ESTIMATING DSGE-MODEL-CONSISTENT TRENDS FOR USE IN FORECASTING

Jean-Philippe Cayen, Marc-André Gosselin, and Sharon Kozicki*

Bank of Canada

Abstract: The workhorse DSGE model used for monetary policy evaluation is designed to capture business cycle fluctuations in an optimization-based format. It is commonplace to log-linearize models and express them with variables in deviation-from-steady-state format. Structural parameters are either calibrated, or estimated using data pre-filtered to extract trends. Such procedures treat past and future trends as fully known by all economic agents or, at least, as independent of cyclical behaviour. With such a setup, in a forecasting environment it seems natural to add forecasts from DSGE models to trend forecasts. While this may be an intuitive starting point, efficiency can be improved in multiple dimensions. Ideally, behaviour of trends and cycles should be jointly modeled. However, for computational reasons it may not be feasible to do so, particularly with medium- or large-scale models. Nevertheless, marginal improvements on the standard framework can still be made. First, pre-filtering of data can be amended to incorporate structural links between the various trends that are implied by the economic theory on which the model is based, improving the efficiency of trend estimates. Second, forecast efficiency can be improved by building a forecast model for model-consistent trends. Third, decomposition of shocks into permanent and transitory components can be endogenized to also be model-consistent. This paper proposes a unified framework for introducing these improvements. Application of the methodology validates the existence of considerable deviations between trends used for detrending data prior to structural parameter estimation and model-consistent estimates of trends, implying the potential for efficiency gains in forecasting. Such deviations also provide information on aspects of the model that are least coherent with the data, possibly indicating model misspecification. Additionally, the framework provides a structure for examining cyclical responses to trend shocks, among other extensions.

JEL: E3 Prices, Business Cycles, and Fluctuations; E52 Monetary Policy; C32 Time Series Models.

Keywords: Bayesian estimation, New Keynesian model.

*Authors’ addresses are: Jean-Philippe Cayen, Senior Analyst, jcayen@bank-banque-canada.ca; Marc-André Gosselin, Principal Researcher, mgosselin@bank-banque-canada.ca; Sharon Kozicki, Research Director, skozicki@bank-banque-canada.ca; All authors are in the Research Department, Bank of Canada, 234 Wellington St, Ottawa ON K1A 0G9, Canada. Thanks to Stephen Murchison for useful comments and suggestions. The views expressed in this paper are those of the authors. No responsibility for them should be attributed to the Bank of Canada. Draft: October 2, 2008.
1 Introduction and Motivation

"...for any degree of theoretical coherence, the degree of empirical coherence should be maximized...” Pagan (2003a)

Recognizing the important insights of the Lucas critique, the standard approach to monetary policy evaluation, including assessment of “optimal monetary policy”, uses dynamic stochastic general equilibrium (DSGE) model specifications. Such specifications embed optimizing behaviour of economic agents with rational expectations in a stochastic environment and enable separation of structural parameters of the model economy from policy responses. In the typical implementation, sufficient frictions are added to the economic structure so that the model simulations are able to replicate key data properties such as correlations and impulse responses.

In the context of policy evaluation, it is most common to express DSGE models in a format to only accommodate business cycle fluctuations. Log-linearized versions of the models describe the evolution of the deviation of macro variables from their “steady states” and, if estimated, the corresponding real-world data is usually detrended. Trend behaviour is generally not at all addressed in model simulations, with trends implicitly treated as fully known by all agents or, at least, treated as independent of the cyclical behaviour.

With a shift in the use of DSGE models, from policy evaluation to forecasting, trend behaviour becomes considerably more important. For instance, the typical DSGE model may be able to generate a consistent forecast of the level of activity relative to trend, but there will generally also be interest in the future trend itself. While deterministic treatments of trends as in Smets and Wouters (2007) and Dib, Gammoudi, and Moran (2006) are consistent with independence of trend and cycle and also straightforward to forecast at any horizon, they are inconsistent with the large empirical literature that

---

1 See for instance, the calibrated model specifications of Woodford (2003).

2 An early example is Rotemberg and Woodford (1997).

3 The Bank of Canada and the central banks of Norway and Sweden have started to use DSGE models for forecasting purposes.
generally supports stochastic trends or, in the very least, deterministic specifications with unanticipated real-time structural breaks. With stochastic trends, an additional challenge is that the latest structural shocks (deviations of the most recently published data from their projected values as in the last projection) will be a combination of unobserved structural trend shocks and unobserved cyclical disturbances. Trend estimation is critical in this setup because it determines the nature of the starting-point shocks to be explained by the DSGE model.

One drawback of the standard methodology of pre-detrending data is that trends are estimated outside of the model using deterministic specifications or arbitrary filtering techniques, implying that the decomposition between trend and cyclical shocks is subjective. A second disadvantage is that this approach does not fully exploit common trend restrictions implied by the economic theory on which the model is based, introducing inefficient trend estimates. A third drawback is that trends used during estimation of structural model parameters may not coincide with trends that maximize consistency of the model with the data. Indeed, when structural parameters are calibrated or estimated using Bayesian techniques with detrended data, it will generally be the case that trends used for detrending data will deviate from model-consistent trends, leading to inefficient forecasts.

Thus, one objective of this project is to endogenize the treatment of trends in DSGE structures. In the proposed approach, we layer structural trends on the solution of a log-linearized DSGE model, cast the system in its state-space representation, and perform the stochastic detrending of the model variables with the Kalman filter. While the

---

4 Stochastic trends or unexpected structural breaks for a wide set of macroeconomic variables are well documented in the literature. See Demers (2003), Lalonde et al. (2003), Roberts (2001), and Laubach and Williams (2003), among others. Some researchers have begun to introduce stochastic trends in their DSGE forecasting models. For instance, Edge, Kiley, and Laforte (2008) estimate a model that assumes a stochastic trend process for technology. See also Adolfson et al (2007).

5 A misspecified trend model may lead to a bad estimation of the trend-gap decomposition at the end of history, which may in turn result in poor forecast accuracy. For example, Cayen and van Norden (2005) show that the revisions of real-time estimations of the output gap for Canada can be large and highly persistent, in particular when potential output is modelled as a deterministic trend.

6 Note that the methodology is sufficiently general to be applied to other types of forecasting models that rely on "stationarized" data, such as VARs.
methodology allows for the simultaneous estimation of trends and structural parameters of the DSGE model, this project takes the structural model parameters as given, either from a calibration exercise or estimated in a first step, and focuses instead on estimation of variances of permanent and transitory structural shocks given structural parameters. As opposed to the literature that seeks to modify the DSGE structure so as to allow for one or two stochastic trends in the model, our starting point is to estimate trends while taking the DSGE structure as given.

The methodology allows for the joint estimation of the variance of permanent stochastic shocks for multiple series and provides a structure to introduce links between the various trends that are implied by the economic theory on which the DSGE model is based, improving the efficiency of trend estimates. For instance, if trends in series are treated individually, then permanent shocks to, say, productivity won’t automatically influence other real variables such as consumption. Cross-trend restrictions can ensure that trends that theory says should move together, will move together in forecasts and simulations. As another example, with cross-trend restrictions, an exogenous increase in, for example, the commodity price trend should generally require a decrease in the trend of non-commodity prices if monetary policy is targeting price stability in the long run. The Kalman structure also provides a forecasting model for model-consistent trends, leading to greater forecast efficiency, and provides a framework for estimating model-consistent decompositions of shocks into permanent and transitory components.

The next section proposes a methodology for joint modeling of trends and cycles in a DSGE framework. The approach is applied to a small estimated DSGE model from the literature in section 3. Although the main benefits of the methodology accrue to medium-

7 Joint estimation of the unobserved trends and unobserved transitory structural shocks is likely to be plagued with identification problems, an issue that will be discussed in more detail later.

8 For instance, Edge, Laubach and Williams (2003) introduce trends in sector-specific productivity processes such that the relative price of investment becomes non-stationary and real investment and consumption can grow at different rates. Similarly, Chang, Doh, and Schorfeide (2006) present a model in which hours worked have a stochastic trend generated by a non-stationary labour-supply shock.

9 Of course, misspecification of the DSGE model could result in misleading conclusions regarding the relative importance of the trend versus the cycle.
to large-scale models where multiple variables share common trends, this application shows how the approach may provide information on model misspecification and guidance on means to improve forecast performance. The final section concludes with a discussion of future work, which will focus on the application of the methodology to the Bank of Canada’s principal projection model of the Canadian economy, ToTEM.10

2 Methodology

The issue with the standard detrending approach is that trends used for detrending data will generally deviate from model-consistent trends, especially when models are calibrated or estimated using Bayesian techniques with detrended data. Consider the following simple example. In a first step, data, y, is detrended using trend, μ, to generate the detrended series, \( \tilde{y} = y - \mu \), which, without loss of generality, is also assumed to have a sample mean of zero. In the second step, model parameters are estimated using detrended data. Suppose the model is \( \tilde{y} = X\beta + \epsilon \), where β is the parameter to be estimated, X contains the regressors (assumed to have a sample mean of zero to simplify the analysis for this example), and \( \epsilon \) are iid \( N(0, \sigma^2) \) errors. The least squares estimate of \( \beta \) is \( \hat{\beta} = (X'X)^{-1}X'\tilde{y} \). The Bayesian estimate of \( \beta \) with prior \( N(b, s^2) \) is \( \hat{\beta} = wb + (1 - w)\hat{\beta} \) where \( w = \sigma^2/(s^2 + \sigma^2) \). Solving backwards, the trend that is consistent with the least-squares estimated model (\( \tilde{y} = X\hat{\beta} \)) is \( \hat{\mu} = y - X\hat{\beta} = \mu \), the trend used to initially detrend the data y. By contrast, the trend that is consistent with the Bayesian-estimated model (\( \tilde{y} = X\hat{\beta} \)) is \( \hat{\mu} = y - X\hat{\beta} = \mu + wX(b - \hat{\beta}) \). Thus, unless the prior mean used during estimation is equal to the least-squares estimate of \( \beta \), the trend that is consistent with the Bayesian-estimated model will deviate from the trend used to estimate the model. In a forecasting environment, it would therefore be inefficient to use forecasts of \( \mu \) combined with forecasts of \( \tilde{y} \) constructed using the Bayesian estimates of \( \beta \).

To address the issues mentioned in the introduction, we propose an approach to endogenize the treatment of trends in a DSGE structure. The starting point is the

---

backward-looking solution of a log-linearized DSGE model:

\[ \hat{x}_t = C_s \hat{x}_{t-1} + H_s u_t \]  

(1)

where \( \hat{x}_t \) is an \( n \times 1 \) vector of stationary variables, and \( E[u_t u'_t] = G_1 \) for \( t = \tau \) and \( 0_n \) otherwise. This model solution is assumed to include expressions for observable and unobservable \( \hat{x}_t \).

The \( n \) elements of \( \hat{x}_t \) can be divided into two subsets. One subset of \( n_1 \) elements of \( \hat{x} \), includes variables expressed in deviation-from-steady-state (or deviation-from-trend) format, where the steady states or trends, \( \mu_t \), may be time-varying and the level variable, \( x_{1,t} = \hat{x}_{1,t} + \mu_t \) is observable (possibly with measurement error). Defining the \( n_1 \times n \) selector matrix \( S_1 \) of 0s and 1s, \( \hat{x}_{1,t} = S_1 \hat{x}_t \). The remaining \( n - n_1 \) elements of \( \hat{x}_t \), \( \hat{x}_{2,t} = S_2 \hat{x}_t \), are possibly unobservable, stationary state variables, where \( S_2 \) is an \( n - n_1 \times n \) selector matrix. To simplify notation, we assume \( \hat{x}_{2,t} \) are zero mean. Note that \( S \equiv [S'_1 \quad S'_2]' \) satisfies \( S^{-1} = S' \).

It is useful to define the unobserved \( n_x \) “structural” components on which the steady states, \( \mu_t \), of the observable data, \( x_{1,t} \), depend. Let’s call these unobserved structural components \( \xi_{1,t} \). For instance, \( \xi_{1,t} \), may contain the inflation target series (or perceived inflation target series prior to the introduction of explicit numerical price objectives), the equilibrium real rate, etc. Assume that the \( n_1 \) observed steady states, \( \mu_t \), can be expressed as linear combinations of the \( n_x \) unobserved structural components, \( \xi_{1,t} \), as follows:

\[ \mu_t = Q_1 \xi_{1,t} \]  

(2)

where \( Q_1 \) is \( n_1 \times n_x \). Thus, for example, \( \mu_t \) may contain the steady-state of the nominal policy rate which, in turn, may be defined to be the sum of the inflation target and the equilibrium real rate, etc.

The unobserved structural components are assumed to evolve according to:

\[ \xi_{1,t} = T \xi_{1,t-1} + v_{1,t} \]  

(3)

with \( E[v_{1,t} v'_{1,t-1}] = G_2 \) if \( t = \tau \) and \( 0 \) otherwise. As noted by Harvey (1988) and others, this format is sufficiently general to include a variety of ARIMA processes.
All that is required to explicitly introduce trends in the DSGE model is to re-express the model solution for $\hat{x}_t$ replacing deviations, $\hat{x}_{1,t}$, with variables and steady states:

$$
S\hat{x}_t = SC_sS' \begin{bmatrix} S_1 \\ S_2 \end{bmatrix} \hat{x}_{t-1} + SH_su_t \\
= SC_sS' \begin{bmatrix} x_{1,t-1} - \mu_{t-1} \\ \tilde{x}_{2,t-1} \end{bmatrix} + SH_su_t
$$

(4)

After replacing the steady states with unobserved structural components of interest and rearranging, the following expression is obtained:

$$
\begin{bmatrix} x_{1,t} \\ \hat{x}_{2,t} \end{bmatrix} = Q_1\xi_{1,t} + SC_sS' \begin{bmatrix} x_{1,t-1} \\ \hat{x}_{2,t-1} \end{bmatrix} - SC_sS' \begin{bmatrix} 0 \\ 0 \end{bmatrix} + SH_su_t
$$

(5)

Note that $SC_sS'$ and $SH_s$ are easily partitioned to be conformable with $\hat{x}_{1,t}$ and $\hat{x}_{2,t}$:

$$
SC_sS' = \begin{bmatrix} SC_sS'_1 & SC_sS'_2 \\ SC_sS'_1' & SC_sS'_2' \end{bmatrix} \\
SH_s = \begin{bmatrix} S_1H_s \\ S_2H_s \end{bmatrix}
$$

(6)

Using this partition and augmenting the system to also include expressions for the evolution of $\xi_{1,t}$, given in (3), one obtains:

$$
\begin{bmatrix} x_{1,t} \\ \hat{x}_{2,t} \\ \xi_{1,t} \end{bmatrix} = \begin{bmatrix} S_1C_sS'_1 & S_1C_sS'_2 & Q_1T - S_1C_sS'_1Q_1 \\ S_2C_sS'_1 & S_2C_sS'_2 & -S_2C_sS'_1Q_1 \\ 0 & 0 & T \end{bmatrix} \begin{bmatrix} x_{1,t-1} \\ \hat{x}_{2,t-1} \\ \xi_{1,t-1} \end{bmatrix} \\
+ \begin{bmatrix} S_1H_s & Q_1 \\ S_2H_s & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} ut \\ v_{1,t} \end{bmatrix}.
$$

(7)

This system provides the transition (or state) equations in a state-space model.

It is convenient to represent the system using the notation of Hamilton (1994). Thus,
define:

\[
\xi_t = \begin{bmatrix} x_{1,t} \\ \hat{x}_{2,t} \\ \xi_{1,t} \end{bmatrix},
\]

\[
v_t = \begin{bmatrix} S_1 H_s & Q_1 \\ S_2 H_s & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} u_t \\ v_{1,t} \end{bmatrix},
\]

\[
F = \begin{bmatrix} S_1 C_s S_1' & S_1 C_s S_2' & Q_1 T - S_1 C_s S_1' Q_1 \\ S_2 C_s S_1' & S_2 C_s S_2' & -S_2 C_s S_2' Q_1 \\ 0 & 0 & T \end{bmatrix},
\]

\[
Q = \begin{bmatrix} S_1 H_s & Q_1 \\ S_2 H_s & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} G_1 & 0 \\ 0 & G_2 \end{bmatrix} \begin{bmatrix} S_1 H_s & 0 \\ S_2 H_s & 0 \\ 0 & I \end{bmatrix}', \tag{8}
\]

Measurement equations that relate the states to observable data, \( y_t \), are appended to complete the system:

\[
y_t = H' \xi_t + w_t, \tag{9}
\]

where \( w_t \) are interpreted as measurement errors. In some cases, elements of \( y_t \) may correspond exactly with observable elements of \( x_t \)—in which case the corresponding measurement error will have zero variance and the appropriate column of \( H \) will contain one entry equal to unity and remaining entries equal to zero. In addition, this structure is sufficiently general to incorporate the large information set approach of Boivin and Giannoni (2006) in which multiple series provide a noisy measure of the true underlying macro concept implicit in the structural model. Generalizations in which measurement errors \( w_t \) may follow an ARMA(p,q) structure are assumed to be captured through elements of \( \hat{x}_{2,t} \). Thus, \( w_t \) is assumed to be iid \( N(0, R) \). The approach is also flexible enough to accommodate exogenous sources of information or expert judgment to estimate and project the trends, such as an official inflation target series or announced government budgets.

With this notation, the state-space structure of the system can be seen to satisfy the format described in Hamilton (1994):
\[ y_t = H'\xi_t + w_t \]
\[ \xi_{t+1} = F\xi_t + v_{t+1} \]
\[ E[w_t v_t'] = Q \text{ for } t = \tau, 0 \text{ otherwise} \]
\[ E[w_t w_{t'}'] = R \text{ for } t = \tau, 0 \text{ otherwise} \]
\[ E[w_t v_{t'}] = 0. \quad (10) \]

In theory, all parameters of the system, including those relevant for the detrended model \((C_s, H_s, \text{ and } G_1)\) as well as those for the trends \((Q_1, T, \text{ and } G_2)\), can be estimated in one step using this structure. However, the inability to obtain analytical expressions for \(C_s\) in terms of the deep structural parameters may complicate estimation. Alternatively, computational challenges have generally led researchers to opt for Bayesian approaches to estimating deep structural parameters. And, in the case where the \(\hat{x}_{2,t}\) contain ARMA “structural” shocks, it may be difficult to separately identify the structural shock processes from the trends.

Given these difficulties, which are magnified for relatively large-scale models used in central banks, we proceed instead taking the deep structural parameters of the detrended model as given and apply (possibly Bayesian) maximum likelihood techniques to estimate the remaining parameters of the trend processes. Of course, the format is sufficiently general to accommodate known or calibrated entries in \(Q_1, T, \text{ or } G_2\), (to ensure consistency with the long-run elasticities implied by the DSGE model) and could be applied to other types of forecasting models, such as VARs.\(^{11}\) In this sense, we are estimating the trends that are consistent with the model specification for detrended data. Because the Bayesian estimation procedure implies that the means of the posterior distribution of the deep structural parameters of the detrended model may be influenced by both the priors and the data, as illustrated earlier, it will not necessarily be the case that the trend estimates we obtain will equal the original trends used to detrend the data. The larger the differences

\(^{11}\)This procedure of taking some parameters as given and estimating the remaining ones is not uncommon in standard DSGE setups. For example, steady-state parameters related to household’s discount rate and capital depreciation rates are typically calibrated rather than estimated. See for instance the estimated model of Edge, Kiley, and Laforte (2008).
between the two trends, the greater the potential for improvement from accounting for these
divergences during forecasting. In addition, an examination of the similarity of the two sets
of trends could be used to determine which model equations are less effective at explaining
historical variation in the data.

3 Application to a Small DSGE Model

The methodology is tested by applying it to a small DSGE models for which Bayesian
estimates of model parameters are available. As a starting point, we use the rational
expectations version of the simple model of Milani (2007). An advantage of this approach
is that the model can be taken as given, and trends can be estimated given the raw data
and the DSGE model structure.

The model contains equations for a quasi-differenced output gap \( \tilde{y}_t \), inflation \( \pi_t \), the
output gap \( y_t \), the policy rate \( i_t \), an AR(1) demand shock \( r^n_t \), and an AR(1) cost-push
shock \( u_t \). The monetary policy shock \( \epsilon_t \) is assumed to be white noise.

\[
\begin{align*}
\tilde{y}_t &= E_t \tilde{y}_{t+1} - (1 - \beta \eta) \sigma [i_t - E_t \pi_{t+1} - r^n_t] \\
\pi_t &= (1 + \beta \gamma)^{-1} \{ \gamma \pi_{t-1} + \xi_p \omega y_t + [(1 - \beta \eta) \sigma]^{-1} \tilde{y}_t \} + \beta E_t \pi_{t+1} + u_t \\
y_t &= (1 + \beta \eta)^{-1} \{ \tilde{y}_t + \eta y_{t-1} + \beta \eta E_t y_{t+1} \} \\
i_t &= \rho i_{t-1} + (1 - \rho) [\psi_p \pi_t + \psi_y y_t] + \epsilon_t, \quad \epsilon_t \sim iid(0, \sigma^2_\epsilon) \\
r^n_t &= \phi^n r^n_{t-1} + v^n_t, \quad v^n_t \sim iid(0, \sigma^2_{r^n}) \\
u_t &= \phi^u u_{t-1} + v^u_t, \quad v^u_t \sim iid(0, \sigma^2_u) \\
\end{align*}
\]

The first equation is the log-linearized Euler equation for households where \( \beta \) is the
household’s discount factor, \( \eta \) is the habit persistence parameter, and \( \sigma \) measures the
intertemporal elasticity of substitution. The second equation is the Phillips curve that
arises from optimal Calvo price-setting with indexation for non-reoptimizing firms.

\(^{12}\)Our objective is not to criticize Milani’s model but to use his specification for experimenting purposes.
We use the RE version of his model since it is a good representation of a standard monetary DSGE model
from the literature. Note that Milani’s preferred model incorporates low frequency movements attributable
to learning by firms and consumers, however.
represents the degree of indexation to past inflation, $\xi_p$ is inversely related to the degree of price stickiness and $\omega$ denotes the elasticity of the marginal disutility of producing output with respect to an increase in output. The current output gap depends on lagged and expected output gaps, and on the \textit{ex ante} real interest rate. Monetary policy is described by the fourth equation, which is a Taylor rule with partial adjustment, where $\rho$ is the interest-rate smoothing term, and $\psi_x$ and $\psi_y$ are the feedback coefficients to inflation and the output gap. Milani estimates these structural parameters with likelihood-based Bayesian methods on U.S. data for output, inflation, and the nominal interest rate over the period 1960Q1-2004Q2. In Milani, inflation is defined as the annualized quarterly growth rate of the GDP implicit price deflator and the federal funds rate is used as the nominal interest rate. GDP is detrended using the Congressional Budget Office’s (CBO) measure of potential output whereas inflation and interest rates are simply expressed in deviation from their sample mean.

In the first stage of the estimation of model-consistent trends, the structural parameters are replaced by their estimated values and the model is reexpressed in its log-linear backward-looking solution, as in equation (1).$^{13}$ Parameter estimates, as provided by Milani (2007) are included in Table 1. Given these parameter estimates, the $C_s(5 \times 5)$ and $H_s(5 \times 3)$ matrices can be computed.$^{14}$ The model is also modified to accommodate low-frequency components (or trends) in the three observable variables. The trends of inflation and of the real interest rates are assumed to follow a random walk. The trend of output is also assumed to follow a random walk but with a stochastic time-varying growth component, as in Clark (1987) and Laubach and Williams (2003). The trend specification can be expressed in term of the matrices $Q_1$ and $T$, as in equations (2) and (3). Thus, the $\mu_t$ trends are assumed to be related to deeper unobserved structural components through the $3 \times 4$ $Q_1$ matrix as follows:

$^{13}$We use AIM (see Anderson and Moore (1985)) in Troll to compute the solution matrices of the log-linearized model of Milani.

$^{14}$The model we use contains 5 instead of 6 equations as we substitute out the quasi-differenced output gap.
\[
\begin{bmatrix}
\mu_\pi^t \\
\mu_y^t \\
\mu^r_t \\
\end{bmatrix} = 
\begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
1 & 0 & 0 & 1 \\
\end{bmatrix} 
\begin{bmatrix}
\xi_{1,t}^\pi \\
\xi_{1,t}^y \\
\xi_{1,t}^g \\
\xi_{1,t}^r \\
\end{bmatrix}
\]

Here, \(\mu_\pi^t\) is the trend inflation rate, \(\mu^i_t\) is the trend in the nominal interest rate, and \(\mu^y_t\) is the level of potential output. The trends of inflation and output are uniquely determined by their structural counterparts (\(\xi^\pi\) and \(\xi^y\), respectively), whereas the trend of the nominal policy rate is given by the sum of the inflation and real interest rate structural trends, \(\xi^\pi\) and \(\xi^r\), respectively. The evolution of the unobserved structural components \(\xi_1\) is described by the \(4 \times 4\) matrix \(T\):

\[
\begin{bmatrix}
\xi_{1,t}^\pi \\
\xi_{1,t}^y \\
\xi_{1,t}^g \\
\xi_{1,t}^r \\
\end{bmatrix} = 
\begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 \\
\end{bmatrix} 
\begin{bmatrix}
\xi_{1,t-1}^\pi \\
\xi_{1,t-1}^y \\
\xi_{1,t-1}^g \\
\xi_{1,t-1}^r \\
\end{bmatrix} + 
\begin{bmatrix}
v_{1,t}^\pi \\
v_{1,t}^y \\
v_{1,t}^g \\
v_{1,t}^r \\
\end{bmatrix}
\]

(random walk)

(random walk + drift)

(random walk)

(random walk)

(13)

Using time series data for inflation, interest rates and output, and given the model’s log-linear solution for \(C_s\) and \(H_s\), our assumptions on \(Q_1\) and \(T\), and Milani’s estimates of the variance of the dynamic \(u\) shocks (\(G_1\), \(\sigma_\varepsilon\), \(\sigma_r\), and \(\sigma_u\) in Milani), we estimate the variance of the permanent \(v_1\) shocks (\(G_2\)) by maximum likelihood and compute the Kalman filter (we assume no measurement error, i.e. \(R=0\)). Estimates of standard deviations of permanent structural shocks and standard errors of estimates are provided in the columns labeled "Trends Only" of Table 2, where standard deviations of transitory shocks are constrained to equal the posterior mean estimates obtained by Milani (2007). Most parameters are estimated fairly precisely. The standard deviation of permanent shocks to inflation is fairly low at 0.180, meaning that inflation target shocks are concentrated around their average. However, the standard deviation of permanent shocks to the real interest rate is high at 0.800, especially when compared to the standard deviation of transitory monetary policy shocks. Although we obtain a sensible estimate of potential output, we are not able to
identify permanent shocks to the level component.\textsuperscript{15}

Figures 1 to 3 report actual and model-consistent trend values for inflation, output and real interest rates under various specifications.\textsuperscript{16} The estimated trend of inflation (labeled "Trend" in Fig. 1) is smooth and broadly consistent with Kozicki and Tinsley (2005), gradually increasing during the 1960s and early 1970s before peaking slightly above 5 per cent in the early 1980s. The inflation target then starts to decline very slowly towards 2.5 per cent at the end of history. Our finding of a sizable gap between inflation and its model-consistent trend means that the model of Milani is able to explain a substantial portion of the cyclical variations in inflation. However, Milani’s rational expectations model misses the low frequency movements in inflation. The methodology is efficient in recovering the model’s implicit trend in the case of output since our estimate of detrended output is very close to the CBO’s estimate of the output gap that was used by Milani (labeled "Gap" in Fig. 2).

As opposed to Laubach and Williams (2003) and Clark and Kozicki (2005), we find that the real interest rate trend (labeled "Trend" in Fig. 3) follows the data very closely, indicating that real rate movements are dominated by trend variations.\textsuperscript{17} This leaves only a small portion of interest rate fluctuations to be explained by the Taylor rule and raises the possibility that the policy rule in Milani’s model may be misspecified. For instance, the coefficients in the Taylor rule are fixed throughout the period, which may not be realistic given monetary policy regime shifts over the sample. In addition, the assumption that monetary policy shocks are white noise may not be reasonable given persistent interest-rate deviations from the Taylor rule seen over history.\textsuperscript{18} The trend estimate of the real rate

\textsuperscript{15}Maximum likelihood estimates of the standard deviation of the innovations of a variable modelled as a unit-root are often biased towards 0, owing to the pile-up problem discussed in Stock and Watson (1998).

\textsuperscript{16}The mean of inflation and real rates is also reported to reflect the model with constant trends. The Kalman filter can produce a one-sided "filter" estimate and a two-sided "smooth" estimate. We report the "filter" estimate, as it is closer to the trend that would be used in a real-time forecasting exercise.

\textsuperscript{17}This finding is in line with the volatile DSGE estimates of the natural rate from Edge, Kiley, and Laforte (2007) and Neiss and Nelson (2003).

\textsuperscript{18}In fact, the estimated monetary policy shocks are autocorrelated and skewed if we use the CBO output gap and assume constant trends for inflation and nominal interest rates. Taking this into account, English, Nelson and Sack (2002) estimate a monetary policy rule that incorporates temporary but persistent...
probably captures the reaction of the Federal Reserve to exogenous economic events that are not captured by Milani’s model (such as financial market turmoil). Those issues could explain why our interest rate trend estimate appears to capture high- rather than low-frequency movements.

Given the possible confusion between persistent transitory shocks and permanent shocks, we estimate the standard deviation of all shocks simultaneously (see the columns labeled "Trends and Transitory" of Table 2). In this joint estimation exercise, the standard deviation of all transitory shocks declines, as does the standard deviation of potential growth shocks. Compensating for these declines, the standard deviations of permanent shocks to the equilibrium real rate, to the level of potential output, and to the inflation target increase. This highlights the difficulty to identify cyclical and permanent shocks simultaneously, particularly in the case of the real interest rate where the variance of the monetary policy shock converges to zero. Looking at the series labeled "Trend (all)" in Figures 1 to 3, we can see that this joint estimation approach yields very similar results in terms of the real interest rate trend but generates an inflation target series that is more variable.

Because of the inability of the joint estimation to distinguish between monetary policy shocks and permanent shocks to the real interest rate, we estimate the same set of variances with the exception of the standard deviation of permanent shocks to the real interest rate, which we calibrate to the value obtained by Laubach and Williams (2003). The results are presented in the columns labeled "Trends and Transitory with LW estimates for \( \sigma_r \)" of Table 2. The decrease in the standard deviation of permanent shocks to the equilibrium real rate is compensated by increases in the standard deviations of all the transitory shocks and of the permanent shocks to the inflation target. The estimated standard deviation of the monetary policy shocks is still lower than Milani’s estimate, but it is now significantly higher than 0. Interestingly, the standard deviations of the other transitory shocks are now very close to the estimates obtained by Milani. Looking at the series labeled "Trend (LW)"

deviations from the Taylor rule for reasons other than interest rate smoothing.
in Figures 1 to 3, we can see that calibrating $\sigma_r$ produces an estimate of the real interest rate trend that is much more consistent with the literature, but this occurs at the expense of an inflation target series that is even more variable than if all variances are freely estimated. While calibrating remains a question of judgment, relative forecast performance could be used as a criterion for constraining certain variances.

To the extent that the model’s variables contain low frequency movements, incorporating stochastic trends should improve forecast accuracy. To verify this, we compute forecasts of inflation, real interest rates, and the output gap using Milani’s model and our estimates of the variance of permanent and temporary shocks (i.e. using the "Trends and Transitory" variance estimates from Table 2).\(^{19}\) Tables 3 to 5 report the forecast errors at different horizons for the version of the model with constant trends and the one with model-consistent trends. Since we do not know the data-generating process underlying the measure of potential output used by Milani, we cannot construct forecasts for the level of GDP. Therefore, we only consider the forecasting accuracy of CBO’s estimate of the output gap using Milani’s IS curve conditional on constant and model-consistent trend estimates of inflation and interest rates (i.e. we use the output gap directly and constrain the variance of $\xi^y$ and $\xi^g$ to zero in the forecasting exercise).

First, we notice that allowing the trends to vary over time helps to reduce the root mean squared forecast errors of inflation (Table 3). While the differences are small and statistically insignificant for the one quarter ahead forecast, they become much more important as the forecasting horizon increases, with the time-varying trend models clearly outperforming the constant trend model at the 8 quarter horizon. The superiority of the time-varying trends model is confirmed by the Diebold-Mariano tests. This result highlights the importance of accounting for low frequency movements in inflation in DSGE models. Doing so produces a measure of detrended inflation that is more consistent with the DSGE assumption of stationarity and a measure of trend inflation that is more consistent with

---

\(^{19}\)It is not a pure real-time forecasting exercise since we are not using real-time data and we are not re-estimating the parameters of the model at each quarter. The mean of inflation and the nominal interest rate used at the detrending stage (for the model with constant trends), is updated at each quarter, as would be the case in a real-time exercise.
the data. Allowing for a stochastic time-varying trend soaks up part of the persistence in inflation that would otherwise go into the gap and violate the DSGE assumption. By pinning down the long-run profile for the economy more accurately, our treatment of trends also leads to an improvement in longer-term forecasts.

The situation is somewhat different for the forecasts of real interest rates, since none of the two versions of the model systematically outperforms the other at the 1- and 4-quarter horizon (Table 4). Differences in accuracy are more significant for the 8-step ahead forecasts. In this case, the model with a stochastic real interest rate trend dominates by a strong margin over the 1976-1985 and 1996-2004 periods while the opposite is true for the 1963-1975 and 1986-1995 periods. Overall, both models seem to perform equally well in a forecasting environment. The benefits of allowing for a time-varying interest rate trend could be small because deviations of the real interest rate from its mean are relatively close to being stationary and therefore not inconsistent with the DSGE structure. This result might also reflect the possibility that Milani’s monetary policy rule is misspecified, as discussed above. Nevertheless, allowing for a stochastic interest rate trend produces more accurate forecasts of inflation. Indeed, in a separate exercise, we found that inflation forecast errors were larger when setting the variance of $\xi^r$ to zero and using demeaned interest rates. We obtain similar results with respect to forecasts of the output gap (Table 5).

In this forecasting exercise, we did not calibrate the standard deviation of the permanent shocks to the real rate of interest, which means that the standard deviation of monetary policy shocks is 0, as presented in Table 2. When we calibrate this parameter to the value from Laubach and Williams (2003), the forecasting performance is generally half-way between the version of the model with constant trend and the one where we freely estimate this parameter. Forecast performance therefore seems to suggest a small preference for the version of the model where all variances are freely estimated.20

To determine the source of the forecasting errors, we decompose the 4-step ahead errors of inflation and the real interest rate into contributions from the gap component.

20Results are available upon request.
and contributions from the trend (Tables 6 and 7). In the case of inflation, we notice that the improvement in forecasting is generally coming from increased accuracy in terms of both the gap and the trend components. Allowing for a time-varying interest rate trend reduces the volatility of the interest rate gap and clearly improves gap forecasts. This occurs at the expense of the accuracy of the trend component, however, which is poorly predicted due to its high volatility. The relatively significant forecasting errors of the time-varying interest rate trend could also reflect the typical end of sample problems of HP filters, as discussed in Mise, Kim, and Newbold (2005).

4 Conclusion and Next Steps

In the DSGE literature it is commonplace to express models with variables in deviation-from-steady-state format and to estimate or calibrate structural parameters using data pre-filtered to extract trends. With such a setup, in a forecasting environment it seems natural to add forecasts from DSGE models to trend forecasts. While this may be an intuitive starting point, efficiency can be improved in multiple dimensions. First, pre-filtering of data can be amended to incorporate structural links between the various trends that are implied by the economic theory on which the DSGE model is based, improving the efficiency of trend estimates. Second, forecast efficiency can be improved by building a forecast model for model-consistent trends. Third, decomposition of shocks into permanent and transitory components can be endogenized to also be model-consistent.

In this paper, we propose a unified framework for introducing these improvements. Application of the methodology to the small DSGE model estimated by Milani (2007) validates the existence of considerable deviations between trends used for detrending data

---

21 We do not decompose forecast errors of the output gap since they are entirely attributable to the gap component in the current setup. The forecasting error on the trend component of the constant trend model is equal to the difference between the rolling mean and the full sample mean. The rolling mean converges to the sample mean as the rolling window increases, implying that trend forecast errors will mechanically converge to zero in the constant trend model. The forecasting error on the trend component of the time-varying trend model is equal to the difference between the rolling Kalman filter estimate and the filtered series over the full sample.

22 This problem can be reduced by conditioning the filter with exogenous sources of information. See for instance Gosselin and Lalonde (2006).
prior to structural parameter estimation and model-consistent estimates of trends, implying
the potential for efficiency gains in forecasting, particularly in the case of inflation. Such
deviations also provide information on aspects of the model that are least coherent with the
data, possibly indicating model misspecification.

Given its size, the Milani model provided a good environment to experiment with the
methodology. The next step will be to apply the approach to the Bank of Canada’s
policy-analysis and projection model, ToTEM. Using this large-scale multi-sector DSGE
model, we will estimate trends for the components of output, relative prices, inflation, the
exchange rate, and the interest rate. The solution matrices of the log-linearized version
of ToTEM are much larger ($C_s$ is $120 \times 120$ and $H_s$ is $120 \times 32$). Of the model’s 120
behavioural equations, 88 represent unobserved state variables and 32 represent observed
series for which structural shocks are backed with raw data. Using data for these series over
the 1973Q1-2007Q2 period and conditional on the model solution, the historical variance of
the structural shocks and the assumption of a random walk for all structural components,
we can simultaneously estimate 32 model-consistent trends by maximum likelihood along
with the Kalman filter. Preliminary results are encouraging, as the procedure converges
and yields plausible results in many cases. Trend estimates will next be refined with the
introduction of common trend restrictions based on theory. Allowing permanent shocks to
affect the model’s short-run dynamics is left to future research.
References


### Table 1: Bayesian Estimates of Model Parameters from Milani (2007)

<table>
<thead>
<tr>
<th>Description</th>
<th>Bayesian Estimation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean estimate</td>
</tr>
<tr>
<td>Habits</td>
<td>( \eta )</td>
</tr>
<tr>
<td>Discount</td>
<td>( \beta )</td>
</tr>
<tr>
<td>IES</td>
<td>( \phi )</td>
</tr>
<tr>
<td>Indexation</td>
<td>( \gamma )</td>
</tr>
<tr>
<td>Fcn. price stick.</td>
<td>( \xi_p )</td>
</tr>
<tr>
<td>Elast. mc</td>
<td>( \omega )</td>
</tr>
<tr>
<td>Int-rate smooth.</td>
<td>( \rho )</td>
</tr>
<tr>
<td>Feedback Infl.</td>
<td>( \chi_{\pi} )</td>
</tr>
<tr>
<td>Feedback gap</td>
<td>( \chi_{x} )</td>
</tr>
<tr>
<td>Autoregr. dem shock</td>
<td>( \phi_{r} )</td>
</tr>
<tr>
<td>Autoregr. sup shock</td>
<td>( \phi_{u} )</td>
</tr>
<tr>
<td>MP shock</td>
<td>( \sigma_{\epsilon} )</td>
</tr>
<tr>
<td>Demand shock</td>
<td>( \sigma_{r} )</td>
</tr>
<tr>
<td>Supply shock</td>
<td>( \sigma_{u} )</td>
</tr>
</tbody>
</table>

* Prior is a Gamma with mean 1 and standard deviation 0.71.

### Table 2: Estimates of Standard Deviations of Structural Shocks

<table>
<thead>
<tr>
<th>Description</th>
<th>Trends Only</th>
<th>Trends and Transitory</th>
<th>Trends and Transitory with LW estimates for ( \sigma_{r} )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate</td>
<td>SE</td>
<td>Estimate</td>
</tr>
<tr>
<td>Permanent Shocks</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inflation Target</td>
<td>( \sigma_{\pi} )</td>
<td>0.180</td>
<td>0.089</td>
</tr>
<tr>
<td>Potential Level</td>
<td>( \sigma_{y} )</td>
<td>0.000</td>
<td>0.415</td>
</tr>
<tr>
<td>Potential Growth</td>
<td>( \sigma_{g} )</td>
<td>0.024</td>
<td>0.013</td>
</tr>
<tr>
<td>Real Rate</td>
<td>( \sigma_{r} )</td>
<td>0.800</td>
<td>0.078</td>
</tr>
<tr>
<td>Transitory Shocks</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MP shock</td>
<td>( \sigma_{\epsilon} )</td>
<td>0.933 **</td>
<td>0.000</td>
</tr>
<tr>
<td>Demand shock</td>
<td>( \sigma_{r} )</td>
<td>1.067</td>
<td>0.473</td>
</tr>
<tr>
<td>Supply shock</td>
<td>( \sigma_{u} )</td>
<td>1.146</td>
<td>1.088</td>
</tr>
</tbody>
</table>

A ** entry for standard error (SE) in the results indicates that this parameter was constrained during estimation.
### Table 3: Forecast RMSE of Inflation

<table>
<thead>
<tr>
<th>Sample</th>
<th>1-quarter ahead</th>
<th>4-quarters ahead</th>
<th>8-quarters ahead</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Constant Varying</td>
<td>Constant Varying</td>
<td>Constant Varying</td>
</tr>
<tr>
<td>1963-1975</td>
<td>1.342 1.319</td>
<td>2.007 1.822</td>
<td>2.619 2.282**</td>
</tr>
<tr>
<td>1976-1985</td>
<td>1.458 1.455</td>
<td>1.827 1.766</td>
<td>2.414 2.433</td>
</tr>
<tr>
<td>1986-1995</td>
<td>0.838 0.826</td>
<td>0.973 0.841**</td>
<td>1.300 1.123</td>
</tr>
<tr>
<td>1996-2004</td>
<td>0.901 0.883</td>
<td>1.100 0.820**</td>
<td>1.572 1.046**</td>
</tr>
<tr>
<td>1963-2004</td>
<td>1.190 1.176</td>
<td>1.597 1.452*</td>
<td>2.116 1.898</td>
</tr>
</tbody>
</table>

### Table 4: Forecast RMSE of Real Rates

<table>
<thead>
<tr>
<th>Sample</th>
<th>1-quarter ahead</th>
<th>4-quarters ahead</th>
<th>8-quarters ahead</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Constant Varying</td>
<td>Constant Varying</td>
<td>Constant Varying</td>
</tr>
<tr>
<td>1963-1975</td>
<td>0.939 0.947</td>
<td>1.994 2.129</td>
<td>2.238** 2.586</td>
</tr>
<tr>
<td>1986-1995</td>
<td>0.479 0.476</td>
<td>1.295 1.349</td>
<td>1.682* 2.246</td>
</tr>
<tr>
<td>1996-2004</td>
<td>0.508 0.474</td>
<td>1.377 1.304</td>
<td>1.868 1.655*</td>
</tr>
<tr>
<td>1963-2004</td>
<td>0.987 0.992</td>
<td>1.879 1.917</td>
<td>2.348 2.494</td>
</tr>
</tbody>
</table>

### Table 5: Forecast RMSE of the Output Gap

<table>
<thead>
<tr>
<th>Sample</th>
<th>1-quarter ahead</th>
<th>4-quarters ahead</th>
<th>8-quarters ahead</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Constant Varying</td>
<td>Constant Varying</td>
<td>Constant Varying</td>
</tr>
<tr>
<td>1963-1975</td>
<td>1.018 0.988</td>
<td>2.462 2.215**</td>
<td>3.119 2.959</td>
</tr>
<tr>
<td>1963-1975</td>
<td>1.057 1.093</td>
<td>2.089 2.471</td>
<td>2.063 2.728</td>
</tr>
<tr>
<td>1986-1995</td>
<td>0.536 0.492**</td>
<td>1.493 1.236*</td>
<td>1.415 1.283</td>
</tr>
<tr>
<td>1996-2004</td>
<td>0.611 0.588</td>
<td>1.400 1.374</td>
<td>1.573 1.827</td>
</tr>
<tr>
<td>1963-2004</td>
<td>0.861 0.851</td>
<td>1.975 1.943</td>
<td>2.253 2.373</td>
</tr>
</tbody>
</table>

Forecast errors for the output gap are calculated using Milani’s IS curve and model-consistent trend estimates of inflation and interest rates. * and ** denote statistically different forecast errors at the 5 and 10 per cent level based on the Diebold-Mariano test, respectively.
We obtain similar results using Theil’s U statistics, which controls for differences in the variance of the observed series.
Fig. 1: Inflation
Fig. 2: Output Gap
Fig. 3: Real Interest Rates