Zhu H., 2011, "A new perspective on Gaussian dynamic term structure models", Review of Financial Studies 24(3), pp 926–970

[66] Kapetanios G., Mumtaz H., Stevens I. & Theodoridis K., 2012, "Assessing the Economy-wide effects of Quantitative Easing", The Economic Journal, 122, pp 316-347

[67] Kiley M.T., 2014, "The Aggregate Demand Effects of Short- and Long-Term Interest Rates", International Journal of Central Banking, 10, pp 69–104

[68] Kilian L. & Inoue A., 2013, "Inference on impulse response functions in structural VAR models", Journal of Econometrics, 177(1), pp 1-13

[69] Kilian L. & Lütkepohl H., 2016, "Structural Vector Autoregressive Analysis", Cambridge University Press, Cambridge

[70] Kilian L. & Murphy D.P., 2012, "Why Agnostic Sign Restrictions Are Not Enough: Understanding the Dynamics of Oil Market VAR Models", Journal of the European Economic Association, 10(5), pp 1166–1188

[71] Koijen R.S.J., Koulischer F., Nguyen B. & Yogo M., 2016, "Quantitative Easing in the Euro Area: the dynamics of risk exposures and the impact on asset prices", Bank of France Working Paper No. 601

[72] Kortela T., 2016, "A Shadow rate model with timevarying lower bound of interest rates", Bank of Finland Research Discussion Paper 19

[73] Krippner L., 2015, "Zero Lower Bound Term Structure Modeling, A Practitioner's Guide", Applied Quantitative Finance series

[74] Krishnamurthy, A. & Vissing-Jorgensen A., 2011, "The Effects of Quantitative Easing on Interest Rates: Channels and Implications for Policy", Northwestern University

[75] Lemke W., 2006, "Term Structure Modeling and Estimation in a State Space Framework", Springer-Verlag Berlin Heidelberg

[76] Lemke W. & Vladu A.L., 2017, "Below the zero lower bound: a Shadow-Rate Term Structure Model for the Euro Area", ECB Working Paper No. 1991

[77] Lütkepohl H., 2005, "New Introduction to Multiple Time Series Analysis", Springer-Verlag Berlin Heidelberg

[78] Middeldorp M., 2015, "Very much anticipated, ECB QE had a big impact on Asset Prices, even before it was officially announced", Bank of England Staff Blog *Bank Underground*

[79] Monfort, A., Pegoraro, F., Renne, J-P. & Roussellet, G., 2017, "Staying at zero with affine processes: An application to term structure modelling", Journal of Econometrics, 201(2), pp 348-366

[80] Mouabbi S. & Sahuc J-G., 2018, "Evaluating the Macroeconomic Effects of the ECB's Unconventional Monetary Policies", Journal of Money, Credit and Banking, 51, pp 831-858

[81] Nekarda C.J. & Ramey V.A., 2019, "The Cyclical Behavior of the Price-Cost Markup", mimeo

[82] Pawitan Y., 2001, "In All Likelihood: Statistical Modelling and Inference Using Likelihood", Oxford University Press

[83] Ravenna, F. & Walsh C.E., 2006, "Optimal monetary policy with the cost channel", Journal of Monetary Economics 53(2), pp 199–216

[84] Rudebusch G.D., Sack B.P. & Swanson E.T, 2007, "Macroeconomic Implications of Changes in the Term Premium", Federal Reserve Bank of St. Louis Review, 89, pp 241–269

37

[85] Rudebusch G., Swanson, E. & Wu, T., 2012, "The Bond Premium in a DSGE Model with Long-Run Real and Nominal Risks", American Economic Journal: Macroeconomics, 4(1), pp 105-143

[86] Rudebusch G., Swanson, E. & Wu, T., 2006, "The bond market 'conundrum' from a macro-finance perspective", Monetary and Economic Studies 24, pp 83-128

[87] De Santis R.A., 2016, "Impact of the asset purchase programme on Euro Area government bond yields using market news", ECB Working Paper Series No. 1939

[88] Simon D., 2006, "Optimal State Estimation", Wiley and Sons

[89] Sims C.A., 1992, "Interpreting the macroeconomic time series facts: The effects of monetary policy", European Economic Review, 36(5), pp 975–1000

[90] Vasicek O., 1977, "An equilibrium characterization of the term structure", Journal of Financial Economics, 5, pp 177–188

[91] Vayanos D. & Vila J-L., 2009. "A Preferred Habitat Model of the Term Structure of Interest Rates" Financial Markets Group Working paper, London School of Economics

[92] Waggoner D.F. & Zha T., 1999, Conditional Forecasts In Dynamic Multivariate Models, The Review of Economics and Statistics, 81(4), pp 639–651

[93] Weale M. & Wieladek T., 2016, "What are the macroeconomic effects of asset purchases?", Journal of Monetary Economics, 79, pp 81-93

[94] Werner R.A., 1995, "How to create a recovery through Quantitative Monetary Easing", Nihon Keizai Shinbun (Nikkei), pp 26

[95] Woodford M., 2003, "Interest and Prices: Foundations of a Theory of Monetary Policy", Princeton University Press

[96] Wright J., 2006, "The Yield Curve and Predicting Recessions", Finance and Economics Discussion Series, Board of Governors of the Federal Reserve System

[97] Wu J.C. & Xia F.D., 2016, "Measuring the Macroeconomic Impact of Monetary Policy at the Zero Lower Bound", Journal of Money, Credit and Banking, 48(2-3), pp 253-291

[98] Wu J.C. & Xia F.D., 2017, "Time Varying Lower Bound of Interest Rates in Europe", working paper, 2017

# Appendices

## A    Term Structure Model

### A.1    Gaussian Affine Term Structure Model

#### A.1.1    Derivation of the one-period nominal interest rate

First, let us impose a no-arbitrage restriction in a frictionless financial market that implies the use of a stochastic discount factor (SDF) or Pricing Kernel. In this setting, the nominal price of a zero-coupon bond follows the law of one price in the real or physical world $\mathbb{P}$, such that:

$$P_{n,t} = E_t^{\mathbb{P}}(M_{t+1}P_{n-1,t+1}) \tag{A.1}$$

where $M_{t+1}$ is the strictly positive nominal stochastic discount factor and $P_{n,t}$ is the nominal price of a $n$-period zero-coupon bond at time $t$.

In particular, for a one-period zero-coupon bond, I have:

$$P_{1,t} = E_t^{\mathbb{P}}(M_{t+1}) \tag{A.2}$$

Now, I assume that the nominal SDF is conditionally log-normally distributed, so that:

$$E_t^{\mathbb{P}}(M_{t+1}) = e^{\mu} + \frac{\sigma^2}{2} \tag{A.3}$$

where $log(M_{t+1}) \sim \mathcal{N}(\mu, \sigma^2)$

Hence, taking the logarithm in Equation (A.3) gives:

$$log[E_t^{\mathbb{P}}(M_{t+1})] = \mu + \frac{\sigma^2}{2} \tag{A.4}$$

which using $m_{t+1} = log(M_{t+1})$, can be re-written:

$$log[E_t^{\mathbb{P}}(M_{t+1})] = E_t^{\mathbb{P}}(m_{t+1}) + \frac{1}{2}Var_t^{\mathbb{P}}(m_{t+1}) \tag{A.5}$$

Using Equation (A.2) and $r_{1,t} = -log(P_{1,t})$, I obtain:

$$r_{1,t} = -E_t^{\mathbb{P}}(m_{t+1}) - \frac{1}{2}Var_t^{\mathbb{P}}(m_{t+1}) \tag{A.6}$$

where $r_{1,t}$ is the one-period nominal interest rate.

To simplify, I drop the period index for the nominal interest rate, which gives:

$$\boxed{r_t = -E_t^{\mathbb{P}}(m_{t+1}) - \frac{1}{2}Var_t^{\mathbb{P}}(m_{t+1})} \tag{A.7}$$

#### A.1.2    Derivation of the nominal Stochastic Discount Factor

In the framework of Equations (2), (1) and (6), the assumption of no-arbitrage implies the existence of the Radon-Nikodym derivative $\xi_t$[32] defined as the ratio of the risk-neutral probability $\mathbb{Q}$ to

---

[32]See Geiger (2011), Section 3.5.2 for more details.

physical probability $\mathbb{P}$ of **Section 2.1**. I assume it follows a log-normal process, so that:

$$\xi_{t+1} = \xi_t e^{-\frac{1}{2}\Lambda_t'\Lambda_t - \Lambda_t'\varepsilon_{t+1}^{\mathbb{P}}} \tag{A.8}$$

where $\varepsilon_{t+1}^{\mathbb{P}}$ is the same disturbance as in Equation (2).

As the expression for the nominal SDF is:

$$M_{t+1} = e^{-r_t}\frac{\xi_{t+1}}{\xi_t} \tag{A.9}$$

I obtain the essentially affine form of the *log* of the nominal stochastic discount factor, as in Duffee (2002):

$$\boxed{m_{t+1} = -r_t - \frac{1}{2}\Lambda_t'\Lambda_t - \Lambda_t'\varepsilon_{t+1}^{\mathbb{P}}} \tag{A.10}$$

### A.1.3   Nominal bond pricing

$$P_{n,t} = E_t^{\mathbb{P}}(M_{t+1}P_{n-1,t+1}) \tag{A.11}$$

Plug Equations (A.10) and (7):

$$P_{n,t} = E_t^{\mathbb{P}}[exp(-r_t - \frac{1}{2}\Lambda_t'\Lambda_t - \Lambda_t'\varepsilon_{t+1}^{\mathbb{P}})exp(A_{n-1} + B_{n-1}'X_{t+1})] \tag{A.12}$$

Plug Equations (2) and (1):

$$P_{n,t} = E_t^{\mathbb{P}}[exp(-\delta_0 - \delta_1'X_t - \frac{1}{2}\Lambda_t'\Lambda_t - \Lambda_t'\varepsilon_{t+1}^{\mathbb{P}})exp(A_{n-1} + B_{n-1}'(\mu^{\mathbb{P}} + \Theta^{\mathbb{P}}X_t + \Sigma_\varepsilon\varepsilon_{t+1}^{\mathbb{P}})] \tag{A.13}$$

$$P_{n,t} = exp(-\delta_0 - \delta_1'X_t - \frac{1}{2}\Lambda_t'\Lambda_t + A_{n-1} + B_{n-1}'\mu^{\mathbb{P}} + B_{n-1}'\Theta^{\mathbb{P}}X_t)E_t^{\mathbb{P}}[exp((B_{n-1}'\Sigma_\varepsilon - \Lambda_t')\varepsilon_{t+1}^{\mathbb{P}})] \tag{A.14}$$

As $\varepsilon_{t+1}^{\mathbb{P}} \sim \mathcal{N}(0, I_N)$, then $exp(\varepsilon_{t+1}^{\mathbb{P}})$ is log-normally distributed so that:

$$E_t^{\mathbb{P}}[exp(b\varepsilon_{t+1}^{\mathbb{P}})] = exp(\frac{1}{2}bI_Nb') \tag{A.15}$$

with $b = B_{n-1}'\Sigma_\varepsilon - \Lambda_t'$

Then,

$$P_{n,t} = exp(-\delta_0 - \delta_1'X_t - \frac{1}{2}\Lambda_t'\Lambda_t + A_{n-1} + B_{n-1}'\mu^{\mathbb{P}} + B_{n-1}'\Theta^{\mathbb{P}}X_t)...$$
$$exp(\frac{1}{2}(B_{n-1}'\Sigma_\varepsilon - \Lambda_t')Var_t^{\mathbb{P}}(\varepsilon_{t+1}^{\mathbb{P}})(B_{n-1}'\Sigma_\varepsilon - \Lambda_t')') \tag{A.16}$$

And $Var_t^{\mathbb{P}}(\varepsilon_{t+1}^{\mathbb{P}}) = I_N$, so:

$$P_{n,t} = exp(-\delta_0 - \delta_1'X_t - \frac{1}{2}\Lambda_t'\Lambda_t + A_{n-1} + B_{n-1}'\mu^{\mathbb{P}} + B_{n-1}'\Theta^{\mathbb{P}}X_t)...$$
$$exp(\frac{1}{2}B_{n-1}'\Sigma_\varepsilon\Sigma_\varepsilon'B_{n-1} - \frac{1}{2}B_{n-1}'\Sigma_\varepsilon\Lambda_t - \frac{1}{2}\Lambda_t'\Sigma_\varepsilon'B_{n-1} + \frac{1}{2}\Lambda_t'\Lambda_t) \tag{A.17}$$

But $\Lambda'_t \Sigma'_\varepsilon B_{n-1}$ is $\in \mathbb{R}$ so it is also equal to its transpose, which is $B'_{n-1} \Sigma_\varepsilon \Lambda_t$

Therefore, I obtain:

$$P_{n,t} = exp(-\delta_0 - \delta'_1 X_t - \frac{1}{2}\Lambda'_t\Lambda_t + A_{n-1} + B'_{n-1}\mu^\mathbb{P} + B'_{n-1}\Theta^\mathbb{P}X_t)...$$
$$exp(\frac{1}{2}B'_{n-1}\Sigma_\varepsilon\Sigma'_\varepsilon B_{n-1} - B'_{n-1}\Sigma_\varepsilon\Lambda_t + \frac{1}{2}\Lambda'_t\Lambda_t) \quad \text{(A.18)}$$

Now, plug Equation (6):

$$P_{n,t} = exp(-\delta_0 - \delta'_1 X_t + A_{n-1} + B'_{n-1}\mu^\mathbb{P} + B'_{n-1}\Theta^\mathbb{P}X_t)...$$
$$exp(\frac{1}{2}B'_{n-1}\Sigma_\varepsilon\Sigma'_\varepsilon B_{n-1} - B'_{n-1}(\lambda_0 + \lambda_1 X_t)) \quad \text{(A.19)}$$

Plug Equation (7) on the left-hand side:

$$exp(A_n + B'_n X_t) = exp(A_{n-1} - \delta_0 + B'_{n-1}(\mu^\mathbb{P} - \lambda_0) + \frac{1}{2}B'_{n-1}\Sigma_\varepsilon\Sigma'_\varepsilon B_{n-1})...$$
$$exp(B'_{n-1}(\Theta^\mathbb{P} - \lambda_1)X_t - \delta'_1 X_t) \quad \text{(A.20)}$$

By identification:

$$\boxed{A_n = A_{n-1} + B'_{n-1}(\mu^\mathbb{P} - \lambda_0) - \delta_0 + \frac{1}{2}B'_{n-1}\Sigma_\varepsilon\Sigma'_\varepsilon B_{n-1}} \quad \text{(A.21)}$$

$$\boxed{B'_n = B'_{n-1}(\Theta^\mathbb{P} - \lambda_1) - \delta'_1} \quad \text{(A.22)}$$

For the one-period zero-coupon bond, I have:

$$P_{1,t} = E_t^\mathbb{P}(M_{t+1}) \quad \text{(A.23)}$$

$$P_{1,t} = e^{-r_t} \quad \text{(A.24)}$$

$$P_{1,t} = e^{-\delta_0 - \delta'_1 X_t} \quad \text{(A.25)}$$

By identification, $A_1 = -\delta_0$ and $B_1 = -\delta_1$

Also, a zero-coupon bond maturing instantaneously has a price of one. This implies that $A_0 = 0$ and $B_0 = (0,0,0)'$

Hence, by recursion and using the fact that $\Theta^\mathbb{Q} = \Theta^\mathbb{P} - \lambda_1$ we find a solution for Equation (A.22):

$$B'_n = -\delta'_1 \sum_{j=0}^{n-1}(\Theta^\mathbb{P} - \lambda_1)^j \quad \text{(A.26)}$$

$$\boxed{B'_n = -\delta'_1[I_N - (\Theta^\mathbb{Q})^n](I_N - \Theta^\mathbb{Q})^{-1}} \quad \text{(A.27)}$$

Unfortunately, there is no closed-form solution to Equation (A.21).

Finally, yield are obtained using:

$$y_{n,t} = -\frac{1}{n}log(P_{n,t}) \quad \text{(A.28)}$$

## A.2 Shadow Rate Term Structure Model

### A.2.1 Expression of the nominal zero-coupon bond yield

Using Equation (A.1) and the law of iterated expectations under the risk-neutral measure $\mathbb{Q}$, I easily obtain:

$$P_{n,t} = E_t^{\mathbb{Q}}[\prod_{j=1}^{n} M_{t+j}] \tag{A.29}$$

Using $r_{t+j} = -log(M_{t+j+1})$, leads to:

$$P_{n,t} = E_t^{\mathbb{Q}}(e^{-\sum_{j=0}^{n-1} r_{t+j}}) \tag{A.30}$$

Plugging $y_{n,t} = -\frac{1}{n}log(P_{n,t})$, I obtain:

$$y_{n,t} = -\frac{1}{n}log[E_t^{\mathbb{Q}}(e^{-\sum_{j=0}^{n-1} r_{t+j}})] \tag{A.31}$$

Now, I use $E_t[log(X)] = log[E_t(X)] - \frac{1}{2}Var_t[log(X)]$:

$$y_{n,t} = -\frac{1}{n}[E_t^{\mathbb{Q}}(-\sum_{j=0}^{n-1} r_{t+j}) + \frac{1}{2}Var_t^{\mathbb{Q}}(-\sum_{j=0}^{n-1} r_{t+j})] \tag{A.32}$$

$$y_{n,t} = \frac{1}{n}[\sum_{j=0}^{n-1} E_t^{\mathbb{Q}}(r_{t+j}) - \frac{1}{2}Var_t^{\mathbb{Q}}(\sum_{j=0}^{n-1} r_{t+j})] \tag{A.33}$$

Now for tractability, let us focus on the expression of the *one*-period forward rate in $n$ periods:

$$f_{n,n+1,t} = (n+1)y_{n+1,t} - ny_{n,t} \tag{A.34}$$

$$f_{n,n+1,t} = \sum_{j=0}^{n} E_t^{\mathbb{Q}}(r_{t+j}) - \frac{1}{2}Var_t^{\mathbb{Q}}(\sum_{j=0}^{n} r_{t+j}) - \sum_{j=0}^{n-1} E_t^{\mathbb{Q}}(r_{t+j}) + \frac{1}{2}Var_t^{\mathbb{Q}}(\sum_{j=0}^{n-1} r_{t+j}) \tag{A.35}$$

$$f_{n,n+1,t} = E_t^{\mathbb{Q}}(r_{t+n}) - \frac{1}{2}(Var_t^{\mathbb{Q}}(\sum_{j=0}^{n} r_{t+j}) - Var_t^{\mathbb{Q}}(\sum_{j=0}^{n-1} r_{t+j})) \tag{A.36}$$

First, I express the first term on the right-hand side of Equation (A.36):

$$E_t^{\mathbb{Q}}(r_{t+n}) = E_t^{\mathbb{Q}}[max(\underline{r}_{t+n}, s_{t+n})] \tag{A.37}$$

$$E_t^{\mathbb{Q}}(r_{t+n}) = E_t^{\mathbb{Q}}\left[max\left(\frac{\underline{r}_{t+n} - E_t^{\mathbb{Q}}(s_{t+n})}{\sigma_{s,n}^{\mathbb{Q}}}, \frac{s_{t+n} - E_t^{\mathbb{Q}}(s_{t+n})}{\sigma_{s,n}^{\mathbb{Q}}}\right)\right]\sigma_{s,n}^{\mathbb{Q}} + E_t^{\mathbb{Q}}(s_{t+n}) \tag{A.38}$$

$$E_t^{\mathbb{Q}}(r_{t+n}) = [Prob_t^{\mathbb{Q}}\left(\frac{s_{t+n} - E_t^{\mathbb{Q}}(s_{t+n})}{\sigma_{s,n}^{\mathbb{Q}}} < \frac{\underline{r}_{t+n} - E_t^{\mathbb{Q}}(s_{t+n})}{\sigma_{s,n}^{\mathbb{Q}}}\right) \times \left(\frac{\underline{r}_{t+n} - E_t^{\mathbb{Q}}(s_{t+n})}{\sigma_{s,n}^{\mathbb{Q}}}\right) + ...$$

$$Prob_t^{\mathbb{Q}}\left(\frac{s_{t+n} - E_t^{\mathbb{Q}}(s_{t+n})}{\sigma_{s,n}^{\mathbb{Q}}} \geq \frac{\underline{r}_{t+n} - E_t^{\mathbb{Q}}(s_{t+n})}{\sigma_{s,n}^{\mathbb{Q}}}\right) \times ...$$

$$E_t^{\mathbb{Q}}\left(\frac{s_{t+n} - E_t^{\mathbb{Q}}(s_{t+n})}{\sigma_{s,n}^{\mathbb{Q}}}\middle|\frac{s_{t+n} - E_t^{\mathbb{Q}}(s_{t+n})}{\sigma_{s,n}^{\mathbb{Q}}} \geq \frac{\underline{r}_{t+n} - E_t^{\mathbb{Q}}(s_{t+n})}{\sigma_{s,n}^{\mathbb{Q}}}\right)\sigma_{s,n}^{\mathbb{Q}} + E_t^{\mathbb{Q}}(s_{t+n})$$

$$\tag{A.39}$$

Now, assume that the *shadow rate* is conditionally normally distributed, so that:

$$\frac{s_{t+n} - E_t^{\mathbb{Q}}(s_{t+n})}{\sigma_{s,n}^{\mathbb{Q}}} \sim \mathcal{N}(0, I_N) \tag{A.40}$$

Moreover, $Prob(x < \alpha) = \Phi(\alpha)$ and $Prob(x \geq \alpha)E_t(x \mid x \geq \alpha) = \phi(\alpha)$ for $x \sim \mathcal{N}(0,1)$, where $\Phi$ is the standard normal cumulative distribution function and $\phi$ is the standard normal probability density function. Consequently, Equation (A.39) becomes:

$$E_t^{\mathbb{Q}}(r_{t+n}) = [\Phi\left(\frac{\underline{r}_{t+n} - E_t^{\mathbb{Q}}(s_{t+n})}{\sigma_{s,n}^{\mathbb{Q}}}\right)\left(\frac{\underline{r}_{t+n} - E_t^{\mathbb{Q}}(s_{t+n})}{\sigma_{s,n}^{\mathbb{Q}}}\right) + ...$$

$$\phi\left(\frac{\underline{r}_{t+n} - E_t^{\mathbb{Q}}(s_{t+n})}{\sigma_{s,n}^{\mathbb{Q}}}\right)]\sigma_{s,n}^{\mathbb{Q}} + E_t^{\mathbb{Q}}(s_{t+n}) \tag{A.41}$$

Using the fact that $\Phi(-x) = 1 - \Phi(x)$ and $\phi(-x) = \phi(x)$, we obtain:

$$E_t^{\mathbb{Q}}(r_{t+n}) = \underline{r}_{t+n} + \sigma_{s,n}^{\mathbb{Q}} \times \left[\Phi\left(\frac{E_t^{\mathbb{Q}}(s_{t+n}) - \underline{r}_{t+n}}{\sigma_{s,n}^{\mathbb{Q}}}\right)\left(\frac{E_t^{\mathbb{Q}}(s_{t+n}) - \underline{r}_{t+n}}{\sigma_{s,n}^{\mathbb{Q}}}\right) + \phi\left(\frac{E_t^{\mathbb{Q}}(s_{t+n}) - \underline{r}_{t+n}}{\sigma_{s,n}^{\mathbb{Q}}}\right)\right] \tag{A.42}$$

$$E_t^{\mathbb{Q}}(r_{t+n}) = \underline{r}_{t+n} + \sigma_{s,n}^{\mathbb{Q}} \times g\left(\frac{E_t^{\mathbb{Q}}(s_{t+n}) - \underline{r}_{t+n}}{\sigma_{s,n}^{\mathbb{Q}}}\right) \tag{A.43}$$

where $g(x) = x\Phi(x) + \phi(x)$

Besides,

$$E_t^{\mathbb{Q}}(s_{t+n}) = \delta_0 + \delta_1' E_t^{\mathbb{Q}}(X_{t+n}) \tag{A.44}$$

And under the probability measure $\mathbb{Q}$, iterating Equation (2) gives:

$$X_{t+n} = \sum_{j=0}^{n-1} \Theta^{\mathbb{Q}j}\mu^{\mathbb{Q}} + \Theta^{\mathbb{Q}n}X_t + \sum_{j=0}^{n-1} \Theta^{\mathbb{Q}j}\Sigma_\varepsilon \varepsilon_{t+n-j}^{\mathbb{Q}} \tag{A.45}$$

Hence, the conditional expectation of the *shadow rate* $s_{t+n}$ is:

$$E_t^{\mathbb{Q}}(s_{t+n}) = \delta_0 + \delta_1'(\sum_{j=0}^{n-1} \Theta^{\mathbb{Q}j}\mu^{\mathbb{Q}} + \Theta^{\mathbb{Q}n}X_t) \tag{A.46}$$

Also, using Equation (1) for $s_t$ and Equation (A.45), I find the standard deviation of the *shadow rate* $s_{t+n}$:

$$\sigma_{s,n}^{\mathbb{Q}} = [Var_t^{\mathbb{Q}}(s_{t+n})]^{1/2} = [\delta_1' Var_t^{\mathbb{Q}}(X_{t+n})\delta_1]^{1/2}$$

$$\sigma_{s,n}^{\mathbb{Q}} = [\sum_{j=0}^{n-1} \delta_1'\Theta^{\mathbb{Q}j}\Sigma_\varepsilon \Sigma_\varepsilon' \Theta^{\mathbb{Q}j\prime}\delta_1]^{1/2} \tag{A.47}$$

Define $\bar{a}_n = \delta_0 + \delta_1'\left(\sum_{j=0}^{n-1} \Theta^{\mathbb{Q}j}\right)\mu^{\mathbb{Q}}$

$a_n = \bar{a}_n - \frac{1}{2}\delta_1'\left(\sum_{j=0}^{n-1} \Theta^{\mathbb{Q}j}\right)\Sigma_\varepsilon \Sigma_\varepsilon'\left(\sum_{j=0}^{n-1} \Theta^{\mathbb{Q}j}\right)'\delta_1$

$b_n' = \delta_1'\Theta^{\mathbb{Q}n}$

$\sigma_{s,n}^{\mathbb{Q}} = [\sum_{j=0}^{n-1} \delta_1'(\Theta^{\mathbb{Q}})^j\Sigma_\varepsilon \Sigma_\varepsilon'(\Theta^{\mathbb{Q}})^{j\prime}\delta_1]^{1/2}$

Thus,

$$E_t^{\mathbb{Q}}(r_{t+n}) = \underline{r}_{t+n} + \sigma_{s,n}^{\mathbb{Q}} \times g\left(\frac{\bar{a}_n + b_n' X_t - \underline{r}_{t+n}}{\sigma_{s,n}^{\mathbb{Q}}}\right) \tag{A.48}$$

Note from Equation (A.43) that when $E_t^{\mathbb{Q}}(s_{t+j}) \gg \underline{r}_{t+n}$, $g(x) \to x$ and $E_t^{\mathbb{Q}}(r_{t+n}) \approx E_t^{\mathbb{Q}}(s_{t+n})$ i.e. the expected short-term interest rate is close to the expected *shadow rate* when the *shadow rate* is far above the lower bound $\underline{r}_t$. On the other hand, when $E_t^{\mathbb{Q}}(s_{t+n}) \ll \underline{r}_{t+n}$, $g(x) \to 0$ and $E_t^{\mathbb{Q}}(r_{t+n}) \approx \underline{r}_{t+n}$, i.e. the expected short-term interest rate is close to the lower bound when the *shadow rate* is far below the lower bound.

Now, let us focus on the second term on the right-hand side of Equation (A.36):

$$Var_t^{\mathbb{Q}}(\sum_{j=0}^{n} r_{t+j}) - Var_t^{\mathbb{Q}}(\sum_{j=0}^{n-1} r_{t+j}) = ...$$

$$\sum_{j=0}^{n} Var_t^{\mathbb{Q}}(r_{t+j}) + 2\sum_{i=0}^{n-1}\sum_{j=i+1}^{n} cov_t^{\mathbb{Q}}(r_{t+j}, r_{t+i}) - \sum_{j=0}^{n-1} Var_t^{\mathbb{Q}}(r_{t+j}) - 2\sum_{i=0}^{n-2}\sum_{j=i+1}^{n-1} cov_t^{\mathbb{Q}}(r_{t+j}, r_{t+i})$$

$$\tag{A.49}$$

$$Var_t^{\mathbb{Q}}(\sum_{j=0}^{n} r_{t+j}) - Var_t^{\mathbb{Q}}(\sum_{j=0}^{n-1} r_{t+j}) = Var_t^{\mathbb{Q}}(r_{t+n}) + 2\sum_{i=0}^{n-1} cov_t^{\mathbb{Q}}(r_{t+n}, r_{t+i}) \tag{A.50}$$

As in Wu & Xia (2016), assume that $Var_t^{\mathbb{Q}}(r_{t+n}) \approx Prob_t^{\mathbb{Q}}(s_{t+n} \geq \underline{r}_{t+n})Var_t^{\mathbb{Q}}(s_{t+n})$ and $cov_t^{\mathbb{Q}}(r_{t+n}, r_{t+i}) \approx Prob_t^{\mathbb{Q}}(s_{t+n} \geq \underline{r}_{t+n})cov_t^{\mathbb{Q}}(s_{t+n}, s_{t+i}) \ \forall \ (i,j) \ \in \ [\![0, n-2]\!] \times [\![1, n-1]\!]$

$$Var_t^{\mathbb{Q}}(\sum_{j=0}^{n} r_{t+j}) - Var_t^{\mathbb{Q}}(\sum_{j=0}^{n-1} r_{t+j}) \approx ...$$

$$Prob_t^{\mathbb{Q}}(s_{t+n} \geq \underline{r}_{t+n})Var_t^{\mathbb{Q}}(s_{t+n}) + 2\sum_{i=0}^{n-1} Prob_t^{\mathbb{Q}}(s_{t+n} \geq \underline{r}_{t+n})cov_t^{\mathbb{Q}}(s_{t+n}, s_{t+i}) \quad \text{(A.51)}$$

$$Var_t^{\mathbb{Q}}(\sum_{j=0}^{n} r_{t+j}) - Var_t^{\mathbb{Q}}(\sum_{j=0}^{n-1} r_{t+j}) \approx ...$$

$$Prob_t^{\mathbb{Q}}(s_{t+n} \geq \underline{r}_{t+n})[\sum_{j=0}^{n} Var_t^{\mathbb{Q}}(r_{t+j}) - \sum_{j=0}^{n-1} Var_t^{\mathbb{Q}}(r_{t+j}) + ...$$

$$2\sum_{i=0}^{n-1}\sum_{j=i+1}^{n} cov_t^{\mathbb{Q}}(s_{t+j}, s_{t+i}) - 2\sum_{i=0}^{n-2}\sum_{j=i+1}^{n-1} cov_t^{\mathbb{Q}}(s_{t+j}, s_{t+i})] \quad \text{(A.52)}$$

$$Var_t^{\mathbb{Q}}(\sum_{j=0}^{n} r_{t+j}) - Var_t^{\mathbb{Q}}(\sum_{j=0}^{n-1} r_{t+j}) \approx Prob_t^{\mathbb{Q}}(s_{t+n} \geq \underline{r}_{t+n})[Var_t^{\mathbb{Q}}(\sum_{j=0}^{n} s_{t+j}) - Var_t^{\mathbb{Q}}(\sum_{j=0}^{n-1} s_{t+j})] \text{ (A.53)}$$

$$Var_t^{\mathbb{Q}}(\sum_{j=0}^{n} r_{t+j}) - Var_t^{\mathbb{Q}}(\sum_{j=0}^{n-1} r_{t+j}) \approx Prob_t^{\mathbb{Q}}\left(\frac{s_{t+n} - E_t^{\mathbb{Q}}(s_{t+n})}{\sigma_{s,n)}^{\mathbb{Q}}} \geq \frac{\underline{r}_{t+n} - E_t^{\mathbb{Q}}(s_{t+n})}{\sigma_{s,n}^{\mathbb{Q}}}\right) \times ...$$

$$[Var_t^{\mathbb{Q}}(\sum_{j=0}^{n} s_{t+j}) - Var_t^{\mathbb{Q}}(\sum_{j=0}^{n-1} s_{t+j})] \quad (A.54)$$

$$Var_t^{\mathbb{Q}}(\sum_{j=0}^{n} r_{t+j}) - Var_t^{\mathbb{Q}}(\sum_{j=0}^{n-1} r_{t+j}) \approx [1 - Prob_t^{\mathbb{Q}}\left(\frac{s_{t+n} - E_t^{\mathbb{Q}}(s_{t+n})}{\sigma_{s,n)}^{\mathbb{Q}}} < \frac{\underline{r}_{t+n} - E_t^{\mathbb{Q}}(s_{t+n})}{\sigma_{s,n}^{\mathbb{Q}}}\right)] \times ...$$

$$[Var_t^{\mathbb{Q}}(\sum_{j=0}^{n} s_{t+j}) - Var_t^{\mathbb{Q}}(\sum_{j=0}^{n-1} s_{t+j})] \quad (A.55)$$

$$Var_t^{\mathbb{Q}}(\sum_{j=0}^{n} r_{t+j}) - Var_t^{\mathbb{Q}}(\sum_{j=0}^{n-1} r_{t+j}) \approx [1 - \Phi\left(\frac{\underline{r}_{t+n} - E_t^{\mathbb{Q}}(s_{t+n})}{\sigma_{s,n}^{\mathbb{Q}}}\right)] \times ...$$

$$[Var_t^{\mathbb{Q}}(\sum_{j=0}^{n} s_{t+j}) - Var_t^{\mathbb{Q}}(\sum_{j=0}^{n-1} s_{t+j})] \quad (A.56)$$

$$Var_t^{\mathbb{Q}}(\sum_{j=0}^{n} r_{t+j}) - Var_t^{\mathbb{Q}}(\sum_{j=0}^{n-1} r_{t+j}) \approx \Phi\left(\frac{E_t^{\mathbb{Q}}(s_{t+n}) - \underline{r}_{t+n}}{\sigma_{s,n}^{\mathbb{Q}}}\right) \times ...$$

$$[Var_t^{\mathbb{Q}}(\sum_{j=0}^{n} s_{t+j}) - Var_t^{\mathbb{Q}}(\sum_{j=0}^{n-1} s_{t+j})] \quad (A.57)$$

$$Var_t^{\mathbb{Q}}(\sum_{j=0}^{n} r_{t+j}) - Var_t^{\mathbb{Q}}(\sum_{j=0}^{n-1} r_{t+j}) \approx \Phi\left(\frac{E_t^{\mathbb{Q}}(s_{t+n}) - \underline{r}_{t+n}}{\sigma_{s,n}^{\mathbb{Q}}}\right) \times ...$$

$$[Var_t^{\mathbb{Q}}(s_{t+n}) + 2\sum_{j=0}^{n-1} cov_t^{\mathbb{Q}}(s_{t+n}, s_{t+j})] \quad (A.58)$$

Now recall that: $s_{t+j} = \delta_0 + \delta_1' X_{t+j}$ and $\sigma_{s,n}^{\mathbb{Q}} = [\sum_{j=0}^{n-1} \delta_1'(\Theta^{\mathbb{Q}})^j \Sigma_\varepsilon \Sigma_\varepsilon' (\Theta^{\mathbb{Q}})^{j\prime} \delta_1]^{1/2}$

$$Var_t^{\mathbb{Q}}(\sum_{j=0}^{n} r_{t+j}) - Var_t^{\mathbb{Q}}(\sum_{j=0}^{n-1} r_{t+j}) \approx \Phi\left(\frac{E_t^{\mathbb{Q}}(s_{t+n}) - \underline{r}_{t+n}}{\sigma_{s,n}^{\mathbb{Q}}}\right) \times ...$$

$$(\sum_{j=0}^{n-1} \delta_1'(\Theta^{\mathbb{Q}})^j \Sigma_\varepsilon \Sigma_\varepsilon' (\Theta^{\mathbb{Q}})^{j\prime} \delta_1 + 2\sum_{j=0}^{n-1} cov_t^{\mathbb{Q}}(\delta_1' X_{t+j}, \delta_1' X_{t+i}) \quad (A.59)$$

Using Equation (A.45), we obtain:

$$Var_t^{\mathbb{Q}}(\sum_{j=0}^{n} r_{t+j}) - Var_t^{\mathbb{Q}}(\sum_{j=0}^{n-1} r_{t+j}) \approx \Phi\left(\frac{E_t^{\mathbb{Q}}(s_{t+n}) - \underline{r}_{t+n}}{\sigma_{s,n}^{\mathbb{Q}}}\right) \times ...$$
$$(\sum_{j=0}^{n-1} \delta_1'(\Theta^{\mathbb{Q}})^j \Sigma_\varepsilon \Sigma_\varepsilon' (\Theta^{\mathbb{Q}})^{j\,\prime} \delta_1 + 2\delta_1 \sum_{j=0}^{n-1} cov_t^{\mathbb{Q}}(\sum_{k=0}^{j-1} \Theta^{\mathbb{Q}k} \Sigma_\varepsilon \varepsilon_{t+j-k}, \sum_{q=0}^{i-1} \Theta^{\mathbb{Q}q} \Sigma_\varepsilon \varepsilon_{t+i-q}) \delta_1' \quad \text{(A.60)}$$

$$Var_t^{\mathbb{Q}}(\sum_{j=0}^{n} r_{t+j}) - Var_t^{\mathbb{Q}}(\sum_{j=0}^{n-1} r_{t+j}) \approx \Phi\left(\frac{E_t^{\mathbb{Q}}(s_{t+n}) - \underline{r}_{t+n}}{\sigma_{s,n}^{\mathbb{Q}}}\right) \times ...$$
$$(\sum_{j=0}^{n-1} \delta_1'(\Theta^{\mathbb{Q}})^j \Sigma_\varepsilon \Sigma_\varepsilon' (\Theta^{\mathbb{Q}})^{j\,\prime} \delta_1 + 2\delta_1 \sum_{j=0}^{n-1} \sum_{k=0}^{j-1} \sum_{q=0}^{i-1} \Theta^{\mathbb{Q}k} \Sigma_\varepsilon cov_t^{\mathbb{Q}}(\varepsilon_{t+j-k}, \varepsilon_{t+i-q}) \Sigma_\varepsilon' \Theta^{\mathbb{Q}q\,\prime} \delta_1' \quad \text{(A.61)}$$

$$Var_t^{\mathbb{Q}}(\sum_{j=0}^{n} r_{t+j}) - Var_t^{\mathbb{Q}}(\sum_{j=0}^{n-1} r_{t+j}) \approx \Phi\left(\frac{E_t^{\mathbb{Q}}(s_{t+n}) - \underline{r}_{t+n}}{\sigma_{s,n}^{\mathbb{Q}}}\right) \times ...$$
$$(\sum_{j=0}^{n-1} \delta_1'(\Theta^{\mathbb{Q}})^j \Sigma_\varepsilon \Sigma_\varepsilon' (\Theta^{\mathbb{Q}})^{j\,\prime} \delta_1 + 2\delta_1 \sum_{j=0}^{n-1} \sum_{q=0}^{i-1} \delta_1' \Theta^{\mathbb{Q}n-i+q} \Sigma_\varepsilon \Sigma_\varepsilon' \Theta^{\mathbb{Q}q\,\prime} \delta_1' \quad \text{(A.62)}$$

$$Var_t^{\mathbb{Q}}(\sum_{j=0}^{n} r_{t+j}) - Var_t^{\mathbb{Q}}(\sum_{j=0}^{n-1} r_{t+j}) \approx \Phi\left(\frac{E_t^{\mathbb{Q}}(s_{t+n}) - \underline{r}_{t+n}}{\sigma_{s,n}^{\mathbb{Q}}}\right) \times ...$$
$$\delta_1'\left(\sum_{j=0}^{n-1} (\Theta^{\mathbb{Q}})^j\right) \Sigma_\varepsilon \Sigma_\varepsilon' \left(\sum_{j=0}^{n-1} (\Theta^{\mathbb{Q}})^j\right)' \delta_1 \quad \text{(A.63)}$$

So in summary,

$$Var_t^{\mathbb{Q}}(\sum_{j=0}^{n} r_{t+j}) - Var_t^{\mathbb{Q}}(\sum_{j=0}^{n-1} r_{t+j}) \approx \Phi\left(\frac{\bar{a}_n + b_n' X_t - \underline{r}_{t+n}}{\sigma_{s,n}^{\mathbb{Q}}}\right) \times (\bar{a}_n - a_n) \quad \text{(A.64)}$$

I now plug Equations (A.48) and (A.64) into Equation (A.36):

$$f_{n,n+1,t} \approx \underline{r}_{t+n} + \sigma_{s,j}^{\mathbb{Q}} \times g\left(\frac{\bar{a}_n + b_n' X_t - \underline{r}_{t+n}}{\sigma_{s,n}^{\mathbb{Q}}}\right) + \Phi\left(\frac{\bar{a}_n + b_n' X_t - \underline{r}_{t+n}}{\sigma_{s,n}^{\mathbb{Q}}}\right) \times (a_n - \bar{a}_n) \quad \text{(A.65)}$$

$$f_{n,n+1,t} \approx \underline{r}_{t+n} + \sigma_{s,j}^{\mathbb{Q}} \times g\left(\frac{\bar{a}_n + b_n' X_t - \underline{r}_{t+n}}{\sigma_{s,n}^{\mathbb{Q}}}\right) + \frac{\partial g\left(\frac{\bar{a}_n + b_n' X_t - \underline{r}_{t+n}}{\sigma_{s,n}^{\mathbb{Q}}}\right)}{\partial \bar{a}_n} \times (a_n - \bar{a}_n) \quad \text{(A.66)}$$

$$\boxed{f_{n,n+1,t} \approx \underline{r}_{t+n} + \sigma_{s,n}^{\mathbb{Q}} \times g\left(\frac{a_n + b_n' X_t - \underline{r}_{t+n}}{\sigma_{s,n}^{\mathbb{Q}}}\right)}$$

where $g(x) = x\Phi(x) + \phi(x)$

$a_n = \delta_0 + \delta_1' \left( \sum_{j=0}^{n-1} \Theta^{\mathbb{Q}j} \right) \mu^{\mathbb{Q}} - \frac{1}{2}\delta_1' \left( \sum_{j=0}^{n-1} \Theta^{\mathbb{Q}j} \right) \Sigma_\varepsilon \Sigma_\varepsilon' \left( \sum_{j=0}^{n-1} \Theta^{\mathbb{Q}j} \right)' \delta_1$

$b_n' = \delta_1' \Theta^{\mathbb{Q}n}$

$\sigma_{s,n}^{\mathbb{Q}} = [\sum_{q=0}^{n-1} \delta_1'(\Theta^{\mathbb{Q}})^q \Sigma_\varepsilon \Sigma_\varepsilon'(\Theta^{\mathbb{Q}})^{q'}\delta_1]^{1/2}$

Note that I have $a_n = A_n - A_{n+1}$ and $b_n' = B_n' - B_{n+1}'$ as $f_{n,n+1,t}^{GATSM} = a_n + b_n' X_t$, where $A_n$ and $B_n$ are the coefficients from the GATSM in Equations (A.21) and (A.22).

Besides, when $\underline{r}_{t+n} \to -\infty$, $g(x) \to x$ and $f_{n,n+1,t} \to f_{n,n+1,t}^{GATSM}$

Finally, using the fact that $f_{0,1,t} = y_{1,t} = r_t$ I retrieve the expression of the $n$-period nominal zero-coupon bond yield $y_{n,t}$:

$$y_{n,t} = \frac{1}{n} \sum_{j=0}^{n-1} f_{j,j+1,t}$$

## A.3  Transformation of the factors

To prove Equation (15), let us start with Equation (2):

$$X_t = \mu^{\mathbb{P}} + \Theta^{\mathbb{P}} X_{t-1} + \Sigma_\varepsilon \varepsilon_t^{\mathbb{P}} \tag{A.67}$$

Now, apply weights $W$ to the model-implied yields $Y_t$ to obtain the model-implied principal components $\mathcal{P}_t$:

$$\mathcal{P}_t = WY_t = W(A + B'X_t) = WA + WB'X_t \tag{A.68}$$

where $A \in \mathbb{R}^J$ and $B \in \mathcal{M}_{N \times J}(\mathbb{R})$ contain coefficients $A_n$ and $B_n$ from the GATSM (see **Section 2.1**), for the $J$ picked maturities, as defined in **Appendix A.4.1**

Let us plug Equation (A.67) into Equation (A.68):

$$WA + WB'X_t = WA + WB'(\mu^{\mathbb{P}} + \Theta^{\mathbb{P}}X_{t-1} + \Sigma_\varepsilon \varepsilon_t^{\mathbb{P}}) \tag{A.69}$$

$$\mathcal{P}_t = WA + WB'\mu^{\mathbb{P}} + WB'\Theta^{\mathbb{P}}X_{t-1} + WB'\Sigma_\varepsilon \varepsilon_t^{\mathbb{P}} \tag{A.70}$$

$$\mathcal{P}_t = WA + WB'\mu^{\mathbb{P}} + WB'\Theta^{\mathbb{P}}(WB')^{-1}WB'X_{t-1} + WB'\Sigma_\varepsilon \varepsilon_t^{\mathbb{P}} \tag{A.71}$$

$$\mathcal{P}_t = WA + WB'\mu^{\mathbb{P}} + WB'\Theta^{\mathbb{P}}(WB')^{-1}WA - WB'\Theta^{\mathbb{P}}(WB')^{-1}WA + ...$$
$$WB'\Theta^{\mathbb{P}}(WB')^{-1}WB'X_{t-1} + WB'\Sigma_\varepsilon \varepsilon_t^{\mathbb{P}} \tag{A.72}$$

$$\mathcal{P}_t = WA + WB'\mu^{\mathbb{P}} - WB'\Theta^{\mathbb{P}}(WB')^{-1}WA + ...$$
$$+ WB'\Theta^{\mathbb{P}}(WB')^{-1}(WA + WB'X_{t-1}) + WB'\Sigma_\varepsilon \varepsilon_t^{\mathbb{P}} \tag{A.73}$$

Let us plug Equation (A.67) at time $t = t - 1$. We obtain:

$$\mathcal{P}_t = \mu_{\mathcal{P}}^{\mathbb{P}} + \Theta_{\mathcal{P}}^{\mathbb{P}} \mathcal{P}_{t-1} + \Sigma_{\varepsilon,\mathcal{P}} \varepsilon_t^{\mathbb{P}} \tag{A.74}$$

where $\mu_{\mathcal{P}}^{\mathbb{P}} = WA + WB'\mu^{\mathbb{P}} - WB'\Theta^{\mathbb{P}}(WB')^{-1}WA \in \mathbb{R}^N$, $\Theta_{\mathcal{P}}^{\mathbb{P}} = WB'\Theta^{\mathbb{P}}(WB')^{-1} \in \mathcal{M}_{N \times N}(\mathbb{R})$ and $\Sigma_{\varepsilon,\mathcal{P}} = WB'\Sigma_\varepsilon \in \mathcal{M}_{N \times N}(\mathbb{R})$

Equation (16) is proved similarly.

Besides, we give the expression of the *shadow rate* in the rotated world. To do so, let us start with Equation (1):

$$s_t = \delta_0 + \delta_1' X_t \tag{A.75}$$

$$X_t = \delta_1'^{-1}(s_t - \delta_0) \tag{A.76}$$

let us plug Equation (A.67) reverted:

$$(WB')^{-1}(\mathcal{P}_t - WA) = \delta_1'^{-1}(s_t - \delta_0) \tag{A.77}$$

$$s_t = \delta_0 - \delta_1'(WB')^{-1}WA + \delta_1'(WB')^{-1}\mathcal{P}_t \tag{A.78}$$

$$s_t = \delta_{0,\mathcal{P}} + \delta_{1,\mathcal{P}}'\mathcal{P}_t \tag{A.79}$$

where $\delta_{0,\mathcal{P}} = \delta_0 - \delta_1'(WB')^{-1}WA \in \mathbb{R}$ and $\delta_{1,\mathcal{P}} = [(WB')^{-1}]'\delta_1 \in \mathbb{R}^N$

Finally, zero-coupon bond yields in the GATSM framework are given by:

$$Y_t = A + B'X_t \tag{A.80}$$

where $A \in \mathbb{R}^J$ and $B \in \mathcal{M}_{N \times J}(\mathbb{R})$ contain coefficients $A_n$ and $B_n$ from the GATSM (see **Section 2.1**), for the $J = 12$ picked maturities, as defined in **Appendix A.4.1**

To obtain the expression in the rotated world, let us plug Equation (A.67):
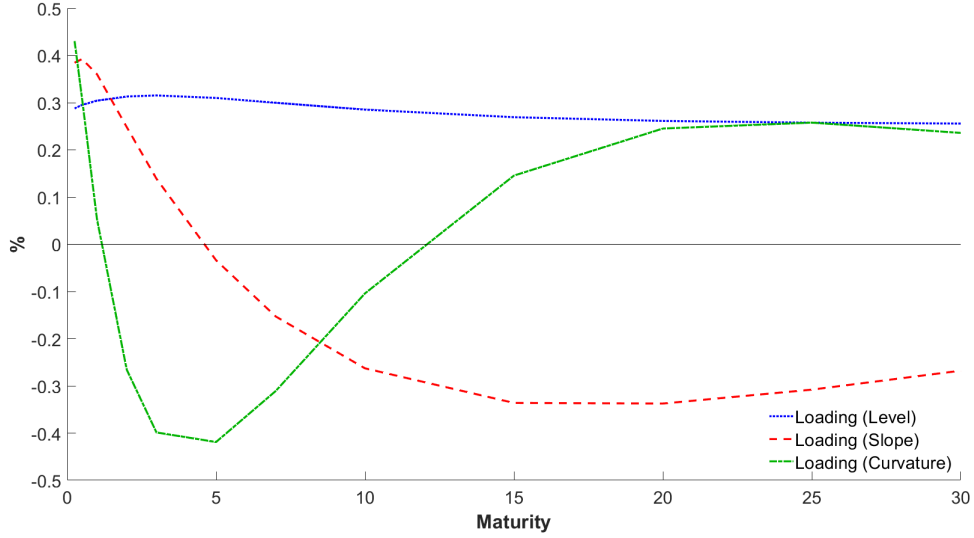
$$Y_t = A + B'(WB')^{-1}(\mathcal{P}_t - WA) \tag{A.81}$$

$$Y_t = A - B'(WB')^{-1}WA + B'(WB')^{-1}\mathcal{P}_t \tag{A.82}$$

$$Y_t = A_{\mathcal{P}} + B_{\mathcal{P}}\mathcal{P}_t \tag{A.83}$$

where $A_{\mathcal{P}} = [I_n - B'(WB')^{-1}W]A \in \mathbb{R}^J$ and $B_{\mathcal{P}} = B'(WB')^{-1} \in \mathcal{M}_{N \times J}(\mathbb{R})$

Figure 13: Loadings on observed yields



Loadings of observed yields from 2004:M9 to 2018:M12

Loadings presented on **Figure 13** have very classic shapes. The first loading is quasi flat as it is usually the case for the *level*, the second is downward slopping as for the *slope* and the third has a humped shape that is similar to the loading on a *curvature* factor.

## A.4 Estimation

### A.4.1 Kalman Filter for the GATSM

The procedure presented here is similar to the one detailed in Lemke (2006), Section 5.[33]

In the Gaussian Affine Term Structure framework of **Section 2.1**, my state-space model is composed of a linear transition Equation (2) and a linear measurement Equation (17), because of the linear nature of $y_{n,t}$ in Equation (8).

Transition equation:
$$X_t = \mu^{\mathbb{P}} + \Theta^{\mathbb{P}} X_{t-1} + \Sigma_\varepsilon \varepsilon_t^{\mathbb{P}} \tag{A.84}$$

where $X_t = (X_t^1, X_t^2, X_t^3)'$, $\mu^{\mathbb{P}} = (\mu_1^{\mathbb{P}}, \mu_2^{\mathbb{P}}, \mu_3^{\mathbb{P}},)'$, $\varepsilon_t^{\mathbb{P}} = (\varepsilon_t^{1\mathbb{P}}, \varepsilon_t^{2\mathbb{P}}, \varepsilon_t^{3\mathbb{P}})'$ with $\varepsilon_t^{i\mathbb{P}} \overset{iid}{\sim} \mathcal{N}(0,1)$ for $i \in [\![1,3]\!]$

$$\Theta^{\mathbb{P}} = \begin{pmatrix} \theta_{11}^{\mathbb{P}} & \theta_{12}^{\mathbb{P}} & \theta_{13}^{\mathbb{P}} \\ \theta_{21}^{\mathbb{P}} & \theta_{22}^{\mathbb{P}} & \theta_{23}^{\mathbb{P}} \\ \theta_{31}^{\mathbb{P}} & \theta_{32}^{\mathbb{P}} & \theta_{33}^{\mathbb{P}} \end{pmatrix}, \ \Sigma_\varepsilon = \begin{pmatrix} \sigma_{\varepsilon 11} & 0 & 0 \\ \sigma_{\varepsilon 21} & \sigma_{\varepsilon 22} & 0 \\ \sigma_{\varepsilon 31} & \sigma_{\varepsilon 32} & \sigma_{\varepsilon 33} \end{pmatrix}$$

Measurement equation:
$$Y_t^\circ = A + B' X_t + \Sigma_\varsigma \varsigma_t \tag{A.85}$$

where $Y_t^\circ = (y_{1,t}^\circ, y_{2,t}^\circ, y_{4,t}^\circ, y_{8,t}^\circ, y_{12,t}^\circ, y_{20,t}^\circ, y_{28,t}^\circ, y_{40,t}^\circ, y_{60,t}^\circ, y_{80,t}^\circ, y_{100,t}^\circ, y_{120,t}^\circ)' \in \mathbb{R}^{12}$

---

[33]See equivalently Hamilton (1994), Section 13, Lütkepohl (2005), Section 18, Durbin & Koopman (2012), Section 4 and 7 or Krippner (2015), Section 3.2.

$$A = (-A_1, -\tfrac{1}{2}A_2, -\tfrac{1}{4}A_4, -\tfrac{1}{8}A_8, -\tfrac{1}{12}A_{12}, -\tfrac{1}{20}A_{20}, -\tfrac{1}{28}A_{28}, ..., -\tfrac{1}{120}A_{120})' \in \mathbb{R}^{12}$$

$$B = (-B_1, -\tfrac{1}{2}B_2, -\tfrac{1}{4}B_4, -\tfrac{1}{8}B_8, -\tfrac{1}{12}B_{12}, -\tfrac{1}{20}B_{20}, -\tfrac{1}{28}B_{28}, ..., -\tfrac{1}{120}B_{120}) \in \mathcal{M}_{3\times12}(\mathbb{R})$$

$$\Sigma_\varsigma \in \mathcal{M}_{12\times12}(\mathbb{R}) \text{ with } \Sigma_\varsigma = diag(\sigma_\varsigma, ..., \sigma_\varsigma),$$

$\varsigma_t = (\varsigma_{1,t}, \varsigma_{2,t}, \varsigma_{4,t}, \varsigma_{8,t}, \varsigma_{12,t}, \varsigma_{20,t}, \varsigma_{28,t}, \varsigma_{40,t}, \varsigma_{60,t}, \varsigma_{80,t}, \varsigma_{100,t}, \varsigma_{120,t})' \in \mathbb{R}^{12}$
with $\varsigma_{n,t} \sim \mathcal{N}(0,1)$ for $n \in \{1, 2, 4, 8, 12, 20, 28, 40, 60, 80, 100, 120\}$

Moreover, parameters $\delta_0$ and $\delta_1$ from Equation (1) are such that:

$$\delta_0 = 0, \ \delta_1 = (1, 1, 1)'$$

While parameters from the risk-neutral probability measure of Equation (3), $\mu^{\mathbb{Q}}$ and $\Theta^{\mathbb{Q}}$ are such that:

$$\mu^{\mathbb{Q}} = (\mu_1^{\mathbb{Q}}, 0, 0)'$$

$\Theta^{\mathbb{Q}}$ is in ordered real Jordan form:[34]

$$\Theta^{\mathbb{Q}} = \begin{pmatrix} \theta_1^{\mathbb{Q}} & \rho_{12} & 0 \\ 0 & \theta_2^{\mathbb{Q}} & \rho_{23} \\ 0 & 0 & \theta_3^{\mathbb{Q}} \end{pmatrix} \tag{A.86}$$

where $\theta_1^{\mathbb{Q}}, \theta_2^{\mathbb{Q}}, \theta_3^{\mathbb{Q}}$ are real eigenvalues and $|\theta_1^{\mathbb{Q}}| > |\theta_2^{\mathbb{Q}}| > |\theta_3^{\mathbb{Q}}|$.
$\rho_{12} = 1$ if $\theta_1^{\mathbb{Q}} = \theta_2^{\mathbb{Q}}$ and 0 otherwise. $\rho_{23} = 1$ if $\theta_2^{\mathbb{Q}} = \theta_3^{\mathbb{Q}}$ and 0 otherwise.

Finally, parameters $\lambda_0$ and $\lambda_1$ from Equation (6) are left unrestricted.[35]

I now stack all the parameters to be estimated in a vector $\Omega \in \mathbb{R}^{23}$:

$$\Omega = (\mu_1^{\mathbb{Q}}, \mu_1^{\mathbb{P}}, \mu_2^{\mathbb{P}}, \mu_3^{\mathbb{P}}, \theta_1^{\mathbb{Q}}, \theta_2^{\mathbb{Q}}, \theta_3^{\mathbb{Q}}, \theta_{11}^{\mathbb{P}}, \theta_{12}^{\mathbb{P}}, \theta_{13}^{\mathbb{P}}, \theta_{21}^{\mathbb{P}}, \theta_{22}^{\mathbb{P}}, \theta_{23}^{\mathbb{P}}, \theta_{31}^{\mathbb{P}}, \theta_{32}^{\mathbb{P}}, \theta_{33}^{\mathbb{P}}, ...$$
$$\sigma_{\varepsilon 11}, \sigma_{\varepsilon 21}, \sigma_{\varepsilon 22}, \sigma_{\varepsilon 31}, \sigma_{\varepsilon 32}, \sigma_{\varepsilon 33}, \sigma_\varsigma)' \tag{A.87}$$

Under the Gaussian assumption, the distribution of $Y_t^\circ$ conditional on the sequence of observations $\mathcal{Y}_{t-1}^\circ = (Y_0^\circ, Y_1^\circ, ..., Y_{t-1}^\circ)'$ is multivariate normal such that:

$$Y_t^\circ \mid \mathcal{Y}_{t-1}^\circ \sim \mathcal{N}((A + B'\hat{X}_{t|t-1}), F_t) \tag{A.88}$$

where $F_t = B'P_{t|t-1}B + \Sigma_\varsigma \Sigma_\varsigma'$

with $P_{t|t-1}$ is the Mean Squared Error (MSE) matrix:

$$P_{t|t-1} = E_{t-1}[(X_t - \hat{X}_{t|t-1})(X_t - \hat{X}_{t|t-1})'] \tag{A.89}$$

In this framework, the conditional density of $Y_t$ can be written as:

$$p(Y_t^\circ \mid \mathcal{Y}_{t-1}^\circ; \Omega) = \frac{exp[-\tfrac{1}{2}(Y_t^\circ - A - B'\hat{X}_{t|t-1})'F_t^{-1}(Y_t^\circ - A - B'\hat{X}_{t|t-1})]}{(2\pi)^{J/2} \mid F_t \mid^{1/2}} \tag{A.90}$$

---

[34]See Joslin, Singleton & Zhu (2011), Appendix C
[35]Still, they remain restricted by Equations (4) and (5).

with $J = 12$ in my case.

Hence, the sample log-likelihood function is:

$$\mathcal{L}(Y_t^\circ, \Omega) = \sum_{t=1}^{T} log[p(Y_t^\circ \mid \mathcal{Y}_{t-1}^\circ; \Omega)] \tag{A.91}$$

$$\mathcal{L}(Y_t^\circ, \Omega) = -\frac{JT}{2}log(2\pi) - \frac{1}{2}\sum_{t=1}^{T} log \mid F_t \mid -...$$

$$\frac{1}{2}\sum_{t=1}^{T}(Y_t^\circ - A - B'\hat{X}_{t|t-1})'F_t^{-1}(Y_t^\circ - A - B'\hat{X}_{t|t-1}) \tag{A.92}$$

Let us now specify the Kalman filter algorithm:

*I start at $i = 0$, where $i$ is the number of iterations.*

**Step 0: Setup**
$\Omega = \Omega_0$
$i = i + 1$

**Step 1: Initialization**
Set $\hat{X}_{0|0} = \hat{X}_0$ and $P_{0|0} = P_0$
t=1

**Step 2: Prediction $1 \leq t \leq T$**
1) $\hat{X}_{t|t-1} = \mu^{\mathbb{P}} + \Theta^{\mathbb{P}}\hat{X}_{t-1|t-1}$
2) $P_{t|t-1} = \Theta^{\mathbb{P}}P_{t-1|t-1}\Theta^{\mathbb{P}'} + \Sigma_\varepsilon\Sigma_\varepsilon'$
3) $\hat{Y}_{t|t-1} = A + B'\hat{X}_{t|t-1}$
4) $F_t = B'P_{t|t-1}B + \Sigma_\varsigma\Sigma_\varsigma'$

**Step 3: Updating at $t$**
1) $K_t = P_{t|t-1}BF_t^{-1}$
2) $\hat{X}_{t|t} = \hat{X}_{t|t-1} + K_t(Y_t^\circ - \hat{Y}_{t|t-1})$
3) $P_{t|t} = P_{t|t-1} - K_tB'P_{t|t-1}$

**Step 4: Loop**
If $t < T$ then $t = t + 1$ and go to **Step 2**, else go to **Step 5**

**Step 5: Compute the log-likelihood**
Compute the log-likelihood $\mathcal{L}_i(Y_t^\circ, \Omega_0)$ using Equation (A.92)

**Step 6: Compare the log-likelihood**
If $\mid \mathcal{L}_i(Y_t^\circ, \Omega_0) - \mathcal{L}_{i-1}(Y_t^\circ, \Omega_0) \mid \leq \epsilon$, where $\epsilon = 10^{-8}$ is the convergence criterion, then stop. Else go to **Step 0**

Starting values for a stationary process $X_t$ are the unconditional mean and variances:[36]

---

[36]See Hamilton (1994), Chapter 13 page 378

$$\hat{X}_{0|0} = (I_N - \Theta^{\mathbb{P}})^{-1}\mu^{\mathbb{P}} \text{ and } P_{0|0} = vec^{-1}([I_{3\times3} - (\Theta^{\mathbb{P}} \otimes \Theta^{\mathbb{P}})]^{-1}vec(\Sigma_\varepsilon \Sigma'_\varepsilon))$$

I employ the *Quasi-Newton* optimization algorithm included in the function "fminunc" available in MatLab Optimization Toolbox.[37]

### A.4.2 Extended Kalman Filter for the SRTSM

The procedure presented here is similar to the one detailed in Wu & Xia (2016).[38]

In the Shadow Rate Term Structure framework of **Section 2.2**, my state-space model is composed of a linear transition Equation (2) and a non-linear measurement Equation (17), because of the non-linear nature of $y_{n,t}$ in Equation (12).

In particular, transition Equation (A.84) from **Appendix A.4.1** remains valid, but the measurement Equation is now:

$$Y_t^\circ = f(X_t) + \Sigma_\varsigma \varsigma_t \tag{A.93}$$

where $Y_t^\circ = (y_{1,t}^\circ, y_{2,t}^\circ, y_{4,t}^\circ, y_{8,t}^\circ, y_{12,t}^\circ, y_{20,t}^\circ, y_{28,t}^\circ, y_{40,t}^\circ, y_{60,t}^\circ, y_{80,t}^\circ, y_{100,t}^\circ, y_{120,t}^\circ)' \in \mathbb{R}^{12}$

$f$ is the non-linear function that links the $n$-period nominal zero-coupon yield to the state vector as detailed in Equation (12).

$\Sigma_\varsigma \in \mathcal{M}_{12\times12}(\mathbb{R})$ with $\Sigma_\varsigma = diag(\sigma_\varsigma, ..., \sigma_\varsigma)$ and

$\varsigma_t = (\varsigma_{1,t}, \varsigma_{2,t}, \varsigma_{4,t}, \varsigma_{8,t}, \varsigma_{12,t}, \varsigma_{20,t}, \varsigma_{28,t}, \varsigma_{40,t}, \varsigma_{60,t}, \varsigma_{80,t}, \varsigma_{100,t}, \varsigma_{120,t})' \in \mathbb{R}^{12}$
with $\varsigma_{n,t} \sim \mathcal{N}(0,1)$ for $n \in \{1, 2, 4, 8, 12, 20, 28, 40, 60, 80, 100, 120\}$

Furthermore, following the same parametrization as in the GATSM (**Appendix A.4.1**), let us stack all the parameters to estimate in a vector $\Omega \in \mathbb{R}^{23}$ (Equation (A.87)).

Under the Gaussian assumption, the distribution of $Y_t^\circ$ conditional on the sequence of observations $\mathcal{Y}_{t-1}^\circ = (Y_0^\circ, Y_1^\circ, ..., Y_{t-1}^\circ)'$ is also multivariate normal such that:

$$Y_t^\circ \mid \mathcal{Y}_{t-1}^\circ \sim \mathcal{N}(f(\hat{X}_{t|t-1}), F_t) \tag{A.94}$$

where $P_{t|t-1}$ is the Mean Squared Error (MSE) matrix as defined in Equation (A.89),
$F_t = H_t P_{t|t-1} H_t' + \Sigma_\varsigma \Sigma'_\varsigma$

The matrix $H_t \in \mathcal{M}_{3x32}(\mathbb{R})$ is the Jacobian of the yield function $Y_t^\circ$ from Equation (A.93) with respect to the state vector $X_t$:

$$H_t' = \left.\frac{\partial Y_t^\circ}{\partial X_t}\right|_{X_t = \hat{X}_{t|t-1}} \tag{A.95}$$

---

[37]See Lütkepohl (2005), Section 12.3.2 or Pawitan (2001), Section 12 for a discussion concerning the optimization algorithms.

[38]See Krippner (2015), Section 4.2, Simon (2006), Section 13.2.3 or Grewal & Andrews (2008), Section 7.2.3.

$$H_t' = \begin{pmatrix} \partial y_{1,t}^\circ/\partial x_{1,t} & \partial y_{1,t}^\circ/\partial x_{2,t} & \partial y_{1,t}^\circ/\partial x_{3,t} \\ \partial y_{2,t}^\circ/\partial x_{1,t} & \partial y_{2,t}^\circ/\partial x_{2,t} & \partial y_{2,t}^\circ/\partial x_{3,t} \\ \partial y_{4,t}^\circ/\partial x_{1,t} & \partial y_{4,t}^\circ/\partial x_{2,t} & \partial y_{4,t}^\circ/\partial x_{3,t} \\ \partial y_{8,t}^\circ/\partial x_{1,t} & \partial y_{8,t}^\circ/\partial x_{2,t} & \partial y_{8,t}^\circ/\partial x_{3,t} \\ \partial y_{12,t}^\circ/\partial x_{1,t} & \partial y_{12,t}^\circ/\partial x_{2,t} & \partial y_{12,t}^\circ/\partial x_{3,t} \\ \partial y_{20,t}^\circ/\partial x_{1,t} & \partial y_{20,t}^\circ/\partial x_{2,t} & \partial y_{20,t}^\circ/\partial x_{3,t} \\ \partial y_{28,t}^\circ/\partial x_{1,t} & \partial y_{28,t}^\circ/\partial x_{2,t} & \partial y_{28,t}^\circ/\partial x_{3,t} \\ \partial y_{40,t}^\circ/\partial x_{1,t} & \partial y_{40,t}^\circ/\partial x_{2,t} & \partial y_{40,t}^\circ/\partial x_{3,t} \\ \partial y_{60,t}^\circ/\partial x_{1,t} & \partial y_{60,t}^\circ/\partial x_{2,t} & \partial y_{60,t}^\circ/\partial x_{3,t} \\ \partial y_{80,t}^\circ/\partial x_{1,t} & \partial y_{80,t}^\circ/\partial x_{2,t} & \partial y_{80,t}^\circ/\partial x_{3,t} \\ \partial y_{100,t}^\circ/\partial x_{1,t} & \partial y_{100,t}^\circ/\partial x_{2,t} & \partial y_{100,t}^\circ/\partial x_{3,t} \\ \partial y_{120,t}^\circ/\partial x_{1,t} & \partial y_{120,t}^\circ/\partial x_{2,t} & \partial y_{120,t}^\circ/\partial x_{3,t} \end{pmatrix} \tag{A.96}$$

with $X_t = (x_{1,t}, x_{2,t}, x_{3,t})'$. Note that I have dropped the $|_{X_t = \hat{X}_{t|t-1}}$ for convenience.

$H_t = (H_{1,t}, ..., H_{n,t}, ..., H_{120,t}) \in \mathcal{M}_{3\text{x}32}(\mathbb{R})$

The elements of $H_t$ are calculated using Equation (12), for $n \in \{1, 2, 4, 8, 12, 20, 28, 40, 60, 80, 100, 120\}$ and $i \in \{1, 2, 3\}$:

$$\frac{\partial y_{n,t}^\circ}{\partial x_{i,t}} = \frac{1}{n} \frac{\partial \sum_{j=0}^{n-1} f_{j,j+1,t}}{\partial x_{i,t}} \tag{A.97}$$

$$\frac{\partial y_{n,t}^\circ}{\partial x_{i,t}} = \frac{1}{n} \frac{\partial \sum_{j=0}^{n-1} \left( r_{t+j} + \sigma_{s,j}^{\mathbb{Q}} \times g\left( \frac{a_j + b_j' X_t - r_t}{\sigma_{s,j}^{\mathbb{Q}}} \right) \right)}{\partial x_{i,t}} \tag{A.98}$$

Denote $\theta_j = \frac{a_j + b_j' X_t - r_t}{\sigma_{s,j}^{\mathbb{Q}}}$:

$$\frac{\partial y_{n,t}^\circ}{\partial x_{i,t}} = \frac{1}{n} \frac{\partial \sum_{j=0}^{n-1} \sigma_{s,j}^{\mathbb{Q}} \times (\theta_j \Phi(\theta_j) + \phi(\theta_j))}{\partial x_{i,t}} \tag{A.99}$$

$$\frac{\partial y_{n,t}^\circ}{\partial x_{i,t}} = \frac{1}{n} \sum_{j=0}^{n-1} \sigma_{s,j}^{\mathbb{Q}} \frac{\partial (\theta_j \Phi(\theta_j) + \phi(\theta_j))}{\partial x_{i,t}} \tag{A.100}$$

$$\frac{\partial y_{n,t}^\circ}{\partial x_{i,t}} = \frac{1}{n} \sum_{j=0}^{n-1} \sigma_{s,j}^{\mathbb{Q}} \left( \frac{\partial \theta_j}{\partial x_{i,t}} \Phi(\theta_j) + \theta_j \frac{\partial \Phi(\theta_j)}{\partial x_{i,t}} + \frac{\partial \phi(\theta_j)}{\partial x_{i,t}} \right) \tag{A.101}$$

Using the Chain rule, we obtain:

$$\frac{\partial y_{n,t}^\circ}{\partial x_{i,t}} = \frac{1}{n} \sum_{j=0}^{n-1} \sigma_{s,j}^{\mathbb{Q}} \left( \frac{\partial \theta_j}{\partial x_{i,t}} \Phi(\theta_j) + \theta_j \frac{\partial \Phi(\theta_j)}{\partial \theta_j} \times \frac{\partial \theta_j}{\partial x_{i,t}} + \frac{\partial \phi(\theta_j)}{\partial \theta_j} \times \frac{\partial \theta_j}{\partial x_{i,t}} \right) \tag{A.102}$$

$$\frac{\partial y_{n,t}^\circ}{\partial x_{i,t}} = \frac{1}{n} \sum_{j=0}^{n-1} \sigma_{s,j}^{\mathbb{Q}} \left( \frac{b_j'}{\sigma_{s,j}^{\mathbb{Q}}} \Phi(\theta_j) + \theta_j \phi(\theta_j) \times \frac{b_j'}{\sigma_{s,j}^{\mathbb{Q}}} + \frac{\partial \phi(\theta_j)}{\partial \theta_j} \times \frac{b_j'}{\sigma_{s,j}^{\mathbb{Q}}} \right) \tag{A.103}$$

Given that:

$$\frac{\partial \Phi(\theta_j)}{\partial \theta_j} = \phi(\theta_j) = \frac{1}{\sqrt{2\pi}} exp(-\frac{1}{2}\theta_j^2) \tag{A.104}$$

53

Then we have that:

$$\frac{\partial \phi(\theta_j)}{\partial \theta_j} = -\frac{1}{\sqrt{2\pi}} \theta_j \times exp(-\frac{1}{2}\theta_j^2) = -\theta_j \phi(\theta_j) \tag{A.105}$$

Hence Equation (A.103) becomes:

$$\frac{\partial y_{n,t}^\circ}{\partial x_{i,t}} = \frac{1}{n} \sum_{j=0}^{n-1} b_j' \left(\Phi(\theta_j) + \theta_j \phi(\theta_j) - \theta_j \phi(\theta_j)\right) \tag{A.106}$$

$$\frac{\partial y_{n,t}^\circ}{\partial x_{i,t}} = \frac{1}{n} \sum_{j=0}^{n-1} b_j' \Phi\left(\frac{a_j + b_j' X_t - \underline{r}_{t+j}}{\sigma_{s,j}^\mathbb{Q}}\right) \tag{A.107}$$

As the case $j = 0$ is problematic, taking the limit gives:

$$H_{n,t} = \frac{1}{n} \left(\delta_1' + \sum_{j=1}^{n-1} b_j \Phi\left(\frac{a_j + b_j' X_t - \underline{r}_{t+j}}{\sigma_{s,j}^\mathbb{Q}}\right)\right)$$

for $n \in \{1, 2, 4, 8, 12, 20, 28, 40, 60, 80, 100, 120\}$

As in the GATSM case (**Appendix A.4.1**), we can express the sample log-likelihood function:

$$\mathcal{L}(Y_t^\circ, \Omega) = -\frac{JT}{2} log(2\pi) - \frac{1}{2} \sum_{t=1}^{T} log \mid F_t \mid -...$$

$$\frac{1}{2} \sum_{t=1}^{T} (Y_t^\circ - f(\hat{X}_{t|t-1}))' F_t^{-1} (Y_t^\circ - f(\hat{X}_{t|t-1})) \tag{A.108}$$

with $J = 12$ in my case.

Let us now specify the extended Kalman filter algorithm:

**Step 0, 1, 4, 5, 6** are similar to the ones for the Kalman filter (**Appendix A.4.1**). However, **Step 2** and **Step 3** become:

**Step 2: Prediction $1 \leq t \leq T$**
1) $\hat{X}_{t|t-1} = \mu^\mathbb{P} + \Theta^\mathbb{P} \hat{X}_{t-1|t-1}$
2) $P_{t|t-1} = \Theta^\mathbb{P} P_{t-1|t-1} \Theta^{\mathbb{P}'} + \Sigma_\varepsilon \Sigma_\varepsilon'$
3) $\hat{Y}_{t|t-1} = f(\hat{X}_{t|t-1})$
4) $F_t = H_t' P_{t|t-1} H_t + \Sigma_\varsigma \Sigma_\varsigma'$ with $H_t' = \frac{\partial Y_t^\circ}{\partial X_t}\Big|_{X_t = \hat{X}_{t|t-1}}$

**Step 3: Updating at $t$**
1) $K_t = P_{t|t-1} H_t F_t^{-1}$
2) $\hat{X}_{t|t} = \hat{X}_{t|t-1} + K_t(Y_t^\circ - \hat{Y}_{t|t-1})$
3) $P_{t|t} = P_{t|t-1} - K_t H_t' P_{t|t-1}$

## A.5 Results

Table 6: Maximum Likelihood Estimation

| Parameter | SRTSM |
|:---:|:---:|
| $\mu_{1,\mathcal{P}}^{\mathbb{Q}}$ | 0.753 |
| $\mu_{2,\mathcal{P}}^{\mathbb{Q}}$ | -0.713 |
| $\mu_{3,\mathcal{P}}^{\mathbb{Q}}$ | 0.205 |
| $\mu_{1,\mathcal{P}}^{\mathbb{P}}$ | 2.921 |
| $\mu_{2,\mathcal{P}}^{\mathbb{P}}$ | -1.524 |
| $\mu_{3,\mathcal{P}}^{\mathbb{P}}$ | 0.698 |
| $\theta_{11,\mathcal{P}}^{\mathbb{Q}}$ | 1.003 |
| $\theta_{12,\mathcal{P}}^{\mathbb{Q}}$ | -0.127 |
| $\theta_{13,\mathcal{P}}^{\mathbb{Q}}$ | -0.155 |
| $\theta_{21,\mathcal{P}}^{\mathbb{Q}}$ | 0.006 |
| $\theta_{22,\mathcal{P}}^{\mathbb{Q}}$ | 0.931 |
| $\theta_{23,\mathcal{P}}^{\mathbb{Q}}$ | -0.171 |
| $\theta_{31,\mathcal{P}}^{\mathbb{Q}}$ | 0.000 |
| $\theta_{32,\mathcal{P}}^{\mathbb{Q}}$ | 0.002 |
| $\theta_{33,\mathcal{P}}^{\mathbb{Q}}$ | 1.009 |
| $\theta_{11,\mathcal{P}}^{\mathbb{P}}$ | 1.045 |
| $\theta_{12,\mathcal{P}}^{\mathbb{P}}$ | 0.130 |
| $\theta_{13,\mathcal{P}}^{\mathbb{P}}$ | -3.060 |
| $\theta_{21,\mathcal{P}}^{\mathbb{P}}$ | -0.021 |
| $\theta_{22,\mathcal{P}}^{\mathbb{P}}$ | 0.917 |
| $\theta_{23,\mathcal{P}}^{\mathbb{P}}$ | 1.498 |
| $\theta_{31,\mathcal{P}}^{\mathbb{P}}$ | 0.011 |
| $\theta_{32,\mathcal{P}}^{\mathbb{P}}$ | 0.035 |
| $\theta_{33,\mathcal{P}}^{\mathbb{P}}$ | 0.285 |
| $\sigma_{\varepsilon 11,\mathcal{P}}$ | -0.104 |
| $\sigma_{\varepsilon 12,\mathcal{P}}$ | 0.155 |
| $\sigma_{\varepsilon 13,\mathcal{P}}$ | 0.358 |
| $\sigma_{\varepsilon 21,\mathcal{P}}$ | 0.026 |
| $\sigma_{\varepsilon 22,\mathcal{P}}$ | -0.147 |
| $\sigma_{\varepsilon 23,\mathcal{P}}$ | 0.136 |
| $\sigma_{\varepsilon 31,\mathcal{P}}$ | -0.032 |
| $\sigma_{\varepsilon 32,\mathcal{P}}$ | -0.011 |
| $\sigma_{\varepsilon 33,\mathcal{P}}$ | -0.009 |
| $\sigma_{\varsigma}$ | 0.101 |
| $\delta_{0\mathcal{P}}$ | 0.068 |
| $\delta_{11,\mathcal{P}}$ | 0.296 |
| $\delta_{12,\mathcal{P}}$ | 0.429 |
| $\delta_{13,\mathcal{P}}$ | 0.258 |
| **Log-Likelihood** | **1292** |
| **RMSE** | **0.178** |

Table 7: Models Fit

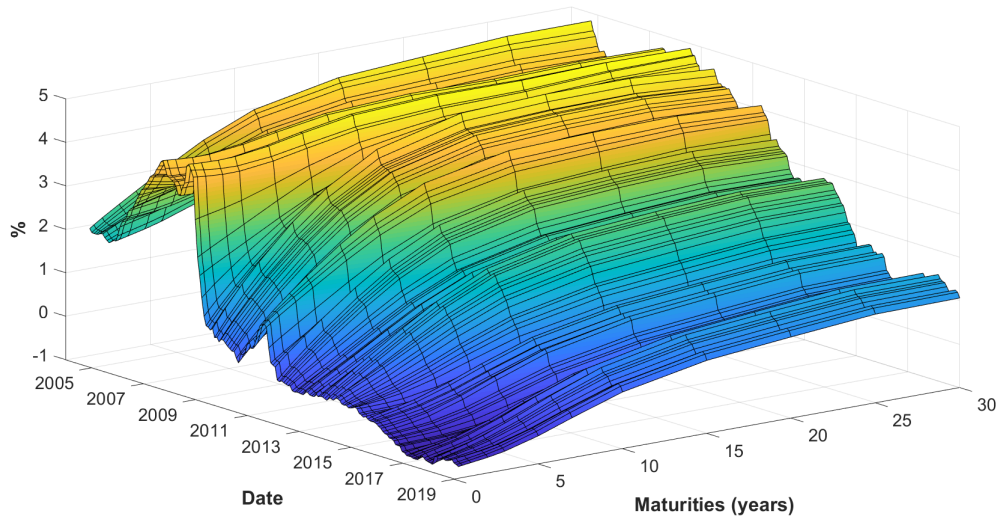| Model | GATSM | SRTSM (A) | SRTSM (B) |
|---|---|---|---|
| **LLH** | 1256 | 1271 | 1292 |
| **RMSE** | 0.214 | 0.179 | 0.178 |

Sample period: 2004M09-2018M12
LLH stands for Log-Likelihood
RMSE stands for Root Mean Square Error
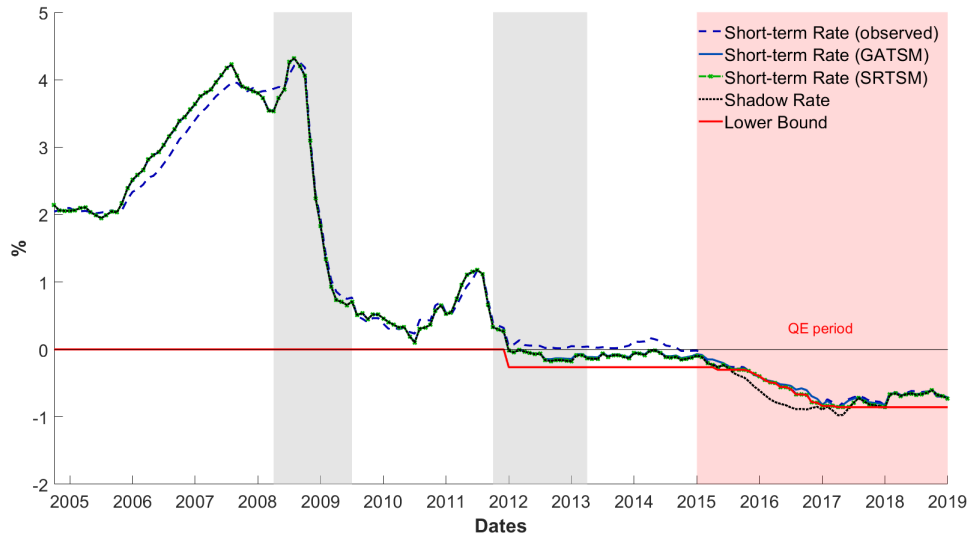
Figure 14: Yield Curves
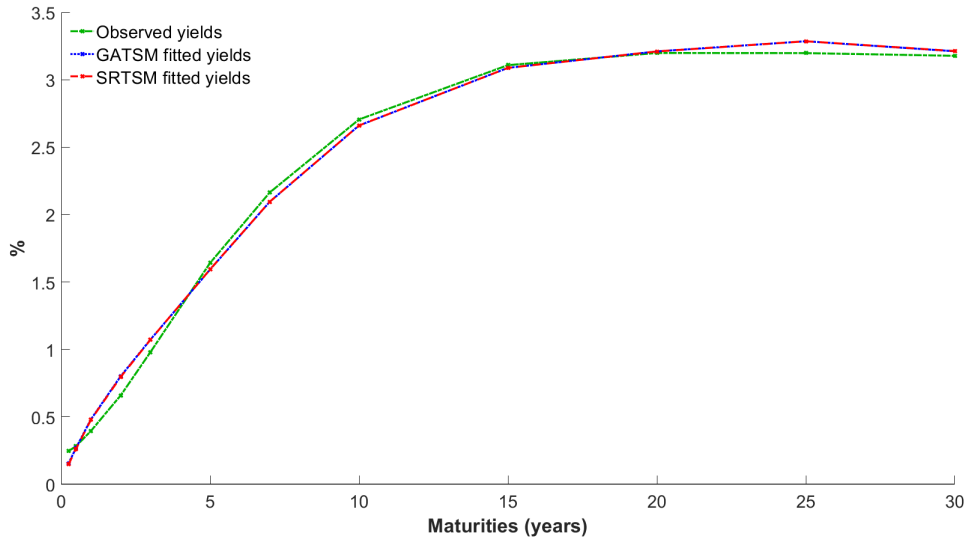


Median yield curves over 2004:M9-2018:M12

Figure 15: Shadow Rate Term Structure Fitted Yield Surface



Fitted Yield surface from 2004:M9 to 2018:M12

Figure 16: Shadow Rate



Shadow rate from 2004:M9 to 2018:M12
Solid grey area corresponds to a recession period as defined by the Center for Economic Policy Research (CEPR)
Solid red area corresponds to the QE period

Figure 17: Factors - SRTSM
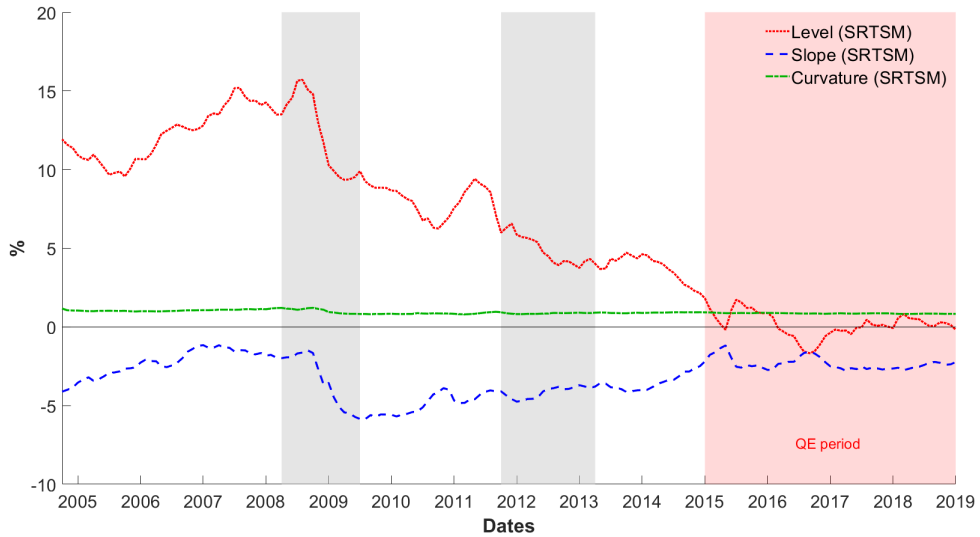


Rotated factors
Solid grey area corresponds to a recession period as defined by the Center for Economic Policy Research (CEPR)
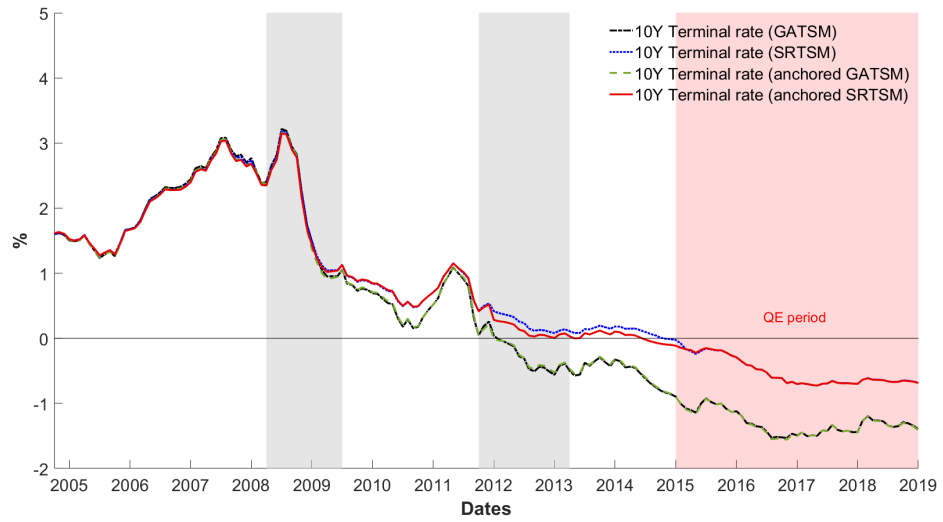Solid red area corresponds to the QE period

Figure 18: Terminal rates



Terminal rates between 2004:M9 and 2018:M12
Solid grey areas correspond to recession periods as defined by the Center for Economic Policy Research (CEPR)
Solid red area corresponds to the QE period

Figure 19: Daily terminal rate during months of key announcement dates
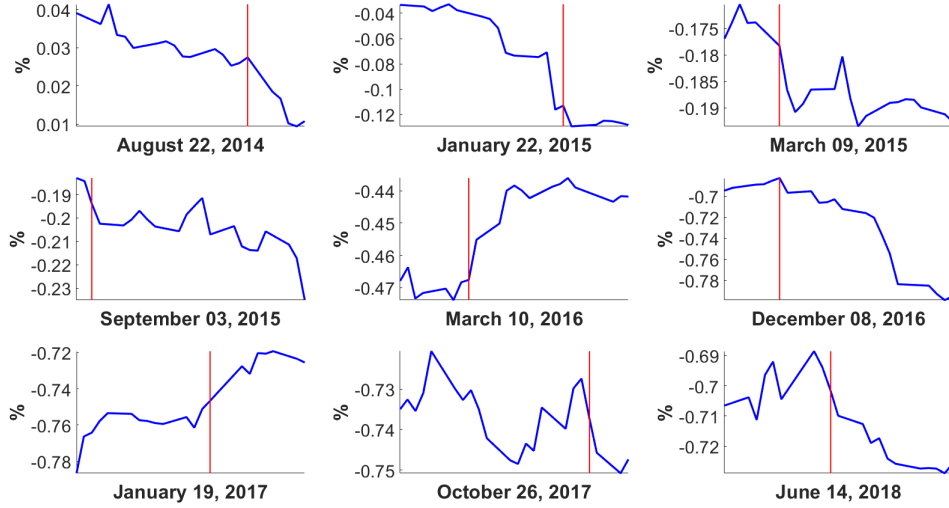


Terminal rate in the month of the major QE-related announcements
Red vertical bar corresponds to announcement

Figure 20: Daily 10-year yield during months of key announcement dates
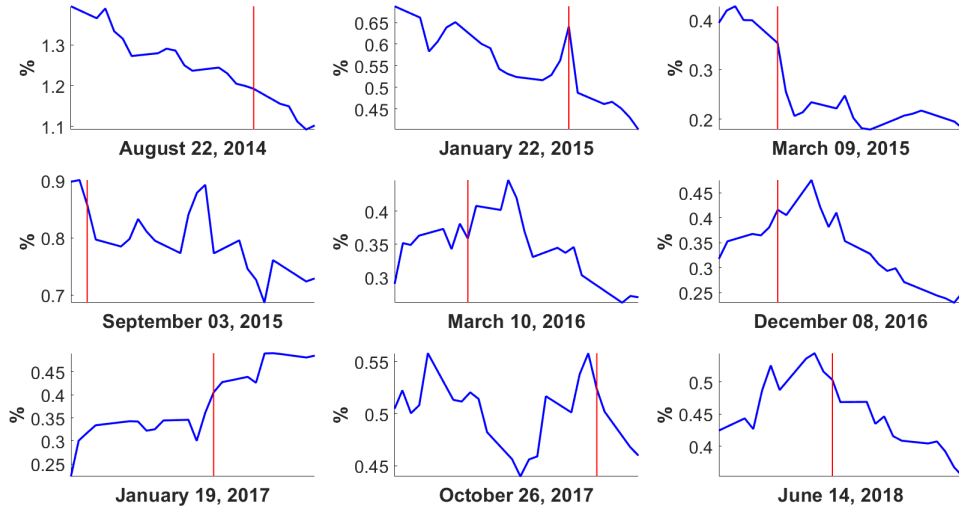


10-year yield in the month of the major QE-related announcements
Red vertical bar corresponds to announcement

Figure 21: Fall in the curves following QE announcement



Figure 22: Fall in the term premium following QE announcement

# B  B-SVAR approach

## B.1  Structural indentification

### B.1.1  Traditional sign restrictions

In the framework of Equations (23), identification is reached by orthogonalizing the shocks to isolate their impact on the endogenous variables. To do so, let us first build $u_t$, a v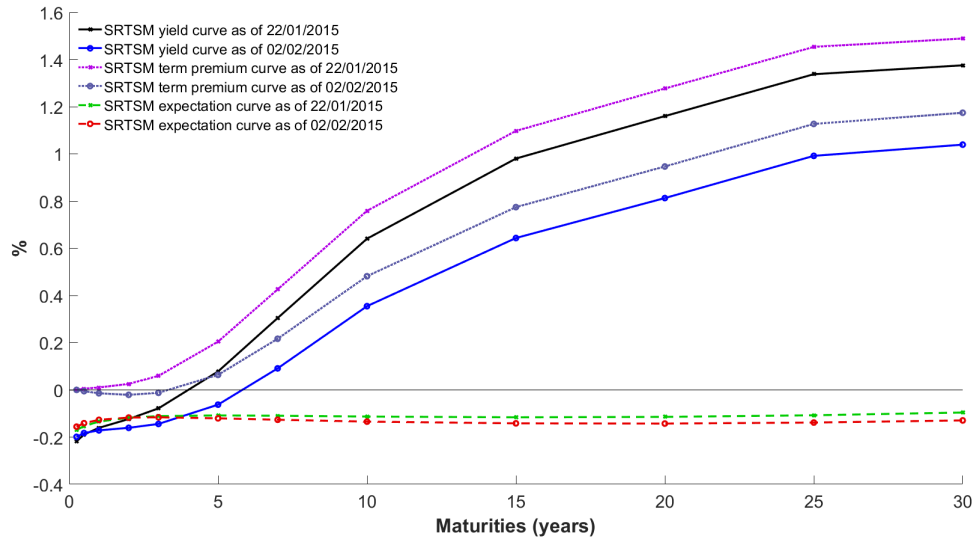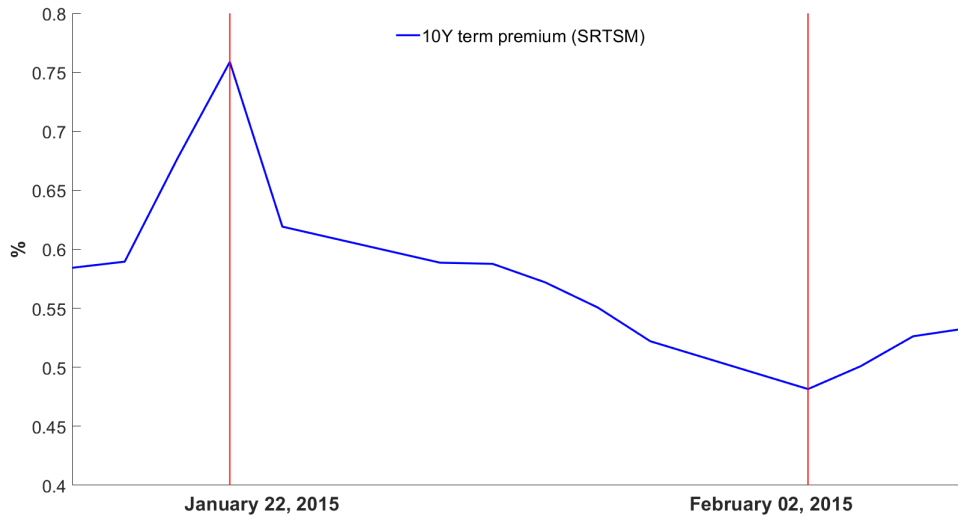ector of uncorrelated (though not standardized) shocks containing the same information as in $\epsilon_t$. Then, let us build $v_t$ so that a one unit shock on $v_{i,t}$ corresponds to a one standard deviation shock on $u_{i,t}$. In the following, I detail the process for the case of an SVAR with three endogenous variables for parsimony, but the procedure remains similar for the more general case of $N$ variables.

Let us first write the Cholesky decomposition of $\Sigma_\epsilon$: $\Sigma_\epsilon = PP' = ADA'$ with
$\epsilon_t = Au_t = AD^{1/2}v_t = Pv_t$, where $A \in \mathcal{M}_{3\mathrm{x}3}(\mathbb{R})$ is a lower triangular matrix with ones on its diagonal, $P \in \mathcal{M}_{3\mathrm{x}3}(\mathbb{R})$ is a lower triangular and non-singular matrix, $D$ is a diagonal matrix such that $Var_t(u_t) = D$ and $v_t$ is the vector of uncorrelated and standardized shocks, such that $\Sigma_v = I_3$.

Therefore:

$$P = \begin{pmatrix} \sqrt{\mathbb{V}(\epsilon_{1,t})} & 0 & 0 \\ a_{21}\sqrt{\mathbb{V}(\epsilon_{1,t})} & \sqrt{\mathbb{V}(\epsilon_{2,t})} & 0 \\ a_{31}\sqrt{\mathbb{V}(\epsilon_{1,t})} & a_{32}\sqrt{\mathbb{V}(\epsilon_{2,t})} & \sqrt{\mathbb{V}(\epsilon_{3,t})} \end{pmatrix} \tag{B.1}$$

where the elements of $P$ are recovered using $u_{1t} = \epsilon_{1t}$ and $u_{jt} = \epsilon_{jt} - \sum_{k=1}^{j-1} a_{jk}u_{kt} \ \forall \ j \in [\![1,3]\!]$, such that $cov(u_{jt}, u_{it}) = 0 \ \forall \ i < j$.

I now define $Q'$, a random orthonormal matrix such that $Q'Q = QQ' = I_3$ and $\epsilon_t = Pv_t = PQQ'v_t = PQw_t$, where $B = PQ$ is called the implied structural impact multiplier matrix with $B \in \mathcal{M}_{3\mathrm{x}3}(\mathbb{R})$ and $w_t$ is the vector of structural shocks from Equation (24). To build the orthogonal matrix $Q$, I employ the Householder transformation approach.[39] That is, I define $W \in \mathcal{M}_{3\mathrm{x}3}(\mathbb{R})$, whose elements are independent and identically distributed (i.i.d) draws from a standard normal distribution. I then use a QR-decomposition by defining $W = QR$ with $R \in \mathcal{M}_{3\mathrm{x}3}(\mathbb{R})$ an upper triangular matrix. If $W$ is invertible and the diagonal elements of $R$ are positive, the decomposition is unique. In practice, I generate the rotation matrix $Q$ until the implied structural impact multiplier matrix $B = PQ$ that satisfies my sign restrictions.

### B.1.2  Zero restrictions

If a zero restriction is added on impact on the Impulse Response Function, the procedure to draw matrix $Q$ is modified. One has to draw the columns of Q recursively following the below algorithm:

**Step 0: initialization**
Set j=1 and gather the impulse response functions in a matrix $\Psi = \widetilde{C}_0$ where $\widetilde{C}_0 = C_0P$ is the orthogonalized IRF at period 0.

**Step 1: Iteration j**
Create a selection matrix $Z_j \in \mathcal{M}_{1\mathrm{x}3\mathrm{p}}(\mathbb{R})$ for $\Psi$ where $Z_j = \varnothing$ if no zero restriction is applied to shock $j$. Then, create a matrix $R_j = (Z_j\Psi, Q'_{j-1})'$ where $Q_{j-1}$ is composed of the columns of $Q$

---

[39]See Kilian & Lütkepohl (2016) for a discussion on the different methods available for identification using sign restrictions.

already drawn with $Q_0 = \varnothing$.

**Step 2: Iteration j**
If $R_j \neq \varnothing$, build a matrix $M_{j-1}$ that is a non-zero orthonormal basis of $R_j$.

**Step 3: Iteration j**
Draw a random vector $y_j$ of size $n = 3$ from a standard normal distribution.

**Step 4: Iteration j**
If $R_j \neq \varnothing$, define column $j$ of Q: $q_j = M'_{j-1}(M_{j-1}y_j/||M'_{j-1}y_j||)$. If $R_j = \varnothing$ then $q_j = y_j/||y_j||$.

**Step 5**
Repeat **Step 1** to **Step 4** until $j = 3$. Then $Q = (q_1, q_2, q_3)$.

Moreover, I now present the way to impose a zero response for $h$ periods of one of the endogenous variables, following a sock. I follow Baumeister & Benati (2013) by *zeroing-out* all coefficients in the equation of that variable, except its own lags.

Especially, in the case of a three-variable SVAR, we have:

$$\begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix}^{-1} \begin{pmatrix} \upsilon_{1,t} \\ \upsilon_{2,t} \\ \upsilon_{3,t} \end{pmatrix} = \sum_{k=1}^{p} \begin{pmatrix} z_{11,k} & z_{12,k} & z_{13,k} \\ z_{21,k} & z_{22,k} & z_{23,k} \\ z_{31,k} & z_{32,k} & z_{33,k} \end{pmatrix} \begin{pmatrix} \upsilon_{1,t-k} \\ \upsilon_{2,t-k} \\ \upsilon_{3,t-k} \end{pmatrix} + \begin{pmatrix} w_t^1 \\ w_t^2 \\ w_t^3 \end{pmatrix} \tag{B.2}$$

The equation for $\upsilon_{3,t}$ is:

$$\upsilon_{3,t} = b_{31}\left( \sum_{k=1}^{p} z_{11,k}\upsilon_{1,t-k} + \sum_{k=1}^{p} z_{12,k}\upsilon_{2,t-k} + \sum_{k=1}^{p} z_{13,k}\upsilon_{3,t-k} + w_t^1 \right)$$

$$+ b_{32}\left( \sum_{k=1}^{p} z_{21,k}\upsilon_{1,t-k} + \sum_{k=1}^{p} z_{22,k}\upsilon_{2,t-k} + \sum_{k=1}^{p} z_{23,k}\upsilon_{3,t-k} + w_t^2 \right)$$

$$+ b_{33}\left( \sum_{k=1}^{p} z_{31,k}\upsilon_{1,t-k} + \sum_{k=1}^{p} z_{32,k}\upsilon_{2,t-k} + \sum_{k=1}^{p} \upsilon_{3,t-k} + w_t^3 \right)$$

If one wants to force the IRF of $\upsilon_{3,t}$ to structural shock $w_t^2$ at zero, I set $b_{31}, b_{32}, z_{13,k}$ and $z_{23,k}$ to zero and compute the IRF using the Wold decomposition with this modified Equation for period 1 to $h$ and using the original Equation from period $h + 1$ on.

### B.1.3   Narrative sign restrictions

In the following, I present the mathematical expressions of the *developed* Narrative sign restrictions that it is possible to impose on the structural shocks and on the historical decompositions. I start by the plain restrictions given in Antolín-Díaz & Rubio-Ramírez (2018) and then provide my developed version.

**Class I** $\oplus$ restriction on structural shock $w_t^3$ at date $t_s$ takes the form:

$$S_{t_s} F(\Theta)e_3 > 0 \tag{B.3}$$

where $e_3$ is the third column of $I_3$

$F \in \mathcal{M}_{3(p+1)\times 3}(\mathbb{R})$ is such that $F = (B^{-1}, Z_1, ..., Z_p)'$ contains the structural parameters from Equation (24). Besides, $S_{t_s} \in \mathcal{M}_{1\times 3(p+1)}(\mathbb{R})$ is a matrix of full row rank such that:

$$S_{t_s} = \begin{pmatrix} v_{1,t_s} & v_{2,t_s} & v_{3,t_s} & \cdots & -v_{1,t_s-p} & -v_{2,t_s-p} & -v_{3,t_s-p} \end{pmatrix} \tag{B.4}$$

**Class I** $\ominus$ restriction is simply $S_{t_s} F(\Theta) e_3 < 0$.

**Class II Type 1** restriction on the historical decomposition of $v_3$ at date $t_s$ takes the form:

$$SF(\Theta)e_3 > 0 \tag{B.5}$$

where $e_3$ is the third column of $I_3$

$F \in \mathcal{M}_{3\times 3}(\mathbb{R})$ is such that:

$$F = \begin{pmatrix} 0 & |H_{v_3,w^2,t_s}| - |H_{v_3,w^1,t_s}| & |H_{v_3,w^3,t_s}| - |H_{v_3,w^1,t_s}| \\ |H_{v_3,w^1,t_s}| - |H_{v_3,w^2,t_s}| & 0 & |H_{v_3,w^3,t_s}| - |H_{v_3,w^2,t_s}| \\ |H_{v_3,w^1,t_s}| - |H_{v_3,w^3,t_s}| & |H_{v_3,w^2,t_s}| - |H_{v_3,w^3,t_s}| & 0 \end{pmatrix} \tag{B.6}$$

where $H_{j,w^i,t_s}$ is defined in Equation (27)

In addition, $S \in \mathcal{M}_{(3-1)\times 3}(\mathbb{R})$ is a matrix of full row rank such that $S = \sum_{j=1}^{3-1} e_{j,3-1} e'_{j,3}$, where $e_{j,i}$ is the $j^{th}$ column of $I_i$

**Class II Type 2** restriction takes the form:

$$SF(\Theta)e_3 > 0 \tag{B.7}$$

where $e_3$ is the third column of $I_3$

$F = (F_1, F_2, F_3) \in \mathcal{M}_{3\times 3}(\mathbb{R})$ is such that:

$$F_1 = \begin{pmatrix} |H_{v_1,w^1,t_s}| - |H_{v_1,w^2,t_s}| - |H_{v_1,w^3,t_s}| \\ |H_{v_2,w^1,t_s}| - |H_{v_2,w^2,t_s}| - |H_{v_2,w^3,t_s}| \\ |H_{v_3,w^1,t_s}| - |H_{v_3,w^2,t_s}| - |H_{v_3,w^3,t_s}| \end{pmatrix}$$

$$F_2 = \begin{pmatrix} |H_{v_1,w^2,t_s}| - |H_{v_1,w^1,t_s}| - |H_{v_1,w^3,t_s}| \\ |H_{v_2,w^2,t_s}| - |H_{v_2,w^1,t_s}| - |H_{v_2,w^3,t_s}| \\ |H_{v_3,w^2,t_s}| - |H_{v_3,w^1,t_s}| - |H_{v_3,w^3,t_s}| \end{pmatrix}$$

$$F_3 = \begin{pmatrix} |H_{v_1,w^3,t_s}| - |H_{v_1,w^2,t_s}| - |H_{v_1,w^1,t_s}| \\ |H_{v_2,w^3,t_s}| - |H_{v_2,w^2,t_s}| - |H_{v_2,w^1,t_s}| \\ |H_{v_3,w^3,t_s}| - |H_{v_3,w^2,t_s}| - |H_{v_3,w^1,t_s}| \end{pmatrix} \tag{B.8}$$

$S = e'_3$ is the transpose of the third column of $I_3$

Now, what follows constitute my *developed* Narrative restrictions:

**Class I Type 1** restriction takes the form:

$$|S_{t_s} F(\Theta)e_3| > |S_{t_s} F(\Theta)e_i| \quad \forall\, i \,\in\, [\![1,2]\!] \tag{B.9}$$

**Class I Type 2** is:

$$|S_{t_s} F(\Theta)e_3| > |S_{t_s} F(\Theta)e_1| + |S_{t_s} F(\Theta)e_2| \tag{B.10}$$

**Class I Category A** restriction is:

$$|S_{t_s}F(\Theta)e_3| > |S_{t_j}F(\Theta)e_3| \ \ \forall\, j \ \in \ [\![1,T]\!] \setminus \{s\} \tag{B.11}$$

With $F$ and $S$ as in Equation (B.3)

Moreover, **Class II** is $\oplus$ if $H_{v_3,w^3,t_s} > 0$ and $\ominus$ if $H_{v_3,w^3,t_s} < 0$.

Finally, **Class II Category A** is expressed by:

$$|H_{v_3,w^3,t_s}| > |H_{v_3,w^3,t_j}| \ \ \forall\, j \ \in \ [\![1,T]\!] \setminus \{s\} \tag{B.12}$$

## B.2  Bayesian Estimation of the VAR

In this subsection, I present the method to conduct a Bayesian estimation of a VAR with Natural Conjugate (Normal-Wishart) priors over the reduced-form parametrization. Calculations are inspired from Canova (2007), Dieppe, Legrand & Van Roye (2016) and Kilian & Lütkepohl (2016).

By the Bayes theorem we have:

$$p_1(\theta \mid \mathcal{Y}_T) = \frac{g(\theta, \mathcal{Y}_T)}{p(\mathcal{Y}_T)} \tag{B.13}$$

And,

$$p_1(\theta \mid \mathcal{Y}_T) = \frac{\mathcal{L}(\theta)p_0(\theta)}{p(\mathcal{Y}_T)} \tag{B.14}$$

where $p_1(\theta \mid \mathcal{Y}_T)$ is the posterior density, $g(\theta, \mathcal{Y}_T)$ is the joint density of the sample and the parameters, $\mathcal{L}(\theta)$ is the likelihood, $p_0(\theta)$ is the prior density and $p(\mathcal{Y}_T)$ is the marginal density of my sample.

We also have that:

$$p(\mathcal{Y}_T) = \int_{-\infty}^{+\infty} \mathcal{L}(\theta)p_0(\theta)d\theta \tag{B.15}$$

To find expressions for the densities, I first re-write Equation (24) in a reduced form:[40]

$$\Upsilon_t = \sum_{k=1}^{p} \Phi_k \Upsilon_{t-k} + \epsilon_t \tag{B.16}$$

where $\epsilon \sim \mathcal{N}(0, \Sigma_\epsilon)$

I now write the stacked form of the model, for $0 \le t \le T$:

$$\Upsilon = Z\Phi + E \tag{B.17}$$

where $\Upsilon \in \mathcal{M}_{T \times 3}(\mathbb{R})$, $Z \in \mathcal{M}_{T \times 3p}(\mathbb{R})$, $\Phi \in \mathcal{M}_{3p \times 3}(\mathbb{R})$ and $E \in \mathcal{M}_{T \times 3}(\mathbb{R})$

Equation (B.17) can also be written:

$$x = (I_3 \otimes Z)\beta + \epsilon \tag{B.18}$$

where $\epsilon \sim \mathcal{N}(0, \Sigma_\epsilon \otimes I_T)$,

---

[40]Let us assume for simplicity that there is no exogenous variables.

$x = (v_{1,1}, ..., v_{1,T}, v_{2,1}, ..., v_{2,T}, v_{3,1}, ..., v_{3,T})'$ a $3 \times T$ vector,

$\beta = vec(\Phi) = vec(\Phi_1, ..., \Phi_p) = (\Phi_{11,1}, ..., \Phi_{33,1}|...|\Phi_{11,p}, ..., \Phi_{33,p})'$ a $3^2 \times p$ vector,

$\epsilon = (\epsilon_{v_{1,1}}, ..., \epsilon_{v_{1,T}}, \epsilon_{v_{2,1}}, ..., \epsilon_{v_{2,T}}, \epsilon_{v_{3,1}}, ..., \epsilon_{v_{3,T}})'$ a $3 \times T$ vector,

And $Z \in \mathcal{M}_{T \times 3p}(\mathbb{R})$ such that:

$$
Z = \begin{pmatrix}
v_{1,0} & v_{2,0} & v_{3,0} & \cdots & v_{1,-p+1} & v_{2,-p+1} & v_{3,-p+1} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
v_{1,T-1} & v_{2,T-1} & v_{3,T-1} & \cdots & v_{1,T-p} & v_{2,T-p} & v_{3,T-p}
\end{pmatrix}
$$

where I assume that the observations begin at $t = -p + 1$ with $p$ the number of lags.

### B.2.1 Normal-Wishart prior

In the above framework, the prior density $p_0(\theta)$ from Equation (B.14) is actually a joint density of the VAR coefficients $\beta$ and $\Sigma_\epsilon$, so that $p_0(\theta) = p_0(\beta, \Sigma_\epsilon)$. As it is not clear what such a joint distribution would be, I follow the standard process that consists in assuming that the parameters are independent so that the joint density becomes the product of the individual densities:

$$p_0(\theta) = p_0(\beta) \times p_0(\Sigma_\epsilon) \tag{B.19}$$

I assume that $\beta$ follows a multivariate Normal distribution:

$$\beta \sim \mathcal{N}(\beta_0, \Sigma_\epsilon \otimes \Omega_0) \tag{B.20}$$

where $\beta_0$ is a $3^2 p$ vector of Minnesota type. That is, elements of $\beta_0$ are 0 except for the first own lag of endogenous variables that are set to $0.8$[41] I give an example of my model in the case of two lags:

$$\beta_0 = (0.8, 0, 0, 0, 0, 0, 0, 0.8, 0, 0, 0, 0, 0, 0, 0.8, 0, 0, 0)'$$

And where $\Omega_0 \in \mathcal{M}_{3p \times 3p}(\mathbb{R})$

$\Omega_0$ also follows a Minnesota prior. It is a diagonal matrix with variance $\sigma^2_{ij,p} = \left(\frac{1}{\sigma_j^2}\right)\left(\frac{\lambda_1}{p^{\lambda_3}}\right)^2$ for own and cross-lags of endogenous variables, where $\sigma_j^2$ the OLS residual variance of individual AR regressions for each endogenous variable, hyperparameter $\lambda_1$ is an overall tightness parameter, $p$ is the lag considered and hyperparameter $\lambda_3$ is a scaling coefficient ensuring an harmonic decay. In the case of two lags, we have:

$$
\Omega_0 = \begin{pmatrix}
\left(\frac{\lambda_1}{\sigma_1}\right)^2 & 0 & 0 & 0 & 0 & 0 \\
0 & \left(\frac{\lambda_1}{\sigma_2}\right)^2 & 0 & 0 & 0 & 0 \\
0 & 0 & \left(\frac{\lambda_1}{\sigma_3}\right)^2 & 0 & 0 & 0 \\
0 & 0 & 0 & \left(\frac{\lambda_1/2^{\lambda_3}}{\sigma_1}\right)^2 & 0 & 0 \\
0 & 0 & 0 & 0 & \left(\frac{\lambda_1/2^{\lambda_3}}{\sigma_2}\right)^2 & 0 \\
0 & 0 & 0 & 0 & 0 & \left(\frac{\lambda_1/2^{\lambda_3}}{\sigma_3}\right)^2
\end{pmatrix}
$$

---

[41] In the original Minnesota prior first own lag of endogenous variables are set to one as they are supposed to have a unit root. Here we remain in a stationary framework.

where I set $\lambda_1 = 0.01$, $\lambda_3 = 1$ and $\Sigma_\epsilon$ is the OLS VAR diagonal variance-covariance matrix.

The kernel for $\beta_0$ is therefore:

$$p_0(\beta) = \propto exp\{-\frac{1}{2}(\beta - \beta_0)'(\Sigma_\epsilon \otimes \Omega_0)^{-1}(\beta - \beta_0)\} \tag{B.21}$$

Secondly, I assume that $\Sigma_\epsilon$ follows an inverse Wishart distribution:

$$\Sigma_\epsilon \sim \mathcal{IW}(S_0, \nu_0) \tag{B.22}$$

where $S_0 \in \mathcal{M}_{3\times3}(\mathbb{R})$ is the prior scale matrix and $\nu_0$ is the prior degrees-of-freedom.

I follow Dieppe, Legrand & Van Roye (2016) in the way they are defined:

$$S_0 = (\nu_0 - 3 - 1)\begin{pmatrix} \sigma_{11} & 0 & 0 \\ 0 & \sigma_{22} & 0 \\ 0 & 0 & \sigma_{33} \end{pmatrix}$$

where $\nu_0 = 3 + 2$ and where $\sigma_{jj}$ is the residual variance of AR model for endogenous variable $j$.

In this framework, the kernel for $\Sigma_\epsilon$ is:

$$p_0(\Sigma_\epsilon) \propto |\Sigma_\epsilon|^{-(\nu_0+3+1)/2} exp\{-\frac{1}{2}tr(\Sigma_\epsilon^{-1}S_0)\} \tag{B.23}$$

### B.2.2   Expression of the Likelihood

The associated likelihood for Gaussian model (B.17) and (B.18):

$$\mathcal{L}(\Phi, \Sigma_E) = \frac{1}{(2\pi)^{3T/2}|\Sigma_E|^{1/2}}exp\left[-\frac{(\Upsilon - Z\Phi)'\Sigma_E^{-1}(\Upsilon - Z\Phi)}{2}\right] \tag{B.24}$$

$$\mathcal{L}(\beta, \Sigma_\epsilon) = \frac{exp[-\frac{1}{2}(\upsilon - (I_3 \otimes Z)\beta)'(\Sigma_\epsilon^{-1} \otimes I_T)(\upsilon - (I_3 \otimes Z)\beta)]}{(2\pi)^{3T/2}|\Sigma_\epsilon \otimes I_T|^{1/2}} \tag{B.25}$$

Focusing on (B.25), I now develop the argument inside the exponential function:

$$(\upsilon - (I_3 \otimes Z)\beta)'(\Sigma_\epsilon^{-1} \otimes I_T)(\upsilon - (I_3 \otimes Z)\beta)$$

$$= (\upsilon - (I_3 \otimes Z)\beta)'(\Sigma_\epsilon^{-1/2} \otimes I_T)(\Sigma_\epsilon^{-1/2} \otimes I_T)(\upsilon - (I_3 \otimes Z)\beta)$$

$$= [(\Sigma_\epsilon^{-1/2} \otimes I_T)'(\upsilon - (I_3 \otimes Z)\beta)]'(\Sigma_\epsilon^{-1/2} \otimes I_T)(\upsilon - (I_3 \otimes Z)\beta) \tag{B.26}$$

As $\Sigma_\epsilon$ is symmetric, so is $(\Sigma_\epsilon^{-1/2} \otimes I_T)$. Therefore, $(\Sigma_\epsilon^{-1/2} \otimes I_T)' = (\Sigma_\epsilon^{-1/2} \otimes I_T)$

Hence,

$$(\upsilon - (I_3 \otimes Z)\beta)'(\Sigma_\epsilon^{-1} \otimes I_T)(\upsilon - (I_3 \otimes Z)\beta)$$

$$= [(\Sigma_\epsilon^{-1/2} \otimes I_T)\upsilon - (\Sigma_\epsilon^{-1/2} \otimes I_T)(I_3 \otimes Z)\beta]'[(\Sigma_\epsilon^{-1/2} \otimes I_T)\upsilon - (\Sigma_\epsilon^{-1/2} \otimes I_T)(I_3 \otimes Z)\beta]$$

$$= [(\Sigma_\epsilon^{-1/2} \otimes I_T)\upsilon - (\Sigma_\epsilon^{-1/2} \otimes Z)\beta]'[(\Sigma_\epsilon^{-1/2} \otimes I_T)\upsilon - (\Sigma_\epsilon^{-1/2} \otimes Z)\beta] \tag{B.27}$$

Since $(\Sigma_\epsilon^{-1/2} \otimes I_T) \in \mathcal{M}_{3T\times3T}(\mathbb{R})$ and $(I_3 \otimes Z) \in \mathcal{M}_{3T\times3*3p}(\mathbb{R})$

Now, let us define $W = \Sigma_\epsilon^{-1/2} \otimes Z$ and $w = (\Sigma_\epsilon^{-1/2} \otimes I_T)\upsilon$

Then,

$$(\upsilon - (I_3 \otimes Z)\beta)'(\Sigma_\epsilon^{-1} \otimes I_T)(\upsilon - (I_3 \otimes Z)\beta) \tag{B.28}$$

$$= (w - W\beta)'(w - W\beta)$$

$$= (\beta - \hat{\beta})'W'W(\beta - \hat{\beta}) + (w - W\hat{\beta})'(w - W\hat{\beta})$$

where

$$\hat{\beta} = (W'W)^{-1}W'w = (\Sigma_\epsilon^{-1} \otimes Z'Z)^{-1}(\Sigma_\epsilon^{-1} \otimes Z)'\upsilon$$

With these notations,

$$(\Sigma_\epsilon^{-1/2} \otimes I_T)\upsilon - (\Sigma_\epsilon^{-1/2} \otimes Z)\beta$$

$$= (\Sigma_\epsilon^{-1/2} \otimes I_T)\upsilon - (\Sigma_\epsilon^{-1/2} \otimes Z)\hat{\beta} + (\Sigma_\epsilon^{-1/2} \otimes Z)(\hat{\beta} - \beta) \tag{B.29}$$

Therefore plugging (B.29) into (B.27) and using (B.2.2) leads to a new expression for (B.28):

$$= [(\Sigma_\epsilon^{-1/2} \otimes I_T)\upsilon - (\Sigma_\epsilon^{-1/2} \otimes Z)\beta]'[(\Sigma_\epsilon^{-1/2} \otimes I_T)\upsilon - (\Sigma_\epsilon^{-1/2} \otimes Z)\hat{\beta}] + \ldots$$

$$(\hat{\beta} - \beta)'(\Sigma_\epsilon^{-1} \otimes Z'Z)(\hat{\beta} - \beta)$$

let us now re-write expression (B.25):

$$\mathcal{L}(\beta, \Sigma_\epsilon) = (2\pi)^{-3T/2}|\Sigma_\epsilon \otimes I_T|^{-1/2}exp[-\frac{1}{2}[(\Sigma_\epsilon^{-1/2} \otimes I_T)\upsilon - (\Sigma_\epsilon^{-1/2} \otimes Z)\hat{\beta}]'[(\Sigma_\epsilon^{-1/2} \otimes I_T)\upsilon - \ldots$$

$$(\Sigma_\epsilon^{-1/2} \otimes Z)\hat{\beta}] - \frac{1}{2}(\hat{\beta} - \beta)'(\Sigma_\epsilon^{-1} \otimes Z'Z)(\hat{\beta} - \beta)] \tag{B.30}$$

let us develop the first part inside the exponential function:

$$\mathcal{L}(\beta, \Sigma_\epsilon) = (2\pi)^{-3T/2}|\Sigma_\epsilon \otimes I_T|^{-1/2}exp\{-\frac{1}{2}(\beta - \hat{\beta})'(\Sigma_\epsilon^{-1} \otimes Z'Z)(\beta - \hat{\beta})\} \times \ldots$$

$$exp\{-\frac{1}{2}[\upsilon'(\Sigma_\epsilon^{-1} \otimes I_T)\upsilon - \upsilon'(\Sigma_\epsilon^{-1} \otimes Z)\hat{\beta} - \hat{\beta}'(\Sigma_\epsilon^{-1} \otimes Z')\upsilon + \hat{\beta}'(\Sigma_\epsilon^{-1} \otimes Z'Z)\hat{\beta}]\} \tag{B.31}$$

But $\hat{\beta}'(\Sigma_\epsilon^{-1} \otimes Z')\upsilon$ is a scalar so it is equal to its transpose:

$$\mathcal{L}(\beta, \Sigma_\epsilon) = (2\pi)^{-3T/2}|\Sigma_\epsilon \otimes I_T|^{-1/2}exp\{-\frac{1}{2}(\beta - \hat{\beta})'(\Sigma_\epsilon^{-1} \otimes Z'Z)(\beta - \hat{\beta})\} \times \ldots$$

$$exp\{-\frac{1}{2}[\upsilon'(\Sigma_\epsilon^{-1} \otimes I_T)\upsilon - 2\hat{\beta}'(\Sigma_\epsilon^{-1} \otimes Z')\upsilon + \hat{\beta}'(\Sigma_\epsilon^{-1} \otimes Z'Z)\hat{\beta}]\} \tag{B.32}$$

$$\mathcal{L}(\beta, \Sigma_\epsilon) = (2\pi)^{-3T/2}|\Sigma_\epsilon \otimes I_T|^{-1/2}exp\{-\frac{1}{2}(\beta - \hat{\beta})'(\Sigma_\epsilon^{-1} \otimes Z'Z)(\beta - \hat{\beta})\} \times \ldots$$

$$exp\{-\frac{1}{2}[\upsilon'(\Sigma_\epsilon^{-1} \otimes I_T)\upsilon - 2\hat{\beta}'(I_3 \otimes Z)'(\Sigma_\epsilon^{-1} \otimes I_T)\upsilon + \hat{\beta}'(\Sigma_\epsilon^{-1} \otimes Z')(I_3 \otimes Z)\hat{\beta}]\} \tag{B.33}$$

$$\mathcal{L}(\beta, \Sigma_\epsilon) = (2\pi)^{-3T/2}|\Sigma_\epsilon \otimes I_T|^{-1/2}exp\{-\frac{1}{2}(\beta - \hat{\beta})'(\Sigma_\epsilon^{-1} \otimes Z'Z)(\beta - \hat{\beta})\} \times \ldots$$

$$exp\{-\frac{1}{2}[\upsilon'(\Sigma_\epsilon \otimes I_T)^{-1}\upsilon - 2[(I_3 \otimes Z)\hat{\beta}]'(\Sigma_\epsilon \otimes I_T)^{-1}\upsilon + \hat{\beta}'(I_3 \otimes Z)'(\Sigma_\epsilon \otimes I_T)^{-1}(I_3 \otimes Z)\hat{\beta}]\}$$

$$\tag{B.34}$$

let us now use the fact that:

$$((vec(X) - vec(M))'(V \otimes U)^{-1}(vec(Y) - vec(N)) = tr[V^{-1}(X - M)'U^{-1}(Y - N)] \quad \text{(B.35)}$$

Hence Equation (B.34) becomes:

$$\mathcal{L}(\beta, \Sigma_\epsilon) = (2\pi)^{-3T/2}|\Sigma_\epsilon \otimes I_T|^{-1/2}exp\{-\frac{1}{2}(\Phi - \hat{\Phi})'(\Sigma_\epsilon^{-1} \otimes Z'Z)(\Phi - \hat{\Phi})\} \times ...$$

$$exp\{-\frac{1}{2}[tr(\Sigma_\epsilon^{-1}\Upsilon'I_T\Upsilon) - 2tr(\Sigma_\epsilon^{-1}(Z\hat{\Phi})'I_T\Upsilon) + tr(\Sigma_\epsilon^{-1}\hat{\Phi}'Z'I_TZ\hat{\Phi})]\} \quad \text{(B.36)}$$

$$\mathcal{L}(\beta, \Sigma_\epsilon) = (2\pi)^{-3T/2}|\Sigma_\epsilon \otimes I_T|^{-1/2}exp\{-\frac{1}{2}(\Phi - \hat{\Phi})'(\Sigma_\epsilon^{-1} \otimes Z'Z)(\Phi - \hat{\Phi})\} \times ...$$

$$exp\{-\frac{1}{2}[tr(\Upsilon\Sigma_\epsilon^{-1}\Upsilon' - 2\Upsilon\Sigma_\epsilon^{-1}\hat{\Phi}'Z' + Z\hat{\Phi}\Sigma_\epsilon^{-1}\hat{\Phi}'Z')]\} \quad \text{(B.37)}$$

$$\mathcal{L}(\beta, \Sigma_\epsilon) = (2\pi)^{-3T/2}|\Sigma_\epsilon \otimes I_T|^{-1/2}exp\{-\frac{1}{2}(\Phi - \hat{\Phi})'(\Sigma_\epsilon^{-1} \otimes Z'Z)(\Phi - \hat{\Phi})\} \times ...$$

$$exp\{-\frac{1}{2}[tr(\Sigma_\epsilon^{-1}(\Upsilon - Z\hat{\Phi})'(\Upsilon - Z\hat{\Phi})]\} \quad \text{(B.38)}$$

Now, as $|A \otimes B| = |A|^m|B|^n$ for $(A, B) \in \mathcal{M}_{n \times n}(\mathbb{R}) \times \mathcal{M}_{m \times m}(\mathbb{R})$

Then, $|\Sigma_\epsilon \otimes I_T|^{-1/2} = |\Sigma_\epsilon|^{T/2} = |\Sigma_\epsilon|^{-3p/2}|\Sigma_\epsilon|^{-(T-3p)/2} = |\Sigma_\epsilon|^{-3p/2}|\Sigma_\epsilon|^{-[(T-3(p+1)-1)+3+1]/2}$

Also, $(\Sigma_\epsilon^{-1} \otimes Z'Z) = [(\Sigma_\epsilon \otimes (Z'Z)^{-1}]^{-1}$

Therefore, Equation (B.38) becomes:

$$\mathcal{L}(\beta, \Sigma_\epsilon) = (2\pi)^{-3T/2}|\Sigma_\epsilon|^{-3p/2}exp\{-\frac{1}{2}(\Phi - \hat{\Phi})'[(\Sigma_\epsilon \otimes (Z'Z)^{-1}]^{-1}(\Phi - \hat{\Phi})\} \times ...$$

$$|\Sigma_\epsilon|^{-[(T-3(p+1)-1)+3+1]/2}exp\{-\frac{1}{2}[tr(\Sigma_\epsilon^{-1}(\Upsilon - Z\hat{\Phi})'(\Upsilon - Z\hat{\Phi})]\} \quad \text{(B.39)}$$

$$\mathcal{L}(\beta, \Sigma_\epsilon) = (2\pi)^{-3T/2}(2\pi)^{3p/2}|(Z'Z)|^{-3/2}\mathcal{MVN}(\hat{\Phi}, \Sigma_\epsilon \otimes (Z'Z)^{-1}) \times 2^{3\nu/2}|\hat{S}|^{-\nu/2}\Gamma_3\left(\frac{\nu}{2}\right)\mathcal{IW}(\hat{S}, \nu)$$

$$\text{(B.40)}$$

where $\Gamma$ is the multivariate Gamma function

Thus, expression (B.25) can be expressed in terms of a multivariate Normal distribution and an inverse Wishart distribution:

$$\mathcal{L}(\beta, \Sigma_\epsilon) \propto |\Sigma_\epsilon|^{-3p/2}exp\{-\frac{1}{2}(\Phi - \hat{\Phi})'[\Sigma_\epsilon \otimes (Z'Z)^{-1}]^{-1}(\Phi - \hat{\Phi})\} \times ...$$

$$|\Sigma_\epsilon|^{-[(T-3(p+1)-1)+3+1]/2}exp\{-\frac{1}{2}[tr(\Sigma_\epsilon^{-1}(\Upsilon - Z\hat{\Phi})'(\Upsilon - Z\hat{\Phi})]\} \quad \text{(B.41)}$$

$$\mathcal{L}(\beta, \Sigma_\epsilon) \propto \mathcal{MVN}(\hat{\Phi}, \Sigma_\epsilon \otimes (Z'Z)^{-1}) \times \mathcal{IW}(\hat{S}, \nu) \quad \text{(B.42)}$$

where $\hat{S} = (\Upsilon - Z\hat{\Phi})'(\Upsilon - Z\hat{\Phi})$ and $\nu = T - 3(p+1) - 1$

### B.2.3 Posterior estimates

let us start by re-writing Equation (B.41):

$$\mathcal{L}(\beta, \Sigma_\epsilon) \propto |\Sigma_\epsilon|^{-3p/2} exp\{-\frac{1}{2}(\beta - \hat{\beta})'[\Sigma_\epsilon \otimes Z'Z^{-1}]^{-1}(\beta - \hat{\beta})\} \times ...$$
$$exp\{-\frac{1}{2}[tr(\Sigma_\epsilon^{-1}(\Upsilon - Z\hat{\Phi})'(\Upsilon - Z\hat{\Phi})]\} \quad \text{(B.43)}$$

Now, let us plug Equations (B.21), (B.23) and (B.43) into (B.14) gives the joint posterior distribution for $\beta$ and $\Sigma_\epsilon$:

$$p_1(\beta, \Sigma_\epsilon \mid \mathcal{Y}_T) \propto |\Sigma_\epsilon|^{-3p/2} exp\{-\frac{1}{2}(\beta - \beta_0)'(\Sigma_\epsilon \otimes \Omega_0)^{-1}(\beta - \beta_0)\} \times ...$$
$$|\Sigma_\epsilon|^{-(\nu_0+3+1)/2} exp\{-\frac{1}{2}tr(\Sigma_\epsilon^{-1}S_0)\} \times ...$$
$$exp\{-\frac{1}{2}(\beta - \hat{\beta})'[\Sigma_\epsilon \otimes (Z'Z)^{-1}]^{-1}(\beta - \hat{\beta})\} \times ...$$
$$exp\{-\frac{1}{2}[tr(\Sigma_\epsilon^{-1}(\Upsilon - Z\hat{\Phi})'(\Upsilon - Z\hat{\Phi})]\} \quad \text{(B.44)}$$

$$p_1(\beta, \Sigma_\epsilon \mid \mathcal{Y}_T) \propto ...$$
$$|\Sigma_\epsilon|^{-(T+3p+\nu_0+3+1)/2} exp\{-\frac{1}{2}[(\beta-\hat{\beta})'[\Sigma_\epsilon \otimes (Z'Z)^{-1}]^{-1}(\beta-\hat{\beta})) + (\beta-\beta_0)(\Sigma_\epsilon \otimes \Omega_0)^{-1}(\beta-\beta_0)]\} \times ...$$
$$exp\{-\frac{1}{2}tr[\Sigma_\epsilon^{-1}[S_0 + (\Upsilon - Z\hat{\Phi})'(\Upsilon - Z\hat{\Phi})]\} \quad \text{(B.45)}$$

let us develop the argument inside the first exponential:

$$(\beta - \hat{\beta})'(\Sigma_\epsilon^{-1} \otimes Z'Z)(\beta - \hat{\beta})) + (\beta - \beta_0)(\Sigma_\epsilon \otimes \Omega_0)^{-1}(\beta - \beta_0) = ...$$
$$\beta'(\Sigma_\epsilon^{-1} \otimes Z'Z)\beta + \hat{\beta}'(\Sigma_\epsilon^{-1} \otimes Z'Z)\hat{\beta} - 2\beta'(\Sigma_\epsilon^{-1} \otimes Z'Z)\hat{\beta} + ...$$
$$\beta'(\Sigma_\epsilon \otimes \Omega_0)^{-1}\beta + \beta_0'(\Sigma_\epsilon \otimes \Omega_0)^{-1}\beta_0 - 2\beta'(\Sigma_\epsilon \otimes \Omega_0)^{-1}\beta_0 \quad \text{(B.46)}$$

Re-arranging, we obtain:

$$(\beta - \hat{\beta})'(\Sigma_\epsilon^{-1} \otimes Z'Z)(\beta - \hat{\beta})) + (\beta - \beta_0)(\Sigma_\epsilon \otimes \Omega_0)^{-1}(\beta - \beta_0)$$
$$= \beta'[(\Sigma_\epsilon \otimes \Omega_0)^{-1} + \Sigma_\epsilon^{-1} \otimes Z'Z]\beta - 2\beta'[(\Sigma_\epsilon \otimes \Omega_0)^{-1}\beta_0 + (\Sigma_\epsilon^{-1} \otimes Z'Z)\hat{\beta}] + \hat{\beta}'(\Sigma_\epsilon^{-1} \otimes Z'Z)\hat{\beta} + \beta_0'(\Sigma_\epsilon \otimes \Omega_0)^{-1}\beta_0$$
$$= \beta'\Sigma_\epsilon^{-1} \otimes (\Omega_0^{-1} + Z'Z)\beta - 2\beta'\Sigma_\epsilon^{-1} \otimes (\Omega_0^{-1}\beta_0 + Z'Z\hat{\beta}) + \hat{\beta}'(\Sigma_\epsilon^{-1} \otimes Z'Z)\hat{\beta} + \beta_0'(\Sigma_\epsilon \otimes \Omega_0)^{-1}\beta_0$$
$$\text{(B.47)}$$

Where I define $\bar{\Omega} = [\Omega_0^{-1} + Z'Z]^{-1}$ and $\bar{\beta} = \bar{\Omega}[\Omega_0^{-1}\beta_0 + Z'Z\hat{\beta}]$

Then,

$$(\beta - \hat{\beta})'(\Sigma_\epsilon^{-1} \otimes Z'Z)(\beta - \hat{\beta})) + (\beta - \beta_0)(\Sigma_\epsilon \otimes \Omega_0)^{-1}(\beta - \beta_0)$$
$$= \beta'\Sigma_\epsilon^{-1} \otimes \bar{\Omega}^{-1}\beta - 2\beta'\Sigma_\epsilon^{-1} \otimes \bar{\Omega}^{-1}\bar{\beta} + \hat{\beta}'(\Sigma_\epsilon^{-1} \otimes Z'Z)\hat{\beta} + \beta_0'(\Sigma_\epsilon \otimes \Omega_0)^{-1}\beta_0 \quad \text{(B.48)}$$

$$= (\beta - \bar{\beta})'(\Sigma_\epsilon \otimes \bar{\Omega})^{-1}(\beta - \bar{\beta}) - \bar{\beta}(\Sigma_\epsilon \otimes \bar{\Omega})^{-1}\bar{\beta} + \hat{\beta}'(\Sigma_\epsilon^{-1} \otimes Z'Z)\hat{\beta} + \beta_0'(\Sigma_\epsilon \otimes \Omega_0)^{-1}\beta_0 \quad \text{(B.49)}$$

70

Now, let us plug Equation (B.49) into Equation (B.45):

$$p_1(\beta, \Sigma_\epsilon \mid \mathcal{Y}_T) \propto |\Sigma_\epsilon|^{-3p/2} exp\{-\frac{1}{2}(\beta - \bar{\beta})'(\Sigma_\epsilon \otimes \bar{\Omega})^{-1}(\beta - \bar{\beta})\} \times ...$$

$$|\Sigma_\epsilon|^{-(T+\nu_0+3+1)/2} exp\{-\frac{1}{2}[-\bar{\beta}'(\Sigma_\epsilon \otimes \bar{\Omega})^{-1}\bar{\beta} + \hat{\beta}'(\Sigma_\epsilon^{-1} \otimes Z'Z)\hat{\beta} + \beta_0'(\Sigma_\epsilon \otimes \Omega_0)^{-1}\beta_0 + \quad \text{(B.50)}$$

$$tr(\Sigma_\epsilon^{-1}[S_0 + (\Upsilon - Z\hat{\Phi})'(\Upsilon - Z\hat{\Phi}))]\} \quad \text{(B.51)}$$

$$p_1(\beta, \Sigma_\epsilon \mid \mathcal{Y}_T) \propto |\Sigma_\epsilon|^{-3p/2} exp\{-\frac{1}{2}(\beta - \bar{\beta})'(\Sigma_\epsilon \otimes \bar{\Omega})^{-1}(\beta - \bar{\beta})\} \times$$

$$...|\Sigma_\epsilon|^{-(T+\nu_0+3+1)/2} exp\{-\frac{1}{2}tr[\Sigma_\epsilon^{-1}(-\bar{\beta}'\bar{\Omega}^{-1}\bar{\beta} + \hat{\beta}'(Z'Z)\hat{\beta} +$$

$$\beta_0'\Omega_0^{-1}\beta_0 + S_0 + (\Upsilon - Z\hat{\Phi})'(\Upsilon - Z\hat{\Phi}))]\} \quad \text{(B.52)}$$

$$p_1(\beta, \Sigma_\epsilon \mid \mathcal{Y}_T) \propto |\Sigma_\epsilon|^{-3p/2} exp\{-\frac{1}{2}(\beta - \bar{\beta})'(\Sigma_\epsilon \otimes \bar{\Omega})^{-1}(\beta - \bar{\beta})\}|\Sigma_\epsilon|^{-(T+\nu_0+3+1)/2} exp\{-\frac{1}{2}tr(\Sigma_\epsilon^{-1}\bar{S})\}$$
$$\text{(B.53)}$$

Where $\bar{S} = \hat{\beta}'(Z'Z)\hat{\beta} + \beta_0'\Omega_0^{-1}\beta_0 - \bar{\beta}'\bar{\Omega}^{-1}\bar{\beta} + S_0 + (\Upsilon - Z\hat{\Phi})'(\Upsilon - Z\hat{\Phi})) = \Upsilon'\Upsilon + \beta_0'\Omega_0^{-1}\beta_0 - \bar{\beta}'\bar{\Omega}^{-1}\bar{\beta} + S_0$

The conditional posterior distribution of $\Sigma_\epsilon$ is given by:

$$p_1(\Sigma_\epsilon \mid \mathcal{Y}_T) \propto \int_\beta p_1(\beta, \Sigma_\epsilon \mid \mathcal{Y}_T)d\beta \quad \text{(B.54)}$$

$$p_1(\Sigma_\epsilon, \mid \mathcal{Y}_T) \propto \int_\beta |\Sigma_\epsilon|^{-3p/2} exp\{-\frac{1}{2}(\beta - \bar{\beta})'(\Sigma_\epsilon \otimes \bar{\Omega})^{-1}(\beta - \bar{\beta})\}|\Sigma_\epsilon|^{-(T+\nu_0+3+1)/2} exp\{-\frac{1}{2}tr(\Sigma_\epsilon^{-1}\bar{S})\}d\beta$$
$$\text{(B.55)}$$

$$p_1(\Sigma_\epsilon \mid \mathcal{Y}_T) \propto |\Sigma_\epsilon|^{-(T+\nu_0+3+1)/2} exp\{-\frac{1}{2}tr(\Sigma_\epsilon^{-1}\bar{S})\} \int_\beta |\Sigma_\epsilon|^{-3p/2} exp\{-\frac{1}{2}(\beta - \bar{\beta})'(\Sigma_\epsilon \otimes \bar{\Omega})^{-1}(\beta - \bar{\beta})\}d\beta$$
$$\text{(B.56)}$$

Recall that

$$\int_\beta |\Sigma_\epsilon|^{-3p/2} exp\{-\frac{1}{2}(\beta - \bar{\beta})'(\Sigma_\epsilon \otimes \bar{\Omega})^{-1}(\beta - \bar{\beta})\}d\beta = 1 \quad \text{(B.57)}$$

Therefore,

$$p_1(\Sigma_\epsilon \mid \mathcal{Y}_T) \propto |\Sigma_\epsilon|^{-(T+\nu_0+3+1)/2} exp\{-\frac{1}{2}tr(\Sigma_\epsilon^{-1}\bar{S})\} \quad \text{(B.58)}$$

We recognize an inverse Wishart distribution:

$$p_1(\Sigma_\epsilon \mid \mathcal{Y}_T) \sim \mathcal{IW}(\bar{S}, \bar{\nu}_1) \quad \text{(B.59)}$$

where $\bar{S} = \Upsilon'\Upsilon + S_0 + \beta_0'\Omega_0^{-1}\beta_0 - \bar{\beta}'\bar{\Omega}^{-1}\bar{\beta}$
and $\bar{\nu}_1 = T + \nu_0$

Now, conditional posterior distribution of $\beta$ is given by:

$$p_1(\beta \mid \mathcal{Y}_T) \propto \int_{\Sigma_\epsilon} p_1(\beta, \Sigma_\epsilon \mid \mathcal{Y}_T)d\Sigma_\epsilon \quad \text{(B.60)}$$

$$p_1(\beta \mid \mathcal{Y}_T) \propto \int_{\Sigma_\epsilon} |\Sigma_\epsilon|^{-3p/2} exp\{-\frac{1}{2}(\beta - \bar{\beta})'(\Sigma_\epsilon \otimes \bar{\Omega})^{-1}(\beta - \bar{\beta})\}|\Sigma_\epsilon|^{-(T+\nu_0+3+1)/2} exp\{-\frac{1}{2}tr(\Sigma_\epsilon^{-1}\bar{S})\}d\Sigma_\epsilon$$
$$\text{(B.61)}$$

$$p_1(\beta \mid \mathcal{Y}_T) \propto \int_{\Sigma_\epsilon} |\Sigma_\epsilon|^{-3p/2} exp\{-\frac{1}{2}(\beta-\bar{\beta})'(\Sigma_\epsilon \otimes \bar{\Omega})^{-1}(\beta-\bar{\beta})\}|\Sigma_\epsilon|^{-(T+\nu_0+3+1)/2} exp\{-\frac{1}{2}tr(\Sigma_\epsilon^{-1}\bar{S})\}d\Sigma_\epsilon$$

<div align="right">(B.62)</div>

Using Equation (B.35), we have

$$p_1(\beta \mid \mathcal{Y}_T) \propto \int_{\Sigma_\epsilon} |\Sigma_\epsilon|^{-(T+3p+\nu_0+3+1)/2} exp\{-\frac{1}{2}tr\Sigma_\epsilon^{-1}[\bar{S}+(\beta-\bar{\beta})'\bar{\Omega}^{-1}(\beta-\bar{\beta})]\}d\Sigma_\epsilon \qquad \text{(B.63)}$$

However,

$$\int_{\Sigma_\epsilon} \frac{|\bar{S}+(\beta-\bar{\beta})'\bar{\Omega}^{-1}(\beta-\bar{\beta})|^{\frac{3p+\bar{\nu}_1}{2}}}{2^{\frac{3(3p+\bar{\nu}_1)}{2}}\Gamma_3\left(\frac{3p+\bar{\nu}_1}{2}\right)}|\Sigma_\epsilon|^{-(T+3p+\nu_0+3+1)/2} exp\{-\frac{1}{2}tr\Sigma_\epsilon^{-1}[\bar{S}+(\beta-\bar{\beta})'\bar{\Omega}^{-1}(\beta-\bar{\beta})]\}d\Sigma_\epsilon = 1$$

(B.64)

It is the integral of an Inverse-Wishart distribution, therefore:

$$p_1(\beta \mid \mathcal{Y}_T) \propto |\bar{S}+(\beta-\bar{\beta})'\bar{\Omega}^{-1}(\beta-\bar{\beta})|^{-\frac{3p+\bar{\nu}_1}{2}} 2^{\frac{3(3p+\bar{\nu}_1)}{2}}\Gamma_3\left(\frac{3p+\bar{\nu}_1}{2}\right) \qquad \text{(B.65)}$$

Getting rid of the constants leads to:

$$p_1(\beta \mid \mathcal{Y}_T) \propto |\bar{S}+(\beta-\bar{\beta})'\bar{\Omega}^{-1}(\beta-\bar{\beta})|^{-\frac{3p+\bar{\nu}_1}{2}} \qquad \text{(B.66)}$$

Since $\bar{\nu}_1 = T + \nu_0$.

$$p_1(\beta \mid \mathcal{Y}_T) \propto |\bar{S}|^{-\frac{3p+T+\bar{\nu}_0}{2}} |I_3+\bar{S}^{-1}(\beta-\bar{\beta})'\bar{\Omega}^{-1}(\beta-\bar{\beta})|^{-\frac{3p+T+\bar{\nu}_0}{2}} \qquad \text{(B.67)}$$

$$p_1(\beta \mid \mathcal{Y}_T) \propto |I_3+\bar{S}^{-1}(\beta-\bar{\beta})'\bar{\Omega}^{-1}(\beta-\bar{\beta})|^{-\frac{3p+T+\bar{\nu}_0}{2}} \qquad \text{(B.68)}$$

$$p_1(\beta \mid \mathcal{Y}_T) \propto |I_3+\bar{S}^{-1}(\beta-\bar{\beta})'\bar{\Omega}^{-1}(\beta-\bar{\beta})|^{\frac{-(T+\nu_0-3+1)+3+3p-1}{2}} \qquad \text{(B.69)}$$

This is the kernel of a Matrix-student distribution:

$$p_1(\beta \mid \mathcal{Y}_T) \propto MT(\bar{\beta},\bar{S},\bar{\Omega},\nu_2) \qquad \text{(B.70)}$$

with $\nu_2 = T + \nu_0 - 3 + 1$

It can be shown that this is equivalent to drawing the parameter vector $\beta$ from the following multivariate Normal distribution:

$$\boxed{p_1(\beta \mid \Sigma_\epsilon, \mathcal{Y}_T) \sim \mathcal{MVN}(\bar{\beta}, \Sigma_\epsilon \otimes \bar{\Omega}) \qquad \text{(B.71)}}$$

Now that we know the form of the posterior distributions of $\beta$ and $\Sigma_\epsilon$, we need to draw from this conjugate Normal inverse-Wishart posterior distribution over the structural parametrization conditional on the zero, traditional and Narrative sign restrictions. This is done thanks to the following the Algorithm from Arias, Rubio-Ramirez & Waggoner (2018).

**B.2.4 Algorithm: draw conditional on zero, traditional and Narrative sign restrictions**

*Start at $i = 1$, where $i$ is the number of iteration.*

**Step 0: Setup**
Define $n$, the number of draws to reach.

**Step 1: Iteration $i$, draw $\Sigma_\epsilon^{(i)}$**
$\Sigma_\epsilon^{(i)} \sim \mathcal{IW}(\bar{S}, \bar{\nu}_1)$ with $\bar{S} = \Upsilon'\Upsilon + S_0 + \beta_0'\Omega_0^{-1}\beta_0 - \bar{\beta}'\bar{\Omega}^{-1}\bar{\beta}$, $\bar{\nu}_1 = T + \nu_0$, $\bar{\Omega} = [\Omega_0^{-1} + Z'Z]^{-1}$ and $\bar{\beta} = \bar{\Omega}[\Omega_0^{-1}\beta_0 + Z'Z\hat{\beta}]$

**Step 2: Iteration $i$, draw $\beta^{(i)}$**
$\beta^{(i)} \sim \mathcal{MVN}(\bar{\beta}, \Sigma_\epsilon \otimes \bar{\Omega})$ $\Phi^{(i)} = vec^{-1}(\beta^{(i)})$

**Step 3: Iteration $i$**
Draw a random orthonormal matrix $Q^i$ satisfying the zero restrictions (see **Appendix B.1.2**)

**Step 4: Transform reduced form parameters into structural parameters**
From Equation (B.16), $\Upsilon_t = \sum_{k=1}^p \Phi_k \Upsilon_{t-k} + \epsilon_t = \sum_{k=1}^p BZ_k \Upsilon_{t-k} + Bw_t$
$B^i = P^i Q^i$ with $P^i$ a lower triangular matrix from of Cholesky decomposition of $\Sigma_\epsilon^{(i)}$

**Step 5: Structural IRF**
Compute the structural impulse response functions: $\tilde{C}_q^{j,i} = C_q^{j,i} B^i$

**Step 6: Check traditional sign restrictions**
If traditional sign restrictions are respected, go to **Step 7**. Otherwise, go back to **Step 0**

**Step 7: Structural shocks and Historical Decomposition**
Compute structural shocks and Historical Decompositions

**Step 8: Check Narrative sign restrictions**
If Narrative sign restrictions are respected, go to **Step 9**. Otherwise, go back to **Step 0**

**Step 9: Loop**
Keep the draw $(\beta^{(i)}, \Sigma_\epsilon^{(i)})$. If $i < n$, $i = i + 1$ and go to Step 1, else stop.

**Step 10: Resample**
Resample the $n$ successful draws using weights from the zero and Narrative sign restrictions procedure, following Arias, Rubio-Ramirez & Waggoner (2018) and Antolín-Díaz & Rubio-Ramírez (2018).

## B.3 Shock-specific Conditional Forecasts

The algorithm below used to draw shock-specific conditional forecasts relies on Dieppe, Legrand & Van Roye (2016).

Conditions of the type $v_{3,t+h} = \bar{v}_3$ are conditions on the structural shocks $w_t$ from Equation (24). These conditions can be gathered in a linear form and stacked in a linear system:

$$Rw = c \tag{B.72}$$

where $R \in \mathcal{M}_{r \times (3 \times h)}(\mathbb{R})$, with $r \leq 3 \times h$ the number of conditions, is the matrix of linear restrictions containing the relevant coefficients of the Wold decomposition $\bar{C}_{t_s}^j = C_{t_s}^j B^i$ for $t_s \in [\![1,h]\!]$, $w = (w_{t+1}^1, w_{t+1}^2, w_{t+1}^3, ..., w_{t+h}^1, w_{t+h}^2, w_{t+h}^3)' \in \mathbb{R}^{3 \times h}$ is the vector of structural shock at each forecasted date and $c = (\bar{v}_{3,t_1} - \widetilde{v}_{3,t+t_1}, ..., \bar{v}_{3,t_r} - \widetilde{v}_{3,t+t_r})' \in \mathbb{R}^r$, with $t_s \in [\![1,h]\!]$ where $s \in [\![1,r]\!]$ and $r \leq h$ is the vector containing the differences between the conditional values and the unconditional forecasts from Equation (28).

$w^3$ shock is constructive for $v_{3,t_s}$, whereas $w^1$ and $w^2$ are not. Besides, at each date where a condition is set on $v_{3,t_s}$, $v_3$ constitute its own block.

For example, to set a condition at date $t = t + t_s$, Equation (B.72) rewrites:

$$\begin{pmatrix} \widetilde{\phi}_{t_s-1,31} & \widetilde{\phi}_{t_s-1,32} & \widetilde{\phi}_{t_s-1,33} & \cdots & \widetilde{\phi}_{0,31} & \widetilde{\phi}_{0,32} & \widetilde{\phi}_{0,33} & 0 & 0 & 0 & \cdots \end{pmatrix} \begin{pmatrix} w_{t+1}^1 \\ w_{t+1}^2 \\ w_{t+1}^3 \\ \vdots \\ w_{t+t_s}^1 \\ w_{t+t_s}^2 \\ w_{t+t_s}^3 \\ \vdots \\ w_{t+h}^1 \\ w_{t+h}^2 \\ w_{t+h}^3 \end{pmatrix} = (\bar{v}_{3,t_s} - \widetilde{v}_{3,t+t_s}) \tag{B.73}$$

where $(\widetilde{\phi}_{t_s-1,31}, \widetilde{\phi}_{t_s-1,32}, \widetilde{\phi}_{t_s-1,33})$ is the third line of $\bar{C}_{t_s-1}^j$

This simplifies to:

$$\begin{pmatrix} \widetilde{\phi}_{t_s-1,31} & \widetilde{\phi}_{t_s-1,32} & \widetilde{\phi}_{t_s-1,33} & \cdots & \widetilde{\phi}_{0,31} & \widetilde{\phi}_{0,32} & \widetilde{\phi}_{0,33} \end{pmatrix} \begin{pmatrix} w_{t+1}^1 \\ w_{t+1}^2 \\ w_{t+1}^3 \\ \vdots \\ w_{t+t_s}^1 \\ w_{t+t_s}^2 \\ w_{t+t_s}^3 \end{pmatrix} = (\bar{v}_{3,t_s} - \widetilde{v}_{3,t+t_s}) \tag{B.74}$$

Now, as $w^1$ and $w^2$ are non-constructive shocks, I can draw them from $\mathcal{N}(0,1)$ and transfer them to the right-hand-side of Equation (B.74):

$$\begin{pmatrix} 0 & 0 & \widetilde{\phi}_{t_s-1,33} & \cdots & 0 & 0 & \widetilde{\phi}_{0,33} \end{pmatrix} \begin{pmatrix} w_{t+1}^1 \\ w_{t+1}^2 \\ w_{t+1}^3 \\ \vdots \\ w_{t+t_s}^1 \\ w_{t+t_s}^2 \\ w_{t+t_s}^3 \end{pmatrix} = ...$$

$$(\bar{v}_{3,t_s} - \widetilde{v}_{3,t+t_s} - \widetilde{\phi}_{t_s-1,31}w_{t+1}^1 - \widetilde{\phi}_{t_s-1,32}w_{t+1}^2 - ... - \widetilde{\phi}_{0,31}w_{t+t_s}^1 - \widetilde{\phi}_{0,32}w_{t+t_s}^2) \tag{B.75}$$

The next step is to draw the structural shocks $w_t^3 \mid_{Rw=c}$ that satisfy the above conditions. This can be done following Waggoner & Zha (1999):

$$w \sim \mathcal{N}(\bar{w}, \bar{\Lambda}) \tag{B.76}$$
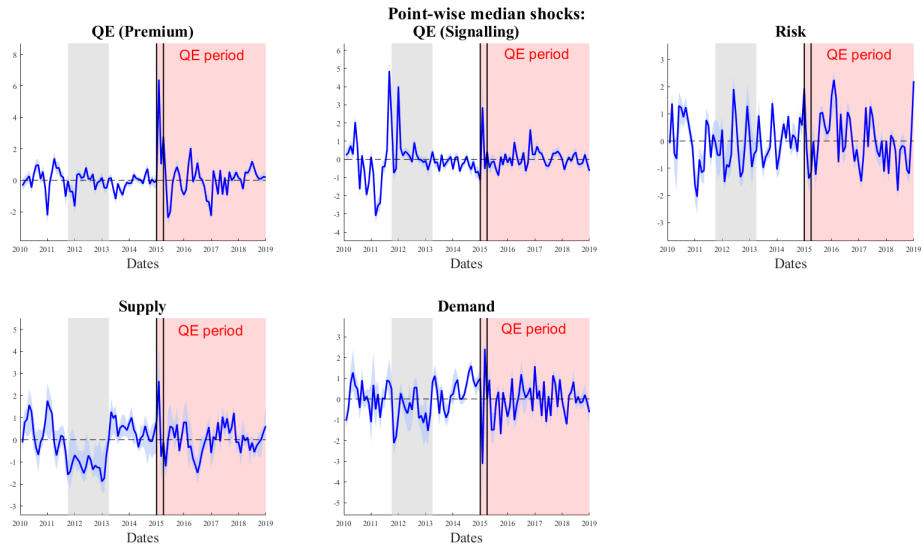
where $\bar{w} = R'(RR')^{-1}c$ and $\bar{\Lambda} = I_3 - R'(RR')^{-1}R$

Draws of $w^1$ and $w^2$ are discarded as they have already been drown.

Then, repeat Equation (B.73) to Equation (B.76) for the next period at which a condition is imposed, taking as known all previous structural shocks (that have already been drawn).

## B.4 Results

Figure 23: Structural shocks - Identification Scheme I



Solid grey area corresponds to a recession period as defined by the Center for Economic Policy Research (CEPR)
Solid red area corresponds to the QE period

Figure 24: Fry & Pagan median target Impulse Response Functions - Identification Scheme I
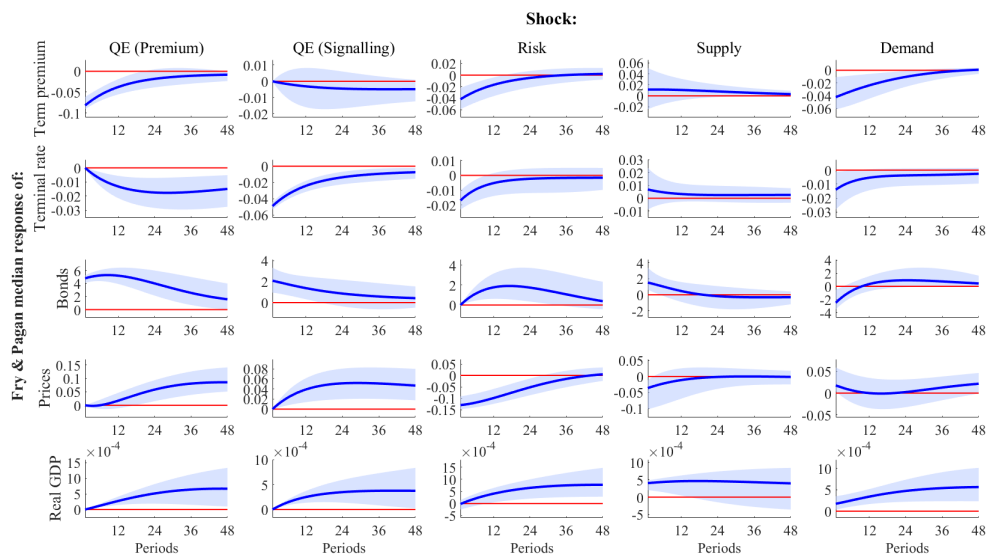


Figure 25: Historical Decomposition of the term premium - Identification Scheme I
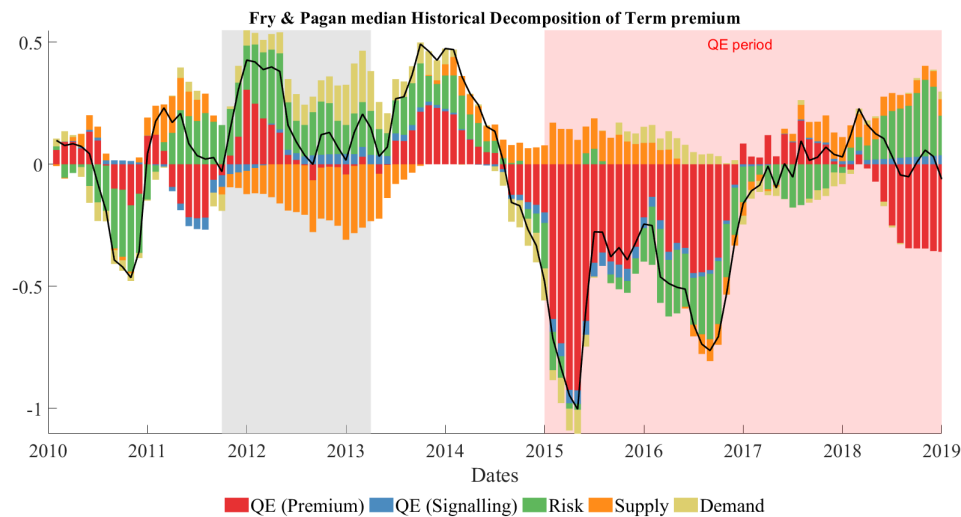


Historical decomposition is expressed in deviation from initial conditions
Solid grey area corresponds to a recession period as defined by the Center for Economic Policy Research (CEPR)
Solid red area corresponds to the QE period

Figure 26: Historical Decomposition of the Terminal rate - Identification Scheme I
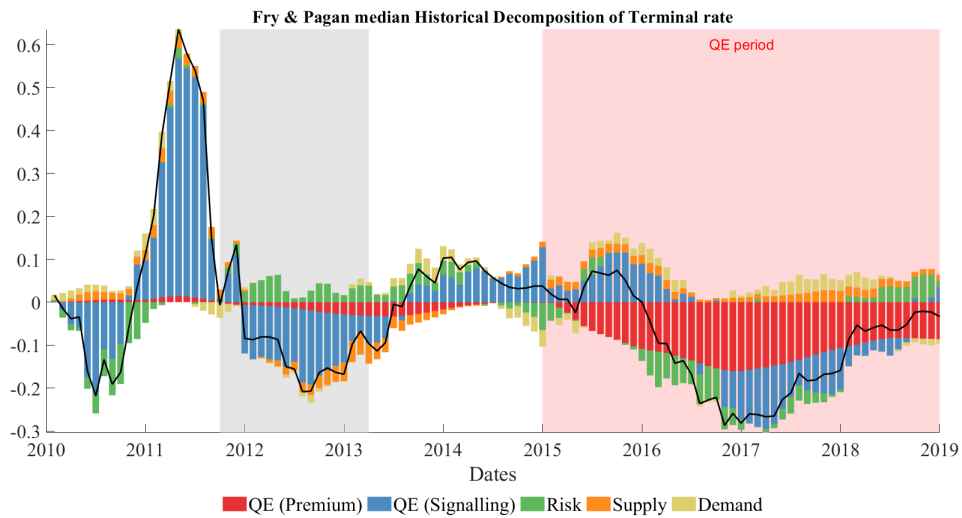
**Fry & Pagan median Historical Decomposition of Terminal rate**



QE (Premium) ■ QE (Signalling) ■ Risk ■ Supply ■ Demand

Historical decomposition is expressed in deviation from initial conditions
Solid grey area corresponds to a recession period as defined by the Center for Economic Policy Research (CEPR)
Solid red area corresponds to the QE period

Figure 27: Counterfactual Real GDP - Identification Scheme I

**Point-wise median counterfactual:**



—— Real GDP
--- Counterfactual Real GDP

Δ=0.0019

77
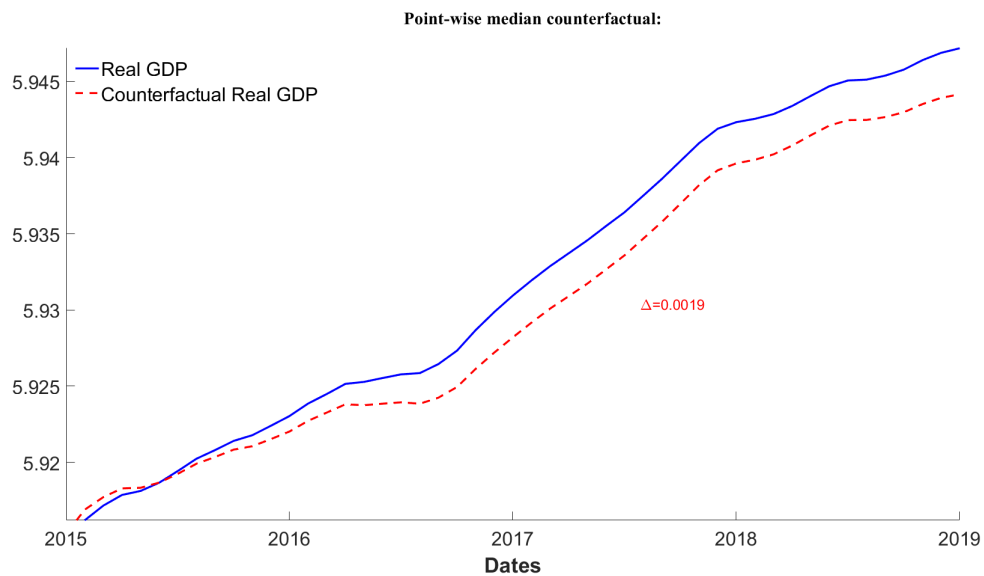
Figure 28: Counterfactual Prices - Identification Scheme I



**Point-wise median counterfactual:**
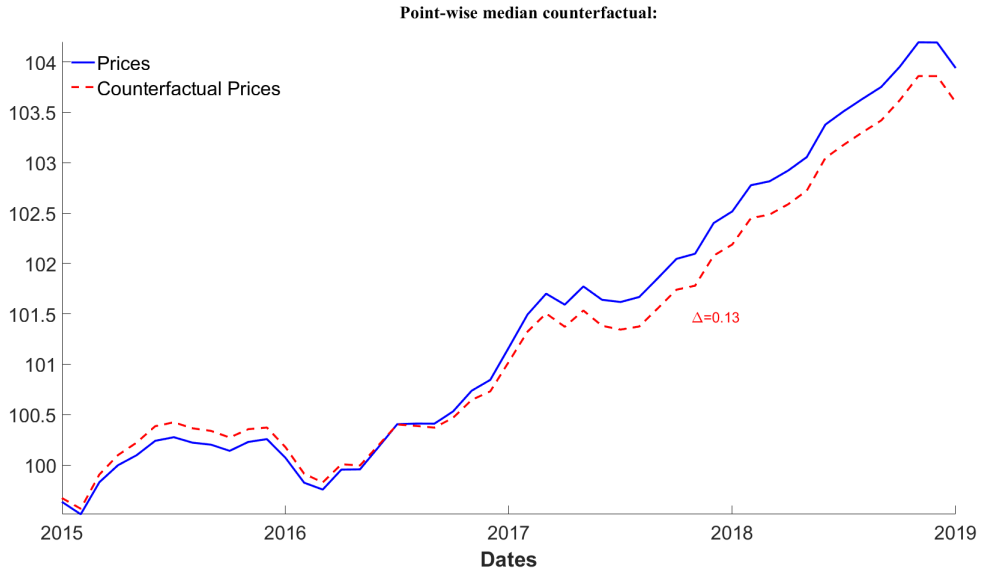
Table 8: Regressions Fit

| Equation | (30) | (31) | (29) |
|----------|------|------|------|
| **RMSE** | 2.67 | 1.55 | 0.31 |

Sample period: 2004Q4-2018Q4
RMSE stands for Root Mean Square Error

# C    DCC-GARCH model

DCC(M,N)-GARCH($P_i$,$Q_i$) model assuming data have been demeaned:

$$\Upsilon_t = H_t^{1/2} z_t \quad \text{with} \quad z_t \overset{iid}{\sim} \mathcal{N}(0, I_3) \tag{C.1}$$

where $\Upsilon_t = (\upsilon_{1,t}, \upsilon_{2,t}, \upsilon_{3,t})'$,

$$H_t = D_t R_t D_t \tag{C.2}$$

where $H_t \in \mathcal{M}_{3\times 3}(\mathbb{R})$ is the conditional variance-covariance matrix of interest,

$$D_t = diag(\sqrt{h_{1t}}, \sqrt{h_{1t}}, \sqrt{h_{3t}}) \tag{C.3}$$

where $D_t \in \mathcal{M}_{3\times 3}(\mathbb{R})$ is a diagonal matrix of standard deviation of $\Upsilon_t$,

$$h_{it} = \alpha_{i0} + \sum_{q=1}^{Q_i} \alpha_{iq} \Upsilon_{i,t-q}^2 + \sum_{p=1}^{P_i} \beta_{ip} h_{i,t-p} \tag{C.4}$$

78

with $i \in [\![1,3]\!]$ is the GARCH Equation,

$$R_t = Q_t^{*-1} Q_t Q_t^{*-1} \tag{C.5}$$

where $R_t \in \mathcal{M}_{3\times3}(\mathbb{R})$ is the conditional correlation matrix and $Q_t^* \in \mathcal{M}_{3\times3}(\mathbb{R})$ is a diagonal matrix with the square root of the diagonal elements of $Q_t$ at the diagonal with

$$Q_t = (1 - \sum_{m=1}^{M} a_m - \sum_{n=1}^{N} b_n)\bar{Q}_t + \sum_{m=1}^{M} a_m \epsilon_{t-1}\epsilon_{t-1}' + \sum_{n=1}^{N} b_n Q_{t-1} \tag{C.6}$$

with $a_m \in \mathbb{R}$, $b_n \in \mathbb{R}$ is the DCC Equation,

$$\bar{Q}_t = \frac{1}{T} \sum_{t=1}^{T} \epsilon_{t-1}\epsilon_{t-1}' \tag{C.7}$$

where $\bar{Q}_t$ is the unconditional variance-covariance matrix of the standardized errors $\epsilon_t$, with

$$\epsilon_t = D_t^{-1}\Upsilon_t \tag{C.8}$$

Here I employ a DCC(1,1)-GARCH(1,1) for simplicity, hence GARCH Equation (C.4) becomes:

$$h_{it} = \alpha_{i0} + \alpha_i \Upsilon_{i,t-1}^2 + \beta_i h_{i,t-1} \tag{C.9}$$

with $i \in [\![1,3]\!]$. DCC Equation (C.6) boils down to:

$$Q_t = (1 - a - b)\bar{Q}_t + a\epsilon_{t-1}\epsilon_{t-1}' + bQ_{t-1} \tag{C.10}$$

with $a \geq 0$, $b \geq 0$ and $a + b < 1$ to ensure that $H_t$ is positive definite.

Last, the log-likelihood of the system of equations is given by:

$$\mathcal{L}(\theta) = -\frac{1}{2} \sum_{t=1}^{T} \left( nln(2\pi) + 2ln(|D_t|) + ln(|R_t|) + \Upsilon_t'D_t^{-1}R_t^{-1}D_t^{-1}\Upsilon_t \right) \tag{C.11}$$

where $\theta$ contains all parameters split into two groups: the ones from Equation (C.9) and the ones from Equation (C.10).

The estimation is performed in three stages by Quasi-maximum likelihood (QMLE). I first estimate the conditional variance parameters from Equation (C.9) and then the correlation parameters from Equation (C.10). Finally, the log-likelihood is maximized conditionally on the previously estimated parameters.[42]

## C.1   Results

---

[42]I thank Kevin Sheppard from the University of Oxford for sharing his code.
It is available at: http://www.kevinsheppard.com/MFE_Toolbox

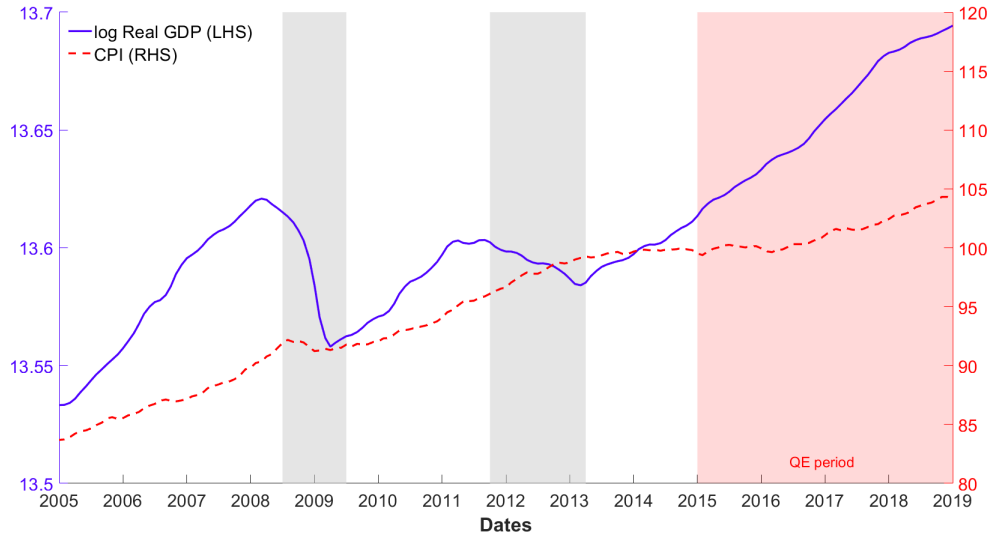Figure 29: Covariance between the term premium and the cost of borrowing
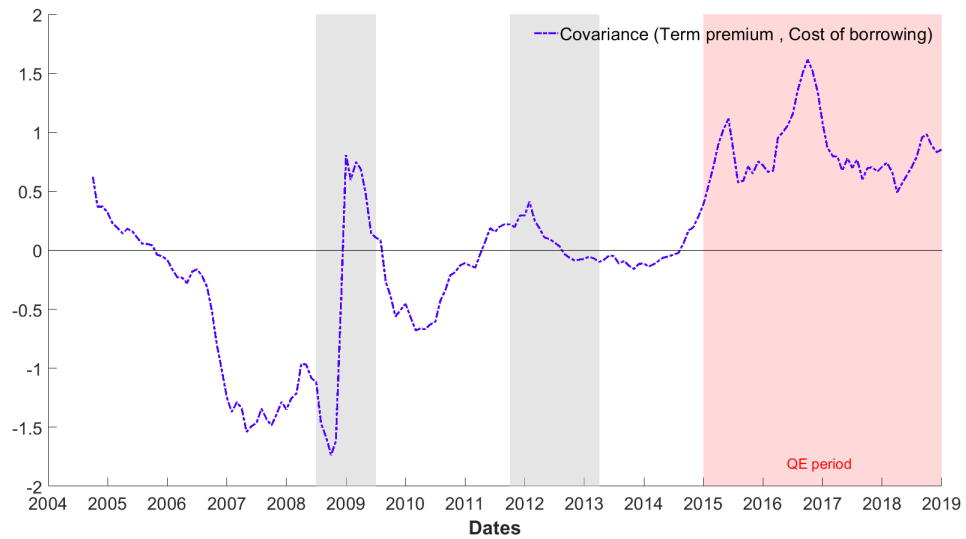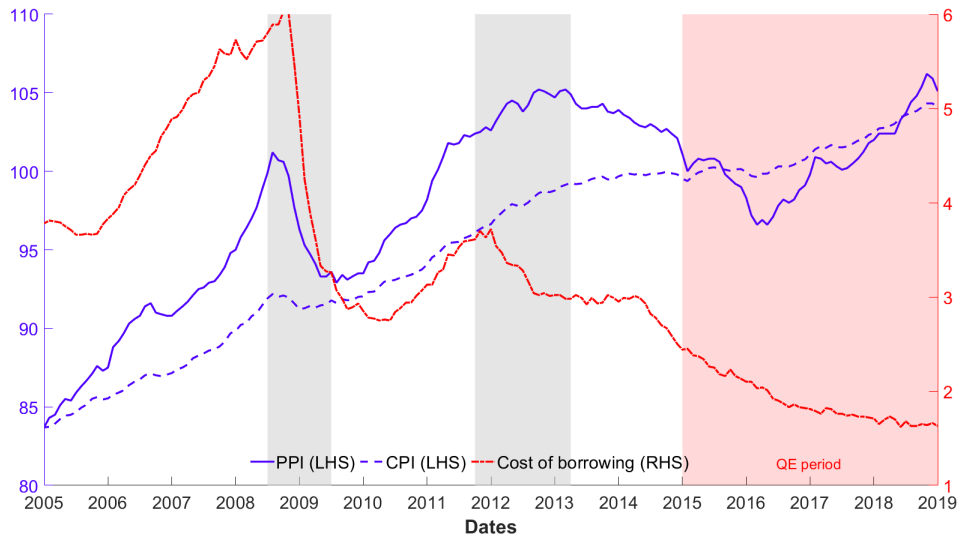
# D   Data

Figure 30: Consumer Price Index and Real GDP



Consumer Price Index (CPI) and log real GDP (interpolated) from 2004:M9 to 2018:M12
Solid grey areas correspond to recession periods as defined by the Center for Economic Policy Research (CEPR)
Solid red area corresponds to the QE period

Figure 31: PPI, CPI and cost of borrowing



Sample period: 2004Q9-2018Q4
Solid grey area corresponds to a recession period as defined by the Center for Economic Policy Research (CEPR)
Solid red area corresponds to the QE period