Semiparametric Bayesian inference for time-varying parameter regression models with stochastic volatility

Dimitrakopoulos Stefanos[1]

Warwick University, UK

Abstract

We develop a Bayesian semiparametric method to estimate a time-varying parameter regression model with stochastic volatility, where both the error distributions of the observations and parameter-driven dynamics are unspecified. We illustrate our methodology with an application to inflation.

Keywords: Dirichlet process, Markov chain Monte Carlo, stochastic volatility, time-varying parameters, inflation

JEL CODE: C11, C14, C15, C22

[1]Correspondence to: Department of Economics, Warwick University, Coventry, CV4 7ES, UK. E-mail address: s.dimitrakopoulos@warwick.ac.uk.

1 Introduction

A vast literature has demonstrated the gains from allowing for time-varying parameters in stochastic volatility models (TVP-SV models), when analyzing (macro)financial data (Primiceri, 2005; Cogley and Sargent, 2005; Stock and Watson, 2007; D'Agostino et al., 2013; Clark and Ravazzolo, 2015). Due to the presence of the stochastic volatility component the likelihood function for this class of models is intractable. As a result, researchers have developed Markov chain Monte Carlo (MCMC) algorithms for estimating the model parameters (see, for example, Nakajima (2011)).

In this paper, we consider two semiparametric extensions of the TVP-SV model, utilising a popular Bayesian prior for modelling unknown distributions, the Dirichlet process (DP) prior (Ferguson, 1973). We first use this prior to model in a flexible way the distribution of the dependent variable's innovation and second, to consider wider class of the distribution of the time-varying parameter's innovation. The resulting semiparametric TVP-SV model is referred to as the S-TVP-SV model. To estimate the model parameters and the unknown distributions, we propose an efficient MCMC algorithm.

The first semiparametric extension has already been applied in the context of standard stochastic volatility models (Jensen and Maheu, 2010; Delatola and Griffin, 2011). The second semiparametric extension is novel and constitutes our main contribution to the Bayesian semiparametric literature on TVP-SV models.

The motivation behind the S-TVP-SV model stems from the empirical literature on inflation modelling. Recently, evidence has been found of non-normality in modelling inflation persistence, leading to increased interest in non-Gaussian (fat-tailed) distributions for modelling inflation dynamics (Lanne and Saikkonen, 2011; Lanne et al., 2012; Chiu et al., 2014; Lanne, 2015). Our point of departure is an autoregressive version of the unobserved components with stochastic volatility (UC-SV) model, proposed by Stock and Watson (2007). Stock and Watson (2007) considered a UC-SV model that decomposed inflation into a trend and a transitory component and assumed fat-tailed error distributions for the observation and state equations to control for outliers.

In this paper, we generalize the approach of Stock and Watson (2007) to account for shocks that may not be symmetrically distributed, as economic systems may react differently in recessions and expansionary periods. Furthermore, if there are different regimes operating within the sample period, a fat-tailed distribution may be inadequate to capture this data characteristic. In our proposed model, each of the unconditional error distributions for the observations and the parameter-driven dynamics is allowed to follow an infinite mixture of normals.

2 Econometric set up

2.1 The TVP-SV model

Consider the following TVP-SV model

$$y_t = \mathbf{x}_t'\boldsymbol{\beta} + \mathbf{z}_t'\boldsymbol{\alpha}_t + \varepsilon_t,\ \varepsilon_t \sim N(\mu, \exp(h_t)),\ t = 1, ..., T, \tag{1}$$

$$\boldsymbol{\alpha}_{t+1} = \boldsymbol{\alpha}_t + \mathbf{u}_t,\ \mathbf{u}_t \sim N(\mathbf{0}, \boldsymbol{\Sigma}),\ t = 0, 1, ..., T-1, \tag{2}$$

$$h_{t+1} = \mu_h + \phi h_t + \eta_t,\ |\phi| < 1,\ \eta_t \sim N(0, \sigma_\eta^2). \tag{3}$$

Equation (1) contains two types of coefficients: the constant coefficient vector, $\boldsymbol{\beta}$, of dimension $k \times 1$ and time-varying coefficients, $\boldsymbol{\alpha}_t$, of dimension $p \times 1$. $\mathbf{x}_t$ and $\mathbf{z}_t$ are the design matrices which do not include an intercept and h_t is the log-volatility at time t.

Equation (2) is a random walk process which is initialized with $\boldsymbol{\alpha}_0 = \mathbf{0}$ and $\mathbf{u}_0 \sim N(\mathbf{0}, \boldsymbol{\Sigma}_0)$, where $N(\cdot, \cdot)$ denotes the normal distribution with the initial state error variance $\boldsymbol{\Sigma}_0$ being known.

The error terms ε_t and η_t are assumed to be independent[1] for all t. The error term

[1] In the context of stochastic volatility models, Jensen and Maheu (2014) assumed that the errors ε_t and η_t are correlated and modelled them nonparametrically, using DP priors.

ε_t follows a normal distribution with mean μ and time-varying variance $\sigma_t^2 = \exp(h_t)$. The dynamics of the log-volatility $h_t = \log(\sigma_t^2)$ are described by equation (3) which is a stationary ($|\phi| < 1$) first-order autoregressive process. This process is initialized with $h_1 \sim N(\mu_h/(1-\phi), \sigma_\eta^2/(1-\phi^2)$. The parameter ϕ is the persistence volatility that measures the degree of autocorrelation in h_t, and σ_η is the standard deviation of the shock to log-volatility.

We assume the following priors over the set of parameters $(\boldsymbol{\beta}, \sigma_\eta^2, \boldsymbol{\Sigma}, \mu_h, \mu)$,

$$\boldsymbol{\beta} \sim N(\boldsymbol{\beta}_0, \mathbf{B}),\ \sigma_\eta^2 \sim \mathcal{IG}(v_a/2, v_\beta/2),\ \boldsymbol{\Sigma} \sim IW(\delta, \Delta^{-1}),$$

$$\mu_h \sim N(\bar{\mu_h}, \bar{\sigma_h}^2), \quad \mu \sim N(\bar{\mu}, \bar{\sigma}^2),$$

where IW and $\mathcal{IG}$ denote the Inverse-Wishart distribution and the inverse gamma distribution, respectively. To guarantee that the persistence parameter ϕ satisfies the stationarity restriction, we assume $(\phi + 1)/2 \sim Beta(\phi_a, \phi_\beta)$.

2.2 Two semiparametric extensions

The advantage of Dirichet process modelling results from its theoretical properties, one of which is the clustering property. A detailed exposition of the statistical properties of the DP prior is given, among others, by Ghosal (2010).

The error term ε_t, is assumed to have an unspecified functional form based on the following Dirichlet process mixture (DPM) model

$$\begin{aligned}
&\varepsilon_t | \vartheta_t, h_t \sim N(\mu_t, \lambda_t^2 \exp(h_t)),\ \vartheta_t = (\mu_t, \lambda_t^2), t = 1..., T,\\
&\vartheta_t \overset{i.i.d}{\sim} G,\\
&G | a, G_0 \sim DP(a, G_0),\\
&G_0 = N(\mu_t; \mu_0, \tau_0 \lambda_t^2) \mathcal{IG}(\lambda_t^2; \tfrac{e_0}{2}, \tfrac{f_0}{2}),\\
&a \sim \mathcal{G}(\underline{c}, \underline{d}),
\end{aligned} \tag{4}$$

where μ_h in the stochastic volatility equation is set to zero for identification reasons. The unspecified functional form of the distribution of ε_t, given in (4), was first proposed by Jensen and Maheu (2010).

According to specification (4), the conditional distribution of ε_t given h_t and ϑ_t is Gaussian with mean μ_t and variance $\lambda_t^2 \exp(h_t)$. The $\vartheta_t = (\mu_t, \lambda_t^2)$ is generated from an unknown distribution G. For the prior base distribution G_0 we assume a conjugate normal-inverse gamma, $N(\mu_t; \mu_0, \tau_0 \lambda_t^2)\mathcal{IG}(\lambda_t^2; \frac{e_0}{2}, \frac{f_0}{2})$. A gamma prior distribution $\mathcal{G}(\underline{c}, \underline{d})$ is placed upon a, which is the precision parameter (positive scalar). As a tends to infinity G converges pointwise to G_0.

One can show that the unconditional distribution of ε_t follows an infinite mixture model with time-varying means and variances. So our DPM model is able to capture asymmetries and multiple modes that may characterize the data.

Furthermore, to capture the uncertainty about the distribution of $\mathbf{u}_t$, we impose on it the following novel flexible structure,

$$\begin{aligned}
&\mathbf{u}_t|\omega_t, \mathbf{\Sigma} \sim N(0, \omega_t^{-1}\mathbf{\Sigma}),\ t = 1, ..., T-1, \\
&\omega_t \overset{i.i.d}{\sim} G_\omega, \\
&G_\omega|a_\omega, G_{0\omega} \sim DP(a_\omega, G_{0\omega} = \mathcal{G}(\tfrac{e_\omega}{2}, \tfrac{e_\omega}{2})), \\
&a_\omega \sim \mathcal{G}(\underline{c}_\omega, \underline{d}_\omega).
\end{aligned} \tag{5}$$

The positive scale parameter ω_t in (5) comes from an unknown discrete distribution G_ω. The Dirichlet process prior in (5) is defined by the parameter a_ω and the base gamma distribution $G_{0\omega}$. As the precision parameter a_ω tends to infinity, G_ω converges pointwise to $G_{0\omega}$. In this case, the unconditional distribution of $\mathbf{u}_t$ is a multivariate Student-t distribution with e_ω degrees of freedom and as e_ω increases the error distribution mimics the Normal distribution. For small values of a_ω the unconditional distribution of $\mathbf{u}_t$ is a finite mixture of multivariate normals, each of which has the same mean. Therefore, our semiparametric approach for the distribution of $\mathbf{u}_t$ can capture the potential clustering in the mixing scalar parameter of the innovation's covariance matrix.

The TVP-SV model combined with the DPM models of (4) and (5) produces the semiparametric TVP-SV model (S-TVP-SV model).

3 Posterior analysis

3.1 The MCMC algorithm for the S-TVP-SV model

Define

$$\boldsymbol{y} = (y_1, ..., y_T),\ \boldsymbol{\alpha} = (\boldsymbol{\alpha}_1, ..., \boldsymbol{\alpha}_T),\ \mathbf{h} = (h_1, ..., h_T),$$
$$\boldsymbol{\theta} = (\vartheta_1, ..., \vartheta_T),\ \vartheta_t = (\mu_t, \lambda_t^2),\ \boldsymbol{\omega} = (\omega_1, ..., \omega_{T-1}).$$

Our MCMC scheme for the semiparametric model consists of two parts. In part I, we update the parameters ($\boldsymbol{\beta}$, $\boldsymbol{\Sigma}$, σ_η^2, $\boldsymbol{\alpha}$, $\mathbf{h}$, ϕ) and recover the error terms $\{\varepsilon_t\}_{t=1}^T$ and $\{\mathbf{u}_t\}_{t=1}^{T-1}$. We sample $\boldsymbol{\alpha}$ using the simulation smoothing algorithm of De Jong and Shephard (1995). To update the volatility vector $\mathbf{h}$ we apply the approach of Chan (2015), which is not based on Kalman-filter methods but on the precision sampler of Chan and Jeliazkov (2009).

Having calculated the error terms $\{\varepsilon_t\}_{t=1}^T$ and $\{\mathbf{u}_t\}_{t=1}^{T-1}$, we update, in part II, the DP parameters $(\boldsymbol{\theta}, \boldsymbol{\omega}, a, a_\omega)$ using marginal methods, since the DP is integrated out; see, for example, Escobar and West (1995) and MacEachern and Müller (1998).

Details of the MCMC algorithm for the semiparametric model along with a simulation study are given in the Online Appendix.

3.2 Posterior predictive density of the error term ε_{T+1}

A key quantity of interest in density estimation and an important feature of Bayesian inference is the posterior predictive density. With respect to the S-TVP-SV we obtain from the sampler the out-of-sample posterior predictive density for the (one-step ahead) error term ε_{T+1} conditional on the data $\boldsymbol{\Omega}_T = (\boldsymbol{y}, \mathbf{X}_T, \mathbf{Z}_T)$, where $\mathbf{X}_T =$

$(\mathbf{x}_1, ..., \mathbf{x}_T)$ and $\mathbf{Z}_T = (\mathbf{z}_1, ..., \mathbf{z}_T)$ which is given by

$$f(\varepsilon_{T+1}|\mathbf{\Omega}_T) = \int f(\varepsilon_{T+1}|\boldsymbol{\theta}^*, h_{T+1}, a)\pi(\boldsymbol{\theta}^*, h_{T+1}, a|\mathbf{\Omega}_T)d\boldsymbol{\theta}^* dh_{T+1}da. \tag{6}$$

$$\approx \tfrac{1}{L}\textstyle\sum_{l=1}^{L} f(\varepsilon_{T+1}|\boldsymbol{\theta}^{*(l)}, h_{T+1}^{(l)}, a^{(l)}),$$

where $\boldsymbol{\theta}^* = (\vartheta_1^*, ..., \vartheta_M^*)'$, $M \leq T$ is the set of unique values from $\boldsymbol{\theta}$, with ϑ_m^*=(μ_m^*, λ_m^{*2}), $m = 1, ..., M$ and M is the number of clusters in $\boldsymbol{\theta}$ (see also Online Appendix for further details). $\boldsymbol{\theta}^{*(l)}$ and $a^{(l)}$ are simulated samples of $\boldsymbol{\theta}^*$ and a respectively and $h_{T+1}^{(l)}$ is a posterior draw generated from $N(\phi^{(l)}h_T^{(l)}, \sigma_\eta^{2(l)})$. L is the number of iterations after the burn-in period. The predictive density of ε_{T+1} conditional on $\boldsymbol{\theta}^{*(l)}, h_{T+1}^{(l)}$ and $a^{(l)}$ is a mixture of a Student-t density and Normal densities, namely,

$$f(\varepsilon_{T+1}|\boldsymbol{\theta}^{*(l)}, h_{T+1}^{(l)}, a^{(l)}) = \frac{a^{(l)}}{a^{(l)} + T} q_t(\varepsilon_{T+1}|\mu_0, (\exp(h_{T+1}^{(l)}) + \tau_0)f_0/e_0, e_0)$$

$$+\tfrac{1}{a^{(l)}+T}\textstyle\sum_{m=1}^{M^{(l)}} n_m^{(l)} N(\varepsilon_{T+1}|\mu_m^{*(l)}, \exp(h_{T+1}^{(l)})\lambda_m^{*2(l)}), \tag{7}$$

where $q_t(.|m, v, u)$ is the Student-t distribution with mean m, degrees of freedom u and scale factor v. The quantity n_m is explained in the Online Appendix.

3.3 Model comparison

We conduct Bayesian model comparison, using the Deviance information criterion (DIC) (Spiegelhalter et al., 2002) and cross-validation predictive densities. Further details on how to implement these methods are provided in the Online Appendix.

4 Empirical application

We use data on US quarterly consumer price index (CPI) inflation from 1948Q1 to 2013Q2. For modelling inflation persistence, we consider the following autoregressive TVP-SV (AR-TVP-SV) model,

$$y_t = \alpha_{1,t} + \alpha_{2,t} y_{t-1} + \varepsilon_t,\ \varepsilon_t \sim N(0, \exp(h_t)),\ t = 1, ..., T,$$

$$\boldsymbol{\alpha}_{t+1} = \boldsymbol{\alpha}_t + \mathbf{u}_t,\ \mathbf{u}_t \sim N(\mathbf{0}, \boldsymbol{\Sigma}),\ t = 0, 1, ..., T-1,$$

$$h_{t+1} = \mu_h + \phi h_t + \eta_t,\ |\phi| < 1,\ \eta_t \sim N(0, \sigma_\eta^2),$$

where $y_t = 400 * log(l_t/l_{t-1})$ denotes the CPI inflation and l_t is the quarterly CPI figure. We plot y_t in Figure 1.

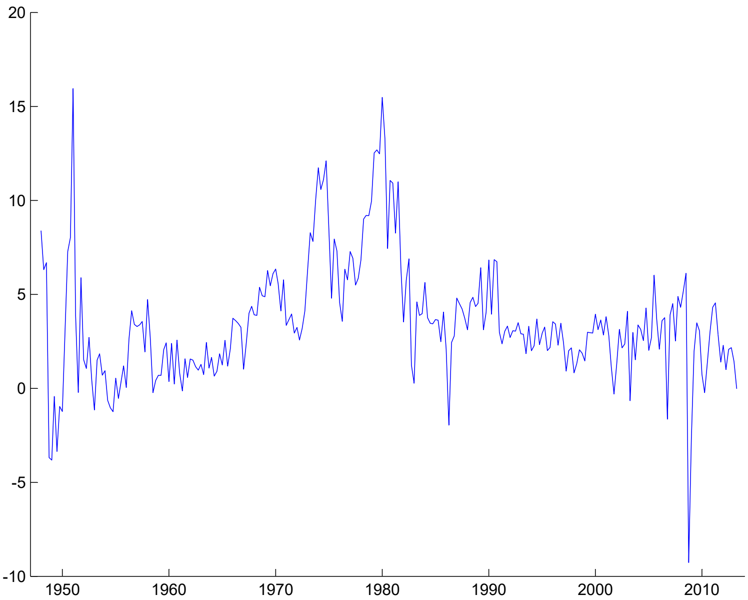

Figure 1: The inflation path from 1948Q1 to 2013Q2.

In the semiparametric version of the AR-TVP-SV model, denoted as the AR-S-TVP-SV model, the error terms ε_t and $\mathbf{u}_t$ follow the DPM models of (4) and (5), respectively[2]. For comparison purposes, we also estimated the AR-TVP-SV model,

[2]A limitation of the semiparametric model is that mixing over the time-varying parameters

with the errors ε_t and $\mathbf{u}_t$ being Student-t distributed. We refer to this model as AR-St-TVP-SV. The St-TVP-SV model is presented in the Online Appendix.

After discarding the first 80000 draws, we run the sampler for 150000 iterations. To monitor convergence we use the CD statistics of Geweke (1992) and the inefficiency factor (IF); see, for example, Chib (2001). For the AR-S-TVP-SV model, we chose the same hyperparameters for the priors as in the simulation study (see Online Appendix).

The estimation results are presented in Table 1. Across all models of Table 1, all the parameters but μ_h are significant. Based on the DIC and CV values (Table 1), the AR-S-TVP-SV model has the best fit to the data. The AR-TVP-SV model is the least preferred model.

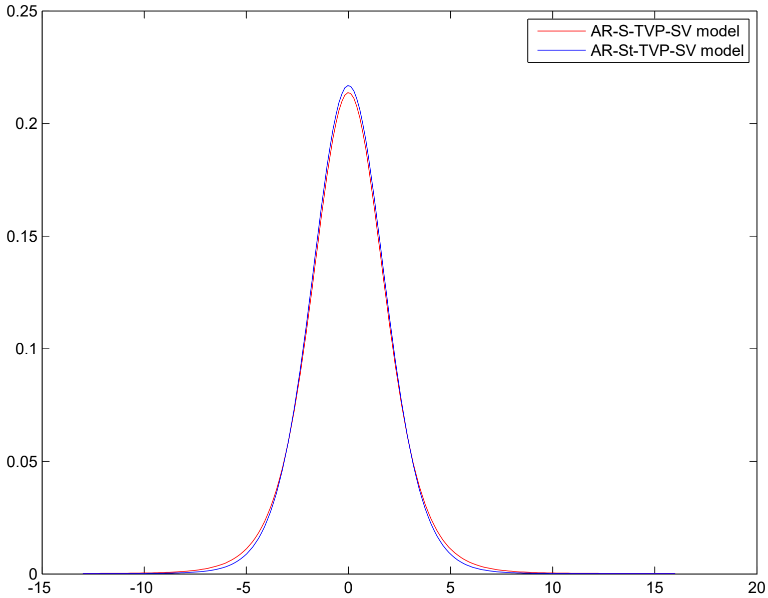

Figure 2: Posterior predictive densities for ε.

The posterior predictive density of the error term ε_t for the AR-S-TVP-SV model, which is plotted against that of the AR-St-TVP-SV model (Figure 2), indicates that the distribution of the dependent's variable innovation is nonnormal (with kurtosis 5.6254 and skewness 1.9735). This empirical finding is supported by the fact that

scaled covariance matrix fails to capture the regime switching behavior of the Sims and Zha (2006) model. A change from one regime's parameter values to another is only possible if the mixture representation of the parameter innovations is mixed over the mean vector of the normal kernel.

the semiparametric model requires $M = 4.3764$ clusters to fit the data (Table 1). The parametric models inflate the volatility parameter σ_η to compensate for the excess kurtosis found in the data; the estimated degrees of freedom $v1$ for the AR-St-TVP-SV is 9.3842.

The path of the posterior estimates of $\exp(h_t)$ obtained from the semiparametric model shows high inflation volatility during the Great Moderation and the Great Recession (Figure 3).

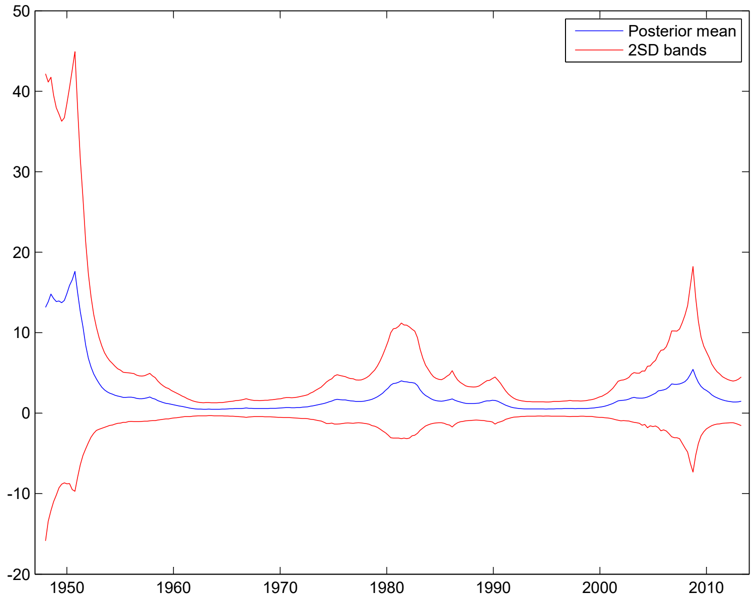

Figure 3: Evolution of $\exp(h_t)$ obtained from the AR-S-TVP-SV model; posterior mean (blue), two standard deviation bands (red).

Also, the AR-S-TVP-SV model highlights some degree of clustering in the mixing scalar parameter of the innovation's covariance; the number of clusters in $\boldsymbol{\omega}$ was found to be M_ω =4.0796 (Table 1)-M_ω is explained in the Online Appendix.

Figure 4 presents the estimates of $\boldsymbol{\alpha}_t$ for the AR-S-TVP-SV model. As can be seen, there is apparent time-variation in these estimates, highlighting the importance of allowing for time-varying parameters. Similar results were produced by the rest of the models (see Online Appendix).

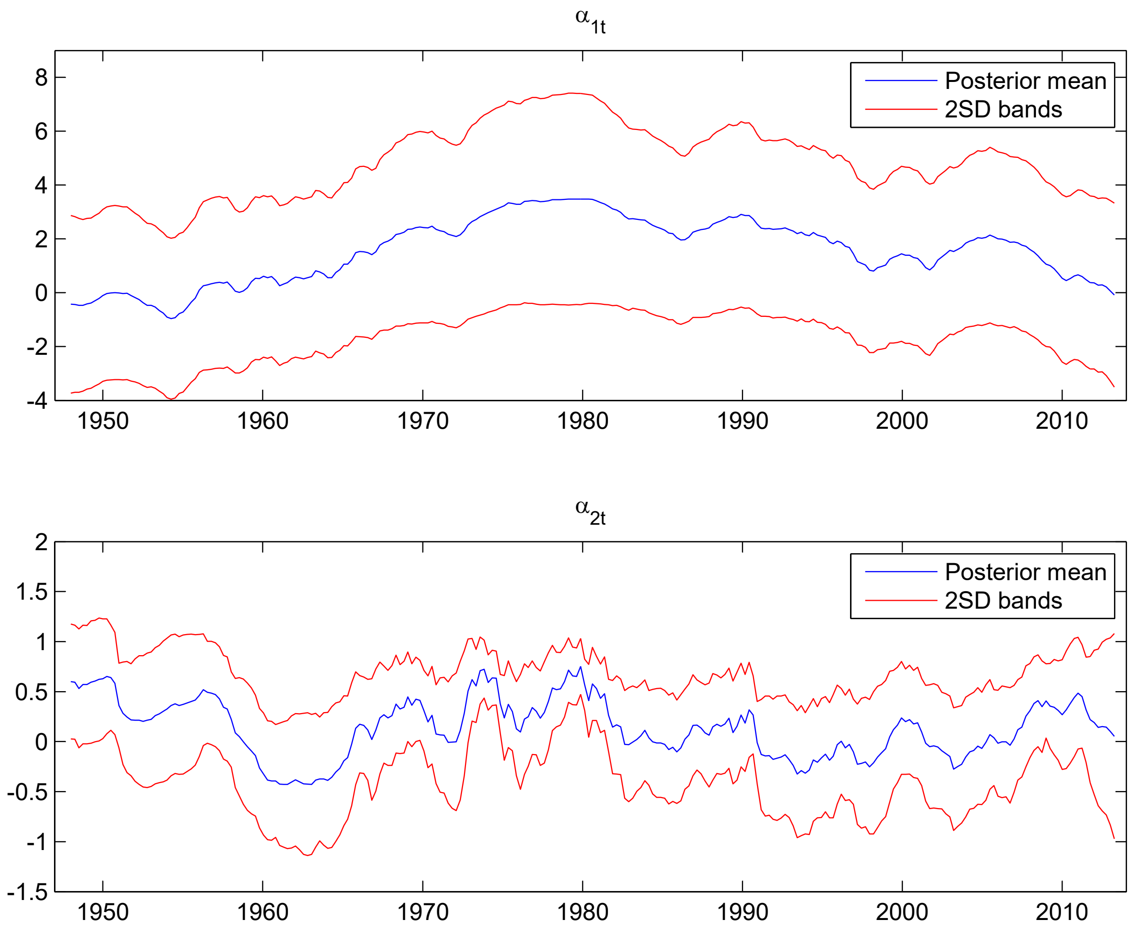

Figure 4: Evolution of $\boldsymbol{\alpha}_t$ obtained from the AR-S-TVP-SV model; posterior mean (blue), two standard deviation bands (red).

5 Conclusions

We proposed a novel Bayesian semiparametric time-varying parameter regression model with stochastic volatility (TVP-SV), where both the error distributions of the observations and parameter-driven dynamics were left unspecified. The Dirichlet process was used as a prior to these unknown distributions. We devised an efficient Markov chain Monte Carlo algorithm to estimate the model parameters and unknown distributions. An autoregressive version of the proposed model was applied to inflation persistence. The empirical results showed that the proposed model had better fit to the data than competing parametric models. For future research it would be interesting to enrich the proposed semiparametric model with a nonparametric leverage effect, as macro shocks that have the greatest effect on the economy are often asymmetrically distributed. Extending the TVP-SV model of this paper in this direction could prove fruitful relative to existing TVP-SV findings.

Table 1: Empirical results

Model	AR-TVP-SV			AR-S-TVP-SV			AR-St-TVP-SV		
	Mean	CD	IF	Mean	CD	IF	Mean	CD	IF
$\mathbf{\Sigma}_{11}$	0.2157*	0.642	63.84	0.2808*	0.771	59.20	0.1544*	0.846	37.68
	(0.1076)			(0.1590)			(0.0573)		
$\mathbf{\Sigma}_{22}$	0.0483*	0.271	17.92	0.0679*	0.017	26.14	0.0358*	0.347	6.09
	(0.0114)			(0.0216)			(0.0070)		
ϕ	0.9586*	0.079	53.77	0.9655*	0.004	47.31	0.9674*	0.000	13.03
	(0.0276)			(0.0232)			(0.0233)		
μ_h	0.3664	0.630	7.48				0.2110	0.933	9.22
	(0.9169)						(0.8628)		
σ_η	0.3964*	0.207	115.85	0.2577*	0.009	113.23	0.2964*	0.010	70.76
	(0.1022)			(0.0830)			(0.0633)		
M				4.3764*	0.333	75.07			
				(2.4396)					
M_ω				4.0796*	0.477	22.67			
				(2.1991)					
$v1$							9.3842*	0.242	142.56
							(14.0543)		
$v2$							65.8133*	0.088	50.84
							(21.5601)		
CV	0.4909			0.5234			0.5122		
DIC	2501.9			1948.2			2018.8		

*Significant based on the 95% highest posterior density interval. Standard errors in parentheses.

References

Chan JCC. 2015. The stochastic volatility in mean model with time-varying parameters: An application to inflation modeling. *Journal of Business & Economic Statistics, forthcoming.*

Chan JCC, Jeliazkov I. 2009. Efficient simulation and integrated likelihood estimation in state space models. *International Journal of Mathematical Modelling and Numerical Optimisation* **1**: 101–120.

Chib S. 2001*. Markov Chain Monte Carlo methods: Computation and Inference.* In Heckman JJ. and Leamer E, (Eds.), Handbook of Econometrics, Volume 5, (pp. 3569-3649). Elsevier.

Chiu CWJ, Haroon M, Pinter G. 2014. Fat-tails in VAR models. *Working paper.*

Clark TE, Ravazzolo F. 2015. Macroeconomic forecasting performance under alternative specifications of time-varying volatility. *Journal of Applied Econometrics* **30**(4): 551–575.

Cogley T, Sargent TJ. 2005. Monetary policy and learning the conquest of US inflation: Learning and robustness to model uncertainty. *Review of Economic Dynamics* **8**(2): 528 – 563.

D'Agostino A, Gambetti L, Giannone D. 2013. Macroeconomic forecasting and structural change. *Journal of Applied Econometrics* **28**(1): 82–101.

De Jong P, Shephard N. 1995. The simulation smoother for time series models. *Biometrika* **82**(2): 339–350.

Delatola EI, Griffin JE. 2011. Bayesian nonparametric modelling of the return distribution with stochastic volatility. *Bayesian Analysis* **6**(4): 1–26.

Escobar M, West M. 1995. Bayesian density estimation and inference using mixtures. *Journal of the American Statistical Association* **90**(430): 577–588.

Ferguson TS. 1973. A Bayesian analysis of some nonparametric problems. *The Annals of Statistics* **1**(2): 209–230.

Geweke J. 1992. *Evaluating the Accuracy of Sampling-Based Approaches to the Calculation of Posterior Moments.* In Bernardo J, Berger J, Dawid A and Smith A, (Eds.), Bayesian Statistics 4, Oxford: Clarendon Press, (pp. 641-649).

Ghosal S. 2010 . *The Dirichlet process, related priors and posterior asymptotics.* in Hjort N, Holmes C, Mller P, Walker SG. (Eds.). Bayesian nonparametrics: principles and practice. Cambridge University Press.

Jensen M, Maheu JM. 2010. Bayesian semiparametric stochastic volatility modelling. *Journal of Econometrics* **157**(2): 306–316.

Jensen M, Maheu JM. 2014. Estimating a semiparametric asymmetric stochastic volatility model with a dirichlet process mixture. *Journal of Econometrics* **178**: 523–538.

Lanne M. 2015. Noncausality and inflation persistence. *Studies in Nonlinear Dynamics and Econometrics* **19**(4): 469481.

Lanne M, Luoto J, Saikkonen P. 2012. Optimal forecasting of noncausal autoregressive time series. *International Journal of Forecasting* **28**(3): 623 – 631.

Lanne M, Saikkonen P. 2011. Noncausal autoregressions for economic time series. *Journal of Time Series Econometrics.*

MacEachern S, Müller P. 1998. Estimating mixture of Dirichlet process models. *Journal of Computational and Graphical Statistics* **7**(2): 223–238.

Nakajima J. 2011. Time-varying parameter VAR model with stochastic volatility: An overview of methodology and empirical applications. *Monetary and Economic Studies* **29**: 107–142.

Primiceri GE. 2005. Time varying structural vector autoregressions and monetary policy. *The Review of Economic Studies* **72**(3): 821–852.

Sims CA, Zha T. 2006. Were there regime switches in U.S. monetary policy? *The American Economic Review* **96**(1): 54–81.

Spiegelhalter D, Best N, Carlin B, Van Der Linde A. 2002. Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **64**(4): 583–639.

Stock JH, Watson MW. 2007. Why has U.S. inflation become harder to forecast? *Journal of Money, Credit, and Banking* **39**: 3–33.