DYNAMIC ECONOMETRIC MODELS
Vol. 7 – Nicolaus Copernicus University – Toruń – 2006

Jacek Kwiatkowski
Nicholas Copernicus University in Toruń

A Bayesian Estimation and Testing of STUR Models With Application to Polish Financial Time-series [1]

1. Introduction

One of the basic assumptions in the (theoretical) finance is that the logarithmic prices of financial series of assets or exchange rates, display random walk-type behavior. Econometric tradition has been to incorporate ARIMA models to capture the dynamics of economic time series. However, recent empirical test results for finance series suggest that they are often processes that have a root that is not constant, but is stochastic. These processes are known as stochastic unit root process (STUR). One of their important property is that they have a root that is time-varying around unity, therefore they can be stationary or explosive. Many empirical results on the identification of STUR processes are encountered in Leybourne, McCabe and Tremayne (1996), Granger and Swanson (1997), Jones and Marriott (1999), Sollis, Leybourne and Newbold (2000), Kwiatkowski and Osińska (2004), Kwiatkowski (2005a and 2005b).

The aim of this paper is to present with the Bayesian estimation and testing of STUR processes, where the random parameter follows first-order stationary autoregressive process. Probably the first attempt to employ the Bayesian inference was presented in Jones and Marriott (1999). In their paper they have used Granger and Swanson (1997) model to derive posterior marginals and summary statistics. This paper is concerned with the STUR model introduced by Leybourne, McCabe and Tremayne (1996), which is computationally less demanding, and easy to implement. The marginal posteriors of parameters and summary statistics can be obtained by Gibbs sampler.

---

[1] This research was supported by Polish Committee of Science grant 2 H02B 015 25; *e-mail: jkwiat@uni.torun.pl.*

The paper is organized as follows. Section 2 presents the stochastic unit root model as well as its Bayesian estimation and testing.. Section 3 provides an empirical application to the stock returns and exchange rates of zloty, for weekly sampling frequencies. Section 4 concludes. Details on the implementation of Gibbs sampler for STUR models are provided in the appendix.

2. The Model and Bayesian Inference

The STUR (*stochastic unit roots*) processes are presented by Leybourne, McCabe and Tremayne (1996) and Granger and Swanson (1997). Consider the following STUR model:

$$y_t = (1+\beta_t)y_{t-1} + \varepsilon_t , \tag{2.1}$$

where y_t denotes an observed process at time *t* and β_t is a first-order stationary autoregressive process:

$$\beta_t = \alpha + \phi_1(\beta_{t-1} - \alpha) + \eta_t . \tag{2.2}$$

Parameter ϕ_1 is the autoregression coefficient, which is a number between -1 and 1. Here ε_t and η_t are white noise processes having zero mean and respective variances σ^2 and ω^2. We also assume that ε_t and η_t are mutually independent.

When $\omega^2 = 0, \phi_1 = 0$ and $\alpha = 0$, y_t follows the random walk process. For $\omega^2 > 0$ and free (ϕ_1, α), we have a process with a unit root in mean, called a stochastic unit root process. The parameters β_t follow an autoregressive mechanism, so the original series tends to possess one unit root in the long run, but in sub-periods may have stationary or explosive roots.

Let assume that y_t is a process with stationary first differences Δy_t:

$$\Delta y_t = \beta_t y_{t-1} + \varepsilon_t , \tag{2.3}$$

$$\beta_t = \alpha + \phi_1(\beta_{t-1} - \alpha) + \eta_t , \tag{2.4}$$

where Δy_t denotes first differences of the observed process y_t at time *t.* Values for autoregressive parameter ϕ_1 lie in the stationary region; $\phi_1 \in (-1,1)$. Stochastic processes $\varepsilon_t \sim N(0,\sigma^2)$ and $\eta_t \sim N(0,\omega^2)$ are assumed to be independent.

Due to the Normality of the unobserved random processes β_t and ε_t, we can express the model in the following structure:

$$\Delta y_t \big| y_{t-1}, \beta_t, \sigma^2 \sim N\left(\beta_t y_{t-1}, \sigma^2\right),$$
$$\beta_t \big| \beta_{t-1}, \alpha, \phi_1, \omega^2 \sim N\left(\alpha + \phi_1(\beta_{t-1} - \alpha), \omega^2\right). \tag{2.5}$$

For the special case of (2.5) in which the stochastic unit root process for y_t follows an for β_t *iid* process we can inference of parameters of interest by putting $\phi_1 = 0$.

Therefore the sampling distribution is:

$$p(\Delta y \mid y_{-1}, \beta, \beta_0, \theta) = \prod_{t=1}^{T} f_N\left(\beta_t \mid \alpha + \phi_1(\beta_{t-1} - \alpha), \omega^2\right) \prod_{t=1}^{T} f_N\left(\Delta y_t \mid \beta_t y_{t-1}, \sigma^2\right), \tag{2.6}$$

where $\Delta y = (\Delta y_1, \Delta y_2, ..., \Delta y_T)'$, $y_{-1} = (y_{0,}, y_1, y_2, ..., y_{T-1})'$, $\theta = \left(\alpha, \phi_1, \omega^2, \sigma^2\right)$, $\alpha \in R$, $\phi_1 \in (-1,1)$, $\omega^2 \in R_+$, $\sigma^2 \in R_+$ and $\beta = (\beta_1, \beta_2, ..., \beta_T)' \in R^T$, and T denotes number of observations, $f_N\left(x \mid c, w^2\right)$ denotes Normal distribution with mean c and variance w^2 [2].

The prior information about all parameters is reflected by the following density:

$$p(\theta) \propto f_N\left(\alpha \mid \mu_\alpha, \sigma_\alpha^2\right) f_N\left(\phi_1 \mid \mu_{\phi_1}, \sigma_{\phi_1}^2\right) f_{Inv-Gam}\left(\sigma^2 \mid a_1, b_1\right) f_{Inv-Gam}\left(\omega^2 \mid a_2, b_2\right), \tag{2.7}$$

where $f_{Inv-Gam}(x \mid a, b)$ means Inverse Gamma distribution with shape parameter a and scale parameter b.

Since the parameter β_t is a part of the model, we can assume that all information about β_t is included in the likelihood (Jones and Marriott, 1999; Jostova and Philipov, 2005). For the autoregression coefficient ϕ_1 the prior density is truncated to the stationary region $(-1,1)$.

Under this prior structure (2.7), the joint posterior density of the parameters is:

$$p(\beta, \theta \mid \Delta y, y_{-1}, \beta_0) \propto \prod_{t=1}^{T} f_N\left(\beta_t \mid \alpha + \phi_1(\beta_{t-1} - \alpha), \omega^2\right) \prod_{t=1}^{T} f_N\left(\Delta y_t \mid \beta_t y_{t-1}, \sigma^2\right)$$
$$\times f_N\left(\alpha \mid \mu_\alpha, \sigma_\alpha^2\right) f_N\left(\phi_1 \mid \mu_{\phi_1}, \sigma_{\phi_1}^2\right) f_{Inv-Gam}\left(\sigma^2 \mid a_1, b_1\right) f_{Inv-Gam}\left(\omega^2 \mid a_2, b_2\right). \tag{2.8}$$

[2] The statistical distributions used in this paper are presented e.g. in Gelman, Carlin, Stern and Rubin (1997).

In order to obtain the posterior marginals and summary statistics for them, we could employ Gibbs sampler algorithm. For STUR model it is quite easy because, the proper prior densities (2.7) leads to standard conditional posteriors. The details of Gibbs sampler for STUR model are included in the appendix.

In the Bayesian approach to comparing models, it is considered useful to employ probabilities to represent degree of belief associated with alternative models. For the STUR model we can test whether the random process follows the first-order autoregressive process or the white noise process. We can also test whether the data can be considered as generated by the STUR or the exact unit roots process.

Using Bayes's theorem, the posterior odds ratio for this problem is given by:

$$\frac{p(M_i|\Delta y)}{p(M_j|\Delta y)}=\frac{p(M_i)}{p(M_j)}\frac{p(\Delta y|M_i)}{p(\Delta y|M_j)},$$

where M_i and M_j are the two models we are comparing. Assigning equal prior model probabilities $p(M_i)=p(M_j)=0.5$, comparison of the models can be summarized by the Bayes factor:

$$B_{ij}=\frac{p(\Delta y|M_i)}{p(\Delta y|M_j)}. \tag{2.9}$$

If this ratio is larger then one, we can say that the data supports model M_i over model M_j.

The practical difficulty in implementing posterior odds ratio is the computation of the marginal data density value $p(\Delta y|M_i)$. For the STUR model this integral is not analytically tractable. One of simple, numerical approaches is to consider Newton and Raftery's (1994) harmonic mean estimator:

$$p(\Delta y|M_i)^{-1}=\frac{1}{K}\sum_{k=1}^{K}p\left(\Delta y|\theta_i^{(k)},M_i\right)^{-1}, \tag{2.10}$$

where the $\theta_i^{(k)}$ are drawn from the posterior using the Markov chain Monte Carlo (MCMC) methods. This estimator is easy to implement but can be quite unstable, because it fails to obey the Gaussian central limit theorem (Carlin and Louis, 2000). Although for many applications the Newton and Raftery (N-R) estimator is stable enough and close to the true value of marginal data density (Osiewalski and Pipień, 2004). In the STUR case, the N-R estimator is unstable because the small conditional likelihood values overly influence the harmonic mean values. Therefore we can only test STUR process with random parameter which follows white noise. In that case we can integrate out analytically the density (2.6) with respect to β_t. The conditional distribution of Δy_t at time t is Normal with mean αy_{t-1} and variance $\sigma^2 + \omega^2 y_{t-1}^2$. Similar approach is used by Nicholls and Quinn (1982) for the likelihood of RCA models.

In order to test autoregression of the random parameter, we can use less formal approach, namely the highest probability density (HPD) interval. This interval contains all a posteriori most likely values of ϕ_1.

3. Application to Polish Financial Time-series

We apply the STUR model to weekly returns on stock and stock indexes listed at the Stock Exchange in Warsaw. We also estimate the STUR model to weekly exchange rates of foreign currencies in zlotys. The weekly stock and exchange rates returns cover almost 5-year sample periods from January 2000 until September 2005. It gives approximately 292 observation. We use the logarithmic transformations of the original series P_t, computed as $y_t = \ln(P_t)$. Diffused but proper joint prior distributions reflects the lack of information about parameters. Values for autoregressive parameter ϕ_1 lie in the stationary region between -1 and 1. Hence, for these parameters, we select a truncated-Normal prior with mean 0 and large variance equal to 10. For the variance σ^2 and ω^2, we use an Inverse Gamma prior with shape and scale parameters equal to 0.01. For the unconditional mean parameter, α, Normal prior with mean 0 and variance equal to 1 is selected. Joint prior structure is expressed by equation (2.7). All models have equal prior probabilities.

We apply the Bayesian methodology for two mutually exclusive and independent models from each other:

$$RW : \Delta y_t = \varepsilon_t ,$$
$$WN : \Delta y_t = \beta_t y_{t-1} + \varepsilon_t ,$$
$$\beta_t = \alpha + \eta_t \ .$$

The Gibbs sampler for the Bayesian analysis of the STUR model is presented in appendix. The logarithms of the Bayes factor in favor of random walk computed by Newton-Raftery for the stock and indexes returns are given in table 1. Table 2 contains logs of Bayes factor in favor of random walk for weekly exchange rates. In order to provide necessary level of accuracy of Newton – Raftery estimator, we simulated 500000 draws. Both tables also show the ranking obtained using this approximation.

Table 1. Decimal logs of Bayes factor in favor of random walk, approximated by Newton – Raftery estimator for indexes WIG , WIG20, MIDWIG, TECHWIG and for stock returns.

Weekly returns	Random walk		STUR with WN	
	Rank	$\log_{10}(B_{RWRW})$	Rank	$\log_{10}(B_{RWWN})$
WIG	1	0.0000	2	95.175
WIG20	1	0.0000	2	55.668
MIDWIG	1	0.0000	2	95.434
TECHWIG	1	0.0000	2	21.697
AMATOR	1	0.0000	2	11.127
BRE	1	0.0000	2	18.424
BZWBK	1	0.0000	2	18.515
DEBICA	1	0.0000	2	26.371
HANDLOWY	1	0.0000	2	35.224
MIESZKO	2	0.0000	1	-2.098
MILLENNIUM	2	0.0000	1	-3.047
OPTIMUS	2	0.0000	1	-12.658
PROCHNIK	1	0.0000	2	1.226
TPSA	1	0.0000	2	8.569
WAWEL	1	0.0000	2	14.587

Notes: Column headed Rank contains the rank of the respective models according to Bayes factor.

The results in tables 1 and 2 show that there is no substantial evidence for the presence of stochastic unit root. Notice that only for the three stock returns, namely MIESZKO. MILLENNIUM and OPTIMUS, Bayes factor supports STUR model over random walk. The results in table 2 suggest that is poor evidence of STUR to weekly exchange rate returns. After estimating the STUR model, it turns out that random unit root model is not very popular for selected financial series.

In order to examine autoregressive behavior of random parameter β_t, we have to analyze posterior distribution of ϕ_1 parameter. The posterior quantile information and other characteristics are summarized in table 3. In the case of these three series there is no evidence, that random parameter follows autoregressive process.

Table 2. Decimal logs of Bayes factor in favor of random walk. approximated by Newton – Raftery estimator for exchange rates: Australian dollar (AUD), Canadian dollar (CAD), Swiss franc (CHF), Czech koruna (CZK), Danish crone (DKK), Euro (EUR), Pound sterling (GBP), Japanese yen (JPY), and US dollar (USD).

Weekly returns	Random walk		STUR with WN	
	Rank	$\log_{10}(B_{RWRW})$	Rank	$\log_{10}(B_{RWWN})$
AUD	1	0.0000	2	16.2361
CAD	1	0.0000	2	15.9253
CHF	1	0.0000	2	20.5405
CZK	1	0.0000	2	40.8721
DKK	1	0.0000	2	8.1587
EUR	1	0.0000	2	29.9665
GBP	1	0.0000	2	37.9542
JPY	1	0.0000	2	16.2075
USD	1	0.0000	2	26.2151

Notes: Column headed Rank contains the rank of the respective models according to Bayes factor.

Table 3. Posterior summaries for autoregression parameter ϕ_1 calculated for MIESZKO, MILLENNIUM and OPTIMUS

	Posterior quantile				Posterior	
Series	0.0025	0.500	0.975	$P(\phi_1 > 0 \mid \Delta y)$	Mean	Standard deviation
MIESZKO	-0.186	0.020	0.221	0.578	0.019	0.104
MILLENNIUM	-0.300	-0.105	0.090	0.141	-0.105	0.099
OPTIMUS	-0.124	0.023	0.174	0.618	0.023	0.076

Table 4 presents the posterior means and standard deviations (in parenthesis) for STUR parameters α, ω^2 and σ^2, calculated for MIESZKO, MILLENNIUM and OPTIMUS, where random parameter follows *iid*.

Table 4. Posterior means and standard deviations in (parentheses) of the coefficient estimates of STUR models with random parameter which follows *iid.* process.

	Parameters		
Series	α	ω^2	σ^2
MIESZKO	-0.0005 (0.0026)	0.0012 (0.0002)	0.0022 (0.0005)
MILLENNIUM	0.0034 (0.1425)	0.0018 (0.0003)	0.0021 (0.0005)
OPTIMUS	0.0021 (0.0847)	0.0009 (0.0001)	0.0019 (0.0005)

4. Conclusion

The paper presents a Bayesian estimation of the stochastic unit root model, where random parameter follows white noise or first-order autoregressive process. The results set out in tables 1 and 2 demonstrate that the STUR model does not improve upon a random walk model, either for weekly returns on stock or exchange rates. Only three of twenty four series exhibit random unit root behavior.

References

Box, G.E.P., Jenkins, G.M. (1976), *Time Series Analysis: Forecasting and Control,* San Francisco, Holden-Day.

Carlin, B.P., Louis, T.A. (2000), *Bayes and Empirical Bayes Methods for Data Analysis,* New York, Chapman & Hall/CRC.

Gelman, A., Carlin, J., Stern, H., Rubin, D. (1997), *Bayesian Data Analysis*. London, Chapman & Hall.

Granger, C.W.J., Swanson, N.R. (1997), An Introduction to Stochastic Unit–root Process, *Journal of Econometrics*, 80, 35-62.

Jones, C.R., Marriott, J.M. (1999), A Bayesian analysis of stochastic unit root models, *Bayesian Statistics*, 6, 785-794.

Jostova, G., Philipov, A. (2005), Bayesian analysis of stochastic betas. *Journal of Financial and Quantitative Analysis*, 40, 4, 747-778.

Kwiatkowski, J. (2005a), Maximum likelihood estimation of stochastic unit root models with GARCH disturbances. *Forecasting Financial Markets, Theory and Applications*, Lodz, 149-157.

Kwiatkowski, J. (2005b), Bayesian analysis of STUR models, *working paper version.*

Kwiatkowski, J., Osińska, M. (2005), Forecasting STUR processes. A comparison to threshold and GARCH models. *Acta Universitatis Lodziensis Folia Oeconomica*, 190, Lodz, 159-176.

Leybourne, S.J., McCabe, B.P.M., Mills, T.C. (1996), Randomized unit root processes for modelling and forecasting financial time series: theory and applications, *Journal of Forecasting*, 15, 253-270.

Leybourne, S.J., McCabe, B.P.M., Tremayne, A.R. (1996), Can economic time series be differenced to stationarity? *Journal of Business and Economic Statistics*, 14, 435-446.

Nicholls, D.F., Quinn, B.G. (1982), *Random Coefficient Autoregressive Models: An Introduction,* New York, Springer-Verlag.

Newton, M.A., Raftery, A.E. (1994), Approximate Bayesian inference by the weighted likelihood bootstrap (with discussion). *Journal of the Royal Statistical Society B*, 56, 3-48.

Osiewalski, J., Pipień, M. (2004), Bayesian comparison of bivariate ARCH-type models for main exchange rates in Poland, *Journal of Econometrics*, 123, 371-391.

Sollis, R., Leybourne, S.J., Newbold, P. (2000), Stochastic unit roots modeling of stock price indices, *Applied Financial Economics*, 10, 311-315.

Appendix

The Gibbs sampler and posterior densities for STUR model

The Gibbs sampler is a Markov chain Monte Carlo method for drawing from a joint posterior distribution by sampling from the conditional distribution. (Gelman, Carlin, Stern and Rubin, 1997). It consists of sampling random variates from Markov chain, such that its stationary distribution is the posterior distribution of the parameter of interest. For our purpose, STUR model is represented by equations (2.3)-(2.4). To apply this approach we need all conditional posterior distributions, given appropriate prior distribution. By assuming prior independence and standard distributions (Normal and Inverse Gamma) for all unknown parameters, the joint prior distribution is given by:

$$p(\theta) \propto f_N\left(\alpha \mid \mu_\alpha, \sigma_\alpha^2\right) f_N\left(\phi_1 \mid \mu_{\phi_1}, \sigma_{\phi_1}^2\right) f_{Inv-Gam}\left(\sigma^2 \mid a_1, b_1\right) f_{Inv-Gam}\left(\omega^2 \mid a_2, b_2\right). \tag{A.1}$$

Having defined joint prior distribution, all conditional posterior distributions have Inverse Chi-square or Normal distribution. Due to standard form of all conditionals it is very easy to sample from posterior distribution, because we can draw directly from Inverse Chi-square and Normal distribution. Applying Bayes theorem we can derive following conditional posterior distributions:

$$p\left(\omega^2 \mid \alpha, \phi_1, \beta, \beta_0\right) = f_{Inv-\chi^2}\left(\omega^2 \mid T + 2a_2, \frac{\sum_{t=1}^{T}\left(\beta_t - \alpha - \phi_1\left(\beta_{t-1} - \alpha\right)\right)^2 + 2b_2}{T + 2a_2}\right), \tag{A.2}$$

$$p\left(\sigma^2 \mid \Delta y, y_{-1}, \beta, \beta_0\right) = f_{Inv-\chi^2}\left(\omega^2 \mid T + 2a_1, \frac{\sum_{t=1}^{T}\left(\Delta y_t - \beta_t y_{t-1}\right)^2 + 2b_1}{T + 2a_1}\right), \tag{A.3}$$

$$p\left(\alpha \mid \phi_1, \omega^2, \beta, \beta_0\right) = f_N\left(\alpha \mid \frac{(1-\phi_1)\sigma_\alpha^2 \sum_{t=1}^{T}\left(\beta_t - \phi_1\beta_{t-1}\right) + \mu_\alpha\omega^2}{T(1-\phi_1)^2\sigma_\alpha^2 + \omega^2}, \left(\frac{\sigma_\alpha\omega}{\sqrt{T(1-\phi_1)^2\sigma_\alpha^2 + \omega^2}}\right)^2\right), \tag{A.4}$$

$$p\left(\phi_1 \mid \alpha, \omega^2, \beta, \beta_0\right) \propto f_N\left(\phi_1 \mid \frac{\mu_{\phi_1}\omega^2 + \sigma_{\phi_1}^2 \sum_{t=1}^{T}\left(\beta_{t-1} - \alpha\right)\left[\beta_t - \alpha\right]}{\sigma_{\phi_1}^2 \sum_{t=1}^{T}\left(\beta_{t-1} - \alpha\right)^2 + \omega^2}, \left(\frac{\sigma_{\phi_1}\omega}{\sqrt{\sigma_{\phi_1}^2 \sum_{t=1}^{T}\left(\beta_{t-1} - \alpha\right)^2 + \omega^2}}\right)^2\right), \tag{A.5}$$

where $f_{Inv-\chi^2}\left(x \mid \nu, s^2\right)$ means scaled Inverse Chi-square distributions with $\nu > 0$ degrees of freedom and scale $s > 0$. Due to stationarity of random process β_t, conditional posterior distribution of autoregression coefficient is truncated to stationary region. The full conditional density for β_t at time t is Normal and can be written as:

$$p(\beta_t \mid \theta, \Delta y_t, y_{t-1}, \beta_{t-1}, \beta_{t+1}) = f_N\left(\beta_t \mid \frac{\sigma^2\left[\alpha(1-\phi_1)^2 + \phi_1(\beta_{t-1}+\beta_{t+1})\right] + \omega^2 \Delta y_t y_{t-1}}{\sigma^2\left(1+\phi_1^2\right) + \omega^2 y_{t-1}^2}, \left(\frac{\sigma\omega}{\sqrt{\sigma^2\left(1+\phi_1^2\right) + \omega^2 y_{t-1}^2}}\right)^2\right) \tag{A.6}$$

for $t = 1, \dots, T-1$
and for the last observation $t = T$

$$p(\beta_t \mid \theta, \Delta y_t, y_{t-1}, \beta_{t-1}) = f_N\left(\beta_t \mid \frac{\sigma^2[\alpha + \phi_1(\beta_{t-1} - \alpha)] + \omega^2 \Delta y_t y_{t-1}}{\sigma^2 + \omega^2 y_{t-1}^2}, \left(\frac{\sigma\omega}{\sqrt{\sigma^2 + \omega^2 y_{t-1}^2}}\right)^2\right). \tag{A.7}$$

These conditionals are similar to conditional posterior distributions derived by Jostova and Philipov (2005) for simple regression linear model with random parameter. Their model has been used to describe the evolution of stochastic betas for US industry portfolios.