FORECASTING AND POLICY ANALYSIS WITH TREND-CYCLE BAYESIAN VARS*

Michal Andrle[†] and Jan Bruha[‡]

Preliminary and Incomplete

First version: Sept 16, 2016

September 6, 2017

*The views expressed herein are those of the authors and should not be attributed to the International Monetary Fund, its Executive Board, or its management. The views should not be attributed to the Czech National Bank and its management either. We would like to thank Grace Bin Li (IMF), Joris de Wind (CPB), Anders Warne (ECB), and particpants of the ECB Seminar for comments and discussions. Our special thanks go to Miroslav Plašil (CNB) for his detailed comments, reviews of the manuscript versions and joint work on system priors.

[†]International Monetary Fund, Research Dept. (corresponding author)

[‡]Czech National Bank, Research Division

Abstract

This paper introduces Trend-Cycle Bayesian VARs (TC-BVARs) for use in macroeconomic forecasting and policy analysis. Economic theory supports the view that trends and cycles are dominated by different shocks and transmission channels. Each variable is decomposed into components, $y_t = \bar{y}_t + \hat{y}_t$. The flexibility of TC-BVARs comes from the fact that the model specifies flexible processes for low-frequency movements (trends) of variables and flexible VAR process for the cyclical frequencies. There is a clear distinction of cycles, trends, or exogenous time-varying policy targets. TC-BVARs benefit from the flexibility of VARs and from careful anchoring of the models' long-run behavior. The state-space form of the model helps to work with missing data, mixed frequencies, and various forms of expert judgment and conditional forecasting. Structural TC-BVARs benefit from less biased reduced-form model specification.

Keywords: VAR, trend, cycle, forecast, system priors, penalized maximum likelihood, sequential Monte Carlo

J.E.L.: E3, E37, C11, C18, C32

Contents Page

Tables

Figures

I. INTRODUCTION

Vector Auto-Regressive (VAR) models are often used for short-term or medium-term forecasting. In this paper we introduce Trend-Cycle Bayesian Vector Auto-Regressive (TC-BVAR) models and make a case for their use in forecasting instead of the VARs. The idea behind TC-BVARs is that the behavior of macroeconomic variables at different frequencies (trends, cyclices) may be dominated by different economic phenomena and modeling them as such is beneficial. Each variable is decomposed into parts, $y_t = \bar{y}_t + \hat{y}_t$. A natural example would be to think of the dynamics of the potential output and the 'output gap' and their different impact on inflation and interest rates. Specifying a flexible TC-BVAR model introduces some estimation challenges due its flexibility. To cope with these challenges we illustrate how Bayesian 'system priors', priors about the overall model properties, can help with the implementation of the model.

The benefits of separating clearly the low-frequency and cyclical dynamics of the data is at least threefold. First, acknowledging different dynamics may significantly de-bias the estimated model and lead to meaningful estimates of underlying transmission channels, as illustrated in Andrle and Bruha (2014). As an example, think of a desinflating economy, where low-frequency continued decline of both inflation (target) and interest rates will dominate the standard VAR model and lead to price puzzles, and others. Second, the trend-cycle specification is very flexible with respect to data transformations, i.e. estimation in levels or in growth-rates. Often, stationarity-inducing transformations lead to implausible constraints, as is the case when transitory shock to policy rate leads to transitory response of output growth, with a permanent change in the level of output. Third, from a purely pragmatic forecasters' point of view, the stationary cyclical dynamics will die out and the medium-term forecast is dominated by the trends (potential output growth, inflation target, or change in the world trade openness) that are good candidates for expert judgment or satellite models, often differ from the historical episodes, and it is too demanding to ask for these to be captured in a one-shock-per-variable VARs.

Economic theory supports the notion that different structural shocks and transmission mechanisms may be dominant at different frequencies of economic time series. Different factors dominate long-run economic growth and economic business cycles. Of course, factors behind long-term growth can and do spill over into cyclical frequencies to some extent, see Aguiar and Gopinath (2007), for instance, and such spillovers should be reflected in structural models. There is a strong tradition of business cycle analysis both in theoretical and empirical economics.

In the econometric literature, the recognition of frequency-specific modeling dates back to Jevons or more recently at least to band-spectrum regression by Engle (1974) and the work by Grether and Nerlove (1970).[1] Further, coherence analysis of macroeconomic time series, however, further supports the view that a distinction between trends and cycles can be a good approximation for flexible time-series models. Forni and others (2005), Stock and Watson (2002), or Andrle, Bruha, and Solmaz (2016) illustrate that macro data features high coherence at cyclical frequencies, and much lower coherence at low frequencies and very high frequencies. Economic cycles can be described with just a small number of common factors, often rather stable in time, even through Great Recession, see Stock and Watson (2012).

Evolution of 'trends' can be highly complex and high-frequencies can be affected by data measurement errors, leaving the cyclical dynamics as primary candidates for VAR analysis. Yes, the distinction of low and cyclical frequencies is sometimes an admission of imperfection of our models and admission about a higher degree of uncertainty about the low-frequency component of the model. This aspect of theory and data has not been reflected in the VAR literature to our knowledge, though it is sometimes the case with structural dynamic stochastic general equilibrium (DSGE). This is a bit paradoxical, as the structural DSGE models possibly can and should strive to address the issue, while it may bee too much to ask from a [S]VAR.

In this paper, the examples provided mostly deal with a simple 'monetary policy' VARs. Simple time-series model of output, inflation, interest rates and its open-economy extension. Hence, the concept discussed here are 'trend' level of output and output gaps, inflation deviation from the inflation target or long-term inflation expectations, or trend real rate of interest. Most concepts, perhaps with the exception of the inflation target and long-term inflation expectations are unobserved and require the TC-BVAR specification or an explicit ad-hoc pre-filtering.

But TC-VARs are not tied to these type of 'monetary policy' models. For instance, models of foreign trade and its impact on the economy, not acknowledging the changes time-varying nature of the trade openness of the economy, driven by evolution of supply-chain management, international trade agreements or trade unions expansions, can significantly affect the cyclical implications of the model. In the same vein, in most labor markets over the World a

[1] "In itself, the division of a time series into several unobserved components is of little significance; it is, rather, that the components are themselves ascribable to separate and distinct groups of causes or influences."(Grether and Nerlove, 1970, pp. 686) "It may be too much to ask of a model that it explain both slow and rapid shifts in the variables, or both seasonal and non-seasonal behavior. It is at least reasonable to test the hypothesis that the same model applies at various frequencies."(Engle, 1974, pp. 5)

version of 'Okun's Law' can be found, while the trend labor force participation, equilibrium unemployment, or secular changes in the level of real wages differs starkly.

Trends and Cycles

More specifically, the TC-BVAR models decompose variables of interest into three, often unobserved, components capturing the trends, cycles, and high-frequency dynamics and model them separately:

$$y_t = \bar{y}_t + \widehat{y}_t + \dot{y}_t. \tag{1}$$

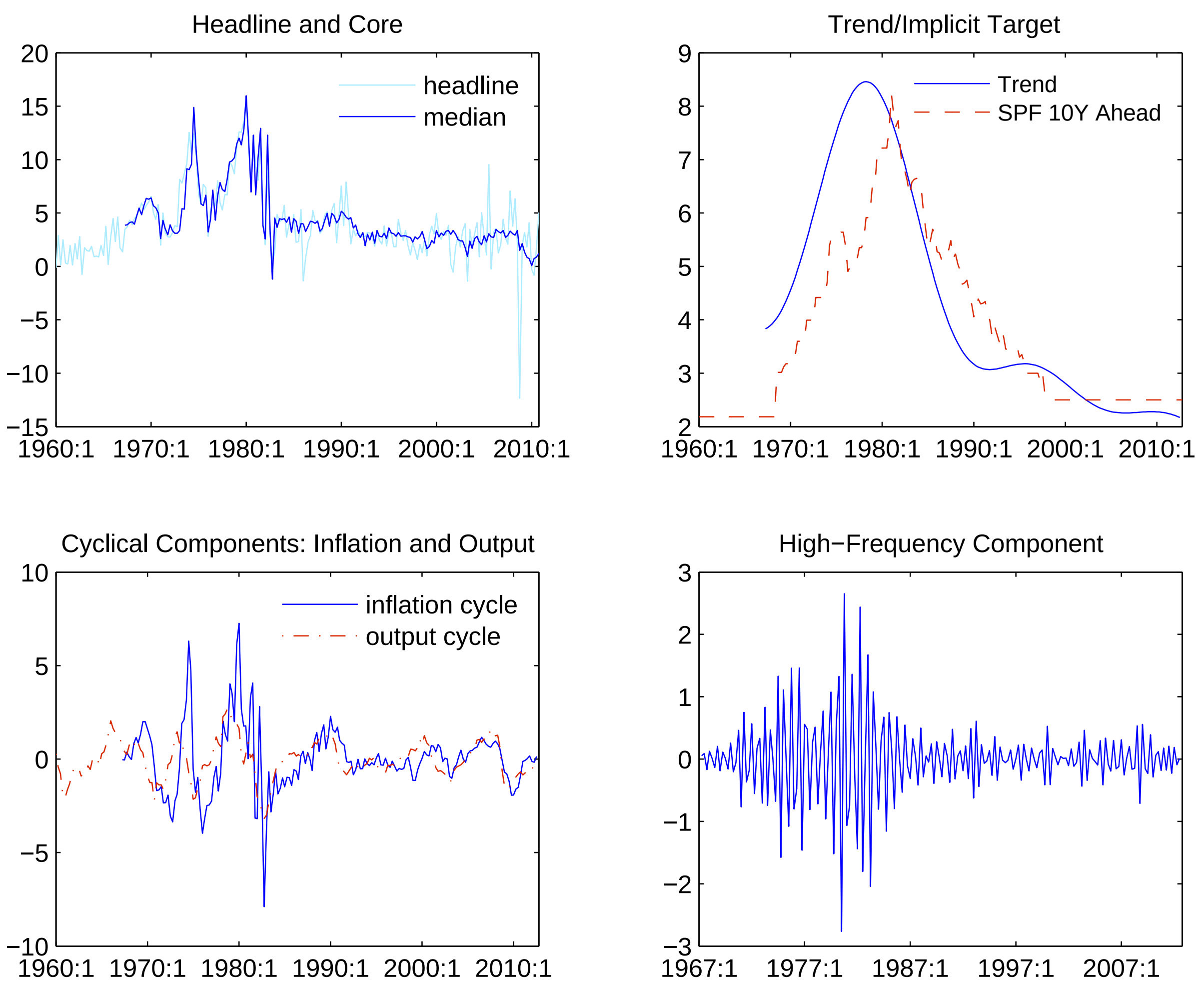
The low-frequency component, $\bar{Y}_t$, can be modeled using a parsimonious version of a well-known local-level model, see or . The unobserved cyclical component $\widehat{y}_t$, is specified as a zero-mean VAR model with k lags. The high-frequency component $\dot{y}_t$ can be omitted or used when the data is particularly noisy at high frequencies, possible due to measurement issues.[2] A white-noise process or a MA(1) process are natural candidates in this case, as discussed below.

Clearly, resulting model is essentially a VARIMA model and the use of prior is crucial for sensible estimates. Formally, the TC-BVAR models take a state-space form, as do linearized structural DSGE models, and so the tools from the technical toolbox available for these models are easy to apply. Specifically, dealing with missing observations, mixed frequencies (quarterly, annual), conditional and unconditional forecasting, and forecasting with expert judgment are available.

To illustrate the importance of the decomposition into components, while avoding frequency-domain formalism, we decomposed U.S. median inflation (in percent, annualized) into the three components. The trend component is frequencies lower than 32 periods, the cyclical frequency is up to 2 periods, and the high-frequency is the residual of band-pass filter. Jointly with the filtered trend inflation we show an inflation-trend measure from the FRB/US obtained from splicing model estimates until late 1970's and Survey of Professional Forecasters ten-years-ahead survey. While the two series differ, they share the same principle. The cyclical component of the output is plotted jointly with cyclical component of inflation. For the most part, only inflation deviation from the long-term expectation should respond to cyclical state of the economy. As one would expect, it is the cyclical component of output that co-moves with deviation of inflation from its long-term expectations, a version of a short-run Phillips curve relationship.

[2] In the spirit of structural time-series models, Harvey (1989), adding seasonal component to the model can be a natural extension of the model, avoiding series-by-series ad-hoc seasonal adjustment.

Figure 1. Frequency Decomposition of U.S. Median Inflation



As an example, in a closed-economy model with output, inflation, and interest rates there are only three stochastic shocks that are asked to explain all the dynamics in a regular VAR. Researchers then face issues of variable transformations, level or growth rates, issues with arbitrarilly estimated or non-existent steady state, etc. With a TC-BVAR for the same economy, a local level trend model is specified for the potential output, inflation target is either observed or inferred from ten-year ahead inflation expectations and trend in real interest rate accommodated, with the flexible VAR modeling the deviation of output from the potential, inflation from the inflation target, and the cyclical frequency of interest rate.

Structual VARs

All the principles and techniques in this paper carry over to estimation and use of Structural VARs, that is, to Structural TC-BVARs. Structural VARs can be easily estimated in a trend-cycle setup and would often benefit from lower misspecification of the reduced form by innaccurate representation of trends or steady-state values of variables. With structural VARs, the structural shocks can be viewed as function of one-step-ahead forecast errors, and usually the covariance matrix of residuals is used for shock identification. Hence, having the reduced-form VAR severely misspecified will dramatically affect any structural inference using the model, as argued in Andrle and Bruha (2014).

Structural VARs, would also benefit from the use of System Priors. In the forecasting applications, the system priors introduced below are primarily aimed at assuring stationarity of the cyclical components, restricting the persistence of the model and oscillatory dynamics, etc. With Structural VARs, the system priors should be based on economic theory, be it about impulse-response functions, variance decomposition, or shock decompositions. All these are function of parameters and the system priors would translate into model estimates.

Estimation with System Priors

Estimation of TC-BVAR is a little bit more involved than estimation of a regular VAR models or BVAR models. The reason is the presence of the unobserved variables and no closed-form solution for the likelihood estimate. Formally, TC-BVAR estimation it is ‘just’ a Bayesian estimation of a state-space model, not really different from estimation of linearized DSGE models. For meaningful estimation of such a flexible model with a great deal of parameters the importance of both shrinkage and economically-meaningful priors is crucial. For this purpose, common priors for VAR models are complemented with ‘system priors’ about the overall behavior of the model (IRFs, frequency-response functions, etc.), making the estimation process simpler.

The part of the paper focused on estimation of the TC-BVARs explains the incorporation of System Priors, but more details are to be found in Andrle and Beneš (2013) or and Andrle and Plašil (2016). The estimate of the posterior mode, with or without the use of homotopy methods, can be a part of Markov Chain Monte-Carlo (MCMC) methods as with DSGE models, often relying on Metropolis-Hastings algorithm. However, the recently introduced Sequential Monte Carlo (SMC) approach of Herbst and Schorfheide (2014) and Herbst and Bognanni (2015) is a particularly suitable approach when the system priors are used. The link between using the homotopy and using tampered distributions in each of the method explores the idea of building the estimate off the prior information.

Toolbox For practitioners, the code to create, estimate, and forecast with the models is available for Matlab. The code is designed for ease of use rather than speed, under the premise that the estimation step will be done a few times, while the forecasting and policy analysis will be a daily activity.[3] Most of the technical details of the state-space representation or the Bayesian estimation with System priors is relegated to the appendix.

Relationship to the Literature

This paper extends our work on trend-cycle VARs we initiated in Bruha, Pierluigi, and Serafini (2011), creating a TC-VAR for the euro-area labor market, and Andrle, Ho, and Garcia-Saltos (2013), constructing a forecasting trend-cycle VAR model for Poland. Andrle and Bruha (2014) make a list of issues with specification of the reduced-form VARs and how it may distort both the forecasting performance and structural shock identification with the VARs. One of the issues explained is the need to acknowledge trends, cycles, steady states, and policy regimes. In this paper, we go a step further, illustrating the model that in our view solves many issues raised and we put our money where our mouth is. Also, we employ System Priors to greatly facilitate the estimation of the flexible model with sensible properties.

One of the first papers in the VAR literature pointing out the need for some medium-term or long-run anchor is Villani (2009) who constructs a VAR with a prior about a constant steady state in a mean-adjusted VAR. The steady-state is determined by past data. The issue is the fact that a constant steady-state assumption may not be feasible for many economies, where the low-frequency components, trends, do vary in time. Sometimes, the time-varying trends are even observed and it is enough just to subtract those from the raw data, as is the case with the inflation targets, for instance. The most related work in is the approach the semi-structural state-space approach of Benes and N'Diaye (2004), essentially a restricted TC-VAR, semi-structural forward-looking models by Carabenciov and others (2008a) where New-Keynesian model is used to specify the cyclical dynamics and economic theory is guiding trend specification. Canova (2014) proposes to link a flexible local-level model with structural DSGE models to link the models to raw data if the permanent shocks are not part of the structural model. The whole literature stands on the shoulders of structural time-series models by Harvey (1989).

Trend-Cycle BVARs are no panacea. If a structural model is available, its use is strongly preferable. But when a VAR is to be estimated, TC-VARs do offer a better-behaved alternative to forecasting and policy analysis due to their well-anchored medium-run dynamics and

[3] The code and examples can be downloaded at: `michalandrle.weebly.com/xx`

possibly less biased cyclical dynamics. TC-BVARs are still just a reduced-form time series methods that can under no circumstances replace the need for structural models for economic analysis. Further, TC-BVARs need enough macroeconomic data of reasonable quality to be estimated, though both shrinkage- and system priors alleviate this burden to some extent.

II. Building a Trend-Cycle VAR

To build a simple trend-cycle VAR we will create a simple closed-economy model for analysis of real output, y_t, inflation, π_t, and nominal short-term interest rate, i_t. In the empirical part of the paper, we will estimate a version of the model using for the United States and discuss the setup of marginal-independent and system priors.

A. A Prototype Closed-Economy TC-VAR

1. Aggregation

In the closed-economy TC-VAR we will specify a simple model for output, inflation, and interest rates. Other variables, like unemployment, are easily added. All variables will be split into three components: trend, cycle, and high-frequency. Trend variables will be denoted by overbar, the cyclical variables with a 'hat':

$$\begin{aligned} y_t &= \bar{y}_t + \hat{y}_t + u_y && (2) \\ \pi_t &= \bar{\pi}_t + \hat{\pi}_t + u_{pi} && (3) \\ i_t &= \max[\hat{i}_t + \bar{i}_t,\ i_{\text{floor},t}] + u_i. && (4) \end{aligned}$$

Note, that the units of variables and variables' transformations matter a lot and the TC-VAR framework is very flexible in this respect. For instance, the cyclical dynamics here is specified in terms of level of output at cyclical frequency with the level of interest rate at cyclical frequency.[4] In a regular VAR, using log-level of GDP and level of interest rates would leave to misspecification. In the same vein, using the growth rates of output and level of interest rates has sever consequences as well.

[4] We do not purse the labels of the 'output gap' or deviation from 'equilibrium' interest rates, etc. but in principle that is how part of the literature thinks of these components.

The reason is, in our simple example, that the economic theory suggests that it should be the cyclical position of the economy, the business cycle, that is linked via Phillips curve relationship with inflation and interest rates, see Fig. 1. Apart from econometric issues, with growth rates (first-difference filter) amplifying high-frequency volatility, the very unfortunate implication of a model with output growth rate is be that a transitory shock to interest rate will result in transitory response of the GDP growth, that is, into a *permanent* change in the level of output. That is quite a strong a priori restriction on the model and at odds with the theory. Yet, often the result of standard VARs specified in growth rates to render the output stationary.[5]

2. Cycles

The business-cycle dynamics is driven by a zero-mean VAR(k) model for all endogenous variables: output, inflation, and interest rate:

$$\mathbf{A0}\begin{bmatrix}\hat{y}_t\\ \hat{\pi}_t\\ \hat{i}_t\end{bmatrix} = \mathbf{A}_1\begin{bmatrix}\hat{y}_{t-1}\\ \hat{\pi}_{t-1}\\ \hat{i}_{t-1}\end{bmatrix}+\cdots+\mathbf{A}_k\begin{bmatrix}\hat{y}_{t-k}\\ \hat{\pi}_{t-k}\\ \hat{i}_{t-k}\end{bmatrix}+\mathbf{C}\begin{bmatrix}e_{\hat{y},t}\\ e_{\hat{\pi},t}\\ e_{\hat{i},t}\end{bmatrix}. \tag{5}$$

The coefficient matrices, $\mathbf{A}_1,\ldots,\mathbf{A}_k$ are the auto-regressive coefficient driving the dynamics of the VAR. Since the model is for cyclical components, they are restricted to be stationary, restricting the eigenvalues of the model. Also, zero intercept translates into zero mean of the cyclical components for any specification of the VAR dynamical coefficients. Medium and long-term forecast is purely driven by the trend specification. In the case of forecasting VARs, when identification of structural shocks need not to be considered, the matrix $\mathbf{A0}$ will be a unit matrix. Assuming that $\mathbf{e}_t \sim N(\mathbf{0},\mathbf{I})$, the loading matrix $\mathbf{C}$ estimates the covariance matrix of shocks, or residuals.

3. Trends

The low-frequency, or trend, specification closes the model. The logarithm of potential output, $\bar{y}_t$, is assumed to follow a local-level model, with the level of output being subject to per-

[5] Similar point can be made about fiscal VARS where transitory changes in tax rates may lead to permanent changes in the level of output.

manent level shocks, $u_{\bar{y},t}$, and growth rate shocks, $u_{g,t}$:

$$\begin{aligned}\bar{y}_t &= \bar{y}_{t-1} + g_t/4 + u_{\bar{y},t} & (6)\\ g_t &= \rho_g g_{t-1} + (1-\rho_g) g_{ss} + u_{g,t}. & (7)\end{aligned}$$

This specification allows the potential output growth rate to be pinned down by the parameter g_{ss}, for $\rho_g < 1$, or become a unit-root process otherwise. It is interesting to note that the process for a well-known Hodrick-Prescott filter is nested in our specification, with the output gap, $\hat{y}_t$, being a white noise, $\rho_g = 1$, and variance of $u_{\bar{y},t}$ being zero. The HP's λ_{hp} parameter is then a relative of the variance of $e_{\hat{y},t}$ and $u_{g,t}$. Understanding the local-level model is important for devising sensible priors for the estimation stage and forecasting overall.

Low frequency of inflation, its 'trend', is assumed to be constant or deterministic path consistent with long-term inflation expectations or inflation target:

$$\bar{\pi}_t = \bar{\pi}_{t-1} + u_{\bar{\pi},t} \qquad \text{and} \qquad E[\pi_{t+j|t}] = \bar{\pi}_t \text{ for } j \to \infty \tag{8}$$

Clearly, in economies with an explicit inflation target, the choice is easy and allows to accommodate time-varying inflation targets for economies undergoing disinflation (e.g. Poland or the Czech Republic). Given the stationary cyclical component, $\hat{\pi}_t$, the long-run forecast of inflation is the trend, or inflation target. In economies without a formal inflation target within the sample, further information about $\bar{\pi}_t$ beyond intrinsic realtionship of cyclical components can be long-term inflation expectations surveys[6] or long-term interest rates, if interest rates are part of the model. Any relevant measures for n-steap ahead inflation forecast are easy to use with the model, defining a new variable of expected n-period-ahead inflation, given the recursive nature of the model.

The trend level of nominal interest rate is determined as a trend level of real interest rate plus the inflation trend (target):

$$\begin{aligned}\bar{i}_t &= \bar{r}_t + \bar{\pi}_t & (9)\\ \bar{r}_t &= \rho_{\bar{r}} \bar{r}_{t-1} + (1-\rho_{\bar{r}}) \bar{r}_{ss} + u_{\bar{r}}. & (10)\end{aligned}$$

The real trend of interest rate is assumed to follow a persistent AR(1) model with a steady-state equilibrium determined by assumptions on the long-run neutral rate of interest. For $\rho_{\bar{r}} = 1$ a unit-root behavior will accommodate significant trends in the real interest rates, which is useful for forecasting with expert judgment on features of the economy that cannot

[6] Survey of Professional forecasters in the U.S. provides a ten-year ahead forecast. Also, the FRB/US variable PTR is a constructed historical measure of inflation expectations.

be modeled under such a simple specification. More sophisticated specifications are possible, for instance, linking the long-term real growth rate or potential output growth to real trend interest rate, and observing long-term interest rates, i_t^N, implied by the model, possibly also incorporating a term-premium process to the model.

Linking expectation variables to the model is easy, given the recursive nature of the model. Given the transition law of motion of the model as $\mathbf{X}_t = \mathbf{T}\mathbf{X}_{t-1} + \mathbf{R}\mathbf{e}_t$, expectation variables are linear combination of state variables and expected state variables $\mathbf{X}_{t+i|t} = \mathbf{T}^i\mathbf{X}_t$.[7]

4. High-Frequency Dynamics

The third component of the modeled macroeconomic variables are high-frequency dynamics. The third component can often times be ignored. For variables with significant high-frequency variation that is hard to explain with the model, especially on monthly or daily frequency, it may help to improve forecasting performance of the model in terms of turning points and cycles.

A simple specification can entail a white-noise process, reflecting measurement errors. An alternative is a MA(1) specification $u_t = \xi_t + \phi\xi_{t-1}$ with $\phi < 0$; in this case, a large transitory shock today to the cyclical dynamics is largely corrected the next period.[8]

[7] When computational speed is not of the essence or when expectational variables are part of the simultaneous system, say in the VAR, the solution algorithms for linear rational expectation models easily apply and simplify the model specification.

[8] This specification may be more useful in forward-looking structural models, where expectations are crucial. For instance, with a persistent inflation dynamics an MA(1) specification alongside with an other shock will create 'short' and 'long' cons-push shocks with dramatically different dynamics, accommodating high-frequency volatility in inflation.

Closed-Economy Model: Summary

[A] Aggregation:

$$y_t = \bar{y}_t + \hat{y}_t + u_{y,t} \tag{11}$$

$$\pi_t = \bar{\pi}_t + \hat{\pi}_t + u_{pi,t} \tag{12}$$

$$i_t = \max[\hat{i}_t + \bar{i}_t,\ i_{\text{floor},t}] + u_{i,t} \tag{13}$$

[B] Cyclical Dynamics:

$$\mathbf{A0}\begin{bmatrix} \hat{y}_t \\ \hat{\pi}_t \\ \hat{i}_t \end{bmatrix} = \mathbf{A}_1 \begin{bmatrix} \hat{y}_{t-1} \\ \hat{\pi}_{t-1} \\ \hat{i}_{t-1} \end{bmatrix} + \cdots + \mathbf{A}_k \begin{bmatrix} \hat{y}_{t-k} \\ \hat{\pi}_{t-k} \\ \hat{i}_{t-k} \end{bmatrix} + \mathbf{C} \begin{bmatrix} e_{\hat{y},t} \\ e_{\hat{\pi},t} \\ e_{\hat{i},t} \end{bmatrix} \tag{14}$$

[C] Trend Component:

$$\bar{y}_t = \bar{y}_{t-1} + g_t/4 + u_{\bar{y},t} \tag{15}$$

$$g_t = \rho_g g_{t-1} + (1-\rho_g) g_{ss} + u_{g,t} \tag{16}$$

$$\bar{\pi}_t = \bar{\pi}_{t-1} + u_{\bar{\pi},t} \quad \text{and} \quad E[\pi_{t+j|t}] = \bar{\pi}_t \text{ for } j \to \infty \tag{17}$$

$$\bar{i}_t = \bar{r}_t + \bar{\pi}_t \tag{18}$$

$$\bar{r}_t = \rho_{\bar{r}} \bar{r}_{t-1} + (1-\rho_{\bar{r}}) \bar{r}_{ss} + u_{\bar{r}} \tag{19}$$

$$i^N_{t|t} = (1/N) \sum_{i=0}^{N} i_{t+i|t} \quad \text{for } N = 4, 20, 40. \tag{20}$$

The TC-VAR model is inherently a state-space model with some observed and some unobserved variables. Hence, all the tools and the flexibility of state-space modeling is available to be used with TC-VARs.

B. Small-Open-Economy TC-BVAR

The SOE TC-BVAR follows the principle of the closed economy model, with a few modifications both to the trend and the cyclical component of the model.

(a) Cyclical component: In the cyclical component, the openness dimension introduces possibly new variables, namely the exchange rate and a set of foreign-country variables–foreign output, inflation, interest rates, commodities etc. As is common in the literature, let us denote the vector home variables as $\mathbf{Y}_t$ and foreign variables by $\mathbf{Y}_t^\star$. The exchange rate must belong to the set of home, or domestic, variables.

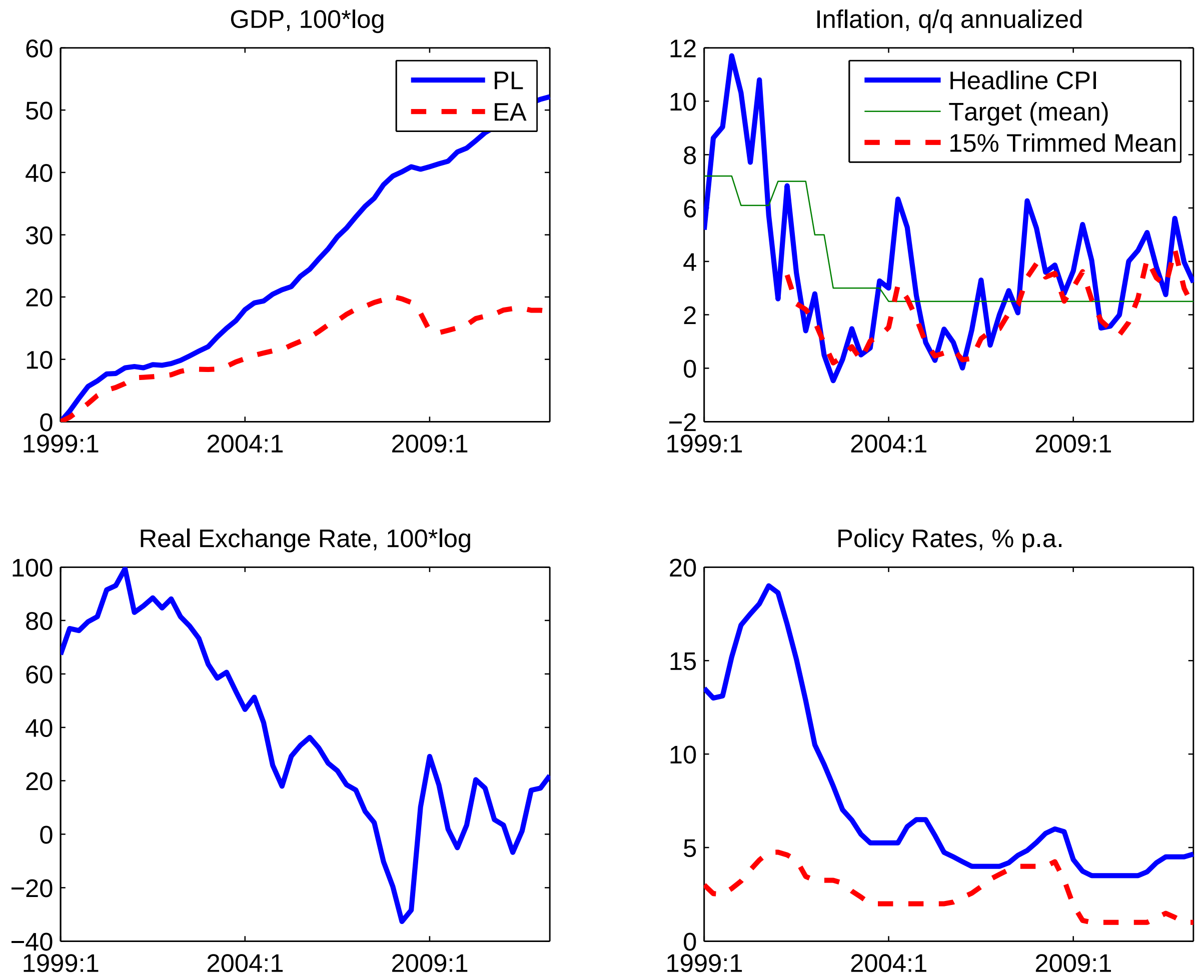
Then, when plausible, the cyclical component of the model clearly must respect the small-open-economy condition that domestic variables do not have effects on the foreign variables, that is:

$$\begin{aligned}\begin{bmatrix}\mathbf{A0}_{11} & \mathbf{A0}_{12}\\ \mathbf{0} & \mathbf{A0}_{22}\end{bmatrix}\begin{bmatrix}\mathbf{Y}_t\\ \mathbf{Y}_t^\star\end{bmatrix} &= \begin{bmatrix}\mathbf{A}_{11,1} & \mathbf{A}_{12,1}\\ \mathbf{0} & \mathbf{A}_{22,1}\end{bmatrix}\begin{bmatrix}\mathbf{Y}_{t-1}\\ \mathbf{Y}_{t-1}^\star\end{bmatrix} + \cdots + \begin{bmatrix}\mathbf{A}_{11,k} & \mathbf{A}_{12,k}\\ \mathbf{0} & \mathbf{A}_{22,k}\end{bmatrix}\begin{bmatrix}\mathbf{Y}_{t-k}\\ \mathbf{Y}_{t-k}^\star\end{bmatrix} + \\ &+ \begin{bmatrix}\mathbf{C}_{11} & \mathbf{C}_{12}\\ \mathbf{0} & \mathbf{C}_{22}\end{bmatrix}\begin{bmatrix}\mathbf{e}_{t-k}\\ \mathbf{e}_{t-k}^\star\end{bmatrix}.\end{aligned}\tag{21}$$

Although the block-zero restrictions on the model in (21) may seem obvious and are present in large part of the literature, there is still abundance of models that fail to acknowledge the small-open-economy assumption in either either the shock loadings or the dynamics of the model. In situations when the small-open-economy assumptions are less clear-cut, a case for a tighter SOE prior can be made for the relevant coefficients.

(b) Trend component: The trend component specification can take many different forms, depending on the nature of the model (monetary policy, trade forecasts, etc.) and the degree of prior restrictions wished to be exercised, but there are some worthy lessons to draw from the literature and economic theory.

Figure 2. Macroeconomic Trends: Poland vs. Euro Area



The trends can be specified more or less flexible, when the increasingly tighter theoretical restrictions (common trends, constant shares, etc.) are tested and if beneficial for the purpose of the model, then retained. The TC-VAR framework is flexible to what extent the theory is embraced. Fig. 2 illustrates how important it is to acknowledge trends in real exchange rate, different long-term growth rates or different inflation targets between the economies and their implied trend of nominal interest rates.

In principle, a flexible specification for trend output of home and foreign economy that are not functionally linked is feasible. A transition economy with higher growth rates (Poland or the Czech Republic) than more advanced countries, the euro area core members, may be cyclically connected but the trend growth rate can be thought of as independent to the process of economic transition. When feasible and with sufficient available, cross-correlation of innovations to trend growth and permanent level shocks can specified, with a prior centered around zero, for instance:

$$y_t^h = y_{t-1}^h + g_t^h/4 + u_{\bar{y},t} + \phi_1 u^{\star}_{\bar{y},t} \tag{22}$$
$$g_t = \rho_g g_{t-1} + (1-\rho_g) g_{ss} + u_{g,t} + \phi_2 u^{\star}_{g,t}. \tag{23}$$

Such specification is, however, making the model even more complex and thus strong priors on ϕ_1, ϕ_2 are recommended. For instance, even semi-structural models usually ignore the cross-country inter-linkages at low frequencies, see Carabenciov and others (2008b) or Andrle and others (2009) for SOE. It is usually not worth deviating from $\phi_1 = \phi_2 = 0$, the model is still very flexible and the forecast horizon is open to expert judgment.

The new home variable is the real exchange rate and the (log) real exchange rate trend, z_t; we assume the nominal exchange rate level is implied from the real exchange rate and inflation process in home and foreign economy. For many countries, the real exchange rates have important low-frequency dynamics. The cyclical component of the real exchange rate is part of the home vector in the VAR model. The specification can follow the local-level model as:

$$z_t = \bar{z}_t + \hat{z}_t \tag{24}$$
$$\bar{z}_t = \bar{z}_{t-1} + g_{z,t}/4 + u_{\bar{z},t} \tag{25}$$
$$g_{z,t} = \rho_{gz} g_{z,t-1} + (1-\rho) g_{z,ss} + u_{gz,t}. \tag{26}$$

Inflation target process is specified as in the case of the closed economy. The same can hold for the real interest rate trend, or it can reflect the arbitrage principle and risk-premium augmented uncovered interest parity (UIP). Note, that in the VAR, the short-term the UIP does not hold, unlike in structural models. To reflect the arbitrage in the trend components one can assume that

$$\bar{r}_t = \bar{r}_t^{\star} + prem_t + (\bar{z}_{t+1|t} - \bar{z}_t), \tag{27}$$

or some less restrictive version of the constraint, possibly ignoring the expectation trends. For the purpose of short-term forecasting with TC-BVARs, especially with high-frequency data, ignoring these constraints is perfectly feasible. The detrending of the real exchange rate itself is the key requirement for the model, with the VAR not explaining all frequencies of the exchange rate.

C. State-Space Form – Costs and Benefits

When all the pieces of the model are put together, the model takes a standard state-space form:

$$\mathbf{Y}_t = \mathbf{K}_1 + \mathbf{Z}\mathbf{X}_t + \mathbf{R}\mathbf{e}_t \tag{28}$$

$$\mathbf{X}_t = \mathbf{K}_2 + \mathbf{T}\mathbf{X}_{t-1} + \mathbf{H}\mathbf{e}_t. \tag{29}$$

The cost of the TC-VARs is that the model is more complex than a simple VAR with all components observed. The benefit of TC-VARs is that all the standard tools, namely the Kalman filter and smoother, can be used with the model, see Harvey (1989) or Durbin and Koopman (2012). Linearized DSGE models have identical state-space reduced-form representation.[9]

Given the state-space form of the model, unconditional or conditional forecasting with the model is easy and well documented in the literature, see Beneš, Binning, and Lees (2008), for instance. Dealing with missing variables and now-casting of selected components of $\mathbf{Y}_t$, is well-documented and simple in the Kalman filter and smoother framework, see Harvey (1989).

When external information or expert judgment is available about the unobserved components of the model in the historical sample, it can be accommodated by the use of auxiliary observations of the information, following Doran (1992), see Andrle (2013b) for details about 'filter expert judgment'.

[9] In principle, TC-BVAR model can be easily estimated using IRIS or Dynare toolbox for Matlab. The publicly available code for this paper is adjusted to be used with the IRIS Toolbox for Matlab, to leverage the reporting and time-series features of the toolbox.

III. TC-BVAR Estimation with System Priors

A. Motivation and Intuition

Estimation of TC-BVAR models can be ill-behaved in short samples with no priors about the parameters, due to the higher number of coefficients and the flexibility of the model. Priors are needed.

We use two types of priors for estimating the TC-VAR models, which can be used for any Bayesian VAR: (i) Bayesian Shrinkage priors Litterman (1986), and (ii) System Priors reflecting more economically-meaningful concepts, see Andrle and Plašil (2016) and Andrle and Beneš (2013).

Bayesian shrinkage is a concept mostly associated with BVARs to fend off over-parameterization of VARs, especially with limited data. Following Litterman (1986) the literature on shrinking VAR parameter values towards a zero (for stationary systems), with parameters loading variables of higher lag-length more tightiely shrunk to zero. Effectively, the dynamical coefficients, $\mathbf{A}_{i,j,k}$, are endowed with a $N(0,\sigma_{i,j,k})$ prior distribution.[10]

System priors may complement or substitute the shrinkage priors for the TC-BVAR models and both the VAR portion and the trend specification of the model may benefit from it. System priors are priors about the computable features of the model, an example being a frequency-specific distribution of the output variance, or speed of convergence of the VAR model to the trend component. In structural VAR specification, the priors may reflect the economic theory priors about the impulse-response function of the model. Some specific examples of system priors are given in the empirical examples below.

B. Formal Implementation

The most general form of estimation with system priors is based on the likelihood function and a *composite prior* $\tilde{p}$ – a result of updating a prior on individual coefficients with a system

[10] Using Normal distribution is the most common option in the literature, but once the model is not specified using conjugate priors and exploring any analytical forms, the options are wider and Bayesian LASSO would be a sensible choice as well .

prior 'likelihood' and observed data likelihood functions:

$$p(\boldsymbol{\theta}|\mathbf{Y}) \propto L(\mathbf{Y}|\boldsymbol{\theta}) \times [S(\mathbf{Z}|\theta) \times p(\boldsymbol{\theta})], \tag{30}$$

$$\propto L(\mathbf{Y}|\boldsymbol{\theta}) \times \tilde{p}(\theta|S(\mathbf{Z})). \tag{31}$$

Here $L(\mathbf{Y}|\boldsymbol{\theta})$ denotes the likelihod function of the VAR model and $p(\boldsymbol{\theta})$ is a proper joint prior the parameters in $\boldsymbol{\theta}$, which includes all dynamical and stochastic ones. $S(\mathbf{Z}|\theta)$ is a 'likelihood' function for the system priors $\mathbf{Z}$.[11]

There are two intuitive ways of looking at (30). As always in the Bayesian estimation, one can view the posterior mode as maximum likelihood estimation with penalization function $p(.)$ or $\tilde{p}(.)$. Another view, is to view (30) as a three-step sequential updating process. In the first step a distribution of coefficients θ is updated, or reweighed, by the system priors $S(\mathbf{Z}|\boldsymbol{\theta})$ to obtain a composite prior $\tilde{p}(.)$ In this sense the system priors play a role of dummy observations, albeit a nonlinear ones. The new composite prior is then further updated by the model and the actual observed data via the likelihood function $L(\mathbf{Y}|\boldsymbol{\theta})$.

Computationally, the approach can be rather straightforward in principle, using either importance sampling (IS), sequential importance sampling (SMC), or Metropolis-Hastings (MH) algorithms, for instance. Given a k-th draw of parameters $\boldsymbol{\theta}_k$ from a suitably chosen proposal density, its posterior probability is obtained by evaluating $p(\boldsymbol{\theta}_k)$, the system prior function $S(\mathbf{Z}|\boldsymbol{\theta}_k)$, and the likelihood function $L(\mathbf{Y}|\boldsymbol{\theta}_k)$.

In particular, two approaches are suggested below for how to proceed with the VAR analysis with system priors, see Appendix for details. Both approaches are motivated by a 'continuation' idea, where one starts with a simple-to-solve problem and gradually moves to the actual, more difficult, problem. In the first approach, one starts directly with numerical search for the posterior mode. This is facilitated by the fact that the 'penalized' likelihood estimation can be initiated using the least-squares (maximum-likelihood) estimates of the model, which is easy to obtain. In the second approach, the SMC, the continuation idea is incorporated in the process of tempering the distribution. One starts with the draws from the marginal prior distribution and gradually adds information from the system priors and the likelihood function of the model. The SMC approach is also very easy to parallelize to speed up the computations.

[11] For instance, the example given in the introduction, a prior on the peak of inflation effect for a 1% demand shock, $z_{\pi,y}$, distributed as $z_{\pi,y} \sim \mathcal{N}(0.5, 0.10)$ and which clearly depends on $\boldsymbol{\theta}$, may help to make the notation of $p(\mathbf{Z}|\boldsymbol{\theta})$ more concrete. The likelihood function, in addition to marginal prior $p(\boldsymbol{\theta})$ is *penalized* when $z_{\pi,y}$ gets into tails of its assumed distribution.

As is usual in the case of Bayesian inference, the devil is in the detail of choosing the suitable proposal density to sample from. For low-dimensional system with a few variables and short lag structure, importance sampling using the prior distribution $p(\boldsymbol{\theta})$ works reasonably well and can easily be parallelized to harness the power of modern computers.

C. Misspecification and Alternatives to Log-Likelihood

The discussion above assumed, as is almost always the case with the VAR models, that the loss function for estimation of the model is the log-likelihood of the state-space form of the model, obtained via one-step-ahead forecast error decomposition, see Durbin and Koopman (2012). This is an optimal approach, if the model is a true data generated process and there is no misspecification.

However, when misspecification is suspected, the log-likelihood is not optimal anymore. Also, depending on the intended use of the model, one-step-ahead forecast error may not be the criterion deemed as the most important for evaluating the model. There is a large literature about the use of multi-step ahead forecast errors as the criterion for specifying empirical models and its link to misspecification, see Tiao and Xu (1993), the insightful frequency-domain considerations by Haywood and Tunnicliffe-Wilson (1997), or recent use for VAR models by Schorfheide (2005) using a plug-in estimator and multi-step iterated approach in Franta (2016), or for DSGE models by Kapetanios, Price, and Theodoridis (2015) and Tonner and Bruha (2014).

The use of Bayesian analysis, and the use of marginal priors and system priors, is consistent with alternative loss function for the estimation of the model motivated by feature matching Y. and Tong (2011) or method-of-moments Andrle (2013a), after appropriate transformation of the loss function, see Chernozhukov and Hong (2003), for instance.[12]

The frequency-specific modeling approach of TC-BVARs may ameliorate the misspecification and the need for alternative loss functions, in our experience. Acknowledging both high-frequency dynamics, or measurement noise, together with low-frequency flexible processes does help increase medium-horizon forecasting performance of the models, in our experience.

[12]The TC-BVAR code base associated with the papers allows for user-defined criterion functions, alongside the default log-likelihood or optinal weighted multi-step-ahead forecast error with weights for each horizon. This follow the insight of Haywood and Tunnicliffe-Wilson (1997) who demonstrate the frequency-specific filter nature of the multi-step forecast errors.

IV. Empirical Applications

A. Baseline TC-BVAR for the United States

In this section a minimal TC-VAR model is estimated for the United States. The TC-BVAR model is estimated as a system, with unobserved trends and cycles using log-likelihood and Bayesian shrinkage and system priors. The model is specified for output, inflation, and interest rates and is mostly illustrative. Despite its simplicity, the model is no 'straw man' and we discuss extensions that would likely lead to increased forecasting accuracy for a real-world application.

In the first step, the initial values for cyclical portion of the model were obtained using a simple BVAR with cyclical components obtained univariate high-pass filters.[13] In the second step, all coefficients are estimated jointly using the likelihood function of the full model, marginal priors on individual parameters and desired shrinkage, as well as system priors on selected properties of the overall model. No homotopy steps were needed for the US model, given the size of the sample and data. Also, the results for the US are satisfactory with the log-likelihood being the criterion function, alongside the priors.

Although we provide a system-wide estimation of the full state-space model in this paper, it is our experience that estimation of the first step with careful calibration of the trend component of the model leads to very good results for practical forecasting work and makes the estimation problem simpler.

For the VAR(k) model, the model has been estimated with $k = 2$, though choosing $k = 1$ leads to comparable results, given the shrinkage effects at the second-order lag. The Bayesian shrinkage prior on the VAR dynamic coefficients are chosen in a rather traditional way: the higher the lag, the stronger shrinkage to zero, see below. Note, there is no constant term in the VAR model for cyclical components.

Selected details on the marginal prior and posterior distributions are in the Appendix, the text mainly motivates the key choices for priors.

Marginal Independent Priors for the Trend Components:

[13] Such a treatment of the initial values for the optimization is not necessary in our model, starting from the prior mode is very effective as well.

The marginal-independent prior distributions for the coefficients driving the trend component of the model reflect our a priori view that the changes in the trend components should be relatively smooth, with potentially long-lived deviation from their respective steady states or steady-state growth rates.

For simplicity, we assume that the long-term growth of the economy is 2.5% a year. The prior for the standard deviation of the permanent level innovation is around 0.1 to 0.05 % of GDP per quarter. The trend growth of output is mean-reverting but a persistent process, with ρ_g prior centered around 0.95. The innovation to annualized trend growth of output has a standard deviation of 0.12 per quarter. Alternative assumptions about the long-term growth can be made, for instance using real-time historical average, long-term growth survey expectations, or setting $\rho_g = 1$ as sometimes seen in the literature.[14]

We make a prior assumption that the trend real rate of interest is a rather smooth (with small variance of innovations), mean-reverting variable. The deviations of the real trend interest rate from the ultimate steady-state value can be protracted, however.[15]

The priors about the variance of the inflation target innovation is irrelevant in our case, as we observe the data for ten-year-ahead inflation expectations, which pin down the inflation target uniquely, setting $\rho_{\bar{\pi}} = 1$.

'Shrinkage' Priors for Cyclical Components

The 'shrinkage' priors follow the bulk of the Bayesian VAR literature, being used to deal with the over-parameterization of the VAR models. Since the VAR describes the cyclical component, required to be stationary, the shrinkage priors on individual coefficients is towards zero. We can assume a 'ridge-regression' type of prior, using Normal distribution prior on the dynamical coefficients, following Litterman (1986), or zero-mean Laplace distribution, following Park and Casella (2008) or Chen and others (2008) for empirical work. For simplicity and comparability with the BVAR literature, Normal distribution is used for the US model. Note, that 'shrinkage' priors do not guarantee the stationarity of the cyclical portion of the model.

[14] Specifying the growth rate of potential output as a unit root significantly increases the potential for real-time revision of the unobserved components estimates. The assumption, however, is very potent with observing long-term expectations from the Survey of Professional Forecasters when available, to discipline the estimates

[15] A posteriori, the estimates approach the upper value of stationarity as close as possible, so in practical application the the unit-root specification should be considered.

The prior for a dynamical coefficient, a_{ijk}, in equation i, loading variable j with a lag k, is given by:

$$a_{ijk} \sim N(0, s^a_{ijk}) \qquad s^a_{ijk} = \lambda_1 \times \frac{\lambda_2}{k^{\lambda_3}} \times \frac{\sigma_i^2}{\sigma_j^2}. \tag{32}$$

The coefficients λ_1 drives the overall level of shrinkage, the degree of tilting the parameter towards zero for all variables at all leads. The coefficient λ_2 is an additional shrinkage for cross-variable relationships, with $\lambda_2 = 1$ when $i = j$. The coefficient λ_3, on the other hand determines the increase in shrinkage as the lag-length increases. The scaling for cross-variable coefficients, σ_i/σ_j, is in principle unknown residual variance. Often, it is approximated by univariate AR regression residuals, in our case based on univariate pre-filtering, or omitted.

The prior for covariance matrix of the one-step-ahead forecast errors is specified following the BVAR literature. However, the system priors will be later affecting the variance of the cyclical component as it plays an important role in the determination of trends and cycles.

For the forecasting, not a structural VAR, the prior for the covariance matrix of the forecast-error terms in the cyclical component, Σ_e, where $\Sigma_e = \mathbf{CC}'$, is determined as Inverted-Wishart prior with a "prior" covariance matrix S_e and degrees of freedom α_s:

$$\Sigma_e \sim \mathrm{IW}(S_e, \alpha_s) \tag{33}$$

and thus the kernel of the prior distribution for Σ_e is given by

$$\Sigma_e \propto |\Sigma_e|^{-(\alpha_s+N+1)/2} \exp\left[-\frac{1}{2}\mathrm{tr}(\Sigma_e^{-1} S_e)\right]. \tag{34}$$

The prior covariance matrix S_e is considered diagonal, with the diagonal elements given by s_i^2. For TC-VARs without structural identification and structural priors about the $\mathbf{C}$ matrix is not distinguished, or can be simply specified by a convenient factorization of the covariance matrix. For small-open-economy models, the Cholesky transformation can be used, for instance.[16]

System priors:

System priors are priors about the features of the model, not individual parameters. System priors can be applied both the cyclical and trend component, mainly as a watchdog of a pri-

[16] For structural TC-VARs with priors about $\mathbf{C}$ the priors will reflect the economics behind the identification and the prior for structural shocks covariance matrix will be diagonal but not necessarily unitary.

ori desirable properties of the forecasting model. Such system priors may not be binding or relevant for most of the parameter space but will kick in when needed.

The key system prior for the TC-BVAR, a dogmatic prior that cannot be updated, is the requirement that the cyclical part of the model be stationary. The rest of the model, the low-frequency portion, need not to be stationary.

To make sure the dynamics of inflation, output, and interest rate is well-behaved in terms of convergence properties, we impose priors on effect of the initial-condition impulse-response function. We want to eliminate strong secondary cycles (or tertiary cycles) and too persistent dynamics, so that inflation and output cyclical components settle timely to ther long-run trends.

To restrict the duration of the response to shocks and initial conditions, we require that the variable, say inflation gap, settles after T_1 periods to zero. Let z denote the ratio of the absolute value of the sum of the response to a shock or initial value up to period T_1 over the sum of absolute value of the response up to period $T_2 > T_1$, with T_2 being significantly larger. Then, we specify a prior $z \sim N(1, \sigma_z)$ with the hyper-parameter σ_z denoting the tightness of the prior restriction. If z deviates significantly from unity, it is an evidence that there is a significant dynamics beyond the horizon T_1. The system prior on z statistic is particularly important as the lag-length of the VAR increases. This prior serves as a backstop for highly infeasible dynamics, so the we choose $T_1 = 20$, five years.

Further, we want to exercise a priori views about the frequency-domain behavior of the model and there are many options to choose from. For instance, Andrle and Plašil (2016) suggest a prior about the fraction of the variance of a variable coming from the business cycle frequencies. In principle, a dominant portion of variance of $\hat{y}_t$ should come from cyclical frequencies and if high-frequency component, $\dot{y}_t$, is specified it is important to restrict it to operate on high frequencies and to restrict its contribution to overall variance of the variable. This is a prior we use for high-frequency component of inflation, $u_{\pi,t}$, specified as a zero-mean white noise, with variance $\sigma^2_{u_\pi}$.

The inflation process is modeled as $\pi_t = \bar{\pi}_t + \hat{\pi}_t + u_{\pi,t}$ and due to the assumption of zero cross-correlation the spectral density of inflation, $S_\pi(\omega)$, and its variance are additive in the three components. The marginal prior for the variance of $u_{\pi,t}$ follows $IG(\mu = 0.1, \sigma = 0.05)$, the contribution to overall variance of $\hat{\pi}_t + u_{\pi,t}$ is limited to ten percent. Alternatively, system prior on contribution to variance at particular frequencies can be specified. However, with the spectrum of $u_{\pi,t}$ being flat, the above priors are sufficient to keep $u_{\pi,t}$ from overwhelming the dynamics of inflation.[17]

The other case for a system priors are prior views directly about *the filter frequency-response function* (FFRF) of the model, $\mathbf{X}(\omega) = \mathbf{\Phi}(\omega)\mathbf{Y}(\omega)$, and implied filter gains.[18] Filter frequency-response function indicates how the estimated unobserved components at different frequencies are formed by the input data. We specify a prior about the cut-off frequency, defined as the frequency where half of the input series' power spectrum is retained and half of the power is eliminated, for output gap and log-level of output and interest rate trend and level of interest rates. For output, we set the cut-off frequency prior as $N(c_y, \sigma_{cy})$, where c_y, in terms of periodicity, corresponds to the cut-off frequency of the univariate Hodrick-Prescott filter with $\lambda_{hp} = 1600$(?).[19] For the interest rate trend, our our prior is to have the trend interest rate rather smooth and to use only very low-frequency dynamics of the observed data to go through, so we specify the cut-off frequency at a periodicity of forty quarters.[20]

For the students of the business cycles wishing to avoid explicit frequency-domain calculations and priors a somewhat crude but very useful way of eliciting priors for the trend-cycle decomposition, motivated by seminal paper on Hodrick-Prescott (Leser) filter. In the baseline HP filter specification, the parameter λ_{hp}, denotes the ratio of the variance of the output gap to variance of the potential output acceleration, that is $\lambda_{hp} = \mathrm{var}(\hat{y}_t)/\mathrm{var}(\Delta\bar{y}_t - \Delta\bar{y}_{t-1})$. In the baseline HP specification, unrealistically, $\hat{y}_t$ is assumed to be white noise and $\Delta\bar{y}_t$ follows a random walk, leading to a poor implicit forecasting performance and thus large real-time revision variance. Specifying a system prior about the ratio above is one way of restricting the behavior of the filter for general models, with implicit prior affecting both dynamical and stochastic coefficients of the model. Priors about features of the filter transfer function, however, can be more nuanced.

Simply put, with system priors a penalty is imposed on such elements in the parameter space, that do not conform to this requirement. Any meaningful statistic that can be computed using

[17] Although the more complex prior is not used, the frequency-distribution of contribution to the variance is checked and reported in the appendix.

[18] See or for frequency-domain analysis of state-space models and theory of linear filters.

[19] In principle, another way of specifying a prior on the shape of the gain of the filter is to measure a distance between a profile of a univariate band-pass or a high-pass filter, say HP filter, and the model's implied gain, etc.

[20] The estimation sample used below does not feature the Great Financial Crisis, and with a rather stiff trend real interest rate, the trend nominal interest rate is not declining much after 2008. This does, expectedly, make the forecast of the policy rate revert quickly towards the trend level during the initial years of the crisis. With an estimation sample including the recession, the posterior estimates imply a higher cut-off frequency of the filter and more flexible trend real rate of interest.

the model can be a candidate for system priors, depending on the analysts prior belief and the structure of the model.[21]

(c) Results: The results for the U.S economy are based on the simplest TC-VAR specification above. The observed variables are real GDP level, measured as the chained index, $100 \times log(GDP)$, FED short-term policy rates, measured in percent p.a., median inflation of Federal Reserve Bank of Cleveland, in percent annualized, and of the Survey of Professional Forecasters' (SPF) ten-year-ahead outlook of consumer inflation.

The inflation target variable is directly linked to the SPF's inflation expectations measure, not smoothed for simplicity. Linking to the model-implied inflation expectations ten years ahead makes no difference, due to the horizon length and the stationarity of the cyclical portion of the model.[22] The inflation expectations are assumed to be unit-root and thus the interpretation of the forecast must, especially in the early portion of the sample, should acknowledge the quasi real-time changing endpoint of the inflation forecast and implications for the level of trend nominal interest rate.

The estimation sample is 1985:Q1–2006:Q4, excluding the Great Recession period. The criterion function for the baseline version of the model is the model's log-likelihood, based on the one-step-ahead forecast error in principle. As the TC-VAR model is most certainly misspecified, the log-likelihood optimality may not hold and other criterion functions may serve better for different applications but the log-likelihood is used for comparability with the VARs in the literature. The appendix illustrates the results with a multi-step ahead forecast criterion.

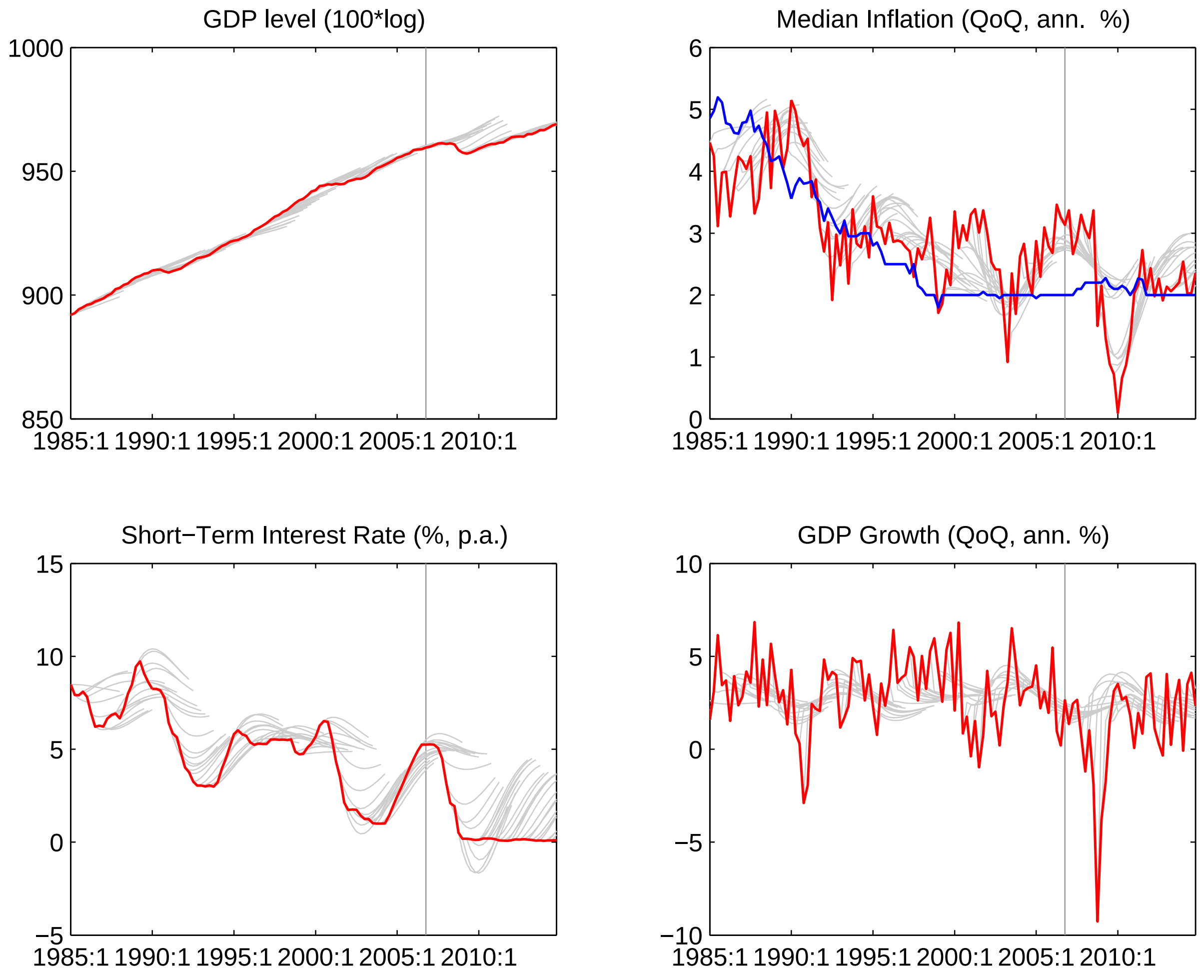
The recursive forecast are unconditional forecasts, with the unobserved states identified real-time using the Kalman smoother up to the initial period of each forecast. These forecast feature no expet judgment and the period 2007:Q1 onwards is truly an out-of-sample exercise.

The results in Fig. 3 and the RMSEs in Tab. 1 suggest that the simple TC-BVAR model has a forecasting performance comparable with the literature, see Carabenciov and others (2008a). Note that the unconditional forecast error of the level of GDP is obviously not defined and thus it is increasing with the forecast horizon.

[21] For details on VARs and SVARs with System Priors, see [Andrle and Plasil, 2017].

[22] Should we measure, say, one-year-ahead inflation expectations with the model either for parameter estimation or just for forecasting, the link would need to be to the model-implied expectations and the results would be affected, of course.

Figure 3. Recursive Forecasts 3Y Ahead: Estimation Sample 1985:Q1–2006:Q4



Note: The breach of zero lower bond on interest rate is not eliminated on purpose for this exercise.

Table 1. RMSEs for the US TC-BVAR

Variable	t+1	t+4	t+8	t+12
GDP level (100*log)	0.54	1.66	2.63	3.51
GDP Growth (QoQ, ann. %)	2.17	2.27	2.19	2.25
GDP Growth (YoY, %)	0.54	1.66	1.59	1.67
Median Inflation (QoQ, ann. %)	0.48	0.52	0.66	0.69
Median Inflation (YoY, %)	0.12	0.36	0.49	0.59
Short-Term Interest Rate (%, p.a.)	0.36	1.14	1.70	1.91

Note: The estimation sample is 1985:Q1–2006:Q4. The RMSE evaluation sample is up to 2017:Q1

Although the paper illustrates mainly the forecasting with TC-BVARs, nothing keeps the analyst from inspecting the analysts the trend and cyclical components of the model in detail and assign interpretation to them, see Fig. 4. Implicitly, the model provides an estimate of the 'output gap' with implicit monetary policy reaction function and Phillips curve nexus embodied in the model. Fig. 7 in the Appendix illustrates the model-implied decomposition of inflation into cyclical and high-frequency component. The presence of high-frequency component of inflation improves the forecasting ability of the model.

The estimate seem to suggest that the short-run Phillips curve in the United States is healthy and alive, as long as the trends and cycles are handled appropriately. Of course, priors matter in the estimation. Policymakers care about measures of 'output gap' in order to assess the cyclical policy of the economy and choose the monetary policy stance, as inflation and unemployment are intertwined with the output gap. Every measure of the output gap is thus meaningful only in a context of a particular model. In our simple model, the measure of output gap is consistent with and co-identified by the inflation dynamics such that the forecast errors are minimized. Our measure of the 'output gap' really does help to forecast inflation.

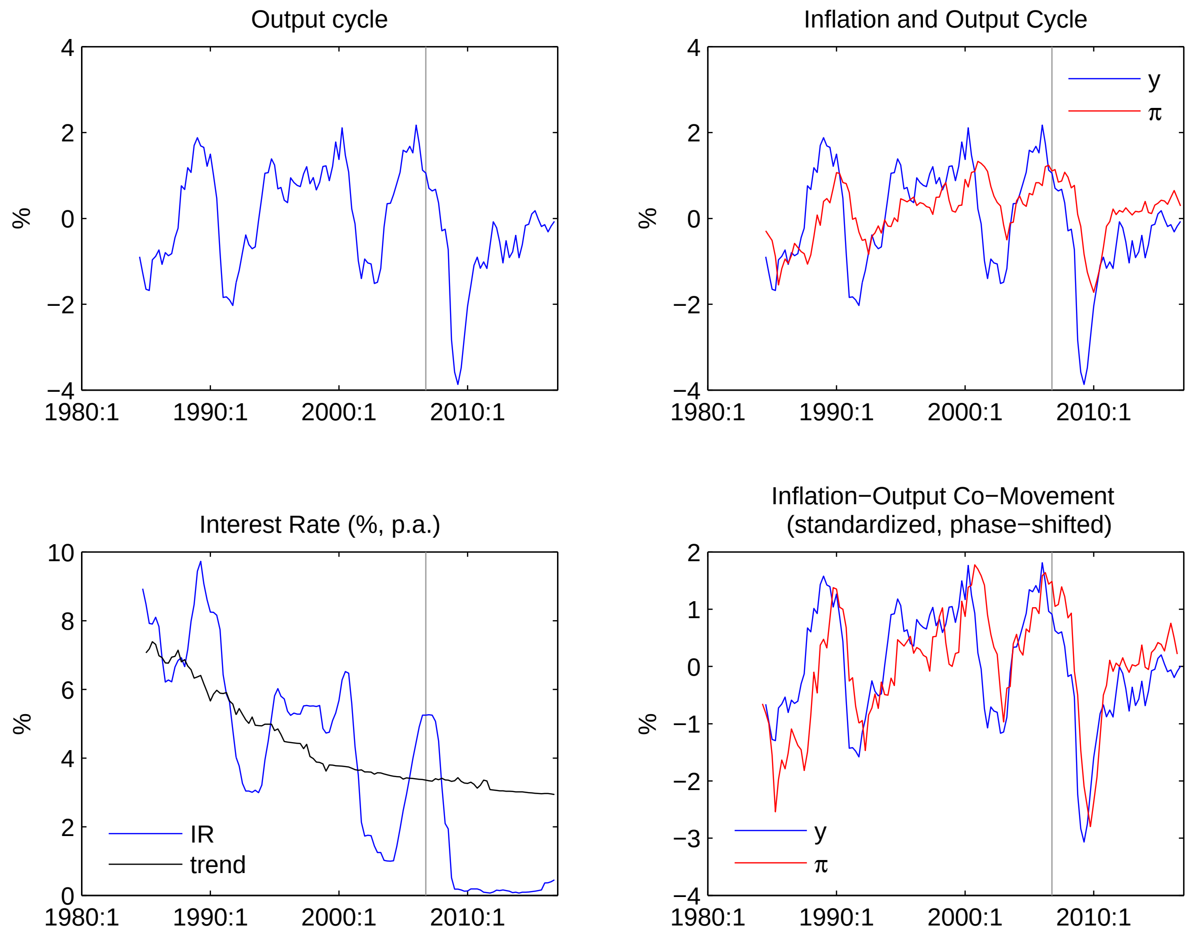
Another feature of the model that helps in deciding about the design of the model is the magnitudes of the real-time revisions of the unobserved components. Fig 5 depicts the quasi real-time estimates of the cyclical position of the economy against the final estimate using the whole sample available. The real-time revisions are kept low due to two important properties of the filter induced by the TC-BVAR model estimated. First, the model features a reasonably high forecasting n-steps-ahead performance, and, second, the implied weights of the filter are not too widely spread out due to the mean-reverting cyclical component and mean-reverting growth rate of the output trends.[23]

Our baseline model also features reporting equations for the yield curve, expressed in terms of expectations theory of the interest rates. The Trend-Cycle VAR implies an expected path of interest rates, using which one-year, five-year, and ten-year ahead interest rates can be computed. Figure 6 depicts the policy rate, the estimate of the five-year and ten-year ahead interest rates, together with the estimate of the term premium for the ten-year zero-coupon bond and comparison with the estimate from the model by Kim and Wright (2005).[24] The estimate of the term premium from the TC-VAR model is comparable with the estimates used by the

[23]For details and comparison to the Hodrick-Prescott filter, see Andrle (2013b)

[24]The Kim and Wright (2005) model estimates are sourced from the St. Louis FRED database, `https://fred.stlouisfed.org/series/THREEFYTP10`

Figure 4. Estimated Trend and Cyclical Components



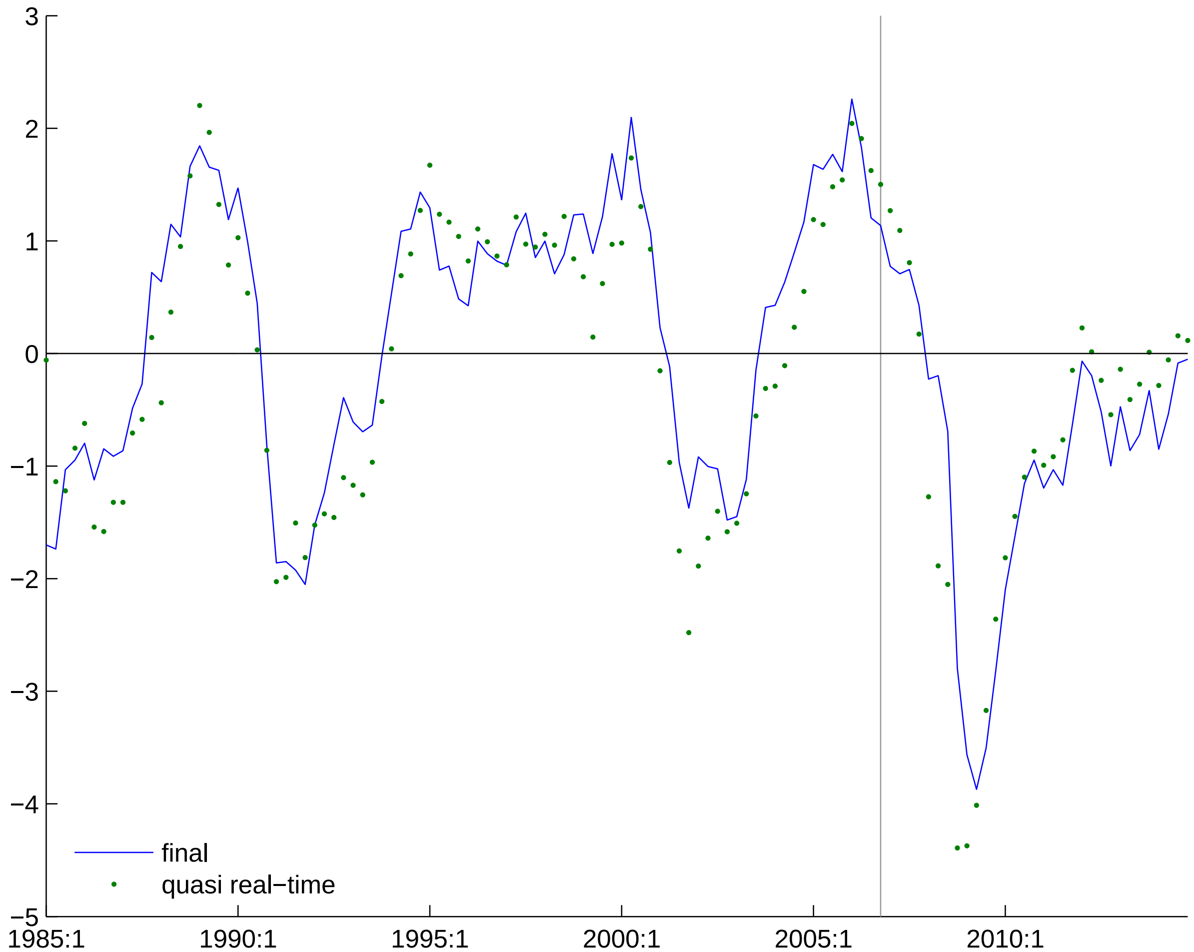
Note: The horizontal line is the end of estimation sample.

Federal Reserve Board, see Canlin, Meldrum, and Rodriguez (2017). After accounting for the term-premium stochastic process, the expectations variables can be linked to the data, complementing or replacing the long-term inflation expectations. The trend-cycle nature of the model is key for realistic model properties, with well-defined, possibly time-varying steady state.

(d) Plausible and Simple Model Extensions: The model can be extended in multiple ways and the alternatives assessed in terms of their forecasting quality.

One possible specification is to make the growth trend GDP to become a unit root as opposed to assumed mean-reverting specification in the current model. Such model may possibly

Figure 5. Final and Quasi Real-Time Estimates of the Outpu Cycle

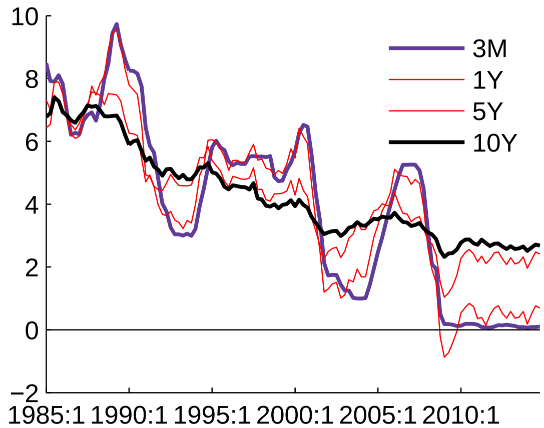


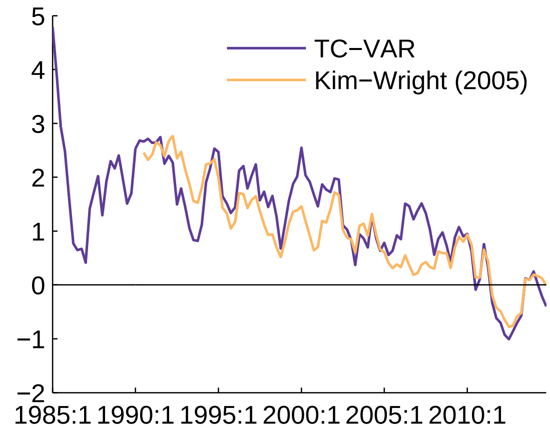
Note: The horizontal line is the end of estimation sample.

better capture the period in early 2000s with the GDP growth above the previous long-run averages with firms and households possibly believing in the new normal with permanently higher GDP growth. Such specification is more flexible, poses greater estimation challenge. Apart from system priors on the behavior of the non-stationary trend growth rate, another possibility is to extract some information from the SPF's survey of ten-year ahead GDP growth and selected quarters of the year, analogous to the SPF's long-term inflation observation.

A priori, there is no reason to believe that observing the SPF's long-term inflation expectations and thus, implicitly, defining inflation target in the model provides a superior forecasting performance as opposed to a more flexible specification with the trend unobserved and smoothed or measuring the long-term inflation expectation with a measurement error, possibly time-varying, and thus only hinting on the plausible dynamics of the inflation target.

Figure 6. Model-Implied Estimate of the Term Premium





Note: Zero-lower bound is not imposed in the simulation of shocks.

The issue arises due to the differences between the survey and the trend obtained from the two-sided band-pass filter. Lastly, the model is easily extended with the yield-curve structure implied by the model and long-term interest rates can be observed to provide information about the long-term real interest rate and long-term inflation.

B. TC-BVAR for Global Oil Price

An example of the model where distinguishing between low-level frequencies and cyclical dynamics is also modeling of oil prices. Specifically, the oil prices could be considered as cyclical fluctuations along the time-varying path of the trend, or equilibrium, oil price as determined mostly by supply-side considerations.

The reason for modeling the trend of the oil prices is that modeling the oil prices in levels or in growth rates in a VAR does not adequately capture the oil price trend growth changes, due to evolution of the supply side of the market.

We will assume that the price of oil, in logarithms, p_t, consist of trend, cycle, and high-frequency component: $p_t = \bar{p}_t + \hat{p}_t + \varepsilon_t^p$.

The VAR model is specified using cyclical components of the oil price, cyclical component of the global production, and the cycle of oil production.

V. Conclusion

This paper has proposed Trend-Cycle Bayesian VARs for forecasting and policy analysis. TC-BVARs describe the trend components and the cyclical components of macroeconomic variables using different specification, the flexible VAR being used for the cyclical component and a form of local-level model for trends.

The notion of frequency-specific models is motivated by economic theory, that acknowledges that long-run growth and economic cycles are often dominated by different shocks and transmission mechanisms. Based in the time domain perspective, TC-VARs allow to model growth and cycles in a more nuanced way.

From the forecasting point of view, the cyclical component of the model will settle towards the trend component of the model, which is well specified and often informed by satellite models. The consequence is that the assessment of uncertainty using the model may result in smaller confidence bands and clear distinction if the uncertainty comes from the business cycle or the trends.

Trend-Cycle BVARs are no panacea. However, they do offered a better-behaved alternative to forecasting with standard BVARs, due to their well-anchored medium-run dynamics and possibly less biased cyclical dynamics. TC-BVARs are still just a reduced-form time series methods that can under no circumstances replace the need for structural models for economic analysis. TC-BVARs, being a flexible over-parameterized models, need enough macroeconomic data of reasonable quality to be estimated, though both shrinkage- and system priors alleviate this burden to some extent.

References

Aguiar, M., and G. Gopinath, 2007, "Emerging Market Business Cycles: The Cycle is the Trend," *Journal of Political Economy*, Vol. 115, No. 1, pp. 69–l02.

Andrle, M., 2013a, "Estimation of Large-Scale Nonlinear Models using Bayesian Simulated Method of Moments with System Priors," International Monetary Fund, mimeograph, March.

———, 2013b, "Understanding DSGE Filters in Forecasting and Policy Analysis," Working Paper 13/98, International Monetary Fund, Washington DC.

Andrle, M., and J. Beneš, 2013, "System Priors: Formulating Priors about DSGE's Models' Properties," Working Paper 13/257, International Monetary Fund, Washington DC.

Andrle, M., and J. Bruha, 2014, "Learning about Monetary Policy using (S)VARs? Some Pitfalls and Some Solutions," Techn. rep., Czech National Bank, http://michalandrle.weebly.com/uploads/1/3/9/2/13921270/andrlebruha_svar_interim_2014_v03.pdf.

Andrle, M., Ch. Freedman, R. Garcia-Saltos, D. Hermawan, D. Laxton, and H. Munandar, 2009, "Adding Indonesia to the Global Projection Model," Working Paper 09/253, International Monetary Fund, Washington DC.

Andrle, M., G. Ho, and R. Garcia-Saltos, 2013, "The Role of Domestic and External Shocks in Poland: Results from an Agnostic Estimation Procedure," IMF Working Paper Series 13/220, Interanational Monetary Fund.

Andrle, M., and M. Plašil, 2016, "System Priors for Econometric Time Series," Working Paper 16/231, International Monetary Fund, Washington DC.

Andrle, Michal, Jan Bruha, and Serhat Solmaz, 2016, "Output and Inflation Co-Movement: An Update on Business-Cycle Stylized Facts," IMF Working Papers 16/241, International Monetary Fund.

Benes, J., and P. N'Diaye, 2004, "A Multivariate Filter for Measuring Potential Output and the NAIRU: Application to the Czech Republic," IMF Working Paper Series 4/45, International Monetary Fund.

Beneš, J., A. Binning, and K. Lees, 2008, "Incorporating Judgment with DSGE Models," Techn. rep., Reserve Bank of New Zealand, DP-2008/10.

Bruha, Jan, B. Pierluigi, and R. Serafini, 2011, "Euro Area Labor Markets – Different Reaction to Shocks?" ECB Working Paper Series 1284, European Central Bank.

Burns, A.F., and W.C. Mitchell, 1946, *Measuring Business Cycles* (New York: NBER).

Canlin, L., A. Meldrum, and M. Rodriguez, 2017, "Robustness of long-maturity term premium estimates," Techn. rep., FEDS Notes, April 03.

Canova, F., 2014, "Bridging DSGE Models and the Raw Data," *Journal of Monetary Economics*, Vol. 67, pp. 1–15.

Carabenciov, I., I. Ermolaev, Ch. Freedman, M. Johnson, M. Juillard, O. Kamenik, D. Korshunov, and D. Laxton, 2008a, "A Small Quarterly Projection Model of the US Economy," Working Paper 08/278 (December), International Monetary Fund.

Carabenciov, I., I. Ermolaev, Ch. Freedman, M. Juillard, O. Kamenik, D. Korshunov, D. Laxton, and J. Laxton, 2008b, "A Small Quarterly Multi-Country Projection Model," Working Paper 08/279, International Monetary Fund, Washington DC.

Chen, M., D. Carlson, A. Zaas, Ch. Woods, G.S. Ginsburg, A. Hero-III J. Lucas, and C. Lawrence, 2008, "The Bayesian Elastic Net: Classifying Multi-Task Gene-Expression Data," Techn. rep., Duke University mimeograph.

Chernozhukov, V., and H. Hong, 2003, "An MCMC approach to classical estimation," *Journal of Econometrics*, , No. 115, pp. 293–346.

Christiano, L.J., and T.J. Fitzgerald, 1999, "The Band Pass Filter," Techn. rep., NBER WP No. 7257.

Doran, H.E., 1992, "Constraining Kalman Filter and Smoothing Estimates to Satisfy Time-Varying Restrictions," *Review of Economics and Statistics*, Vol. 74, pp. 568–572.

Durbin, J., and S.J. Koopman, 2012, *Time Series Analysis by State Space Methods* (Oxford Statistical Science Series (38): Oxford Univ. Press).

Engle, R.F., 1974, "Band Spectrum Regression," *International Economic Review*, Vol. 15, No. 1, pp. 1–ll.

Forni, M., M. Hallin, M. Lippi, and L. Reichlin, 2005, "The Generalized Dynamic Factor Model: One-Sided Estimation and Forecasting," *Journal of the American Statistical Association*, Vol. 100, No. September, pp. 830–840.

Franta, M., 2016, "Iterated Multi-Step Forecasting with Model Coefficients Changing Accross Iterations," Czech National Bank Working Paper No. 5/2016.

Grether, D.M., and M. Nerlove, 1970, "Some Properties of 'Optimal' Seasonal Adjustment," *Econometrica*, Vol. 38, No. 5, pp. 682–703.

Harvey, A.C., 1989, *Forecasting, structural time series models, and the Kalman filter* (Cambridge: Cambridge Univ. Press).

Haywood, J., and G. Tunnicliffe-Wilson, 1997, "Fitting Time Series Models by Minimizing Multistep-Ahead Errors: A Frequency Domain Approach," *Journal of the Royal Statistical Society. Series B (Methodological)*, Vol. 59, No. 1, pp. 237–254.

Herbst, E.P., and M. Bognanni, 2015, "Estimating (Markov-Switching) VAR Models without Gibbs Sampling: A Sequential Monte Carlo Approach," Techn. rep., Finance and Economics Discussion Series, Federal Reserve Board, Washington, D.C.

Herbst, E.P., and F. Schorfheide, 2014, "Sequential Monte Carlo Sampling for DSGE Models," *Journal of Applied Econometrics*, Vol. 15, No. 1, pp. 1–ll.

Kapetanios, G., S. Price, and K. Theodoridis, 2015, "A New Approach to Multi-Step Forecasting Using Dynamic Stochastic General Equilibrium Models," *Economics Letters*, Vol. 136, pp. 237–342.

Kim, D.H., and J.H. Wright, 2005, "An Arbitrage-Free Three-Factor Term Structure Model and the Recent Behavior of Long-Term Yields and Distant-Horizon Forward Rates," Techn. rep., Finance and Economics Discussion Series, Federal Reserve Board, Washington, D.C.

Litterman, R., 1986, "Forecasting with Bayesian Vector Autoregressions – Five Years of Experience," *Journal of Business and Economic Statistics*, , No. 4, pp. 25–38.

Park, T., and G. Casella, 2008, "The Bayesian Lasso," *Journal of American Statistical Association*, Vol. 103, No. June (482), pp. 681–686.

Schorfheide, F., 2005, "VAR Forecasting Under Misspecification," *Journal of Econometrics*, , No. 128, pp. 99–136.

Stock, J.H., and M.W. Watson, 2002, "Forecasting Using Principal Components from a Large Number of Predictors," *Journal of the American Statistical Association*, Vol. 97, pp. 147–162.

———, 2012, "Disentangling the Channels of the 2007–09 Recession," Brookings papers on economic activity, The Brookings Institution.

Tiao, G.C., and D. Xu, 1993, "Robustness of Maximum Likelihood Estimates for Multi-step Predictions: The Exponential Smoothing Case." *Biometrika*, Vol. 80, No. 3, pp. 623–644.

Tonner, J., and J. Bruha, 2014, "The Czech Housing Market Through the Lens of a DSGE Model Containing Collateral-Constrained Households," Czech National Bank Working Paper No. 9/2014.

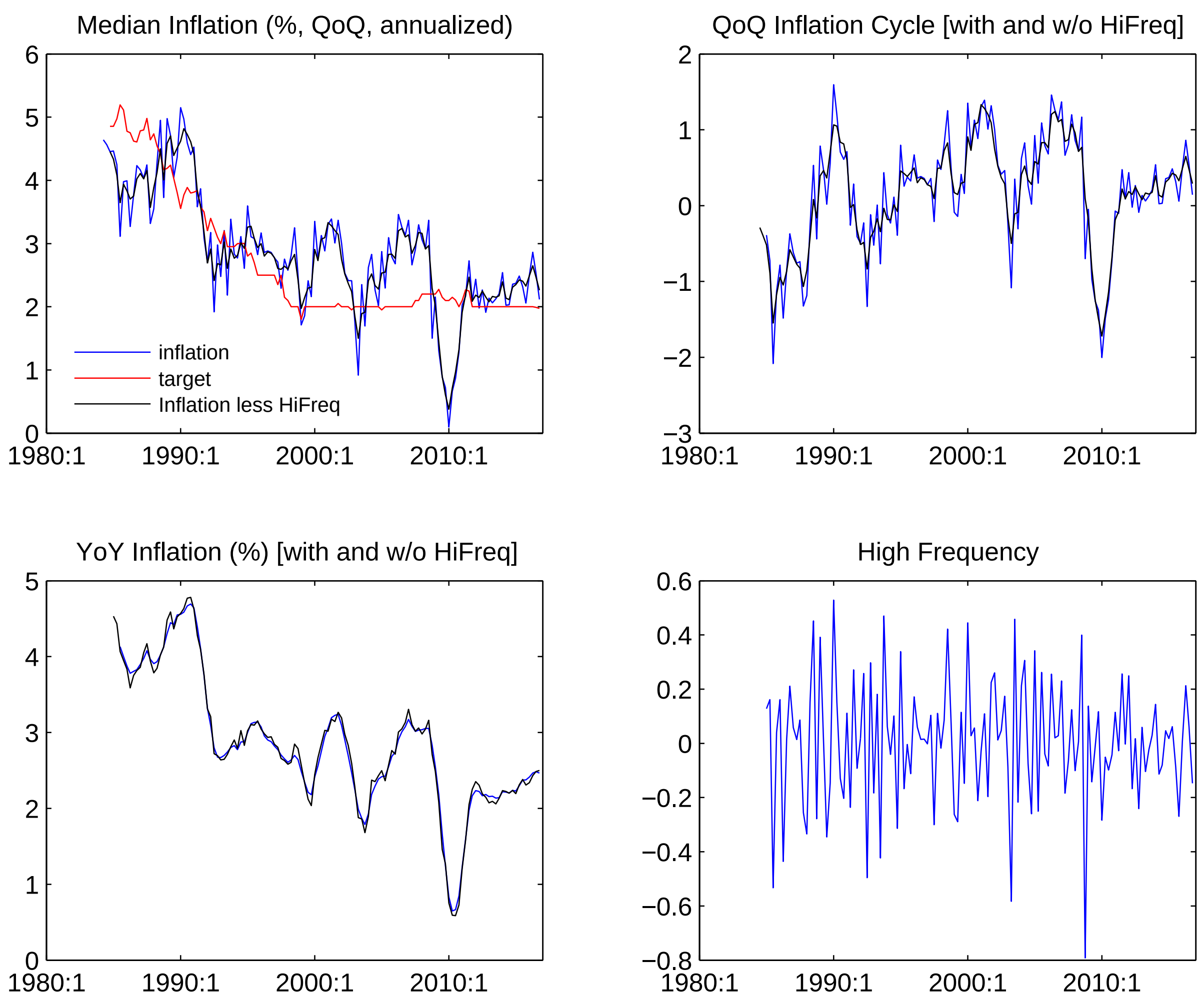
Villani, M., 2009, "Steady-state Priors for Vector Autoregressions," *Journal of Applied Econometrics*, Vol. 24, No. 4, pp. 630–650.

Y., Xia., and H. Tong, 2011, "Feature Matching in Time Series Modeling," *Statistical Science*, , No. 26, pp. 1–26.

VI. APPENDIX

VII. ADDITIONAL RESULTS FOR THE BASELINE U.S. TC-BVAR MODEL

Figure 7. Estimated Components of Inflation



VIII. Estimation Techniques

A. Computing Posterior Mode, with and without Homotopy

One approach to inference for VARs with system priors is to compute the posterior mode and then make use of it when employing inspection of the posterior distribution. Also, estimating the posterior mode can be viewed as a non-Bayesian constrained likelihood estimation, so common in engineering and applied econometrics.

The problem (30) can be rewritten as

$$\log p(\boldsymbol{\theta}|\mathbf{Y}) \propto \lambda \left[\log L(\mathbf{Y}|\boldsymbol{\theta}) + \log S(\mathbf{Z}|\theta)\right] + \log p(\boldsymbol{\theta}),\ \lambda \in [0,1] \tag{35}$$

where λ is a parameter of a continuation (homotopy) problem.

For $\lambda = 1$, the numerical optimization of the posterior is very direct. For complex models, in principle, one could start with a small weight given to the data and/or to system priors and initiate the model from the mode of the prior distribution.Note, where not for the fact that the TC-BVAR model features unobserved trends, a regular BVAR model could use system priors very differently, as the likelihood for a regular VAR is trivial to obtain, and thus one can proceed in reverse, from the likelihood to incorporation of prior information, see [Andrle and Plasil, 2017].

Explicit optimization of (35) is especially feasible for small and medium-sized VAR models, despite the number of parameters growing fast with the number of variables and the lag length. Yet, in comparison with sampling techniques for Bayesian inference and room for parallelization of the optimization, computing gradients, for instance, we view this direct approach as a prime candidate for employment of system priors in practice for small- to medium-sized models.

The posterior mode $\widehat{\boldsymbol{\theta}}_{pm}$ and the inverse Hessian matrix $\widehat{\boldsymbol{\Sigma}}_{pm}$ evaluated at the posterior mode can be used directly for analysis and inference, or for initializing a posterior sampler. **?**, for instance, use $\{\widehat{\boldsymbol{\theta}}_{pm}, \widehat{\boldsymbol{\Sigma}}_{pm}\}$ in multivariate t-distribution to create a proposal distribution for importance sampling algorithm for the posterior of a dynamic stochastic general equilibrium (DSGE) model.

Due its ease of implementation, computing the prior mode and analysis of the posterior distribution using *importance sampling* presents a viable option, despite the well-known drawbacks of the importance sampling, namely with high-dimensional parameter vector $\boldsymbol{\theta}$. Given the posterior mode estimate and $\widehat{\boldsymbol{\Sigma}}$, a t-distribution or Gaussian mixture proposal is 'embarrassingly parallel' and requires little tuning of sampling hyper-parameters, compared with other, more efficient, sampling methods. It is also easy to assess the quality of the posterior distribution approximation, using the measures like perplexity, or effective sample size. See the Appendix for detail of the algorithm.

B. Sequential Monte Carlo Approach to Approximate Posterior

One way of proceeding with the analysis of a VAR with system priors is to follow a set of methods known as sequential Monte Carlo methods. The idea, again, is to start with a relatively simple problem and move towards the actual problem of interested that is more challenging to solve. This time, again, the starting point is sampling from the prior distribution, not the finding the posterior mode. Sequential Monte Carlo (SMC) methods are, in principle, better designed to deal with multi-modal posterior distribution, unlike the importance sampling initiated at one of the posterior modes.

The principle of the approach, is in sync with the previous section on finding the posterior mode with a homotopy, starting from the prior distribution. Sampling from the prior distribution is easy and fast and the information from the system priors and the likelihood function is added gradually to update the proposal distribution.

Sequential Monte Carlo methods are related to importance sampling, as this is the first step. In general, one creates a set of distributions, $\{\pi_n\}$ for $n = 0, \ldots, \mathcal{N}$ as target distributions, starting from π_0 that is easy to sample from. Usually, this is the prior distribution. Let the sequence of target distributions to approximate be given by

$$\pi_n = p(\boldsymbol{\theta}|\mathbf{Y})^{\lambda_n} p(\boldsymbol{\theta})^{(1-\lambda_n)} \qquad \text{with} \qquad \lambda_0 = 0, \lambda_{\mathcal{N}} = 1,\ n = 0, \ldots, \mathcal{N}. \tag{36}$$

The design implies that $\pi_0 = p(\boldsymbol{\theta})$, the marginal prior, and that at the final stage $\pi_{\mathcal{N}}$ should be equal to the posterior distribution, $p(\boldsymbol{\theta}|\mathbf{Y})$. The idea is that for a given 'tempering schedule', $\{\lambda_n\}$, the sequence of distribution is gradual and that the sampled particles approximate well the new distribution. The profile of $\{\lambda_n\}$ is particularly important, as if λ increases too quickly, the likelihood and system priors dominate the marginal priors distribution too fast and the very reason for sequential updating gets lost as the particle impoverishment can be significant.

Details of the algorithm are described in the appendix, draw from the literature, in particular we follow closely Herbst and Schorfheide (2014). The following is just a recap for readers familiar with SMC methods. After an initial draw from a distribution that is easy to draw from, an importance sampling step is carried out to obtain a set of particles $\boldsymbol{\theta}_s$ and their weights W_s. A rejuvenation step follows, where the particles are resampled if the measure of the effective sample size, measuring a quality of the approximation, gets below a certain threshold. The latest step for a given n is a move, where a suitable Markov Chain (MC) kernel, $\mathcal{K}(.,.)$, is used to propagate each particle to ameliorate the degeneracy of particles. The details of the algorithm can be found in the appendix, as it is a straightforward extension of Herbst and Schorfheide (2014).

IX. MONTE-CARLO COMPUTATIONS

A. Importance Sampling Algorithm

1. Use numerical optimization, possibly with a continuation (homotopy) to estimate a posterior mode $\widehat{\boldsymbol{\theta}}_{pm}$ and compute the inverse of Hessian $\widehat{\boldsymbol{\Sigma}}_{pm}$ evaluated at the posterior mode

2. Denote $\upsilon(\boldsymbol{\theta})$ a multivariate t-distribution with a covariance matrix $c \times \widehat{\boldsymbol{\Sigma}}_{pm}$

3. Sample $j = 1, \ldots, N$ draws from $\upsilon(\boldsymbol{\theta})$. An j-th draw is denoted $\boldsymbol{\theta}^{(j)}$

4. Compute the importance weights, $\bar{\omega}_j$, and normalized importance weights, ω_j, using

$$\bar{\omega}_j = \frac{p(\boldsymbol{\theta}|\mathbf{Y},\mathbf{Z})}{\upsilon\left(\boldsymbol{\theta}^{(j)}\right)} = \frac{L(\mathbf{Y}|\boldsymbol{\theta}^{(j)}) \times S(\mathbf{Z}|\boldsymbol{\theta}^{(j)}) \times p(\boldsymbol{\theta}^{(j)})}{\upsilon\left(\boldsymbol{\theta}^{(j)}\right)}, \quad \omega_j = \frac{\bar{\omega}_j}{\sum_{j=1}^{N} \bar{\omega}_j}$$

5. For a function $h(\boldsymbol{\theta})$, approximate the posterior expected value by

$$h(\boldsymbol{\theta}) = \sum_{j=1}^{N} \omega_j \times h(\boldsymbol{\theta}^{(j)})$$

Defficiencies of the standard importance sampling algorithm are rather well known. If the proposal distribution $\upsilon(.)$ is not sufficiently close to the target distribution, the posterior distribution in this case, the approximation can be poor. The larger is the dimension of the parameter vector $\boldsymbol{\theta}$, the more challenging can be the use of importance sampling.

B. Sequential Monte Carlo Algorithms

Sequential Monte Carlo (SMC) algorithms are variants of the importance sampling algorithm, where the idea is to start from a distribution that is easy to draw from, often the prior, and to sequentially develop improved approximation to final target distribution.

We build a sequence of distributions $\{\pi_n\}$ for $n = 0, \ldots, \mathcal{N}$, such as

$$\pi_n = p(\boldsymbol{\theta}|\mathbf{Y})^{\lambda_n} p(\boldsymbol{\theta})^{(1-\lambda_n)} \qquad \text{with} \qquad \lambda_0 = 0, \lambda_{\mathcal{N}} = 1,\ n = 0, \ldots, \mathcal{N}. \tag{37}$$

The algorithm is as follows, drawing from **?**:

1. INITIALIZATION:

   Set $\lambda_0 = 0$ and draw J initial particles from the marginal prior, $p(\boldsymbol{\theta})$:

$$\boldsymbol{\theta}_1^j \sim p(\boldsymbol{\theta}), \qquad W_1^j = 1, \qquad j = 1, \ldots, J.$$

2. LOOP: For $n = 2, \ldots, \mathcal{N}$,

   (a) **Correction step:**

   Reweight particles from the previous stage by updating the weights to reflect distance between the distributions

$$\begin{aligned} w_n^j &= \frac{p(\boldsymbol{\theta}_{n-1}^j|\mathbf{Y})^{\lambda_n} p(\boldsymbol{\theta}_{n-1}^j)^{(1-\lambda_n)}}{p(\boldsymbol{\theta}_{n-1}^j|\mathbf{Y})^{\lambda_{n-1}} p(\boldsymbol{\theta}_{n-1}^j)^{(1-\lambda_{n-1})}} = \left[L(\mathbf{Y}|\boldsymbol{\theta}_{n-1}^j)\times S(\mathbf{Z}|\boldsymbol{\theta}_{n-1}^j)\right]^{\lambda_n-\lambda_{n-1}}, \\ W_n^j &= \frac{w_n^j W^{n-1j}}{\frac{1}{J}\sum_{j=1}^J w_n^j W_{n-1}^j} \qquad \text{for } j = 1, \ldots, J. \end{aligned}$$

   (b) **Selection step:**

   Compute the measure of the approximation quality based on weights' variance, the so called effective sample size, ESS,

$$ESS = J \times \left[\frac{1}{J}\sum_{j=1}^{J}(W_n^j)\right]^{-1}.$$

   If $ESS < J/2$, resample the particles using multinomial or other suitable resampling technique, based on the particle system $\{\boldsymbol{\theta}_{n-1}^j, W_{n-1}^j\}$. Denote the resampled particles $\widehat{\boldsymbol{\theta}}_n^j$ and $W_n^j = 1$, $\forall j$.

   If $ESS \geq J/2$, simply keep the current particle system and set $\widehat{\boldsymbol{\theta}}_n^j = \boldsymbol{\theta}_{n-1}^j$, with W_n^j left unchanged.

   (c) **Mutation step:**

   Propagate the candidate particles, $\{\widehat{\boldsymbol{\theta}}_n^j, W_n^j\}$, using M steps of Metropolis-Hasting algorithm with transition kernel $\boldsymbol{\theta}_n^j \sim K_n(\boldsymbol{\theta}_n^j|\boldsymbol{\theta}_n^j;\xi)$, where ξ is a vector of parameters of the kernel. The kernel and the M-H step are detailed below.

3. STOP:

When $n = \mathcal{N}$ and $\lambda_n = 1$, use the particle system $\{\boldsymbol{\theta}^j_{\mathcal{N}}, W^j_{\mathcal{N}}\}$ to compute the expectation of a funcion $h(\boldsymbol{\theta})$ w.r.t. $\pi_{\mathcal{N}} = p(\boldsymbol{\theta}|\mathbf{Y}, \mathbf{Z})$ using

$$\widehat{h}(\boldsymbol{\theta}) = \sum_{j=1}^{J} h(\boldsymbol{\theta}^j_{\mathcal{N}}) \times W^j_{\mathcal{N}}.$$