D S E

Dipartimento   **Scienze**   **Economiche**

Working Paper

Department of Economics

Ca' Foscari University of Venice

Silvestro Di Sanzo

Output fluctuations persistence: Do cyclical shocks matter?

Output fluctuations persistence: Do cyclical shocks matter?

Silvestro Di Sanzo

Universidad de Alicante

Abstract

The aim of this paper is to identify the different sources of persistence of output fluctuations. We propose an unobserved components model that allows us to decompose GDP series into a trend component and a cyclical component. We let the drift of the trend component to switch between different regimes according to a first-order Markov process. To calculate an appropriate p-value for a test of linearity we propose a bootstrap procedure, which allows for general forms of heteroskedasticity. The performance of the bootstrap is checked by means of a Monte Carlo simulation. Our study concerns the U.S. As suggested by the Endogenous Growth theory, cyclical shocks appear to play an important role on the observed persistence of output. We argue that the traditional explanation of persistence, which is related to Real Business Cycle models with exogenous productivity, is not consistent with our data. We also find that the majority of business cycle fluctuations in the U.S. are due to real shocks.

Keywords
Business cycle; Persistence; Unobserved Components model; First-order Markov process; Wild bootstrap; Monte Carlo.

Address for correspondence:
Silvestro Di Sanzo
Departamento de Fundamentos del Análisis Económico
Edificio Ciencias Sociales
Universidad de Alicante
Ctra. San Vicente del Raspeig, s/n
99-E-03080 Alicante Spain
Office: P2072CCSS
Phone: 965903400 (ext 3349)
Box: 40CCSS
E-mail: silvio@merlin.fae.ua.es

1 Introduction

An important stylized fact of business cycle is the persistence of output fluctuations. With persistence of output fluctuations we refer to a shock to output that is persistent and whose effect does not dissipate in the near future. That is, output does not show a significant tendency to return to its trend level after a shock. In an influential paper, Nelson and Plosser (1982) show that output is better characterized as a nonstationary process with no tendency to revert to a deterministic path. Thus, output has a unit root. These results have been widely confirmed by other authors. For example, Campbell and Mankiw (1987) find that quarterly GNP for G-7 countries is highly persistent. They show that an unexpected change in real GNP of 1 percent should change one's forecast by over a long horizon by over 1 percent. Further, according with these results, Cogley (1990) finds high persistence in output fluctuations for a similar sample of countries. Although much disagreement remains among researchers using postwar quarterly data over which are the shocks that cause business fluctuations, there is a broad agreement on the fact that output fluctuations are highly persistent.

The presence of a stochastic trend in output series is consistent with several economic models. It is possible to divide them into two categories: Real Business Cycle (RBC) and Endogenous Growth (EG) models. In RBC model (see Kydland and Prescott (1982), King *et al.* (1991) or Rebelo (2005) for a complete survey), real shocks are the central driver of business cycle and the only source of persistence in output fluctuations. Conventional RBC theory assumes that GNP per capita follows a random walk with drift, where the drift is exogenously determined by the rate of labor-augmenting technological progress. Only small deviations around a steady state solution are analyzed and transitional growth dynamics are ignored. In this set-up, exogenous permanent shifts in the production function are the sole responsible of the existence of a stochastic trend.

On the other hand, other authors related to EG literature (King *et al.* (1988), Stadler (1990) and Fatás (2000)) notice that if we relax the hypothesis that technological progress is exogenous, the notion of persistent fluctuation may have a very different interpretation. More precisely, if we consider a model where the technological progress depends on economic factors, there are many types of shocks different from a permanent shift in the production function, that can generate persistent fluctuations. In particular, any temporary shock can stimulate learning and cause a rise in the accumulated technical knowledge of the economy shifting the trend of GDP upward. As a consequence, demand shocks, traditionally considered as temporary, can have a permanent impact on the trend of the output.

The two sets of models described above imply the presence of a stochastic trend, so both are able to produce persistent output fluctuations. But while the source of the stochastic trend is due to permanent shifts in the production function in RBC models, it is attributable to cyclical shocks in EG models. Therefore, their implications are different in terms of causes and welfare costs of output fluctuations.

The objective of this paper is to find a stylized fact against which these macroeconomic theories can be measured. We are conscious that it is imprudent to make definitive judgments regarding the above theories on the basis of a unique stylized fact, but we believe that it is interesting to find empirical evidence that allows us to determine if cyclical shocks are an important source of the persistence of output fluctuations.

With this in mind, we propose an unobserved components (UC)[1] model that allows us to decompose GDP series into a trend component and a cyclical component. The former is assumed to follow a random walk with drift, the latter is assumed to be a stationary second-order autoregressive process. We let the drift of the trend component to switch between regimes according to a first-order Markov process as in Hamilton (1989). It is thus possible to capture business cycle asymmetry. By asymmetry we mean that the dynamics of recessions are qualitatively distinct from those of expansions[2]. The *Endogenous Growth Hypothesis*, i.e. cyclical shocks have a significant effect on the trend component of the output, is introduced by allowing cyclical shocks to have a lagged effect on the trend of the output series. This parameterization is at the basis of the hysteresis model[3] of Jaeger and Parkinson (1994), which postulates that the cyclical modifies permanently the trend.

A key problem which arises in actual applications is how to determine the number of regimes of the model. Hamilton's original paper offers a suggestive evidence that the two-state Markov-switching model outperforms linear models in terms of forecasts, but not statistical tests. Thus, the dominant approaches in the empirical literature are to fix the number of regimes *a priori* or to determine them in an informal manner by visual inspection of data plots. Both approaches are evidently subjective, somewhat arbitrary and leave much to be desired.

---

[1]See Harvey (1989), Kim and Nelson (1999) and Clark (1987) for a detailed description of the Unobserved Component models.

[2]See Kim and Murray (2002) and Kim and Pigger (2002) for different strategies about how to incorporate business cycle asymmetries in time series models.

[3]The hysteresis model of Jaeger and Parkinson (1994) was proposed for unemployment to formalize the idea that a rise in cyclical unemployment can lead to a permanent increase in the natural rate.

A more formal statistical procedure for determining the number of regimes is to perform a test of linearity using likelihood ratio tests. Such tests are complicated because usual regularity conditions are not fulfilled due to the presence of nuisance parameters that are not identified under the null hypothesis (see Hansen (1996) and Andrews and Ploberger (1994)). This implies that the scores are identically zero . Hence, standard regularity conditions are not valid, and then conventional statistics do not have an asymptotic standard distribution. In our case, the null hypothesis is a one-regime model versus the alternative of two regimes. In this context, transition probabilities are not identified under the null. Hansen (1992a,b) proposed a theory of hypothesis testing to calculate the p-value of the test of interest under the presence of nuisance parameters in Markov-switching models, but his procedure is computationally burdensome and only delivers p-values which are an upper bound for the true p-values. A less computationally demanding test procedure was proposed by Garcia (1998), but his method is theoretically little attractive since it overlooks the problem of the singular information matrix. Moreover, both procedures ignore the power of the test. Recently, Carrasco *et al.* (2004) propose a test procedure that does not require estimating the model under the alternative. This is a great advantage over competing tests, as those of Garcia (1998) and Hansen (1992a,b), because estimating a Markov-switching model is particularly burdensome. The problem is that Carrasco's test only works with stationary data and in our paper we employ a series which displays a non-stationary behaviour.

To circumvent these difficulties, we propose an alternative methodology to test the number of regimes. We calculate exact p-values from a bootstrap approximation to the null distribution of the likelihood ratio test statistic. To our knowledge, there has so far been no work on the use of such methodology to select the number of regimes in the frame of UC models with the parameters of interest switching between two different regimes according to a first-order Markov process. Once we have determined the number of regimes, we are able to perform a simple test to assess the *Endogenous Growth Hypothesis*. It is clear that if we accept this hypothesis, the empirical evidence suggests that cyclical shocks play an important role in the observed persistence.

The rest of the paper is organized as follows. Section 2 specifies the model and states the hypothesis testing procedure. Section 3 proposes a bootstrap procedure to calculate the p-value of a linearity test within our framework of interest. An empirical application for the U.S. is given in section 4. Section 5 assesses the finite sample performance of the proposed bootstrap procedure by means of a Monte Carlo experiment. The conclusion is provided in the last section. Estimation methods are relegated to Appendix A. Appendix B

contains all the tables and figures.

2   The Model

We decompose the observed series into two additive unobservable components: a nonstationary trend, Y_t^N, and a stationary cyclical component, Y_t^C,

$$Y_t = Y_t^N + Y_t^C.$$

Following Hamilton (1989), the growth rate of the trend component is allowed to switch between different regimes. We consider the cycle modifies also the trend, as Jaeger and Parkinson (1994) postulate in their hysteresis model. Thus, the trend component is defined as a random walk with a switching drift plus a term which captures the effect of the cyclical component on it. That is,

$$Y_t^N = \mu_{S_t} + Y_{t-1}^N + \theta Y_{t-1}^C + \epsilon_t^N,$$

where parameter θ captures the possible impact of the cyclical component on the trend, and $S_t = \{1, 2\}$ indicates the regime of the economy. We consider an economy with two regimes, one related to expansions and the other to recessions. We also assume that S_t is driven by a first-order Markov process with transition probabilities given by

$$\begin{aligned}
\Pr(S_t &= 1 | S_{t-1} = 1) = p_{11}, \\
\Pr(S_t &= 2 | S_{t-1} = 2) = p_{22},
\end{aligned}$$

with $\sum_{j=1}^{2} p_{ij} = 1$ for all i. The transition probability p_{ij} gives the probability that state i will be followed by state j.

The cyclical component is assumed to follow a stationary second-order autoregressive process,

$$Y_t^C = \phi_1 Y_{t-1}^C + \phi_2 Y_{t-2}^C + \epsilon_t^C.$$

The disturbances ϵ_t^N and ϵ_t^C are assumed to be mutually uncorrelated shocks, which are normally distributed with variances σ_N^2 and σ_C^2, respectively. In order to test the number of regimes[4], we perform a linearity test. In our case, the null hypothesis of interest is that of a single regime against the alternative of two regimes, that is,

---

[4]The extension to M states is straightforward.

$$H_0 : \mu_1 = \mu_2 \qquad \text{vs} \qquad H_1 : \mu_1 \neq \mu_2.$$

Note that transition probabilities, p_{11} and p_{22}, are unidentified under the null since any value between 0 and 1 leaves the likelihood function unchanged. This implies that the information matrix is singular under the null, and hence standard regularity conditions are not satisfied. Therefore, the asymptotic null distribution of the likelihood ratio test statistic is not a χ^2 with one degree of freedom. To calculate an appropriate p-value for the test of linearity we propose a bootstrap procedure.

Once the number of regimes has been determined, we are able to test the *Endogenous Growth Hypothesis,* i.e. cyclical shocks have a significant effect on the trend component of output, by simply performing an appropriate test of the relevant null hypothesis, $H_0 : \theta = 0$. If we reject the null, the empirical evidence then suggests that cyclical shocks play an important role in the observed persistence. In fact, if this is the case, as it is pointed out by the EG theory, cyclical shocks produce changes in the resources allocated to growth and, consequently, may have permanent effects on the output level. On the contrary, if we accept the relevant null hypothesis, cyclical shocks do not have a permanent effect on the level of the economy. Therefore, as RBC theory suggests, real shocks are the only source of persistence of output fluctuations. In order to perform a test for H_0 we can employ standard likelihood tests, which are chi-squared distributed with one degree of freedom. The problem of unidentified nuisance parameters under the null vanishes when this test is individually considered.

3 Design of the experiment for computing the bootstrap p-value for the linearity hypothesis test

The objective of this section is to approximate the distribution of the test statistic of interest by a consistent bootstrap procedure. We implement a Wald test statistic. We propose a bootstrap procedure that allows for the presence of general heteroskedasticity.

3.1 Heteroskedastic bootstrap

Our aim is to calculate a bootstrap distribution of the Wald test allowing for the possibility of general heteroskedasticity. In particular, the resampling

we propose is based on the idea of the wild bootstrap, which was studied for the first time by Wu (1986) in the context of variance estimation in heteroskedastic linear models. In our context, it looks like this:

STEP 1: We estimate the Wald test. To compute this test we just need to estimate the model under the alternative hypothesis of nonlinearity. Once the model in Section 2 is put into a state-space form, to estimate it we use the Kim's approximate maximum likelihood algorithm [5]. It is well-known that the likelihood function of state-space models with Markov switching is severely ill-behaved, and it usually has numerous local optima. A nonlinear maximization routine is employed to estimate the model. To implement such a routine we need to set starting values, and their choice can have a dramatic influence upon the maximum which is found. Thus, it is possible that we do not detect the global maximum of the likelihood function. As a result, the Wald test statistic will be an underestimate of the true Wald statistic. To give robustness to our results, we use the Nelder-Mead simplex algorithm to search for a set of parameters near the maximum, and we then estimate the model using Kim's algorithm.

The parameters of interest are $\gamma_1 = (\sigma_N, \sigma_C, \mu_1, \mu_2, \theta, \phi_1, \phi_2, p_{11}, p_{22})'$. We calculate the maximum likelihood estimates $\widehat{\gamma}_1 = (\widehat{\sigma}_N, \widehat{\sigma}_C, \widehat{\mu}_1, \widehat{\mu}_2, \widehat{\theta}, \widehat{\phi}_1, \widehat{\phi}_2, \widehat{p}_{11}, \widehat{p}_{22})'$ and compute the Wald test statistic,

$$W = R\widehat{\gamma}_1(R\widehat{Var}(\widehat{\gamma}_1)R')^{-1}(R\widehat{\gamma}_1)',$$

where $R = (0\ 0\ 1\ -1\ 0\ 0\ 0\ 0\ 0)$ is a vector (1×9), and $\widehat{Var}(\widehat{\gamma}_1)$ is the heteroskedasticity-robust maximum likelihood estimator of the variance-covariance matrix.

STEP 2: We compute the residuals under the null hypothesis of linearity, i.e. $H_0 : \mu_1 = \mu_2$. We construct estimates of the vector of unknown parameters $\lambda_0 = (\sigma_N, \sigma_C, \mu, \theta, \phi_1, \phi_2)'$ and the vector of unobserved variables $(Y_t^N, Y_t^C)'$ using the Kalman filter methodology[6]. Let $\widetilde{\lambda}_0 = (\widetilde{\sigma}_N, \widetilde{\sigma}_C, \widetilde{\mu}, \widetilde{\theta}, \widetilde{\phi}_1, \widetilde{\phi}_2)'$ and $(\widetilde{Y}_t^N, \widetilde{Y}_t^C)'$ denote the maximum likelihood estimates of coefficients and unobserved components, respectively. We compute the residuals

$$\begin{aligned}
\widetilde{\epsilon}_t^C &= \widetilde{Y}_t^C - \widetilde{\phi}_1 \widetilde{Y}_{t-1}^C - \widetilde{\phi}_2 \widetilde{Y}_{t-2}^C; \\
\widetilde{\epsilon}_t^N &= \widetilde{Y}_t^N - \widetilde{\mu} - \widetilde{Y}_{t-1}^N - \widetilde{\theta} \widetilde{Y}_{t-1}^C;
\end{aligned}$$

---

[5]See the Appendix for details about the algorithm proposed by Kim (1994).

[6]See Hamilton (1994) and the Appendix for a detailed description of the Kalman Filter algorithm.

Let T^* denote the sample size for the bootstrap.

STEP 3: We generate B bootstrap samples $Z^*_{T^*b} = \{Z^*_{tb} = (Y^*_t)' : t = 1, ..., T^*\}$, $b = 1, ..., B$. In order to do that, we propose the following algorithm:

I. Generate η^N_t independent and identically distributed variables from a fixed distribution[7] such that $E(\eta^N_t) = 0$ and $E[(\eta^N_t)^2] = E[(\eta^N_t)^3] = 1$. Define $\epsilon^{N*}_t = \widetilde{\epsilon}^N_t \eta^N_t$, where $\widetilde{\epsilon}^N_t$ is the tth residual calculated in Step 2. The bootstrap error ϵ^{N*}_t satisfies $E^*(\epsilon^{N*}_t) = 0$ and $E^*[(\epsilon^{N*}_t)^2] = (\widetilde{\epsilon}^N_t)^2$........ Similarly, generate η^C_t and construct $\epsilon^{C*}_t = \widetilde{\epsilon}^C_t \eta^C_t$.

II. We set the following initial conditions $(Y^{*N}_0, Y^{*C}_0, Y^{*C}_{-1}) = (Y_0, 0, 0)$, that is, the underlying natural component at time 0 is assumed to be equal to the observed value of the GDP series at time 0, whereas the cyclical elements were assumed to be zero. For $t = 1, 2, ..., T^*$, we set $(Y^{*N}_t, Y^{*C}_t) = (\widetilde{Y}^N_t, \widetilde{Y}^C_t)$, that is, unobserved bootstrap components are generated with conditionally fixed design on the estimated unobserved components in Step 2,

$$
\begin{aligned}
Y^{*N}_t &= \widetilde{\mu} + \widetilde{Y}^N_{t-1} + \widetilde{\theta}\widetilde{Y}^C_{t-1} + \epsilon^{N*}_t; \\
Y^{*C}_t &= \widetilde{\phi}_1\widetilde{Y}^C_{t-1} + \widetilde{\phi}_2\widetilde{Y}^C_{t-2} + \epsilon^{C*}_t.
\end{aligned}
$$

III. We next define the bootstrap observations,

$$
Y^*_t = Y^{*N}_t + Y^{*C}_t.
$$

The parameters used to construct the data are the parameter values estimated in Step 2.

STEP 4: Each bootstrap sample $\{Z^*_{Tb} : b = 1, ..., B\}$ is then used to re-estimate the parameters under H_1. Let $\widehat{\gamma}^*_1$ denote the estimator of γ_1 when using the bootstrap sample. We then compute the Wald test statistic associated with the bootstrap sample as

$$
W^* = R\widehat{\gamma}^*_1(R\widehat{Var}(\widehat{\gamma}^*_1)R')^{-1}(R\widehat{\gamma}^*_1)',
$$

STEP 5: Repeating this for $b = 1, ..., B$ gives a sample $\{W^* : b = 1, ..., B\}$ of W values. This sample mimics a random sample of draws of W under the null hypothesis. We compute the bootstrap p-value as $p_B = card(W^* \geq W)/B$, that is the fraction of W^* values that are greater than the observed value W.

We carry out $B = 1000$ bootstrap replications.

---

[7]In particular, the variable η^N_t was sampled from Mammen's (1993, p.257) two-point distribution attaching masses $(5+\sqrt{5})/10$ and $(5-\sqrt{5})/10$ at the points $-(\sqrt{5}-1)/2$ and $(\sqrt{5}+1)/2$, respectively.

4 Empirical results

4.1 Data

The economic series employed is the quarterly real gross domestic product (GDP) for the United States. The time range of the data is 1952:1 to 2005:2. Data series is taken from the OECD *Main Economic indicators* Database. GDP series is obtained as seasonally adjusted and is in natural logs.

4.2 Evidence for unit root

We have imposed that the natural component contains a unit root. Thus, we are interested in testing for a unit root in GDP series. We employ the Phillips-Perron test for unit roots. We obtain that GDP displays a nonstationary behaviour. This result is presented in Table 1.

4.3 Estimation results and hypothesis testing

The result for the linearity test is reported in Table 2. The p-value obtained, 0.000, is calculated following the bootstrap technique described in Section 3. The diagnosis checking of the residuals of the linear model[8] leads us to implement a heteroskedastic bootstrap. According to the bootstrap p-value, we reject the null hypothesis of a linear model at the 1%-significance level. Results concerning the estimated nonlinear model are available in Table 3.

We define asymmetry as the different dynamic of a macroeconomic series during recessions and expansions. Therefore, we are interested in checking if the estimated smooth probabilities coincide with the timing of recessions for the U.S. economy. Figure 1 shows these probabilities along with NBER recession dates. The business cycle peaks derived from the model are quite close to NBER. Thus, our model identifies those recession periods when the growth rate of the trend undergoes changes in its dynamic.

The sign of the growth rate of the trend component is positive during expansions, $\mu_1 = 0.009$, and negative during recessions, $\mu_2 = -0.01$ This result confirms Hamilton's conclusions about real GNP growth is well characterized by recurrent shifts between positive and negative growth periods.

We also obtain that the standard deviation of the trend, $\widehat{\sigma}_N = 0.006$, is larger than the standard deviation of the cycle innovation, $\widehat{\sigma}_C = 0.001$. The

---

[8]The assumptions underlying the errors of the linear model are tested via appropriate autocorrelation, heteroskedasticity and normality test statistics. We find evidence in favour of non-autocorrelation and heteroskedasticity. All results are available from the author upon request.

ratio of standard deviations, $\widehat{\sigma}_N/\widehat{\sigma}_C$, is greater than one. This indicates that most of the output variation can be attributed to the trend component. This result is in contrast to the evidence presented by other studies that find that the majority of output fluctuations in the U.S. are due to cyclical shocks (see Clark, 1987). On the other hand, our result is consistent with the evidence presented by Nelson and Plosser (1982) and Murray and Papanyan (2004).

We next move to the *Endogenous Growth Hypothesis*, i.e. $H_0 : \theta = 0$. We reject the null hypothesis at the 5% significance level, with a p-value equal to 0.000. This means that cyclical shocks have a permanent impact on the trend component of GDP, thus they are an important source of the observed persistence of output fluctuations. Therefore, this persistence cannot be reduced to the presence of exogenous permanent shifts, as it was pointed out in section 1 in relation to RBC theory. Even if real shocks dominate as the source of output fluctuations, these shocks do not work through the mechanisms highlighted in RBC models. Real shocks affect the economy through some Keynesian channel, as the EG theory suggests. The positive sign of parameter θ implies that there is a positive correlation between cyclical shocks and the trend component. This means that a positive cyclical shock leads to a permanent increment in the GDP trend.

We conclude that, if our goal is to study the persistence of output fluctuations, RBC theory may be misleading.

5   Monte Carlo Experiment

This section reports a simple Monte Carlo experiment that is conducted in order to gauge the finite sample performance of the linearity test. The experiments are based upon the model design described in Section 2. The parameter values defining the 11 selected data-generating processes (DGPs) are listed in Table 4.

It is worth noting that DGP_0 corresponds to a single regime model (i.e. $\mu_1 = \mu_2$), whereas from DGP_1 to DGP_{10} we have a two-regime model (i.e. $\mu_1 \neq \mu_2$). This involves generating data from DGP_0 to explore the empirical size of the test, while the remaining ten parameterizations are considered to explore the empirical power of the test. Parameter values for DGP_1 correspond with the point estimates of our model.

Transition probabilities $(p_{11}, p_{22}) = (0.6, 0.4)$ imply that the regime indicator variables $\{S_t\}$ are uncorrelated. In the case of $(p_{11}, p_{22}) = (0.9, 0.9)$ or $(p_{11}, p_{22}) = (0.9, 0.98)$, the two regimes are highly persistent, with the regime corresponding to $S_t = 2$ being almost absorbing in the later case (the stationary distribution of $\{S_t\}$ is $(0.5, 0.5)$ and $(0.1167, 0.8333)$, respectively).

The experiments proceed by generating an artificial time series of length $T + 50$ according to each one of the DGPs under study, with $T \in \{100, 250\}$ and initial values for both the cyclical component and the trend set to zero. The initial value of the Markov chain $\{S_t\}$ was drawn randomly from its stationary distribution, that is, $p(s_t = 1) = (1 - p_{22})/(2 - p_{11} - p_{22})$ and $p(s_t = 2) = (1 - p_{11})/(2 - p_{11} - p_{22})$. The first 50 pseudo-data points are then discarded to minimize the effect of initial conditions and the remaining T points are used to compute the test statistic. For each simulated sample, the p-values were calculated using 250 bootstrap replications.

In Table 5 we report the rejection frequencies in 1000 Monte Carlo replications at the 5% and 10% significance levels. The empirical size of the test is remarkably close to its nominal value even for a small sample size $T = 100$. If the sample size is increased to $T = 250$, the empirical size is more accurate. The power of the test is also quite good. Especially noteworthy is the sensitivity of the results with regard to the magnitude of parameter change. It is evident that the test performance improves as the difference between the values of parameters in the two regimes increases. In addition, the magnitude of the transition probabilities significantly affects the power of the test. More precisely, in both cases, $(p_{11}, p_{22}) = (0.9, 0.9)$ and $(p_{11}, p_{22}) = (0.9, 0.98)$, we observe a high power. A loss in power is observed in the case of $(p_{11}, p_{22}) = (0.6, 0.4)$. As expected, it can be seen for each design point that, as T increases, the power improves considerably.

To sum up, the general conclusion that can be drawn from the simulation findings is that the proposed test performs quite well.

6   Conclusions

We have claimed that one of the important stylized facts of business cycle is the persistence of output fluctuations. The aim of this paper is to identify the source of this observed persistence for the case of the U.S. economy. With this in mind, we propose an unobserved components model that allows us to decompose GDP series into a trend component and a cyclical component. We model the trend component as a random walk with drift. The transitory component is supposed to follow a stationary second-order autoregressive process. Following Hamilton (1989), we allow the growth rate of the trend component to switch between different regimes according to a first-order Markov process.

To determine the number of regimes, we propose a bootstrap procedure that allows for general forms of heteroskedasticity. We investigate the performance of the proposed bootstrap procedure by means of a Monte Carlo

11

simulation. We find that it works correctly. Hypothesis testing suggests that cyclical shocks are an important source of persistence in output fluctuations.

We conclude that the persistence of output fluctuations cannot be simply reduced to the presence of exogenous permanent shifts in the production function, as it is usually assumed by the RBC theory. As the EG theory suggests, cyclical shocks need to be take into account.

References

[1] Andrews, D.W.K. and Ploberger, W. (1994). *Optimal tests when a nuisance parameter is present only under the alternative.* Econometrica 62(6), 1383-1414.

[2] Campbell, J.Y. and Mankiw, N.G. (1987). *Are output Fluctuations Transitory?.* The Quarterly Journal of Economics 102(4), 857-80.

[3] Carrasco, M., Hu, L. and Ploberger, W. (2004). *Optimal test for Markov Switching.* Working paper, University of Rochester.

[4] Clark, P. (1987). The Cyclical Component of U.S. Economic Activity. Quarterly Journal of Economics102, 797-814.

[5] Cogley, T. (1990). *International Evidence on the Size of the Random Walk in Output.* Journal of Political Economy 98(3), 501-18.

[6] Fatás, A. (2000). *Endogenous growth and stochastic trends.* Journal of Monetary Economics 45(1), 107-128.

[7] Garcia, R. (1998). *Asymptotic null distribution of the likelihood ratio test in Markov Switching Models.* International Economic Review 39(3), 763-88.

[8] Hamilton, J.D. (1989). *A new approach to the economic analysis of nonstationary time series and the business cycle.* Econometrica 57(2), 357-384.

[9] Hamilton, J.D. (1994). *Time Series Analysis.* Princeton: Princeton University Press.

[10] Hansen, B.E. (1992a). *The Likelihood Ratio Test under Nonstandard Conditions: Testing the Markov Switching Model of GNP.* Journal of Applied Econometrics 7(5), S61-82.

[11] Hansen, B.E. (1992b). *Erratum: The Likelihood Ratio Test under Non-standard Conditions: Testing the Markov Switching Model of GNP.* Journal of Applied Econometrics, 11(2), 195-98.

[12] Hansen, B. E. (1996). *Inference when a nuisance parameter is not identified under the null hypothesis.* Econometrica 64(2), 413-30.

[13] Harvey, A.C. (1989). *Forecasting, structural time series models and the Kalman filter.* Cambridge University Press.

[14] Jaeger, A. and Parkinson, M. (1994). *Some evidence on Hysteresis in Unemployment Rates.* European Economic Review 38, 329-42.

[15] Kalman, R.E. (1960). *A New Approach to Linear Filtering and Prediction Problems.* Journal of Basic Engineering, Transactions of the ASME, Series D, 82, 35-45.

[16] Kim, C.J. (1994). *Dynamic linear models with Markov-switching.* Journal of Econometrics 60(1-2), 1-22.

[17] Kim, C.J. and Murray, C. (2002). *Permanent and Transitory Nature of Recessions.* Empirical Economics 27(2), 149-162.

[18] Kim, C. J. and Nelson, C.R. (1999). *State-Space Models with Regime-Switching: Classical and Gibbs-Sampling Approaches with Applications.* 2nd ed. Cambridge: The MIT press.

[19] Kim, C.J. and Piger, J. (2002). *Common Stochastic Trends, Common Cycles, and Asymmetry in Economic Fluctuations.* Journal of Monetary Economics 49(6), 1189-1211.

[20] King, R.G., Plosser, C.I. and Rebelo, S.T. (1988). *Production, Growth and Business Cycles II. New Directions.* Journal of Monetary Economics 21(2-3), 309-41.

[21] King, R.G., Plosser, C.I., Stock, J.H. and Watson, M.W. (1991). *Stochastic Trends and Economic Fluctuations.* American Economic Review 81(4), 819-40.

[22] Kydland, F.E. and Prescott, E.C. (1982). *Time to Build and Aggregate Fluctuations.* Econometrica 50(6), 1345-70.

[23] Mammen, E. (1993). *Bootstrap and Wild Bootstrap for High-Dimensional Linear Models.* The Annals of Statistics 21, 255-85.

[24] Murray, C.J. and Papanyan, S. (2004). *The relative importance of Permanent and Transitory components in Macroeconomic Time series.* Mimeo, University of Houston.

[25] Nelson, C.R. and Plosser, C.I. (1982). *Trends and Random Walks in Macroeconomic Time Series.* Journal of Monetary Economics 10(2), 139-62.

[26] Rebelo, S.T. (2005). *Real Business Cycle Models: Past, Present, and Future.* NBER Working Paper No. W11401.

[27] Stadler, G.W. (1990). *Business cycle models with Endogenous Technology.* American Economic Review 80(4), 763-78.

[28] Wu, C.F.J. (1986). *Jackknife, Bootstrap and Other Resampling Methods in Regression Analysis.* The Annals of Statistics 14(4), 1261-95.

Appendix A: Estimation procedures

In this appendix we present different filters which have been proposed in the relevant literature for estimating the sort of model described in Section 2. Firstly, we examine the Kalman filter, which allows us to estimate the linear model. Secondly, we present Kim's approach, which combines unobserved components models with the Markov-switching methodology.

The Kalman Filter

In 1960, R.E. Kalman published a famous paper describing a recursive solution to the discrete data linear filtering problem. Since that time, due in large part to advances in digital computing, the Kalman filter has been the subject of extensive research and applications, particularly in the area of autonomous or assisted navigation.

The Kalman filter[9] is a set of mathematical equations that provides an efficient recursive computational procedure to estimate the state of a process, in a way that minimizes the mean squared error (MSE). The filter is very powerful in several aspects: it supports estimations of past, present, and even future states, and it can do so even when the precise nature of the modeled system is unknown.

To start with, consider an $(n \times 1)$ vector of observed variables at date t, y_t. These observable variables are related to a possibly unobserved $(r \times 1)$ vector h_t, known as the *state vector*, via a *measurement equation*,

$$y_t = H'h_t + A'x_t + w_t, \tag{1}$$

where H' and A' are matrices of parameters of dimension $(n \times r)$ and $(n \times k)$, respectively, x_t is an $(k \times 1)$ vector containing exogenous or lagged dependent variables, and the $(n \times 1)$ vector w_t is a white noise disturbance vector with covariance matrix given by:

$$E(w_t w'_\tau) = \left\{ \begin{array}{ll} R & for\ t = \tau \\ 0 & otherwise \end{array} \right. ,$$

where R is an $(n \times n)$ matrix.

Despite the fact that the variables of h_t are, in general, not observable, they are known to be generated by a first-order Markov process,

$$h_t = Fh_{t-1} + \Pi'x_t + v_t, \tag{2}$$

---

[9]See Hamilton (1994, Chapter 13) and Harvey (1989, Chapter 3) for a more detailed description of the Kalman filter.

where F and Π' are matrices of parameters of dimension $(r \times r)$ and $(r \times k)$, respectively. The $(r \times 1)$ vector v_t is a withe noise disturbance vector:

$$E(v_t v_\tau') = \begin{cases} Q & for\ t = \tau \\ 0 & otherwise \end{cases},$$

where Q is an $(r \times r)$ matrix. Equation (2) is the *transition equation*.

The disturbances v_t and w_t are assumed to be uncorrelated at all lags:

$$E\ (v_t w_\tau') = 0\ \ for\ all\ t\ and\ \tau.$$

Further assumptions on measurement and transition disturbances are as follows: i) they are uncorrelated with the exogenous variables; ii) they are assumed to be normally distributed in order to calculate the likelihood function.

The state space form (SSF) that represents the dynamics of the univariate time series y_t is composed of equations (1) and (2). The information set at time $t-1$ is given by matrix $\Psi_{t-1} \equiv (y_{t-1}', y_{t-2}', ...y_1', x_{t-1}', x_{t-2}', ..., x_1')'$. Note that there are two set of unknowns: the parameters of the model in H', A', R, F, Π' and Q (these matrices will be referred as the system matrices) and the elements of the state vector h_t........ The goal of the Kalman filter procedure is to form a forecast of the unobserved state vector at time t based on the information at date $t-1$. For now, we will assume that the particular numerical values of system matrices are known. Let $\widehat{h}_{t|t-1}$ denote the linear forecast of the state vector h_t based on (x_t, Ψ_{t-1}), and $P_{t|t-1}$ denote the MSE matrix associated with this forecast.

Because the filter is a recursion, it is started assuming initial values for the mean and variance of the state variables, $\widehat{h}_{1|0}$ and $P_{1|0}$, respectively. Then, we can conduct the Kalman filter in four major steps. First, we calculate the one-period-ahead forecast of the unobserved state vector and the associated MSE matrix :

$$\begin{aligned} \widehat{h}_{t|t-1} &= E\left[h_t | x_t, \Psi_{t-1}\right] = F\widehat{h}_{t-1|t-1} + \Pi' x_t, \\ P_{t|t-1} &= E\left[(h_t - \widehat{h}_{t|t-1})(h_t - \widehat{h}_{t|t-1})' | \Psi_{t-1}\right] = F P_{t-1|t-1} F' + Q. \end{aligned}$$

The next step is to calculate the one-step forecast of the measurement variable y_t at date $t-1$ knowing information up to and including $t-1$,

$$\widehat{y}_{t|t-1} = E(y_t | x_t, \Psi_{t-1}) = H'\widehat{h}_{t|t-1} + A' x_t. \tag{3}$$

Once the new observation y_t becomes available at date t, we can calculate the forecast error on the observed variable and its MSE:

$$\begin{aligned} \lambda_t &= y_t - \hat{y}_{t|t-1}, \\ \Lambda_t &= E[(y_t - \hat{y}_{t|t-1})(y_t - \hat{y}_{t|t-1})' | \Psi_t] = H' P_{t|t-1} H + R, \end{aligned} \tag{4}$$

and it is possible to update the estimated state vector and its MSE matrix:

$$
\begin{aligned}
\widehat{h}_{t|t} &= E\left[h_t | \Psi_t\right] = \widehat{h}_{t|t-1} + \Phi_t \lambda_t, \\
P_{t|t} &= (I - \Phi_t H') P_{t|t-1},
\end{aligned}
$$

where $\Phi_t = P_{t|t-1} H (\Lambda_t)^{-1}$ is known as the *filter gain* since a certain fraction of the difference between the observable and the predicted state is added to the previous prediction. These last two terms that are generated using the updating equations correspond to the inputs of the next filter iteration.

Hence, if the system matrices were known the Kalman filter would yield as outcome the sequences $\{\widehat{h}_{t|t-1}\}_{t=1}^{T}$ and $\{P_{t|t-1}\}_{t=1}^{T}$. We can view the Kalman filter as a sequential updating procedure that consists of forming a prior guess about the state of nature and then adding a correction to this guess, this correction being determined by how well the guess has performed in predicting the next observation. However, the state space model is not entirely estimated since we usually do not known the parameters of the system matrices. Considering that $\{v_t, w_t\}_{t=1}^{T}$ are Gaussian, then the distribution of y_t conditional on (x_t, Ψ_{t-1}) is Gaussian with mean given by equation (3) and variance given by equation (4). We use the *prediction error decomposition* to construct the logarithm of the distribution function, so that for a Gaussian model it has the form:

$$
\ln f(y_t | x_t, \Psi_{t-1}) = -\frac{N}{2} \ln 2\pi - \frac{1}{2} \ln |\Lambda_t| - \frac{1}{2} \lambda_t' \Lambda_t^{-1} \lambda_t.
$$

Finally, to estimate the parameters of the system matrices we maximize the log-likelihood function

$$
\ln L = \sum_{t=1}^{T} \ln f(y_t | x_t, \Psi_{t-1})
$$

with respect to the underlying unknown parameters using nonlinear optimization techniques.

The Kim Filter

The aim of the Kim filter is to extend the Markov-Switching methodology to a general state-space framework, where measurement and transition equations may both switch between regimes. The state-space representation becomes:

$$
\begin{aligned}
y_t &= H'_{s_t} h_t + A'_{s_t} x_t + w_t, \\
h_t &= F_{s_t} h_{t-1} + \Pi'_{s_t} x_t + \nu_t,
\end{aligned}
$$

with

$$
(w_t, v_t)' \sim N \left(0, \left(\begin{array}{cc} R_{s_t} & 0 \\ 0 & Q_{s_t} \end{array} \right) \right),
$$

where s_t is a latent process which reflects the state of the system. It can take values $1, 2, ...M$, with M standing for the number of states. As in Hamilton (1989), s_t is supposed to follow an M-state first-order Markov process with transition probabilities:

$$
p_{ij} = \Pr\left(s_t = j | s_{t-1} = i, s_{t-2} = k...\right) = \Pr(s_t = j | s_{t-1} = i)
$$

where

$$
\sum_{j=1}^{M} p_{ij} = 1.
$$

The goal of this filter is to construct a forecast of h_t and its associated mean square error, based not only on Ψ_{t-1} but also on the random variable s_t taking value j and s_{t-1} taking value i:

$$
\begin{aligned}
\widehat{h}_{t|t-1}^{(i,j)} &= E\left[h_t | \Psi_{t-1}, s_t = j, s_{t-1} = i\right] \\
P_{t|t-1}^{(i,j)} &= E\left[\left(h_t - \hat{h}_{t|t-1}\right)\left(h_t - \hat{h}_{t|t-1}\right)' | \Psi_{t-1}, s_t = j, s_{t-1} = i\right]
\end{aligned}
$$

The Kim filter proceeds as follows. We first apply the Kalman filter algorithm conditioning on $s_t = j$ and $s_{t-1} = i$:

$$
\begin{aligned}
\widehat{h}_{t|t-1}^{(i,j)} &= F_j \widehat{h}_{t-1|t-1}^{i} + \Pi'_j x_t \\
P_{t|t-1}^{(i,j)} &= F_j P_{t-1|t-1}^{i} F'_j + Q_j
\end{aligned}
$$

$$
\begin{aligned}
\lambda_{t|t-1}^{(i,j)} &= y_t - H'_j \hat{h}_{t|t-1}^{(i,j)} - A'_j x_t \\
\Lambda_t^{(i,j)} &= H'_j P_{t|t-1}^{(i,j)} H_j + R_j
\end{aligned}
$$

$$
\begin{aligned}
\hat{h}_{t|t}^{(i,j)} &= \hat{h}_{t|t-1}^{(i,j)} + \Phi_t^{(i,j)} \lambda_{t|t-1}^{(i,j)} \\
P_{t|t}^{(i,j)} &= \left(I - \Phi_t^{(i,j)} H'_j\right) P_{t|t-1}^{(i,j)}
\end{aligned}
$$

18

with

$$\Phi_t^{(i,j)} = P_{t|t-1}^{(i,j)} H_j [\Lambda_t^{(i,j)}]^{-1}$$

representing the filter gain, and so on.

Each iteration of the Kalman filter produces an M-fold increase in the number of cases to be considered, which will turn the algorithm into something computationally unfeasible even for the simplest case. To do the filter operable, Kim proposes the following approximations:

$$\widehat{h}_{t|t}^j = \frac{\sum_{i=1}^M \Pr(s_{t-1}=i, s_t=j|\Psi_t)\hat{h}_{t|t}^{(i,j)}}{pr\,(s_t=j|\Psi_t)},$$

$$P_{t|t}^j = \frac{\sum_{i=1}^M \Pr(s_{t-1}=i, s_t=j|\Psi_t)\left[P_{t|t}^{(i,j)} + \left(\widehat{h}_{t|t}^j - \hat{h}_{t|t}^{(i,j)}\right)\left(\widehat{h}_{t|t}^j - \hat{h}_{t|t}^{(i,j)}\right)'\right]}{\Pr\,(s_t=j|\Psi_t)}$$

The weighted state vector $\widehat{h}_{t|t}^j$ (and its variance) constitutes the input of the modified Kalman filter for the next iteration. The probabilities are computed using a filter similar to that of Hamilton (1989) and are based on the Markov-switching assumption.

At first step, knowing the regime transition matrix, $\Pr(s_t = j|s_{t-1} = i)$, and the probability of being in state i at $(t-1)$, $\Pr(s_{t-1} = i|\Psi_{t-1})$, it is possible to calculate:

$$\Pr(s_{t-1}=i, s_t=j|\Psi_{t-1}) = \Pr(s_t=j|s_{t-1}=i) * \Pr(s_{t-1}=i|\Psi_{t-1}).$$

This allows us to calculate the joint conditional density function of y_t and (s_t, s_{t-1}):

$$f(y_t, s_{t-1}=i, s_t=j|\Psi_{t-1}) = f(y_t|s_{t-1}=i, s_t=j|\Psi_{t-1}) * \Pr(s_{t-1}=i, s_t=j|\Psi_{t-1}),$$

with

$$f(y_t|s_{t-1}=i, s_t=j, \Psi_{t-1}) = (2\pi)^{-n/2}|\Lambda_t^{(i,j)}|^{-1/2}\exp(-\frac{1}{2}(\lambda_{t|t-1}^{(i,j)})'\Lambda_t^{(i,j)^{-1}}\lambda_{t|t-1}^{(i,j)})$$

standing for the likelihood derived from the model.

We finally obtain the probabilities needed for the last collapsing operation:

$$\Pr(s_{t-1}=i, s_t=j|\Psi_t) = \frac{f(y_t, s_{t-1}=i, s_t=j|\Psi_{t-1})}{f(y_t|\Psi_{t-1})},$$

19

with

$$f(y_t|\Psi_{t-1}) = \sum_{i=1}^{M} \sum_{j=1}^{M} f(y_t, s_{t-1} = i, s_t = j|\Psi_{t-1}),$$

and

$$\Pr(s_t = j|\Psi_t) = \sum_{i=1}^{M} \Pr(s_{t-1} = i, s_t = j|\Psi_t).$$

A direct inference for the unobserved component is:

$$\hat{h}_{t|t} = \sum_{j=1}^{M} \Pr(s_t = j|\Psi_t)\hat{h}_{t|t}^{j}.$$

As in the usual kalman filter, the filter is conducted under the assumption that the parameters to be estimated are constant and known. The log-likelihood is a by-product of the filter and is calculated as follows:

$$\ln L = \sum_{t=1}^{T} \ln f(y_t|\Psi_{t-1}).$$

We maximize the log-likelihood function with respect to the underlying unknown parameters using non-linear optimizations routines.

Appendix B: Tables and Figures

Table 1: Unit Root Tests

Phillips-Perron test on GDP series	
U.S.	0.050

Note 1: For the Phillips-Perron test, we use Mackinnon critical values for rejecting of hypothesis of a unit root. We do not reject the null hypothesis of a unit root at 1%, 5% and 10%.

Table 2: Linearity test

	$H_0 : \mu_1 = \mu_2$
U.S.	$P - value = 0.000^*$

*Significant at 1%.

Table 3: Estimation Results for the U.S.[‡]

NONLINEAR MODEL		
Natural Component Equation	$i = 1$	$i = 2$
μ_i	0.009 (0.001)	-0.010 (0.003)
θ	2.403 (0.256)	
σ^N	0.006 (0.0002)	
Cyclical Component Equation		
ϕ_1	1.356 (0.195)	
ϕ_2	-0.722 (0.128)	
σ^C	0.001 (0.0003)	
Transition Probabilities		
p_{11}	0.960 (0.059)	
p_{22}	0.280 (0.091)	

[‡]standard errors in parentheses.

$Table\ 4$: **Parameter Values for the DGP**

	ϕ_1	ϕ_2	σ_C	μ_1	μ_2	θ	σ_N	p_{11}	p_{22}
DGP_0	0.5	-0.06	0.02	0.2	0.2	0.06	0.5	0	0
DGP_1	1.4	-0.7	0.001	0.009	-0.01	2.4	0.006	0.96	0.28
DGP_2	1.6	-0.6	0.004	0.9	0.7	0.03	0.005	0.6	0.4
DGP_3	1.6	-0.6	0.004	0.9	0.7	0.03	0.005	0.9	0.9
DGP_4	1.6	-0.6	0.004	0.9	0.7	0.03	0.005	0.9	0.98
DGP_5	1.6	-0.6	0.004	0.9	0.5	0.03	0.005	0.6	0.4
DGP_6	1.6	-0.6	0.004	0.9	0.5	0.03	0.005	0.9	0.9
DGP_7	1.6	-0.6	0.004	0.9	0.3	0.03	0.005	0.9	0.98
DGP_8	1.6	-0.6	0.004	0.9	0.3	0.03	0.005	0.6	0.4
DGP_9	1.6	-0.6	0.004	0.9	0.3	0.03	0.005	0.9	0.9
DGP_{10}	1.6	-0.6	0.004	0.9	0.3	0.03	0.005	0.9	0.98

$Table\ 5$: **Monte Carlo Size and Power**

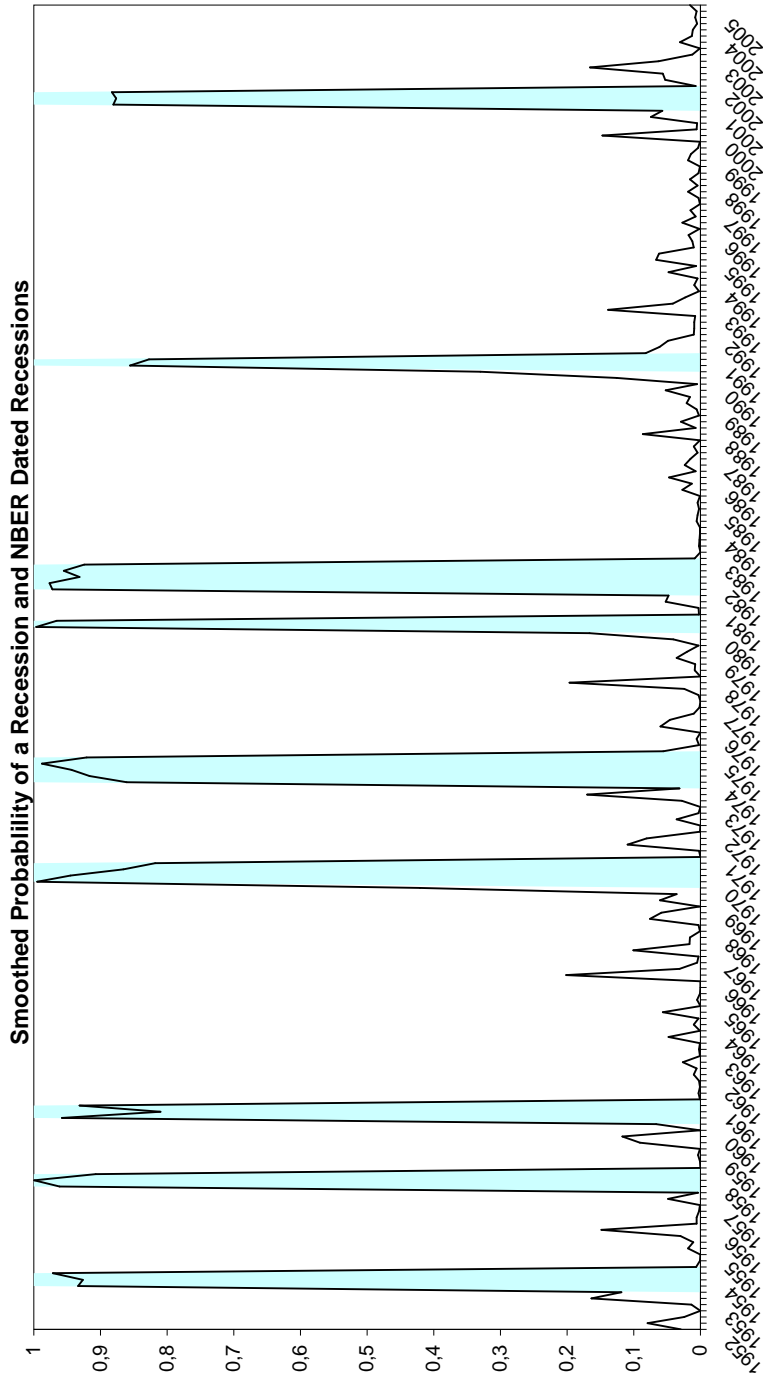|	$T = 100$		$T = 250$	
Level	0.05	0.10	0.05	0.10
DGP_0	0.042	0.092	0.048	0.099
DGP_1	0.690	0.702	0.913	0.920
DGP_2	0.325	0.496	0.422	0.528
DGP_3	0.572	0.633	0.730	0.757
DGP_4	0.555	0.610	0.675	0.700
DGP_5	0.462	0.525	0.605	0.778
DGP_6	0.632	0.679	0.777	0.895
DGP_7	0.619	0.670	0.751	0.770
DGP_8	0.534	0.661	0.731	0.756
DGP_9	0.772	0.899	0.950	0.996
DGP_{10}	0.668	0.695	0.932	0.976

Figure 1: Smooth probability of regime $S_t = 2$ (GDP: 1952:I-2005:II)