Inflation Forecasting in a Changing Environment

Douglas Eduardo Turatti

Universidade Federal de Santa Catarina

Guilherme Valle Moura

Universidade Federal de Santa Catarina

Abstract

This article proposes an efficient estimation procedure for nonlinear state-space models, where some states can be integrated out analytically . One example is the unobserved component stochastic volatility model used in the literature to model inflation. However, because of the difficulties in estimating this model, it has not yet been estimated, and articles that utilize it have always calibrated its parameter. Here the efficient importance sampling procedure is used together with a Rao-Blackwellization step to construct a highly efficient estimation procedure that produces continuous approximations to the likelihood function, greatly enhancing parameter estimation. The estimated model is used to forecast inflation of all G7 countries and showed superior results with respect to *benchmark* models. Moreover, estimates of the unobserved components indicate that the Great Inflation of the 70s was caused by a rise in trend inflation, while random innovations have been dominating the inflation process in the last decades for most of the G7 countries.

1 Introduction

Price stability has become an important mandate of many central banks around the world since the 1980s. It is now widely accepted that decision making becomes more complex in high and persistent inflation scenarios, as inflation may

cloud public confidence as well as economic agents' assessments of future economic activity (Golob 1994). Moreover, low inflation seems to promote growth and support sustainable employment in the long run (Bernanke 2007). Thus, it is not surprising that a lot of effort has been devoted to the development of models that can accurately explain the dynamics and volatility of inflation rates.

However, as pointed out by Stock & Watson (2006), inflation dynamics and volatility have changed over the last decades. Following the Great Inflation of the 1970s, central banks of developed countries have made successful efforts not only to lower, but also to stabilize inflation. These efforts have contributed to the Great Moderation of output volatility (see, for example, Cecchetti et al. 2007).

Therefore, it has been hard to use a single model to forecast inflation in different periods of time. Atkeson & Ohanian (2001) have shown that Phillips curve-type models have performed well until the 1980s, but cannot display the same accuracy in recent decades, when simple statistical models like the first order autoregressive model have been dominating forecasts.

In order to cope with this changing environment, Stock & Watson (2006) have proposed the unobserved component stochastic volatility model (UCSV), which is based on the traditional local level model with stochastic volatilities added to the variances of the shocks. The stochastic volatility implies that the model is a IMA(1,1) model with time varying moving average parameter, which allows it to capture the changes in the inflation process.

Although promising, the UCSV has never been estimated, thus its out-of-sample forecast capabilities could not be tested. Stock & Watson (2006) and Cecchetti et al. (2007) have used this model, but they have chosen to calibrate the value of the unknown parameter because of the difficulties in estimation. The estimation of this model is complicated by the fact that the likelihood function is a multidimensional integral with no closed form solution, thus numerical integration techniques have to be used. Here we propose a maximum simulated likelihood estimation procedure for this model based on Efficient Importance Sampling of Richard & Zhang (2007). Additionally, we use a Rao-Blackwellization step in-

creasing the numerical efficiency of the algorithm. The procedure is fast and provides continuous likelihood approximations, facilitating numerical optimization of the log-likelihood function.

To preview some of our results, we show that the maximum likelihood parameter estimates are statistically different from the calibrated values used in the literature. Moreover, the UCSV forecast performance is shown to be superior in comparison to AR(1) models and to the random walk. The remainder of this article is organized as follows. In Section 2, we describe the model and some of its properties, in Section 3 we present the estimation procedure, Section 4 compares the estimation procedure with the particle filter, and compares the out-of-sample forecast performance of the estimated model compared to benchmarks, and Section 5 concludes with some final remarks.

2 The model

Inflation forecasting is important to reduce uncertainty and facilitate decision making. However, the behavior of inflation has changed over time, compromising its predictability. In order to capture these changes, SW proposed a model which allows for changes in inflation dynamics and suggesting that inflation forecast could be improved. This model decomposes inflation into a stochastic trend component (τ_t), which is the conditional mean of inflation, and a temporary shock (η_t), where the variance of both components follow a random-walk stochastic volatility process. More specifically:

$$
\begin{aligned}
\pi_t &= \tau_t + \eta_t, && (1)\\
\eta_t &= \sigma_{\eta,t}\zeta_{1,t}, && (2)\\
\tau_t &= \tau_{t-1} + \varepsilon_t, && (3)\\
\varepsilon_t &= \sigma_{\varepsilon,t}\zeta_{2,t}; \quad \zeta_t \sim iidN(0, I_2), && (4)\\
\ln\sigma^2_{\eta,t} &= \ln\sigma^2_{\eta,t-1} + v_{1,t}, && (5)\\
\ln\sigma^2_{\varepsilon,t} &= \ln\sigma^2_{\varepsilon,t-1} + v_{2,t}; \quad v_t \sim iidN(0, \gamma I_2). && (6)
\end{aligned}
$$

Due to its characteristics, the model was labeled as the unobserved components stochastic volatility (UCSV) model. This is a generalization of the well-known local-level state space model, where the variances follow a stochastic volatility. Equation (1) characterizes the measurement equation and equations (3), (5) and (6) represent the state transition equations. The model features a single parameter, γ, which controls the variance of both stochastic volatility processes. Differently from Stock & Watson (2006), who calibrated γ based on the full sample, here we propose an efficient estimation procedure that allows the estimation of this parameter.

2.1 Statistical properties of the UCSV model

It is interesting to examine some statistical properties of the UCSV model. First, as its trend component follows a random walk, the model is non-stationary. However, in small samples it can mimic stationary process very well depending on the volatilities of the non-stationary component τ_t and of the stationary component η_t.

As shown by Stock & Watson (2006), another important characteristic of the UCSV specification is its equivalence to a integrated moving average (IMA) process with time-varying parameter,

$$\Delta\pi_t = (1-\theta_t L)a_t, \tag{7}$$

where a_t is serially uncorrelated with mean zero and L is the lag operator. The stochastic volatilities in the UCSV model imply that the moving average coefficient θ_t is time-varying. The IMA model was successfully used by Nelson & Schwert (1977) to forecast inflation in the 70s.

The statistical properties of the model to be discussed next have important implications for its ability to account for structural changes and to forecast inflation. The model capability to mimic stationary and non-stationary processes is one of its main advantages in forecasting inflation. This property can be viewed examining the autocorrelation of the changes in inflation:

$$\Delta\pi_t = \Delta\tau_t + \Delta\eta_t = \varepsilon_t + \eta_t - \eta_{t-1}, \tag{8}$$

this is a first order moving average process with a negative coefficient, thus its first order autocorrelation function will be negative, while higher order autocorrelations will be zero. Moreover, the autocorrelation of $\Delta\pi_t$ depends on the variances of the components, thus the presence of stochastic volatility makes it time-varying:

$$\rho_{\Delta\pi} = \frac{-\sigma^2_{\eta,t}}{2\sigma^2_{\eta,t} + \sigma^2_{\varepsilon,t}}. \tag{9}$$

Note that $-0.5 \leq \rho_{\Delta\pi} \leq 0$ and whenever $\sigma_{\varepsilon,t}$ is large relative to $\sigma_{\eta,t}$, $\rho_{\Delta\pi}$ is close to zero, implying that π_t is highly autocorrelated. On the other hand, whenever $\sigma_{\eta,t}$ is large compared to $\sigma_{\varepsilon,t}$, $\rho_{\Delta\pi}$ is close to -0.5. Thus, in finite samples, the autoregressive coefficient in an AR(1) regression for the level of inflation will vary depending on the ratio $\sigma_\varepsilon/\sigma_\eta$. Following Cecchetti et al. (2007), we performed a Monte Carlo exercise where we run the following regression:

$$\pi_t = \alpha + \rho_\pi \pi_{t-1} + \lambda_t. \tag{10}$$

For a chosen value of $\sigma_\varepsilon/\sigma_\eta$, 5000 samples with 100 observations were drawn

from the data generating process (1)-(6) and 5000 estimates of ρ_π were computed. Figure 1 shows the median, and the 10^{th} and 90^{th} percentile of ρ_π for values of $\sigma_\varepsilon/\sigma_\eta$ ranging from 0.005 to 2.5. From Figure 1 it is clear that the DGP (1)-(6) can mimic positively correlated autoregressive process and even white noises in finite samples depending on the ratio $\sigma_\varepsilon/\sigma_\eta$. Thus, the presence of stochastic volatilities adds a lot of flexibility to this model.

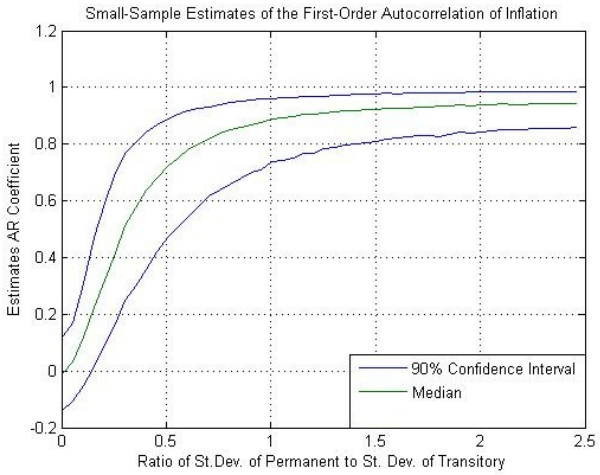


Figure 1: Finite sample estimates of the first-order autoregressive coefficient of inflation.

Another property of the model is that its optimal forecast is equal to the current value of the trend component regardless of the horizon:

$$E_t \pi_{t+k} = \tau_{t|t}, \tag{11}$$

where $\tau_{t|t} = E_t \tau_t$ is the filtered value of the unobserved component τ_t and depends on t, but not on the forecast horizon. This is linked to the concept of anchoring of expectations and it is relatively insensitive to the arrival of new information. Rather, agents appear to stick to their long-run reference of inflation when making their forecasts (see, for example Mishkin 2007).

3 Estimation strategy

The presence of stochastic volatilities both in the measurement equation (1) as well as in the transition equation (3) implies that the state space model is nonlinear. Therefore, the integrals over the unobservable state variables do not posses analytical solutions and the Kalman filter is not applicable. In these cases, the integrals in the expression for the likelihood function of the model have to be approximated numerically. Here we use the Efficient Importance Sampling method (EIS henceforth) proposed by Richard & Zhang (2007) to approximate these integrals via Monte Carlo (MC) integration.

3.1 Efficient Importance Sampling

Efficient Importance Sampling is a simple, generic and highly accurate MC procedure for the evaluation of high dimensional integrals. It is based upon a sequence of auxiliary regressions that can be linear under appropriate conditions. Consider the evaluation of a functional integral

$$G(\gamma) = \int_{X(\gamma)} \varphi(x;\gamma)\, dx, \tag{12}$$

where γ is a vector of parameters, $\varphi = g(x;\gamma)p(x;\gamma)$ such that g is a function which is integrable with respect to a density $p(x;\gamma)$ with support $X(\gamma)$. In our case, $G(\gamma)$ can be seen as the likelihood function and $x = \{\tau_t, \sigma_{\eta_t}, \sigma_{\varepsilon_t}\}_{t=0}^{T}$ as the vector of unobserved variables. Importance sampling implies that it is possible to rewrite (12) using an auxiliary density $m(x|\alpha)$:

$$G(\gamma) = \int_{X(\gamma)} \varphi(x;\gamma)\, dx = \int_{X} \frac{\varphi(x;\gamma)}{m(x;\alpha)} m(x;\alpha) dx \tag{13}$$

and $G(\gamma)$ can be estimated as

$$\widehat{G}(\gamma) = \frac{1}{S}\sum_{i=1}^{S} \frac{\varphi(\tilde{x}^i;\gamma)}{m(\tilde{x}^i;\alpha)}, \tag{14}$$

where $\{x^i\}_{i=1}^S$ denotes a set of i.i.d. random draws from $m(x;\alpha)$, the importance sampler, and α is a set of auxiliary parameters. Under weak regularity conditions, this simulator will converge to $G(\gamma)$ almost surely as S goes to infinity (see, for example, Geweke 1996). Although unbiased, the MC sampling variance of φ on m can be large for naive choices of m and accurate MC estimation of G may require prohibitively large numbers of draws. The MC sampling variance of the ratio φ/m on m is given by

$$\widehat{V}(\alpha;\gamma) = \frac{1}{S}\int_{X(\gamma)} \left[\frac{\varphi(x;\gamma)}{m(x;\alpha)} - G(\gamma)\right]^2 m(x;\alpha)\, dx. \tag{15}$$

The tails of m must not decay faster than the tails of φ in order to ensure that (15) is finite. Moreover, (15) makes clear that to avoid large variance, m should mimic φ as well as possible, minimizing variations of the ratio φ/m. The EIS strategy is thus to construct a density kernel $k(x;\alpha)$ that can mimic $\varphi(x;\gamma)$ in shape and with tails that do not decay faster than the ones from $\varphi(x;\gamma)$. In addition, the integrating constant $\chi(\alpha) = \int k(x;\alpha)\, dx$ should have a known analytical form.

The selection of a class of auxiliary samplers is problem specific, the EIS principle then addresses the problem of selecting a (near) optimal $\alpha(\gamma)$ within the chosen class. This is done solving the following least-squares problem:

$$\widehat{\alpha} = \arg\min_{\alpha} \sum_{i=1}^{S} \left[\ln\varphi(x^i;\gamma) - c_t - \ln k(x^i;\alpha)\right]^2 g(x^i;\gamma), \tag{16}$$

where $\{x^i\}_{i=1}^S$ are draws from an initial sampler $p(x|\alpha_0)$ and c_t accounts for missing constants. Note that the fact that the initial sampler depends itself upon α, imply that we need to compute a converging sequence of least-square problems given by:

$$\widehat{\alpha}_{k+1} = \arg\min_{\alpha} \sum_{i=1}^{S} \left[\ln\varphi(x^i;\gamma) - c_t - \ln k(x^i;\widehat{\alpha}_k)\right]^2 g(x^i;\gamma). \tag{17}$$

Once $\widehat{\alpha}_k$ has converged to a fixed value, we can draw $\{x^i(\widehat{\alpha})\}_{i=1}^S$ from $m(x;\alpha) = k(x;\widehat{\alpha}))/\chi(\widehat{\alpha})$ and compute (14).

In order to ensure the smoothness of (17) for different values of γ, we use the common random number technique, where all successive sequences of *i.i.d.* draws $\{x^{i,k}\}_{i=1}^S$ are generated as non-stochastic transformations of a single set of random draws $\{u^i\}_{i=1}^S$ (see, for example, Richard & Zhang 2007).

3.2 Sequential EIS

However, in high dimensional problems the least-squares problem in (16) is unfeasible and needs to be replaced by a sequence of lower dimensional optimization problems. The following factorization is developed in Richard & Zhang (2007):

$$\varphi(\underline{x}_T;\gamma) = \prod_{i=1}^T \varphi_t(x_t|\gamma) = \prod_{t=1}^T g_t(x_t;\gamma)p_t(x_t|x_{t-1};\gamma), \tag{18}$$

where $\underline{x}_T = \{x_1,\ldots,x_T\}$ Now draws $\{\{x_t^i\}_{t=1}^T\}_{i=1}^S$ from the transition densities (the natural samplers) $p_t(x_t|x_{t-1};\gamma)$, can be taken to solve the following backward sequence of low dimensional least-squares problem going from T to 1

$$\widehat{\alpha}_t^{k+1} = \arg\min_{\alpha_t} \sum_{i=1}^S \left[\ln\left(\varphi_t(x_t^i;\gamma)\cdot\chi_{t+1}(x_t^i;\widehat{\alpha}_{t+1}^k)\right) - c_t - \ln k(x_t^i;\widehat{\alpha}_t^k)\right]^2 g(x_t^i;\gamma) \tag{19}$$

where we transfer the integrating constant $\chi_{t+1}(x_t^i;\widehat{\alpha}_{t+1})$ back to period's t problem achieving a smoother effect, and $\chi_{T+1}(x_T^i) = 1\forall i$. Therefore, (19) selects the $\widehat{\alpha}_t$ that minimizes the distance between $\varphi_t(x_t^i;\gamma)\chi_{t+1}(x_t^i;\widehat{\alpha}_t+1)$ and $k(x_t^i;\alpha_t)$. Note that the EIS approximations constitute a global approximation, as they are obtained through the solution of a least squares problem on the full support $X(\gamma)$ of φ.

The final EIS sampler is given by

$$m(\underset{\sim}{x}_T;\widehat{\alpha}) = \prod_{t=1}^{T} m(x_t;x_{t-1};\widehat{\alpha}_t) = \prod_{t=1}^{T} \frac{k_t}{\chi_t} = \frac{1}{\chi_1}\prod_{t=1}^{T} \frac{k_t}{\chi_{t+1}}. \tag{20}$$

which shows that if all the k_t/χ_{t+1} mimic φ_t, then $m(\underset{\sim}{x}_T;\widehat{\alpha})$ will also mimic $\varphi(x_T|\gamma) = \prod_{t=1}^{T}\varphi_t$.

3.3 Rao-Blackwellized EIS

It is possible to exploit the structure of the state space model in (1)-(6) to take advantage of the conditional linearity of it. Note that given the two stochastic volatilities, the model is a traditional linear and Gaussian local level model (see, for example, Durbin & Koopman 2012, Chapter 2). Thus, it is possible to analytically integrate out the trend component τ using the Kalman filter, reducing the Monte Carlo variation of the already optimized likelihood estimator even further. This procedure is known as Rao-Blackwellization as it is an implication of the Rao-Blackwell Theorem (e.g., see Robert & Casella 2005, p. 130).

In the context of the UC-SV model, (18) for a given period t can be written as:

$$g(\pi_t|x_t;\gamma)p(x_t|x_{t-1};\gamma), \tag{21}$$

where $g(\pi_t|x_t;\gamma)$ is the measurement density represented in (1), and $p(x_t|x_{t-1};\gamma)$ is the trivariate transition density implied by (3), (5) and (6). Equation (21) can be decomposed as:

$$g(\pi_t|\tau_t,\sigma_t;\gamma)p(\tau_t|\sigma_t,x_{t-1};\gamma)p(\sigma_t|\sigma_{t-1};\gamma), \tag{22}$$

where $\sigma_t = [\sigma_{\eta_t},\sigma_{\varepsilon_t}]$. Note that given σ_t, τ_t can be integrated out of (22) using the Kalman filter, yielding

$$g^*(\pi_t|\sigma_t)p(\sigma_t|\sigma_{t-1},\gamma), \tag{23}$$

where $g^*(\pi_t|\sigma_t) \sim N(\tau(\sigma_t)_{t|t-1}, P(\sigma_t)_{t|t-1} + \sigma^2_{\eta_t})$. The predicted value of τ conditional on σ_t, $\tau(\sigma_t)_{t|t-1}$, as well as its variance $P(\sigma_t)_{t|t-1}$ can be computed via the Kalman filter for each value of σ_t.

Choosing the Gaussian class of importance samplers and exploiting the fact that Gaussian kernels are closed under multiplication, we can parameterize the kernel of the importance sampling density of period t as:

$$k(\sigma_t;\alpha_t) = p(\sigma_t|\sigma_{t-1};\gamma)\xi(\sigma_t;\alpha_t), \tag{24}$$

where $\xi(\sigma_t,\alpha_t)$ is a Gaussian kernel given by $\xi(\sigma_t,\alpha_t) = \exp(\sigma_t B_t + \sigma_t C_t \sigma_t')$ and $\alpha_t = \{B_t, C_t\}$. Since the transition density $p(\sigma_t|\sigma_{t-1};\gamma)$ appears both in the importance sampler and in the target integrand (23), it cancels out in the sequence of EIS optimizations, simplifying (19) to an ordinary least square problem:

$$\widehat{\alpha}_t^{k+1} = \arg\min_{\alpha_t} \sum_{i=1}^{S} \left[\ln g^*(\pi_t|\sigma_t) \cdot \chi_{t+1}(\sigma_t^i|\widehat{\alpha}_{t+1}^k) - c_t - \sigma_t^i \widehat{B}_t^k - \sigma_t^i \widehat{C}_t^k \sigma_t^{i'}\right]^2. \tag{25}$$

The optimized vector of auxiliary parameters $\widehat{\alpha}$ allow us to compute the EIS sampler as in (20), which is a global approximation to the target integrand (23). The likelihood integral can then be estimated as

$$\hat{L}(\gamma;\pi_{1:T}) = \frac{1}{N}\sum_{i=1}^{N}\prod_{t=1}^{T}\left[\frac{g^*(\pi_t|\sigma_t^i)\chi_{t+1}(\sigma_t^i|\widehat{\alpha}_{t+1}^k)}{\xi(\sigma_t^i,\widehat{\alpha}_t)}\right] = \frac{1}{N}\sum_{i=1}^{N}\prod_{t=1}^{T} w_t^i. \tag{26}$$

After γ has been estimated based on (26), it is possible to obtain smoothed values of the unobserved components using the normalized importance weights

$$\overline{w}_t^{(i)} = \frac{w_t^{(i)}}{\sum_{i=1}^N w_t^{(i)}}, \quad \text{where} \quad w_t^i = \left[\frac{g^*(\pi_t|\sigma_t^i)\chi_{t+1}(\sigma_t^i|\widehat{\alpha}_{t+1}^k)}{\xi(\sigma_t^i,\widehat{\alpha}_t)}\right] \tag{27}$$

A smoothed estimate of the trend component can be estimated as:

$$E[\tau_t|\pi_{1:T}] \approx \sum_{i=1}^{N} \overline{w}_t^{(i)} \tau(\sigma_t^i)_t. \tag{28}$$

4 Results

The UCSV model was used to forecast inflation of the G7 countries and its predictive power was measured by the root mean square error (RMSE) in comparison to an autoregressive model. The predictions were computed out of the sample based on an expanding window starting from the first quarter of 1990. Moreover, smoothed estimates of the unobserved components are used to shed light on the evolution of the inflation processes of these countries in the last decades.

We estimated the UCSV model using time series of inflation for the G7 countries. The series used are the quarterly consumer price index (CPI) running from the second quarter 1955 to the last quarter of 2012 for all countries, totaling 231 observations. The series were obtained from the OCDE statistics database. The maximum likelihood estimation and state smoothing were performed using the Rao-Blackwellized sequential EIS procedure described in the last session. Table 1 shows the parameter estimates for all countries in the sample, together with their statistical and MC standard errors.

The estimation results presented in Table 1 showed that the estimated values of the parameter vary a lot from country to country and are all between 0.03 and 0.11. Moreover, it is observed that the estimated value for the countries were lower and statistically different from 0.2, the value set by SW for the US and by Cecchetti et al. (2007) for all G7 countries.

After estimating the parameters, the EIS procedure was used to obtain smoothed estimates of the state variables. It is observed from Figure 2 that the volatility of the trend component showed an increase in the 70's and early 80 for all countries except Japan and Germany. After that, trend volatility is reduced to levels even lower than before the Great Inflation of the 70's and remained in those lev-

Table 1: Estimation Results

	Parameter (γ)	Log-Likelihood
Canada	0.0822 (0.0368) [0.0098]	−209.4394
France	0.1107 (0.0440) [0.0099]	−180.9287
Germany	0.0322 (0.0202) [0.0014]	−193.3220
Italy	0.0891 (0.0355) [0.0088]	−201.3819
Japan	0.0633 (0.0312) [0.0018]	−282.5239
U.K.	0.0937 (0.0380) [0.0027]	−203.1276
U.S.	0.1010 (0.0381) [0.0058]	−162.0727

Note: Values in parenthesis denote statistical standard errors, while values in brackets are MC standard errors based on 30 different sets of CRN. The MC sample size used in each EIS estimation was 300.

els ever since. It is noteworthy that the 2008 crisis did not affect this component of inflation in any of the countries. As for the volatility of the transitory innovations, Figure 3, they displayed a moderate rise in the early 80s, when the volatility started to be transferred from the trend to the transitory shock. Estimates for the 90s, the decade of the Great Moderation, show a reduction in the volatility of the trend and shock, what is consistent with price stability over the period. However, since the 2000s the volatility of the transitory component has increased, and the recent financial crisis strongly affected this part of the volatility, especially in the U.S.

Figure 4 shows the plots of the autocorrelation of the first difference of the inflation. They indicate that during the Great Inflation of the 70s, inflation approached a unit root process in all countries except Germany and Japan, countries well known for the strong control of inflation. Finally, the plots of inflation and its estimated trend, Figure 5 (in appendix), highlights the fact that the inflation of the 70s was caused predominantly by the trend component. In that period, the trend was very close to the estimated inflation in all countries, except in those already mentioned above. After this period, and especially more recently, the fluctuation in inflation has been largely caused by the transitory shock.

More specifically, the UC-SV model explains the persistence of inflation dur-

ing the 70s via a high volatility in the trend component in comparison to the volatility of the transitory component. As shown in the Monte Carlo experiment in Figure 1, this allows the model to mimic the behavior of a unit root process. On the other hand, fluctuations in inflation after this period, particularly during the recent financial crisis, are characteristic of stationary processes and the UC-SV model explain this feature via a reduction in the volatility of the trend component and a rise in the volatility of the transitory component, allowing the model to replicate the behavior of a stationary time series.

The likelihood integral of the UC-SV model can also be estimated by the Rao-Blackwellized particle filter (PF) (see Creal 2009, for more details on this version of the standard particle filter). The importance density used by this filter is usually the transition density, thus the likelihood estimate has the following simplified form:

$$\hat{L}(\gamma;\pi_{1:T}) = \frac{1}{N}\sum_{i=1}^{N}\left\{\prod_{t=1}^{T} g^{*}(\pi_t|\sigma_t^i,\gamma)\right\} \tag{29}$$

However, as it is going to be shown bellow, the likelihood estimate of particle filters are discontinuous functions of parameter estimates, rendering as problematic maximum likelihood estimation based on numerical derivative methods (see, for example, DeJong et al. 2013, for problems related to the particle filter).

4.0.1 Likelihood profile

To illustrate the differences between the Rao-Blackwellized sequential EIS procedure used in this paper and the Rao-Blackwellized Particle Filter, the following exercises were performed: a Monte Carlo study to compute the MC variance of each likelihood estimator; and the computation of the likelihood profile around the maximum likelihood value, which evidences the undesirable likelihood discontinuity of particle filters.

The results of the Monte Carlo study are presented in Table 2. The experiment was based on 30 different likelihood evaluations based on different sets of CRNs and the value of the parameter was fixed in its maximum likelihood estimate de-

scribed in Table 1. The Monte Carlo sample size was 300 for the EIS procedure and 5,000 for the particle filter. Standard deviation of the log-likelihood estimates are small when compared to their average value in both cases, although the Monte Carlo sample size of the EIS procedure is more than 15 times smaller, highlighting the accuracy of the EIS procedure with respect to the particle filter.

Table 2: Monte Carlo Experiment Results

	RB-EIS		RB-PF	
	Mean	Std.	Mean	Std.
Canada	−208.9652	0.2133	−208.8289	0.2038
France	−180.7666	0.3139	−180.4073	0.2454
Germany	−193.3902	0.0966	−193.4021	0.1520
Italy	−201.3218	0.4063	−200.9770	0.2079
Japan	−282.4148	0.1128	−282.4083	0.1898
U.K.	−203.3067	0.1902	−203.4005	0.1752
U.S.	−161.9216	0.2253	−161.5717	0.2309

The second experiment compares the behavior of the log-likelihood function for different parameter values using the EIS procedure and the PF. A grid of parameter values containing the maximum likelihood estimate was constructed and the likelihood function was evaluated at those values using the same set of CRNs for each filter. Again, the Monte Carlo sample size was 300 in the EIS case and 5,000 for the PF. The log-likelihood cut presented in Figure 6 shows that the function estimated by the Rao-Blackwellized sequential EIS is smooth with respect to the parameter of the model and strictly convex, while PF estimate is a discontinuous function of the parameters, which would hinder optimization based on numerical gradients.

4.0.2 Out-of-sample forecasting

This section presents results on the out-of-sample predictive power of the UC-SV model in comparison to the first order autoregressive model, labeled AR(1), the random walk (RW) . The comparison will also be made with the local level

model (LL) and the UCSV with γ fixed equal to 0.2. The estimation is done through an expanding window starting in the first quarter of 1990. Forecasts are computed for 1,2 and 4 steps ahead and are all evaluated on the basis of the root mean square error (RMSE). Differences in RMSE of competing models will be compared on the basis of the Diebold-Mariano test for the non-nested models, and (n.d.) test for the local level model. Table 2 (in appendix) confirms the expectation that the UC-SV model would be a good forecasting model and shows that the RMSE for the UC-SV is always smaller than the one for the competing models. Results on the Diebold-Mariano test show that these differences are very often statistically significant.

5 Conclusions

This paper proposes the use of the UC-SV model of Stock & Watson (2006) to forecast inflation of G7 countries. The model breaks down inflation into a permanent and a transitory component, where the variance of both components follow a stochastic volatility process. It is argued that this specification is robust to changes in the inflation process observed in the last decades, making it a strong candidate to forecast inflation of different countries.

However, parameter estimation and filtering in UCSV is complicated because of the stochastic volatility and until now its unique parameter has only been calibrated and never estimated. Here we propose an estimation procedure based on Efficient Importance Sampling of by Richard & Zhang (2007) together with a Rao-Blackwellization step to efficiently estimate the parameter of the UC-SV model. The procedure is fast and provides continuous likelihood approximations, facilitating numerical optimization of the log-likelihood function.

Results shows that the maximum likelihood parameter estimates are statistically different from the calibrated values used in the literature. Moreover, the UC-SV forecast performance is shown to be superior to AR(1) models and to the random walk.

References

(n.d.).

Atkeson, A. & Ohanian, L. E. (2001), 'Are phillips curves useful for forecasting inflation?', *Federal Reserve Bank of Minneapolis Quarterly Review* (25:1), 2–11.

Bernanke, B. (2007), 'Inflation expectations and inflation forecasting', *Speech at the Monetary Economics Workshop of the National Bureau of Economic Research* .

Cecchetti, S., Hooper, P., Kasman, B., Schoenholtz, K. & Watson, M. (2007), 'Understanding the evolving inflation process', *US Monetary Policy Forum* .

Creal, D. (2009), 'A survey of sequential monte carlo methods for economics and finance', *Working Paper, University of Chicago* .

DeJong, D., Dharmarajan, H., Liesenfeld, R., Moura, G. V. & Richard, J. (2013), 'Efficient likelihood evaluation of state-space representations', *Review of Economic Studies* **80(2)**(283), 538–567.

Durbin, J. & Koopman, S. J. (2012), *Time Series Analysis by State Space Methods*, Oxford University Press.

Geweke, J. (1996), Monte carlo simulation and numerical integration, *in* H. M. Amman, D. A. Kendrick & J. Rust, eds, 'Handbook of Computational Economics, vol. I', Elsevier.

Golob, J. (1994), 'Does inflation uncertainty increase with inflation?', *Economic Review* **3**.

Mishkin, F. (2007), 'Inflation dynamics', *National Bureau of Economic Research, Working Paper 13147* .

Nelson, C. R. & Schwert, G. W. (1977), 'Short-term interest rates as predictors of inflation: On testing the hypothesis that the real rate of interest is constant', *American Economic Review* **67**.

Richard, J. & Zhang, W. (2007), 'Efficient high-dimensional importance sampling', *Journal of Econometrics* **141**(2), 1385–1411.

Robert, C. P. & Casella, G. (2005), *Monte Carlo Statistical Methods*, Springer.

Stock, J. & Watson, M. (2006), 'Why has u.s. inflation become harder to forecast', *Journal of Money, Credit and Banking* **39**(1).

Table 2: Forecasting Results

		Root Mean Squared Error				
Country	h	UCSV	AR(1)	RW	LL	UCSV(0.2)
US	1	0.6248	0.6960 −2.1802(0.0292)	0.7711 −2.1936(0.0282)	0.6569 −1.6262**	0.6227 0.9154(0.4074)
	2	0.6521	0.8338 −2.2675(0.0233)	1.0149 −2.2052(0.0274)	0.7046 −1.5005**	0.6530 −0.3710(0.0991)
	4	0.6169	0.6438 −1.1139(0.2652)	0.7358 −1.3061(0.1914)	0.6197 −0.3281	0.6144 1.3524(0.4990)
Country	h	UCSV	AR(1)	RW	LL	UCSV(0.2)
France	1	0.3751	0.5168 −8.0453(0.0000)	0.5111 −4.7424(0.0000)	0.4036 −3.0503**	0.3744 1.1843(0.4905)
	2	0.3713	0.6095 −7.4372(0.0000)	0.5187 −3.6632(0.0002)	0.4021 −1.9317**	0.3701 1.1996(0.4140)
	4	0.3660	0.7600 −7.6221(0.0000)	0.4014 −0.9725(0.3307)	0.3752 −0.6546	0.3652 1.3056(0.3814)
Country	h	UCSV	AR(1)	RW	LL	UCSV(0.2)
Canada	1	0.6125	0.6847 −2.7382(0.0064)	0.7564 −3.2833(0.0010)	0.6099 0.3229	0.6023 1.2243(0.4923)
	2	0.6397	0.7931 −3.1202(0.0018)	0.9429 −2.9575(0.0031)	0.6297 0.8417*	0.6247 1.3720(0.1486)
	4	0.6273	0.7399 −3.1959(0.0013)	0.7402 −1.8578(0.0631)	0.6156 0.7564*	0.6214 0.4643(0.1339)

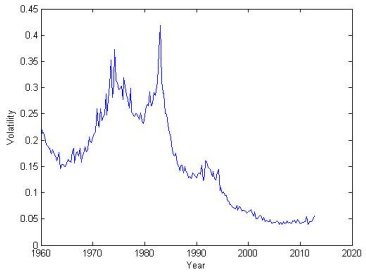
Note: Forecasts are computed using an expanding window initially ending in 1990-1, yielding 48, 47 and 44 out-of-sample forecasts for $h = 1, 2, 4$, which are used to compute the Diebold-Mariano test statistics and their p-values. LL is the local level model. * and ** means statistical significance at 5% e 10% respectively for the (n.d.) test

		Root Mean Squared Error				
Country	h	UCSV	AR(1)	RW	LL	UCSV(0.2)
Italy	1	0.3334	0.3622 −2.0810(0.0374)	0.3372 −0.3095(0.7568)	0.3201 1.2441*	0.3345 −0.6896(0.4356)
	2	0.3825	0.5135 −5.8761(0.0000)	0.4558 −4.2417(0.0000)	0.3632 1.9204**	0.3760 1.0875(0.0494)
	4	0.3652	0.6407 −5.8090(0.0000)	0.3856 −1.0236(0.3088)	0.3677 0.2600	0.3604 1.0216(0.0260)
Country	h	UCSV	AR(1)	RW	LL	UCSV(0.2)
Japan	1	0.5525	0.7835 −7.3017(0.0000)	0.8309 −4.0974(0.0000)	0.5442 2.0186**	0.5525 −0.0032(0.4992)
	2	0.5182	0.8430 −8.2177(0.0000)	0.6171 −1.6469(0.0995)	0.5226 −0.6740	0.5163 0.5651(0.2932)
	4	0.5100	1.0096 −10.1872(0.0000)	0.5437 −0.7130(0.4758)	0.5205 −2.8927**	0.5059 1.3093(0.1128)
Country	h	UCSV	AR(1)	RW	LL	UCSV(0.2)
UK	1	0.2324	0.4714 −10.0210(0.0000)	0.4886 −2.9199(0.0035)	0.2287 0.3617	0.2313 0.7077(0.4872)
	2	0.2215	0.6942 −10.5686(0.0000)	0.6091 −2.8001(0.0051)	0.2264 −0.5505	0.2214 0.0592(0.2123)
	4	0.2014	0.9628 −12.0178(0.0000)	0.8522 −3.3684(0.0007)	0.2286 −2.5810**	0.2012 0.1136(0.2546)

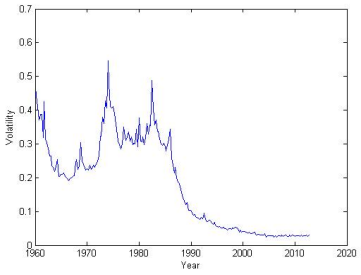
Note: Forecasts are computed using an expanding window initially ending in 1990-1, yielding 48, 47 and 44 out-of-sample forecasts for $h = 1, 2, 4$, which are used to compute the Diebold-Mariano test statistics and their p-values. LL is the local level model.* and ** means statistical significance at 5% e 10% respectively for the (n.d.) test

Country	h	UCSV	AR(1)	RW	LL	UCSV(0.2)
Germany	1	0.4510	0.4816 $-1.8475(0.0646)$	0.5750 $-3.2554(0.0011)$	0.4504 0.0238	0.4580 $-1.0383(0.3496)$
	2	0.4571	0.5023 $-1.8680(0.0617)$	0.5824 $-3.0496(0.0022)$	0.4567 0.0212	0.4654 $-1.3445(0.2612)$
	4	0.4488	0.5248 $-2.4571(0.0140)$	0.4935 $-1.5756(0.1151)$	0.4662 -0.8068^{*}	0.4443 0.8890(0.0505)

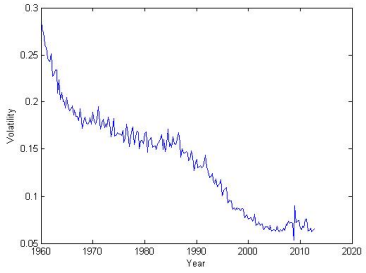
Note: Forecasts are computed using an expanding window initially ending in 1990-1, yielding 48, 47 and 44 out-of-sample forecasts for $h = 1, 2, 4$, which are used to compute the Diebold-Mariano test statistics and their p-values. LL is the local level model. * and ** means statistical significance at 5% e 10% respectively for the (n.d.) test



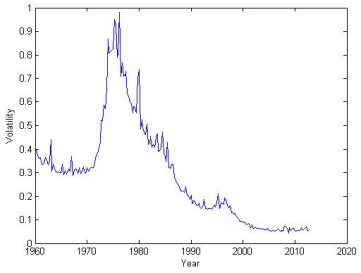
(a) Canada



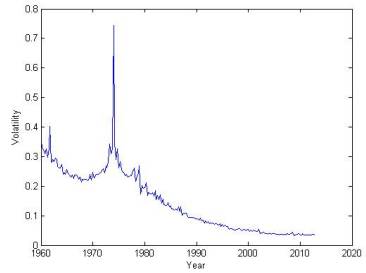
(b) France



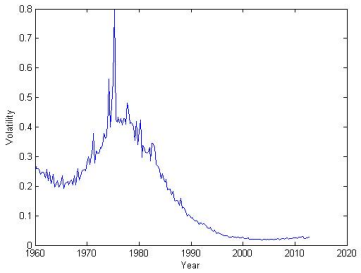
(c) Germany



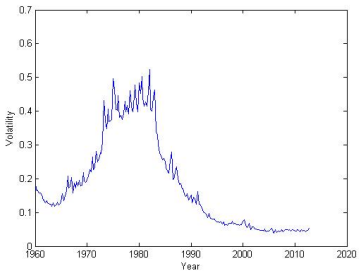
(d) Italy



(e) Japan

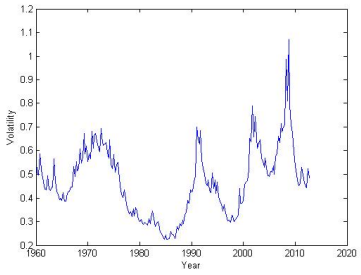


(f) UK

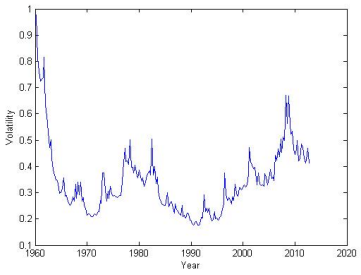


(g) US

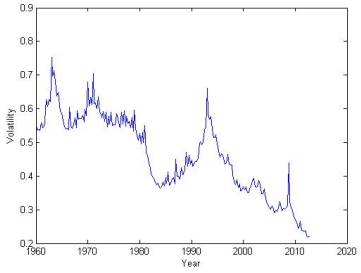
Figure 2: Trend Volatility



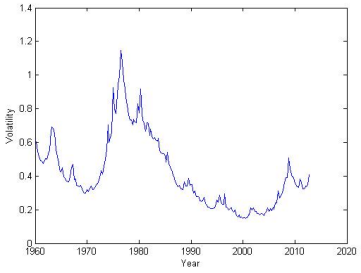
(a) Canada



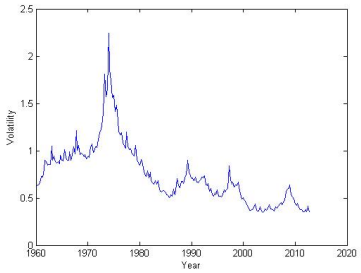
(b) France



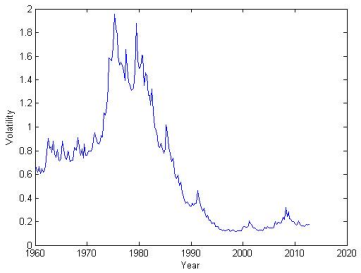
(c) Germany



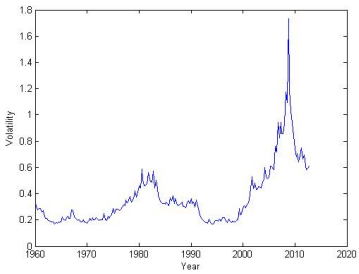
(d) Italy



(e) Japan

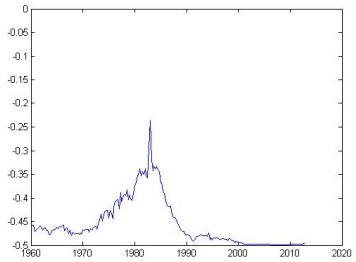


(f) UK

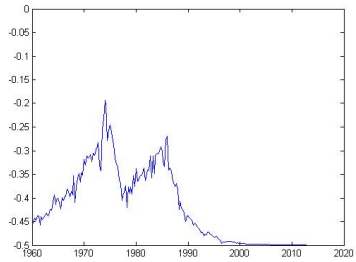


(g) US

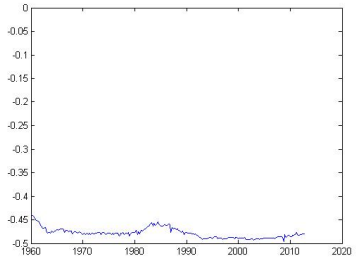
Figure 3: Transitory Volatility



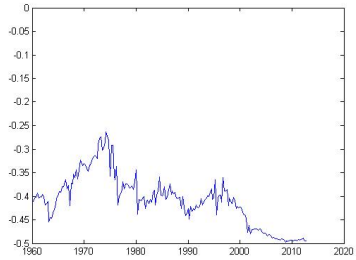
(a) Canada



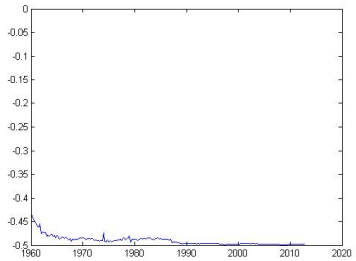
(b) France



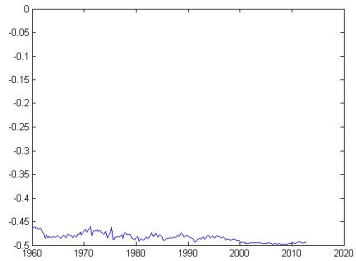
(c) Germany



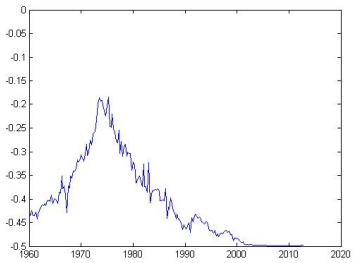
(d) Italy



(e) Japan

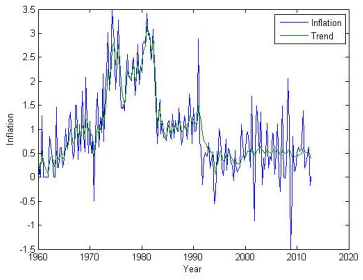


(f) UK

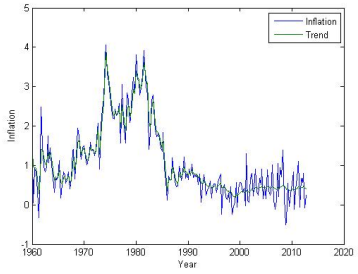


(g) US

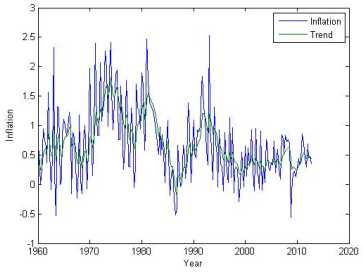
Figure 4: Autocorrelation of the First Difference of the Inflation



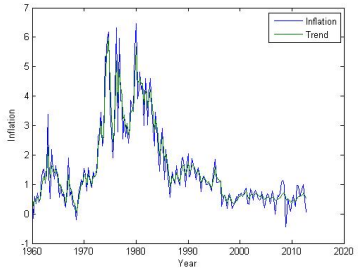
(a) Canada



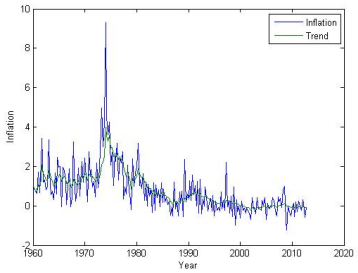
(b) France



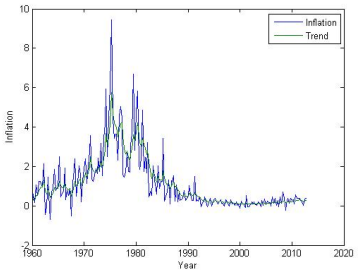
(c) Germany



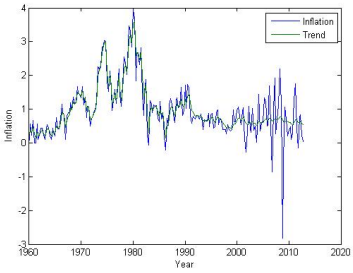
(d) Italy



(e) Japan

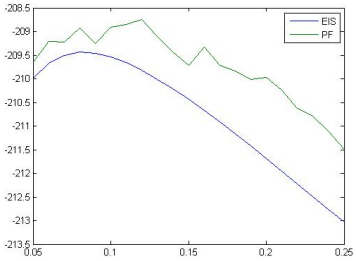


(f) UK

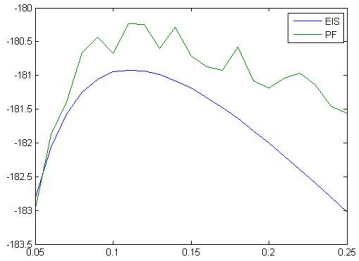


(g) US

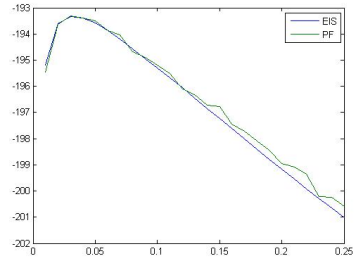
Figure 5: Inflation and its Estimated Trend



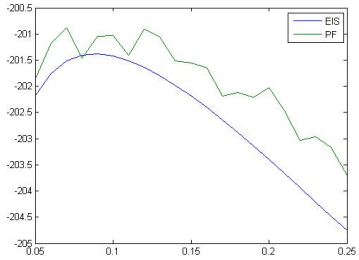
(a) Canada



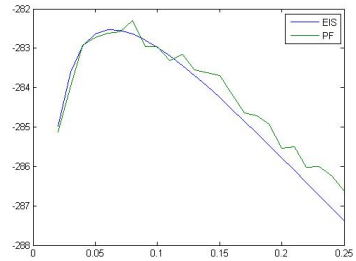
(b) France



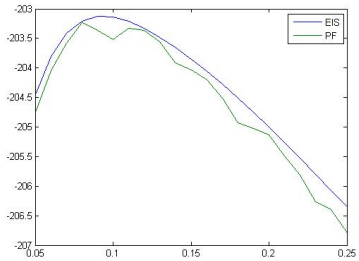
(c) Germany



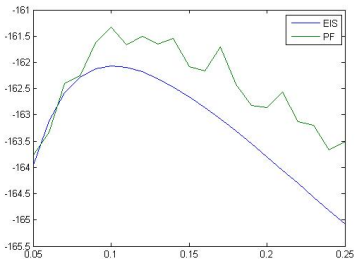
(d) Italy



(e) Japan



(f) UK



(g) US

Figure 6: Log-Likelihood Cuts