Working Paper Series

# Okun's Law: Does the Austrian Unemployment – GDP relationship exhibit Structural Breaks?

Leopold Sögner

Working Paper No. 61
January 2000

Working Paper Series

January 2000

SFB
'Adaptive Information Systems and Modelling in Economics and Management Science'

Vienna University of Economics
and Business Administration
Augasse 2–6, 1090 Wien, Austria

in cooperation with
University of Vienna
Vienna University of Technology

http://www.wu-wien.ac.at/am

This piece of research was supported by the Austrian Science Foundation (FWF) under grant SFB#010 ('Adaptive Information Systems and Modelling in Economics and Management Science').

# Okun's Law: Does the Austrian Unemployment – GDP relationship exhibit Structural Breaks?

**Leopold Sögner**
E-mail: Leopold.Soegner@wu-wien.ac.at
Department of Economics
Vienna University of Economics
and Business Administration
Augasse 2 – 6, A-1090 Vienna, AUSTRIA

December 21, 1999

**Abstract**

*Okun's Law* postulates an inverse relationship between movements of the unemployment rate and the real gross domestic product (GDP). Empirical estimates for US data indicate that a two to three percent GDP growth rate above the natural or average GDP growth rate causes unemployment to decrease by one percentage point and vice versa. In this investigation we check whether this postulated relationship exhibits structural breaks by means of Markov-Chain Monte Carlo methods. We estimate a regression model, where the parameters are allowed to switch between different states and the switching process is Markov. As a by-product we derive an estimate of the current state within the periods considered. Using quarterly Austrian data on unemployment and real GDP from 1977 to 1995 we infer only one state, i.e. there are no structural breaks. The estimated parameters demand for an excess GDP growth rate of 4.16% to decrease unemployment by 1 percentage point. Since only one state is inferred, we conclude that the Austrian economy exhibits a stable relationship between unemployment and GDP growth.

JEL-Classification: D83, D84, G10.

Keywords: Okun's Law, Permutation Sampling, Switching Models.

# 1 Introduction

*Okun's Law* postulates a relationship between movements of the unemployment rate and the real gross domestic product (GDP). Empirical estimates indicate that a two to three percent GDP growth above the natural or average GDP growth causes unemployment to decrease by one percentage point and vice versa (see Blanchard and Fischer (1989) or Romer (1996)). This article investigates whether the Austrian unemployment–GDP relationship exhibits structural changes.

If we are confronted with structural breaks at unknown periods of time, we possibly have to deal with sudden switches between a finite number of states or continuous changes of the parameters at each period. In the latter case Kalman-filter techniques could be used. In this article we restrict our analysis to finite states of the world, where the number of states and the switching periods are unknown. To include regime switches we augment the set of prediction variables by an unobservable variable following a discrete time Markov process. To derive the parameters in a Bayesian model, *Markov-Chain Monte Carlo* methods (MCMC) provide a powerful tool. These sampling techniques take advantage of the hierarchical structure of Bayesian models. After the a-priori and the conditional distribution of the parameters are specified, the parameters are estimated from samples generated from the a-posteriori distribution as described e.g. in Casella and George (1992), Chib and Greenberg (1996) and Robert (1994). In econometrics switching models had already been used by Hamilton (1989), Albert and Chib (1993), Kim (1994), Kaufmann (1997) and Kaufmann (1998).

Since the indices of the switching variable can be permuted without changing the (marginal) likelihood, the general switching model is only identifiable up to permutations in the indices, and therefore not identifiable in a strict sense. A detailed discussion and examples on the problem of identifiability are provided in Frühwirth-Schnatter (1998) and Stephens (1999). By putting constraints on the state specific parameters the model becomes identifiable. Using *permutation sampling*, as suggest by Frühwirth-Schnatter (1998), the parameters are first sampled on the unconstrained set. Secondly, the parameters are resorted as required by the predetermined restriction on the parameter space. The number of states will be estimated by means of the *model likelihood*. Specifically, we use the so called candidate's formula to derive the model likelihood (see Chib (1995) and Frühwirth-Schnatter (1999)).

In this article we regress the annual unemployment growth on the annual GDP growth for quarterly Austrian data from 1976 to 1995. By calculating the model likelihood we infer a model with one state, i.e. Okun's Law is structurally stable. Furthermore a GDP growth of 10.94 billion Austrian Schillings per year is required to derive zero unemployment growth. Okun's law is often expressed as the extra growth rate (above the normal growth rate) required to decrease unemployment by 1% point. This extra growth rate is equal to 4.1577%. This article is organized as follows: Section 2 gives a brief introduction to the estimation method used to derive the model parameters. Section 3 describes the statistical model, and section 4 provides the results of our estimation.

## 2 Markov-Chain Monte Carlo Estimation

This section provides a brief description of MCMC methods. The reader who is already familiar with these tools or only interested in the results could skip to the next section. Further information on MCMC methods is provided in Greene (1997), Casella and George (1992), Albert and Chib (1993), and Robert (1994). For switching models the reader is referred to Hamilton (1989), Chib and Greenberg (1996), or Frühwirth-Schnatter (1998).
First of all, let us consider a Bayesian statistical model with the conditional density $f(x|\theta)$, with the data $x$, the unknown vector of parameters $\theta$, and the prior distribution of parameters, represented by the density $\pi(\theta)$. Given the prior distribution of the parameters and the conditional distribution of the data, the a-posteriori density of $\theta$ can be derived by applying Bayes theorem (see Robert (1994)):

$$\pi(\theta|x) = \frac{f(x|\theta)\pi(\theta)}{\int f(x|\theta)\pi(\theta)d\theta} \ ,$$

where the numerator will be called non-normalized a-postiori density. To estimate the parameters by means of MCMC methods (also called *Bayesian Sampling* or *Gibbs-Sampling*) it is sufficient to work with the non-normalized distribution $\pi(\theta|x) \propto f(x|\theta)\pi(\theta)$, where the symbol $\propto$ stands for "proportional to".
Next, let us include a latent switching variable $I_t$ following a homogenous Markov process in discrete time (see Karlin and Taylor (1975)). The transition probabilities $\eta_{11}, \ldots, \eta_{kk}$ are summarized in the matrix of transition probabilities $\eta$. Each row $\eta_i$ of this matrix provides us with the probabilities $\eta_{il}$ that the process switches from $[I_{t-1} = i]$ to state $[I_t = l]$, $l = 1, \ldots, k$.. Each $\eta_i$ takes values on a k-dimensional simplex $\mathcal{E}$. Since we only observe data $x^N := (x_t)_{t=1}^N$ the corresponding sequence of switching variables is defined by $I^N := (I_t)_{t=1}^N$. Next we augment the parameter space. This results in the *augmented vector of parameters* $\Psi = (\theta, I^N)$, where $\theta$ consists of common parameters $\theta_C$, the state specific model parameters $\theta_I$ and the matrix of the Markov transition probabilities $\eta$.
Due to the hierarchical structure of most Bayesian models the vector of parameters can be derived from successive sampling from the conditional distributions of the parameters. By the convergence properties of ergodic Markov-chains (geometric) convergence to the invariant distribution of $\theta$ is guaranteed (for the regularity conditions see Robert (1994) and Chib and Greenberg (1996)).
At the beginning of our sampling procedures, we have no real prior information to discriminate between the different states. So a permutation of the labels results in the same data likelihood $f(x^N|\Psi)$ for at least two different vectors of parameters. Therefore the unrestricted model is not identifiable. Imposing a restriction $\mathcal{R}$ on $\theta$, i.e. $\theta_1 < ... < \theta_i < ... < \theta_k$, makes the problem identifiable. In this article we follow the method of Frühwirth-Schnatter (1998) called *permutation sampling*, which is very efficient from a computation point of view. The sampling algorithm for sampling periods $j = 1, ..., T$ works as follows, where the $j$-th sample of a parameter will be denoted by the superscript $[j]$ within this article:

---

Step 0: Define a Restriction $\mathcal{R}$: $\theta_1 < ... < \theta_i < ... < \theta_k$.
Step 1: Sampling on the unrestricted set

$$\begin{array}{lll} & \vdots & \\ I^{N,[j]} & \text{from} & \pi(i^N|x^N, \theta^{[j-1]}) \\ \eta^{[j]} & \text{from} & \pi(\eta|x^N, I^{N,[j]}, \theta_C^{[j-1]}, \theta_I^{[j-1]}) \\ \theta_C^{[j]} & \text{from} & \pi(\theta_C|I^{N,[j]}, x^N, \eta^{[j]}, \theta_I^{[j-1]}) \\ \theta_I^{[j]} & \text{from} & \pi(\theta_I|I^{N,[j]}, x^N, \eta^{[j]}, \theta_C^{[j]}) \\ & \vdots & \end{array}$$

Step 2: Sort $\Psi^{[j]}$ according to the restriction on $\theta$.
Step 3: Goto Step 1.

---

The above procedure is repeated until the Markov-chain has reached or is supposed to be near its invariant distribution, resulting in the question when the sampling algorithm should be stopped. Liu *et al.* (1995) and Chib and Greenberg (1996) derived the result that the speed of convergence depends on the correlation of the components. Stopping rules are provided in the Robert (1994) and the articles mentioned above. In this article convergence will be checked as follows: We run the sampler for $T = 3000$ time steps and cut off the first 2000 samples of $\Psi^{[j]}$. Then we take two sub-samples for every particular parameter of the remaining samples of $\Psi^{[j]}$ and compare the distributions of the sub-samples.
*Model Selection:* The number of states $k$ will be estimated by taking the model with the highest model likelihood $L_k(x^N)$. In our analysis $L_k(x^N)$ is derived by means of the so called *candidate's formula* as described in Chib (1995). For a discussion of $L_k(x^N)$ on a restricted parameter space the reader is referred to Frühwirth-Schnatter (1999). Applying this concept to our problem an estimate of the natural logarithm of the model likelihood is derived from: $\log \hat{L}_k(x^N) = \log \hat{f}(x^N|\theta^*) + \log \hat{\pi}(\theta^*) - \log \hat{\pi}(\theta^*|x^N)$, where $\hat{f}(x^N|\theta^*)$ is the estimated marginal likelihood after $I^N$ has been integrated out, $\hat{\pi}(\theta^*)$ is the estimated a-priori density at parameter values $\theta^*$ selected from MCMC output, and $\hat{\pi}(\theta^*|x^N)$ is the estimated posterior density from MCMC output. For $\theta^*$ the mean values of MCMC output will be used in this analysis.

# 3 The Regression Model with Switching

Okun's law states that the annual growth in unemployment $\Delta u_t$ is a function of the growth in real gross domestic product $\Delta GDP_t$. For quarterly data this results in $\Delta u_t := u_t - u_{t-4}$, $\Delta GDP_t := GDP_t - GDP_{t-4}$, and $x^N = (\Delta u_t \ , \ \Delta GDP_t)$. In this article we consider the following model:

$$\begin{aligned} \Delta u_t &= \beta_{0,i} + \beta_{1,i}\Delta GDP_t + \beta_{2,i}\Delta u_{t-1} + \varepsilon_{t,i} \ , \\ &= \beta_i z_t + \varepsilon_{t,i} \end{aligned} \quad (1)$$

where the index $i \in \{1, 2, ..., k\}$ is the label of the state $I_t$. $\beta_{0,i}$ is the intercept if state

| Parameter | a-priori | a-posteriori |
|---|---|---|
| $\eta_i^{[j]}$ | | Dirichlet |
| | $\mathcal{D}(e_{0,i1},...,e_{0,ik})$ | $\mathcal{D}(e_{0,i1}+N_{i1}^{[j]},...,e_{0,ik}+N_{ik}^{[j]})$ |
| $(\sigma_i^2)^{[j]}$ | | Invers Gamma |
| | $\mathcal{IG}(\nu_0, D_0)$ | $\mathcal{IG}(\nu_i^{[j]}, D_i^{[j]})$ |
| $\beta_i^{[j]}$ | | Normal |
| | $\mathcal{N}(b_0, B_0\sigma_i^2)$ | $\mathcal{N}(\kappa_i^{[j]}((Z_i')^{[j]}y^N + B_0^{-1}b_0), \kappa_i^{[j]}(\sigma_i^2)^{[j]})$ |

Table 1: Conjugate Priors for the Switching Model

$[I_t = i]$ is realized. $\beta_{1,i}$ shows the dependence of $\Delta u_t$ on GDP growth $\Delta GDP_t$. $\varepsilon_{t,i}$ is an independent identically distributed (iid.) normal variable with zero mean and variance $\sigma_i^2$, i.e. $\varepsilon_{t,i} \sim iid.\ \mathcal{N}(0,\sigma_i^2)$. Due to autocorrelation in the residuals $\varepsilon_{t,i}$ – in a model without $\beta_{2,i}$ – the lagged variable $\Delta u_{t-1}$ has been added to eliminate autocorrelation in the residuals. $\beta_i$ is the vector of regression parameters, i.e. $\beta_i := (\beta_{0,i},\ \beta_{1,i},\ \beta_{2,i})$. The prediction variables are $z_t := (1\ ,\ \Delta GDP_t\ ,\ \Delta u_{t-1})$, the response variable is $y_t = \Delta u_t$. Within this analysis we also estimate (i) a model where $\sigma_i^2$ is a common parameter ($\sigma_i^2 = \sigma^2$, for $i = 1,\ldots,k$.), (ii) a model where $\beta_{2,i}$ is common ($\beta_{2,i} = \beta_2$, for $i = 1,\ldots,k$.) and (iii) a model where both parameters $\beta_2$ and $\sigma^2$ do not change with the state $I_t$.

**Remark 1** *Hodrick-Presott filtering (HP-filter) techniques could be applied to the GDP time series to eliminate fluctuations at low frequencies, and to emphasize those fluctuations in the range of three to five years (see Cooley (1995)). Since regime switching need not be caused by the detrended business cycles we simply use first differences in (1).*

After the description of the model, we would like to embed this model to the Bayesian sampling scheme, specify the conditional distributions of the parameters, and define a restriction on the parameter space such that the model is identifiable as described in section 2. The parameter vector at the $j$-th sampling step is given by $\theta^{[j]} = (\beta^{[j]}, (\sigma^2)^{[j]}, \eta^{[j]})$, where $\beta^{[j]} := (\beta_1^{[j]},\ldots,\beta_i^{[j]},\ldots,\beta_k^{[j]})$ and $(\sigma^2)^{[j]} := ((\sigma_1^2)^{[j]},\ldots,(\sigma_i^2)^{[j]},\ldots,(\sigma_k^2)^{[j]})$. The augmented parameter vector $\Psi^{[j]}$ is equal to $\theta^{[j]} = (\beta^{[j]}, (\sigma^2)^{[j]}, \eta^{[j]})$ and the sequence of state variables $I^{N,[j]}$. Secondly, *conjugate priors* are used for the sampling algorithm. These distributions share the property that if the a-priori distribution is in the class $\mathcal{C}$, the a-posteriori distribution lies in $\mathcal{C}$. A detailed discussion on the advantages and disadvantages of conjugate prior modeling is provided in Robert (1994). The prior distributions are assumed to fulfill: (i) Independence of the state specific parameters and the switching probabilities $\pi(\theta) = \pi(\beta,\sigma^2)\pi(\eta)$, (ii) $I^N$ is Markov with prior $\pi(i^N|\theta) = \pi(i^N|\eta) = \prod_{i=1}^k \prod_{l=1}^k \eta_{il}^{N_{il}} \pi(i_0)$. $\pi(i_0)$ is the starting distribution where we assume that $\mathbb{P}(I_0 = i) = 1/k$. (iii) $\pi(\beta,\sigma^2)$ and $\pi(\eta)$ are invariant to permutations in the indices. Following Robert (1994) and Frühwirth-Schnatter (1998) the distributions required for the regression model (1) are described in Table 1.
The parameters are defined as follows: $N_{il}^{[j]} := \#(I_t = l | I_{t-1} = i)$ is the number of jumps from $[I_{t-1} = i]$ to $[I_t = l]$ in $I^{N,[j]}$. The assumption that the states are

multinomial $\mathcal{M}_k$ of degree $k$ results in a Dirichlet distribution of the vectors of transition probabilities $\eta_1, \ldots, \eta_i, ..., \eta_k$. Concerning the variance terms $\sigma_i^2$, $i = 1, ..., k.$, the parameters of the inverse gamma distribution are given by $\nu_i^{[j]} := \nu_0 + 0.5 N_i^{[j]}$ and $D_i^{[j]} := D_0 + 0.5 \sum_{t=1}^N S_{t,i}^{N,[j]} (y_t - \beta_i^{[j-1]} z_t)^2$, where $N_i^{[j]} := \#(I_t = i)$ is the frequency the chain has hit state $i$. $S^{N,[j]}$ is a $(N \times k)$ matrix. The $w$-th element, $w = 1, \ldots, k.$, of the $t$-th row of $S_N^{[j]}$, $t = 1, \ldots, N.$, is equal to 1 if $[I_t = w]$ and zero if $[I_t \neq w]$. $S_i^{N,[j]}$ is the $i$-th column of $S^{N,[j]}$ sampled at step $j$, while $S_t^{N,[j]}$ is the $t$-th row of this matrix. $S_{t,i}^{N,[j]}$ is the $i$-th element of $S_t^{N,[j]}$. For the model with common $\sigma^2$ and the model with common $\sigma^2$ and $\beta_2$, $N_i^{[j]}$ is equal to the number of observations (i.e. $N_i^{[j]} = N$) while $D_i^{[j]} = D^{[j]} = D_0 + 0.5 \sum_{i=1}^k \left( \sum_{t=1}^N S_{t,i}^{N,[j]} (y_t - \beta_i^{[j-1]} z_t)^2 \right)$, since all residuals $y_t - \beta_i^{[j-1]} z_t$, $t = 1, \ldots, N.$, are used to calculate $D^{[j]}$. Considering the distribution of $\beta_i^{[j]} = (\beta_{0,i}^{[j]}, \beta_{1,i}^{[j]}, \beta_{2,i}^{[j]})$, we define $Z_i^{[j]} = S_i^{N,[j]} z^N$, where the $t$-th row of $Z_i^{[j]}$ is equal to $z_t$ if $[I_t = i]$, and it is equal to a vector of zeros in all states $[I_t \neq i]$. I.e. by means of $S_i^{N,[j]}$ we project on the data where the states $[I_t = i]$ occur in $I^{N,[j]}$. Last but not least, $\kappa_i^{[j]} =: ((Z_i')^{[j]} Z_i^{[j]} + B_0^{-1})^{-1}$. For the models with common $\beta_2$, this parameter is sampled from a normal distribution $\mathcal{N}(\kappa_2((\tilde{x}^N)' \tilde{y}^{N,[j]} + B_{2,0} b_{2,0}), \kappa_2 (\sigma_i^2)^{[j]})$, where $\tilde{y}_t^{[j]} := y_t - \sum_{i=1}^k S_{t,i}^{N,[j]} \left( \beta_{0,i}^{[j-1]} - \beta_{1,i}^{[j-1]} \Delta GDP_t \right)$ for $t = 1, \ldots, N.$, $\tilde{y}^{N,[j]} = (\tilde{y}_t^{[j]})_{t=1}^N$, $\tilde{x}_t$ is the lagged variable $\Delta u_{t-1}$, $\tilde{x}^N = (\tilde{x}_t)_{t=1}^N$, $b_{2,0}$ and $B_{2,0}$ define the prior of $\beta_2$ and $\kappa_2^{[j]} = \kappa_2 = ((\tilde{x}^N)' \tilde{x}^N - B_{2,0}^{-1})^{-1}$. The prior of $\beta_2$ is normal as well. The state dependent parameters $\beta_{0,i}$ and $\beta_{1,i}$ are derived in the same was as we did for vector $\beta_i$. The restriction $\mathcal{R}$ on the parameters is given by size of the regression term $\beta_{1,i}^{[j]}$, i.e. $\beta_{1,1}^{[j]} < \beta_{1,2}^{[j]} < ... < \beta_{1,k}^{[j]}$.

# 4 Results

This section presents the MCMC estimates. The data used are 80 observations of quarterly Austrian real GDP and unemployment from 1976 to 1995 resulting in $N = 75$. First differences $\Delta u_t$ and $\Delta GDP_t$ are required to derive stationary time series. $\Delta u_t$ is measured in percents, while $\Delta GDP_t$ is measured in billions of Austrian Schillings.

*Prior Distribution:* The prior distribution of the parameters is the following: $e_{0,il} = 4$ for $i = l$ and $e_{0,il} = 3$ for all $i \neq l$. Since a $\mathcal{D}(1, ..., 1)$ prior does not result in convergence of the transition probabilities $\eta_{il}$ and the variances $\sigma_i^2$, higher $e_{0,ii}$ are necessary. This prior assumption is nevertheless relatively vague compared to some other investigations, for example McCulloch and Tsay (1994) use a $\mathcal{D}(45, 2)$ prior in a model with two states. It is worth noting that altering this prior assumption does not result in major differences in the estimates of $\beta$ and $\sigma^2$. For the parameters of the state specific priors we use the ordinary least squares estimator and the corresponding variance of the residuals $s_y^2$ in the following way: $b_0 = (0.2682,\ -0.0246,\ 0.7110)$, $B_0 = diag(1, 1, 1)$, $\nu_0 = 2$, $D_0 = s_y^2$, where $s_y^2 = 0.0611$ and $diag(.)$ stands for diagonal matrix. For the common parameter models $B_0 = diag(1, 1)$, $B_{2,0} = 1$, $b_0 = (0.2682,\ -0.0246)$ and $b_{2,0} = 0.7110$. Thus, we use relatively vague priors on the switching probabilities $\eta$ and informative priors on the state specific parameters $\beta$ and $\sigma^2$. Since we claim an inversely related interdependence in the data, we use this additional information in our sampler.

| $k$ | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| $\log \hat{L}_k(x^N)$ | -114.145 | -118.408 | -121.373 | -121.910 | -122.446 |
| $SD(\log \hat{L}_k(x^N))$ | (0.017) | (0.066) | (0.251) | (1.068) | (3.092) |

Table 2: Estimated Model log-Likelihoods $\hat{L}_k(y^N)$.

| Common parameter | $\sigma^2$ | $\beta_2$ | $\sigma^2$, $\beta_2$ |
|---|---|---|---|
| $\log \hat{L}_2(x^N)$ | -118.043 | -119.299 | -116.888 |
| $SD(\log \hat{L}_2(x^N))$ | (0.067) | (0.058) | (0.057) |

Table 3: $\hat{L}_2(y^N)$ with Common Parameters.

**Remark 2** *In contrast to this informative prior assumption on the state specific parameters vague priors can be used. Nevertheless, the question arises why we do not use already existing information on the parameters, since Okun's Law is a stylized fact often discussed in economic literature. E.g. using the OLS estimator for the parameter vector $b_0$ reflects this argument. For the switching variable and the corresponding probabilities no prior information is available. This is the reason why we use a relatively vague prior on $\eta$.*

*Model Selection:* Considering the estimates of the model likelihood we derive that $k = 1$ is the most plausible estimate of $k$. The logarithms of the model likelihoods and its standard deviations ($SD(.)$) for the models where all parameters are state dependent are presented in Table 2. From Table 2 we infer a model with only one state. Since the estimates of $\beta_{2,i}$ and $\sigma_i^2$ are very similar in a model with two states we additionally check the model likelihood for a two-state model with common $\sigma^2$, a two-state model with common $\beta_2$ and a two-state model with common $\sigma^2$ and $\beta_2$. The estimates of the model likelihood are presented in Table 3. From these calculations we conclude that $k = 1$, i.e. we only infer one state without any structural breaks or outliers.

**Remark 3** *Considering the estimated model likelihoods results in a model with one state. We additionally checked the impacts on $k$ when altering the priors. Using very diffuse priors, e.g. $\eta_i \sim \mathcal{D}(1, ..., 1)$ and/or $b_0 = (0, 0, 0)$ results in $k = 1$, where a diffuse prior on $\eta$ resulted in bad convergence properties of the transition probabilities $\eta$ and the variances $\sigma^2$. Also very strong priors on $\beta$ do not change the inference on $k$. E.g. $B_0 = diag(0.01, 0.01, 0.01)$ changes the model likelihood while the order of the model likelihoods do not change. Also very strong priors on $\eta$, e.g. $\mathcal{D}(45, 2, ..., 2)$, does not change the result.*

*Parameter Estimates:* Table 4 presents the parameter estimates from MCMC output and the standard deviations of these parameters. Considering the growth in GDP required to result in a $\Delta u_t$ of zero the estimates require $\Delta GDP_t = 10.4933$ billion Austrian Schillings. Next we investigate the unemployment–GDP relationship as stated in textbooks, where a 2%–3% increase in GDP above the *normal growth* is supposed to cause unemployment

| $\beta_0$ | $\beta_1$ | $\beta_2$ | $\sigma^2$ |
|---|---|---|---|
| 0.2703 | -0.0247 | 0.7118 | 0.0612 |
| (0.0350) | (0.0036) | (0.0357) | (0.0093) |

Table 4: Estimated Parameters

to decline by 1% point (for US data, see Romer (1996)). We express the normal growth by the mean growth rate ($1/N \ \sum_{t=1}^{N}(\Delta GDP_t/GDP_{t-4}) \ \cdot \ [100\%]$), which is equal to 2.3276% for the underlying data. Since the autoregressive variable $\Delta u_{t-1}$ has been included in the regression model (1) we derive the annual excess growth rate necessary to decrease the unemployment rate by 1% point per year, i.e. if the *annual GDP growth rate* ($\Delta GDP_t/GDP_{t-4} \ \cdot \ [100\%]$) stays at this level unemployment declines annually by 1% point. The estimates demand for excess growth rates of 4.4528%, where the mean GDP ($1/N \ \sum_{t=1-4}^{N} GDP_t$) was inserted for $GDP_{t-4}$ to calculate the extra growth rate.
Additionally, we want to investigate the necessary excess growth rate to decrease unemployment by 1% point in one particular period. Considering the regression model this question depends on the unemployment rate of the last period. Nevertheless, $\Delta GDP_t$ is easily derived if we insert the mean unemployment growth ($1/N \ \sum_{t=1}^{N} \Delta u_t = 0.2737$) into (1). This results in an excess growth rate of 4.1577%.

# 5 Conclusions

In this article we checked the dependence of unemployment growth on annual real GDP growth for quarterly Austrian data. Using Bayesian methods we infer one state, i.e. there are neither structural breaks nor outliers. The estimates require a GDP growth of 10.94 billion Austrian Schillings per year to derive zero unemployment growth. Since Okun's law is often expressed as the extra growth rate (above normal growth) required to decrease unemployment by 1% point, we expressed normal growth by the mean growth rate. The estimates demand for an extra growth rate of 4.16%.
Last but not least we want to discuss the impacts of and the conclusions from these estimates on economic policy. Since we have not detected any structural breaks in Austria's Okun's law, the unemployment–GDP relationship is stable for Austrian data. This implies that economic policy did not succeed in decoupling unemployment growth from GDP growth. Moreover, no persistent effects from migration on Okun's Law can be observed. Therefore, neither changes in the political system, governmental employment programs nor migration have caused a systematic change in Austria's Okun's Law.

# References

James H. Albert, Siddhartha Chib (1993). Bayes inference via gibbs sampling of autoregressive time series subject to markov mean and variance shifts. *Journal of Business and Economic Statistics*, 11(1).

Olivier J. Blanchard, Stanley Fischer (1989). *Lectures on Macroeconomics.* MIT-Press, Cambridge.

George Casella, Edward I. George (1992). Explaining the gibbs sampler. *The American Statistician*, 46(3):167–174.

Siddhartha Chib, Edward Greenberg (1996). Markov chain monte carlo simulation methods in econometrics. *Econometric Theory*, 12:409–431.

Siddhartha Chib (1995). Marginal likelihood from the gibbs output. *Journal of the American Statistical Association*, 90(432):1313–1321.

Thomas F. Cooley (1995). *Frontiers in Business Cycle Research.* Princeton University Press, Princeton, New Jersey.

Sylvia Frühwirth-Schnatter (1998). Mcmc estimation of classical and dynamic switching and mixture models. *mimeo, Department of Statistics, Vienna University of Economics and Business Administration.*

Sylvia Frühwirth-Schnatter (1999). Model likelihoods for switching and mixture models. *mimeo, Department of Statistics, Vienna University of Economics and Business Administration.*

William H. Greene (1997). *Econometric Analysis.* Perentice Hall, New Jersey, 3 edition.

James D. Hamilton (1989). A new approach to the economic analysis of non-stationary time series and the business cycle. *Econometrica*, 57(2).

Samuel Karlin, Howard M. Taylor (1975). *A First Course in Stochastic Processes.* Harcourt Brace and Company, Boston, 2 edition.

Silvia Kaufmann (1997). Measuring business cycles with a dynamic markov switching factor model. *Working Paper No. 9710, Department of Economics, University of Vienna.*

Silvia Kaufmann (1998). Bayes inference in common markov switching trends models. *Working Paper No. 9809, Department of Economics, University of Vienna.*

Chang-Jin Kim (1994). Dynamic linear models with markov-switching. *Journal of Econometrics*, 60(1–22).

J.S. Liu, W. W. Wong, A. Kong (1995). Covariance structure of the gibbs sampler with applications to the comparison of estimators and augmentation schemes. *Biometrika*, 1(27–40).

R.E. McCulloch, R.S. Tsay (1994). Statistical analysis of economic time series via markov switching models. *Journal of Time Series Analysis*, 15:523–539.

Christian P. Robert (1994). *The Bayesian Choice*. Springer, New York.

David Romer (1996). *Advanced Macroeconomics*. McGraw-Hill, New York.

Matthew Stephens (1999). Dealing with mulitimodal posteriors and non-identifiability in mixture models. *mimeo, Department of Statistics, Oxford, U.K.*