

Department of Economics

Nonlinear Modelling of Autoregressive Structural Breaks in a US Diffusion Index Dataset

George Kapetanios and Elias Tzavalis

Working Paper No. 537

May 2005

ISSN 1473-0278



Queen Mary
University of London

Nonlinear Modelling of Autoregressive Structural Breaks in a US Diffusion Index Dataset

G. Kapetanios and E. Tzavalis
Queen Mary, University of London

November 29, 2004

Abstract

This paper applies a new model of structural breaks developed by Kapetanios and Tzavalis (2004) to investigate if there exist structural changes in the mean reversion parameter of US macroeconomic series. Ignoring such type of breaks may lead to spurious evidence of unit roots in the autoregressive parameters of economic series. Our model specifies that both the timing and size of breaks are stochastic. We apply the model to a variety of macroeconomic and finance series from the US.

Keywords: Structural Breaks, State Space Model, Nonlinearity

JEL: E32, C13, C22

1 Introduction

The study of structural breaks has attracted considerable interest in the econometric literature in the past 25 years. Breaks are by definition sudden events which change the structure of the econometric model under consideration. The occurrence of structural breaks appears as one of the (if not the) most significant cause of forecasting failure as identified by the work of Hendry among others (see, e.g., Clements and Hendry (1999)). If methods that provide information about the occurrence of structural breaks were available they would be of great help to empirical econometric and economic research. Most of the work on structural breaks is concentrated on detecting the presence of structural breaks and estimating the location of the breaks, as well as the values of the model parameters prior to and following breaks. However, these studies take the occurrence of breaks as given and not viewed as part of model specification. There is little attention in the literature on a discussion of generating sources of structural breaks.

Modelling structural breaks involves setting up a nonlinear model of some form, or other. Available nonlinear econometric models, however, do not appear to capture essential characteristics of structural breaks such as rarity. In addition, they do not allow for stochastic changes in the magnitude of the structural breaks. For example, time varying coefficient models do not allow for discontinuous type of changes in the structural parameters, while nonlinear models such as threshold and Markov switching models [see Lin and Terasvirta (1994), and Hamilton (1989), respectively] assume that the structural parameter changes of fixed magnitude. Clearly, the discontinuity of the breaks requires some sort of trigger that initiates a break. The use of thresholds clearly provides a paradigm for modelling such triggers. A related question then concerns the variables that underlie the threshold behaviour. These variables can be thought of as the regulators of the frequency of the breaks. By their very nature, the occurrence of breaks presupposes the occurrence of some extreme event such as the event of a variable taking an extreme value. The property of a stochastic nature of the magnitude of structural breaks requires the presence of a mechanism for specifying the parametrisation of the model after each break. The set of parameters should be neither finite nor predetermined prior to the break. As a result such a mechanism should be partly characterised by randomness. The above two prerequisites for modelling structural breaks form the basis of the Kapetanios and Tzavalis (2004) model of structural breaks. This model

suggests that structural breaks are driven endogenously by larger than a threshold parameter structural errors, as part of model specification. Thus, this model enables us to identify potential generating sources of structural breaks.

In this paper, we apply an extension of that model of structural breaks with the aim of examining whether there exist structural breaks in the mean reversion parameter of autoregressive models of macroeconomic series. Answering this question may have interesting implications, as evidence suggests that the persistence of major macroeconomic series changes over time. Ignoring such type of breaks may also lead to spurious evidence of unit roots. Furthermore, from an economic point of view it may be interesting to examine if changes in business conditions or monetary regimes can cause shifts in the mean reversion parameters, in the first place. Our model enables us to investigate the above question without *a priori* fixing the number of the breaks in the autoregressive coefficient or the magnitude of the structural break changes. Both of them are stochastic and can be determined endogenously by the data.

The paper is organised as follows: Section 2 presents our modelling procedure of structural breaks and discuss its estimation. Section 3 presents a small Monte Carlo study to appraise the performance of our model. Section 4 provides an empirical application of our model to a number of US macroeconomic series. Finally, Section 5 concludes.

2 Modelling Structural Breaks in autoregressive coefficients

We focus our study on the demeaned autoregressive model of lag order one i.e. AR(1), as this model has attracted a considerable amount of interest in the literature to examine the persistency or nonlinearity of the mean reversion coefficient. We consider the following parameterization of the AR(1) model which allows for structural break type of changes in the mean reversion coefficient

$$y_t = \beta_t y_{t-1} + \epsilon_{1,t} \tag{1}$$

where

$$\beta_t = \beta_{t-1} + I(|\epsilon_{1,t-1}| > r)\epsilon_{2,t-1} \quad (2)$$

$$\epsilon_{1,t} = \rho_1\epsilon_{1,t-1} + \eta_{1,t} \text{ and} \quad (3)$$

$$\epsilon_{2,t} = \rho_2\epsilon_{2,t-1} + \eta_{2,t}, \quad (4)$$

where $\eta_{1,t}$ and $\eta_{2,t}$ are two IID innovations which can be allowed to be contemporaneously correlated. Model (1) constitutes a standard time-varying coefficient, state space model where the mean reversion coefficient is a state variable. $I(\mathcal{A}_t)$ is an indicator function taking the value 1, if the event \mathcal{A}_t , defined as $|\epsilon_{1,t-1}| > r$ where r is a threshold parameter, occurs, and 0 otherwise. This function can capture abrupt discontinuous changes in the mean reversion coefficient β_t , in line with the common perception of structural breaks referred to in the literature. Both the timing and the magnitude of a break in β_t are stochastic in nature. The timing is controlled by the error term $\epsilon_{1,t}$, while the magnitude by the error term by $\epsilon_{2,t}$. The stochastic nature of the magnitude of the break distinguishes our parameterization of structural breaks from other models of structural breaks, such as the Markov Chain, regime switching model of Hamilton (1989) and other time varying threshold models, see, e.g., Lin and Terasvirta (1994). The latter assume that the size of the structural changes in parameters is fixed.

As it stands, model (1) constitutes a nonlinear autoregressive model where the structural changes in β_t are nonlinear functions of the error term $\epsilon_{1,t}$. The autoregressive structure of $\epsilon_{1,t}$ implies that the changes in β_t are associated with bigger than r shifts in the level of the series y_t .¹ This specification of $\epsilon_{1,t}$ also enables forecasting the timing of a possible future break changes based on the information set \mathcal{I}_{t-1} , defined as $\mathcal{I}_{t-1} = (y_{t-1}, \dots, y_0)$. However, if the autoregressive coefficient of the process of $\epsilon_{1,t}$ becomes $\rho_1 = 0$, then $\epsilon_{1,t} = \eta_{1,t}$ which means that the breaks in β_t are driven by the large innovations in the level of the series y_t [see Kapetanios and Tzavalis (2004)], which may have an interesting economic interpretation. In this case, our model can forecast only the magnitude of a future structural break in β_t through the autoregressive processes assumed for the error term $\epsilon_{2,t}$. Finally, if both $\rho_1 = 0$ and $\rho_2 = 0$, then model (1) can not forecast either the timing or the size of a change in β_t . In this case, the model can be only used to track the changes in β_t , ex post.

¹Obviously, this structure of $\epsilon_{1,t}$ (as well as that of $\epsilon_{2,t}$) can be extended to a p^{th} order AR model, or an ARMA(p,q) model. The AR(1) model of $\epsilon_{1,t}$ was chosen for presentational convenience.

Model (1) can nest a familiar model often used, in practice, to describe the level shifts in many macroeconomic series. This is the autoregressive model with a unit root in the autoregressive component. It can be obtained from (1), if the variance of the innovation $\eta_{2,t}$, denoted $\sigma_{\eta_{2,t}}^2$, is $\sigma_{\eta_{2,t}}^2 = 0$ and $\beta_0 = 1$. In this case, the autoregressive coefficient of model (1) becomes $\beta_t = \beta_0 = 1$, for all t . Then, our model becomes a nonstationary one including a stationary component given by the process underlying $\epsilon_{1,t}$. This makes an application of the model to macroeconomic data appealing because, as aptly noted in the literature, ignoring nonlinearities in the functional form of β_t may lead to spurious inference of a unit root in the series y_t . On this front, our model is in the spirit of Engle and Smith (1999) model which considers shifts in β_t to be taken as a non-linear functions of the error term $\epsilon_{1,t}$, given by $\beta_t = \beta_{t-1} + q_{t-1}(\epsilon_{1,t-1}, \gamma)\epsilon_{1,t-1}$, where $q_t(\epsilon_{1,t}, \gamma) = \frac{\epsilon_{1,t}^2}{\gamma + \epsilon_{1,t}^2}$ is a continuous function of the error term $\epsilon_{1,t}$.² The main difference of our modelling approach from that of Engle and Smith's is that we are concerned with level shifts in β_t (and, hence, y_t) which are abrupt and rare in line with the common perceptions of breaks, whereas Engle and Smith consider smooth changes in β_t which occur at every period.

Estimation of this model is tricky. But it can be made simpler by a subtle change in the specification. More specifically, we know that a state space model whose coefficients are stochastic functions of lags of the observed series y_t is still a conditionally Gaussian state space model on which the Kalman filter can be applied. As our model stands now it is not conditionally Gaussian since the coefficients depend on past states which are not functions of the lags of the observed process. Estimates of the states can of course be obtained by nonlinear filtering which is however both cumbersome and difficult to set up. Nevertheless, the optimal minimum mean squared estimates of the state (assuming known hyperparameters) at time $t - 1$ conditional on data available at time $t - 1$, i.e. the filtered estimates of the state obtained via the Kalman filter are, by definition, functions of the lags of the observed process and therefore if we specify the autoregressive coefficient as the state variable

$$\beta_t = \beta_{t-1} + I(|\hat{\epsilon}_{1,t-1|t-1}| > r)\epsilon_{2,t-1}, \quad (5)$$

we have a conditionally Gaussian model. On this we can apply the Kalman filter. In the above specification of the state variable β_t , the change is that, instead of $\epsilon_{1,t-1}$, we specify

²This model is primarily focused on the investigation of whether shocks on the level of economic series have a permanent effect.

that $\hat{\epsilon}_{1,t-1|t-1}$ enters the time-varying coefficient of the transition equation of the model, now given by (5). In other words, the forecastable part of the state controls the timing of the break. Replacing (2) by (5), someone can view this specification of our model as a convenient respecification needed to simplify enormously the estimation of the states based on the information set \mathcal{I}_{t-1} . From this point of view, model (1) with (5) can be thought of as an approximation to Model (1). This specification can be given the economic justification that structural breaks occur intertwined with the expectations formation process of economic agents about the possibility of a future structural break based on an estimate of $\hat{\epsilon}_{1,t-1|t-1}$ relative to the threshold parameter r .

We now give details on the estimation of this model. At first assume that the threshold parameter r is known. Then, we can assume Gaussianity for $\eta_{1,t}$ and $\eta_{2,t}$ and use the standard Kalman filter. More specifically, as Harvey (1989) explains, a conditionally Gaussian model can be constructed. To see this we consider the general state space model

$$y_t = \mathbf{X}'_t \beta_t + u_t, \quad u_t \sim i.i.d.(0, \sigma_t^2) \quad t = 1, \dots, T \quad (6)$$

$$\beta_t = \mathbf{A}_t \beta_{t-1} + \varepsilon_t \quad \varepsilon, \mathbf{t} \sim i.i.d.(0, \Sigma_{\varepsilon,t}) \quad (7)$$

where, in connection to model (1), $\mathbf{X}_t = (y_{t-1}, 1, 0)$, $\beta_t = (\beta_t, \epsilon_{1,t}, \epsilon_{2,t})$, $\varepsilon_t = (0, \eta_{1,t}, \eta_{2,t})$, $\sigma_t^2 = 0$, and $\mathbf{A}_t = \begin{pmatrix} 1 & I(|\hat{\epsilon}_{1,t-1|t-1}| > r) & 0 \\ 0 & \rho_1 & 0 \\ 0 & 0 & \rho_2 \end{pmatrix}$. The optimality of the Kalman filter crucially depends on assuming that the measurement and transition equation errors are normally distributed. We abstract from issues arising from the estimation of the parameters of the models and concentrate on the estimation of the state variable conditional on the parameters being known. We denote the estimator of β_t conditional on the information set up to and including time t by \mathbf{b}_t . We denote the covariance matrix of \mathbf{b}_t by \mathbf{P}_t . The estimator of β_t conditional on the information set up to and including time $t - 1$ is denoted by $\mathbf{b}_{t|t-1}$. Its covariance matrix is denoted by $\mathbf{P}_{t|t-1}$. The Kalman filter comprises sequential application of two sets of equations which recursively deliver the estimates of the state variable and their covariance matrix. The filter is initialised by specifying the estimate of the state and its covariance matrix at the start of the sample. The two sets of equations are given by

$$\mathbf{b}_{t|t-1} = \mathbf{A}_t \mathbf{b}_{t-1} \quad (8)$$

$$\mathbf{P}_{t|t-1} = \mathbf{A}_t \mathbf{P}_{t-1} \mathbf{A}'_t + \Sigma_{\varepsilon,t} \quad (9)$$

which are known as the prediction equations, and

$$\mathbf{b}_t = \mathbf{b}_{t|t-1} + \mathbf{P}_{t|t-1} \mathbf{x}_t \left(\frac{y_t - \mathbf{x}_t' \mathbf{b}_{t|t-1}}{f_t} \right) \quad (10)$$

$$\mathbf{P}_t = \mathbf{P}_{t|t-1} - \mathbf{P}_{t|t-1} \mathbf{x}_t \left(\frac{1}{f_t} \right) \mathbf{x}_t' \mathbf{P}_{t|t-1} \quad (11)$$

which are known as the updating equations. f_t is given by $\mathbf{X}_t' \mathbf{P}_{t|t-1} \mathbf{X}_t + \sigma_t^2$. The log-likelihood of the model may be easily written in terms of the prediction errors. It is given by

$$-\frac{T}{2} \log 2\pi - \frac{1}{2} \sum_{t=1}^T \log f_t - \frac{1}{2} z_t^2 \quad (12)$$

where the prediction errors z_t are given by $y_t - \mathbf{X}_t' \mathbf{b}_{t|t-1}$. The loglikelihood can be used to estimate any unknown parameters. Another set of recursions may be used to obtain estimates of the states and the covariances of these estimates conditional not only on currently available data but on all the data. These are the smoothed estimates of the states.

Usually, $\Sigma_{\epsilon,t}$ and \mathbf{A}_t are assumed to change deterministically with time and are therefore known. However this is not necessary. As long as they depend only on the information set \mathcal{I}_{t-1} , a conditionally Gaussian model can be constructed. The above Kalman filter equations remain unchanged. Further, smoothing and any other standard Kalman filter related evaluations can be carried out. The loglikelihood can be maximised to estimate any hyperparameters. This assumes that r is known. But this is easy to solve. A grid can be constructed for possible values of r . Then, the model can be estimated for every point of the grid and the point giving the maximum likelihood over the grid can be adopted as the estimate for the threshold parameter.

The consistency of the threshold parameter r estimate is discussed and proved in Kapetanios and Tzavalis (2004). An integral part of this discussion has to be the stationarity of the data y_t generated by the model. As the model stands it is not stationary. To see that simply note that the variance of the process β_t is $O(T)$. In fact, the process β is a random walk process. Even if β_t were a stationary process, further strict conditions would need to be satisfied to get covariance stationarity of y_t in the case where $x_t = y_{t-1}$. In the case where x_t is a strictly exogenous stationary process then stationarity of β_t would suffice for stationarity of y_t . Stationarity of y_t is a desirable property irrespective of the proof of consistency of the threshold parameter estimate. There are a number of possible modifications that can be

imposed on β_t to make it stationary. An easy first condition is that

$$\beta_t = \beta_{t-1} + I(|\beta_{t-1}^*| < \beta)I(|\epsilon_{1,t-1}| > r)\epsilon_{2,t-1} \quad (13)$$

where $\beta_{t-1}^* = \beta_{t-1} + I(|\epsilon_{1,t-1}| > r)\epsilon_{2,t-1}$. For obvious reasons we set $\beta = 1$. Of course, our previous comments on conditionally Gaussian models apply since a model with such a specification for β_t would not be conditionally Gaussian and would therefore require filtering using a nonlinear Kalman filter which, as we discussed earlier, is difficult to implement. The alternative we have suggested would then become

$$\beta_t = \beta_{t-1} + I(|\hat{\beta}_{t-1|t-1}^*| < \beta)I(|\hat{\epsilon}_{1,t-1|t-1}| > r)\epsilon_{2,t-1} \quad (14)$$

$$\hat{\beta}_{t-1|t-1}^* = \hat{\beta}_{t-1|t-1} + I(|\hat{\epsilon}_{1,t-1|t-1}| > r)\hat{\epsilon}_{2,t-1|t-1} \quad (15)$$

This specification makes the process β_t stationary and bounded between -1 and 1. Hence, it makes the process y_t stationary. Further restrictions could be placed on the process so that if the bound β is exceeded the process returns to some prespecified level. We do not advocate a particular mechanism for making the process β_t stationary. We simply wish to indicate that there exist specifications which give both a stationary β_t process and a conditionally Gaussian state space model amenable to analysis via the Kalman filter. The exact specification of the process may be left to the empirical researcher depending on his priors on the particular issue at hand.

3 A Monte Carlo Study

In this subsection, we carry out a small scale Monte Carlo study to investigate the performance of our model to track structural breaks adequately, especially in samples where either the number of observations or the number of breaks is relatively small. It is reasonable to expect that the state variable which drives the breaks is hard to carry inference on given that there are only a few observations which will contain information about the occurrence of the breaks.

We abstract from parameter estimation and concentrate on the estimation of the state variables assuming that the parameters of the model are known. We do this for three reasons. Firstly, we know that threshold estimation is difficult even for standard threshold models (see e.g. Kapetanios (2000)). In particular, threshold estimation is slow to improve when extra

observations are added to the sample, despite the superconsistency result of Chan (1993). Secondly, if the Kalman filter is shown to perform well we can reasonably expect that the performance of the threshold estimator will be similar to the case of standard threshold models. Finally, by the nature of the model, the choice of the threshold has to be restricted to extreme values of the threshold variable, as our model practically dictates the choice of the threshold value.

We simulate the model given by (1), with (14), (15), (3) and (4). For simplicity of simulation, we use $\epsilon_{1,t-1}$ rather than $\hat{\epsilon}_{1,t-1|t-1}$ in the indicator function in (14) and (15). However, for estimation $\hat{\epsilon}_{1,t-1|t-1}$ is used.

We set the parameters of the model at the following values: $\rho_1 = \rho_2 = 0.5$, $\sigma_{\eta_1}^2 = 1$ and $\sigma_{\eta_2}^2 = 0.01$. For the threshold parameter, we consider two cases: $r = 2.5$ and $r = 2.93$. In the first case, our model implies that the event $\{|\epsilon_{1,t-1}| > r\}$ occurs quite frequently, approximately every 33 periods, while in the second occurs more rarely, approximately every 92 periods. For the first case, we consider large and small samples of $T = 1000$ and $T = 200$ observations, respectively, while for the second we consider only large samples. The reason that we do not consider small samples for the second case is that the number of breaks is just too small to be picked up in small samples, as $T = 200$. This implies that we conduct in total three set of experiments. These can indicate the size of sample and the number of breaks per sample for which the performance can be considered as satisfactory.

We estimate the state β_t conditional on information available at time $t - 1$, known as one-step ahead forecast of β_t , on information at time t , known as filter estimate and, finally, on all the available information up to time T , known as smoothed estimate. The last two information sets upon which the estimates of β_t are derived are defined as $\mathcal{I}_t = (y_t, y_{t-1}, \dots, y_0)$ and $\mathcal{I}_T = (y_T, \dots, y_t, \dots, y_0)$, respectively. To evaluate the performance of our model, we calculate the average correlation coefficients between the above estimates of β_t and the true state over 500 replications and the three experiments. These estimates are given in Table 1. The estimates reveal that our model can track the breaks in β_t , adequately enough. Following the discussion in Kapetanios and Tzavalis (2004), at this point we would like to note that, intuitively, the performance of the model will improve if we allow for a higher number of breaks per T .

To obtain a better feel of the ability of our model to track the breaks in β_t sufficiently, in Figures 1-3 we present the smoothed estimates of β_t against its true values, together with their confidence intervals calculated at the 90% significance level. In every figure we present three panels. Each panel presents the estimated values of β_t which correspond to the 25%, 50% and 75% quantiles of the empirical distribution of the correlations between the true state and the smoothed estimate of the state. Inspection of this figure indicates that our model can track the true breaks reasonably well. The smoothed estimates closely follow the true values of β_t , across all time-points of the sample, with a high degree of precision.

4 Empirical Application

In this section, we present an empirical application of our model with the aim of examining if there are structural shifts in the mean reversion coefficient of major macroeconomic series. In particular, we analyse a selected number of macroeconomic series from the diffusion index dataset of Stock and Watson (2002) spanning a long interval of monthly observations from 1959M1 to 1998M12 [see Data Appendix for details on the series used]. The series that have been chosen represent all groups of variables representing real activity, money, price, survey and financial variables. All series used in our empirical analysis have been transformed as described in Stock and Watson (2002) to achieve stationarity according to standard unit root tests. Further, they are demeaned and normalised to unit variance.

For the above data set, we estimated our model (1), with (14), (15), (3) and (4), where the initial value of β_t , denoted β_0 , was estimated via maximum likelihood along with the other parameters of the model ρ_1 , ρ_2 , $\sigma_{\epsilon_1}^2$ and $\sigma_{\epsilon_2}^2$. The variance of $\beta_{0|0}$, at time-point $t = 0$, needed as initial value for the Kalman filter is set to 0, while the variances of $\epsilon_{1,0|0}$ and $\epsilon_{2,0|0}$ are set to $\sigma_{\epsilon_1}^2$ and $\sigma_{\epsilon_2}^2$ respectively. The lag order p of the autoregressive processes for $\epsilon_{1,t}$ and $\epsilon_{2,t}$ was set to one. As we will see from the estimation results there is little evidence of any serial correlation in $\epsilon_{1,t}$. To verify this result we conduct misspecification tests for serial correlation in the estimates of $\epsilon_{1,t}$.

In the estimation procedure of our model, an important question relates to the choice of r . As discussed earlier, this can be estimated through a grid search procedure. However, this estimation problem is a difficult one. Threshold estimation is at best difficult (see e.g.,

Kapetanios (2000)). As by the nature of our model, the threshold parameter should be large, this problem may be proved much more difficult. Experimentation with our dataset has suggested that a grid search is likely to cause convergence problems to the maximum likelihood optimisation routine for a number of series of our data set.

Thus, to provide a unified treatment across all series of our data set, we fixed the threshold parameter in the following way.³ First, we estimated a standard AR(1) model for each series and, given the estimate of the autoregression coefficient denoted ϕ , we normalise each series such that it has variance $\frac{1}{1-\phi^2}$. This implies that the residuals of the normalised series have variance equal to 1. Based on these normalised series, in the next step we set the threshold parameter in such a way that if the residual was normally distributed a break would occur 1% of the time or every 100 observations. For monthly data this implies a break every about 8 years. The threshold parameter which corresponds to this is 2.55 since a standard normal variate will exceed this value, in absolute value, with probability which is approximately 1%. As the residuals are not normally distributed, we present the proportion of time that we would have breaks for each series, in Table 2. However, it is important to note that the event of $|\epsilon_{1,t}|$ exceeding r does not necessarily imply an observed break. The Kalman filter provides the optimal estimate of the break size which may very well be zero, implying essentially that although the trigger event has occurred no break can be justified from the data.

In Table 3, we report estimates for the parameters of our model, together with their standard errors reported in parentheses. The stars in parentheses indicate a number of instances where the variance of ρ_2 and $\sigma_{\epsilon_2}^2$ cannot be obtained, as the Hessian of the likelihood is close to singularity. This implies that these two parameters which specify the properties of the break magnitudes are difficult to identify. This is reasonable given that the data only provide information on the break sizes when breaks occur. The results of the table necessitate a number of comments. Firstly, we see that in a majority of series the estimate of ρ_1 is not significantly different from zero. It appears that the AR(1) specification is sufficient to account for all serial correlation in the data. The variance of $\epsilon_{1,t}$ is close to one as expected given the specification of the model. The variance of $\epsilon_{2,t}$ may appear to be small (which is expected) but the standard errors suggest that the null hypothesis of it being equal to zero cannot be rejected. However, this conclusion is wrong. Under the null hypothesis that

³This approach was motivated by work on outliers, see Kapetanios and Tzavalis (2004) and Balke and Fomby (1993).

parameter will lie on the boundary of its space as it cannot be negative. One cannot use standard testing procedures to carry out this test.

For the cases where the hypothesis $\rho_1 = 0$ cannot be rejected we note that $\epsilon_{1,t} = \eta_{1,t}$ which implies that breaks in β_t are driven by white noise innovations to the level of y_t . Note that this happens for series related to changes in business conditions or monetary regime changes, such as CPI (consumer price index), IP (industrial production), PCE (personal consumption) and MS2 (M2 money stock).

To see if the results of Table 3 are sensitive to a potential misspecification of our model, such as serial correlation or other types of neglected nonlinearities, we report, in Table 4, probability values of misspecification test statistics for serial correlation (denoted SC) (LM test for up to a fourth lag) and nonlinearity (denoted NONL) in the estimates of $\epsilon_{1,t}$, obtained by our model. The nonlinearity test is that suggested by Teräsvirta, Lin, and Granger (1993) and we use one lag of $\epsilon_{1,t}$. The results of this table support the specification of our model, for most of the series. The only exceptions are with the CPI, MH, I, NAPM2, PCEPR2 and AHE series for which the NONL test rejects the null hypothesis of linearity at a significance level of 5%.

To track the changes of the mean reversion coefficient, over the sample, in Figures 4-6 we graphically present both filter and smoothed estimates of the state variable β_t . The first conclusion that can be drawn by inspecting the plots is that there are substantial changes in β_t , across the sample. For most of the series whose mean reversion coefficient changes was found to be determined by large innovations, i.e. $\rho_1 = 0$, the movements in β_t are cyclical and seem to follow changes in business conditions [see IP, PCE, MW], or monetary regime changes [see CPI, MS2, MS, IR, SPR] accruing after the changes of the Federal Reserve's money markets operating procedures [see Tzavalis and Wickens (1996)]. In particular the period in the late 70's and early 80's seems to be a particularly volatile period for the persistence of the set of series we consider. CPI, MH, NAPM1, FER, IR1, BY, SPR, MS2, MS3, DIR, PPI, PCEPR1, PCEPR2, and AHE are some of the series that underwent a considerable change in their mean reversion coefficient in that period.

5 Conclusion

Structural breaks have received plenty of attention in the econometric literature in the last couple of decades. However, most of the work has concentrated on detecting breaks and estimating model where the parameter changes are assumed to be fixed. In this paper we suggested a model of structural breaks in the mean reversion coefficient of autoregressive time series models whose both the timing and magnitude of changes are allowed to be stochastic in nature. We assume that these changes are driven by changes related to the levels of the series themselves, or the series innovations.

Monte Carlo results show that the performance of our model is adequate in tracking true structural break type of changes in the mean aversion coefficient of the series.

As an empirical illustration, we apply our model to a number of US macroeconomic series with the aim to investigate if there are apparent changes in the mean reversion coefficient of the series. The results of this exercise have shown that there is a substantial number of series whose mean reversion coefficient seems to considerably change cyclically, over time. These changes seems to be driven by large innovations in the series which can be attributed to changes in business conditions or monetary regime changes. A particular period of break activity seems to be the late 70's and early 80's possibly related to the drastic changes in terms of monetary policy that took place in the US during the period.

Table 1: Monte Carlo results

(r, T)	(2.5, 200)	(2.5, 1000)	(2.93, 1000)
nb	6	30	11
Corr. Coeff. (Filter)	0.209	0.648	0.441
Corr. Coeff. (Forecast)	0.203	0.643	0.437
Corr. Coeff. (Smoothed)	0.364	0.794	0.641

Notes: nb denotes the number of breaks per T

Table 2: Proportion of time $|\epsilon_{1,t}| > 2.55$

Variable	$Prob(\epsilon_{1,t} > 2.55)$
CPI	0.029
IP	0.027
IP2	0.029
MW	0.019
PCE	0.017
PCE2	0.021
PCE3	0.029
HA	0.019
MH	0.023
I	0.025
NAPM1	0.021
NAPM2	0.031
NO	0.031
MFG	0.073
FER	0.031
IR	0.038
BY	0.042
SPR	0.031
MS2	0.036
MS3	0.040
MS	0.029
MB	0.034
DIR	0.040
PPI	0.034
PCEPR1	0.130
PCEPR2	0.170
AHE	0.197

Table 3: Parameter Estimates for the Model

Variable	$\sigma_{\epsilon_2}^2$	$\sigma_{\epsilon_1}^2$	ρ_1	ρ_2	β_0
CPI	0.03 _(*)	0.95 _(0.06)	-0.11 _(0.07)	0.47 _(*)	-0.33 _(0.16)
IP	0.02 _(*)	0.96 _(0.06)	-0.10 _(0.08)	0.39 _(*)	0.70 _(0.14)
IP2	0.03 _(0.03)	0.91 _(0.06)	-0.39 _(0.06)	0.00 _(0.07)	0.31 _(0.12)
MW	0.04 _(*)	0.97 _(0.06)	-0.18 _(0.12)	0.48 _(*)	0.39 _(0.21)
PCE	0.03 _(0.06)	0.96 _(0.06)	-0.18 _(0.12)	0.29 _(2.09)	-0.17 _(0.17)
PCE2	0.03 _(0.03)	0.95 _(0.06)	-0.38 _(0.07)	0.00 _(0.06)	0.32 _(0.15)
PCE3	0.00 _(0.00)	0.99 _(0.06)	-0.14 _(0.09)	0.14 _(0.62)	-0.22 _(0.14)
HA	0.03 _(0.03)	0.82 _(0.05)	0.29 _(0.11)	-0.11 _(1.93)	0.09 _(0.14)
MH	0.18 _(0.14)	0.93 _(0.06)	-0.22 _(0.11)	0.02 _(0.06)	0.06 _(0.12)
I	0.09 _(0.08)	0.92 _(0.06)	-0.09 _(0.08)	0.02 _(0.10)	-0.18 _(0.13)
NAPM1	0.10 _(0.08)	0.94 _(0.06)	-0.07 _(0.13)	0.00 _(*)	0.05 _(0.05)
NAPM2	0.00 _(0.01)	0.97 _(0.06)	-0.23 _(0.07)	0.19 _(0.74)	-0.21 _(0.13)
NO	0.01 _(0.02)	0.96 _(0.06)	-0.14 _(0.09)	0.65 _(0.82)	-0.23 _(0.15)
MFG	0.02 _(0.00)	0.90 _(0.06)	-0.28 _(0.05)	0.08 _(*)	0.73 _(0.12)
FER	0.26 _(0.20)	0.92 _(0.06)	0.21 _(0.07)	-0.00 _(0.08)	0.12 _(0.11)
IR	0.24 _(0.16)	0.92 _(0.06)	0.15 _(0.07)	-0.00 _(*)	0.17 _(0.11)
BY	0.02 _(0.04)	0.93 _(0.06)	0.27 _(0.07)	-0.50 _(0.41)	0.32 _(0.13)
SPR	0.00 _(0.00)	0.98 _(0.06)	-0.15 _(0.07)	-0.21 _(1.76)	-0.24 _(0.11)
MS2	0.02 _(0.01)	0.96 _(0.06)	-0.10 _(0.07)	0.00 _(0.07)	-0.38 _(0.10)
MS3	0.02 _(0.02)	0.88 _(0.06)	-0.42 _(0.07)	-0.84 _(0.12)	-0.52 _(0.08)
MS	0.02 _(0.05)	0.92 _(0.06)	-0.22 _(0.07)	0.67 _(0.92)	-0.49 _(0.09)
MB	0.00 _(0.01)	0.95 _(0.06)	-0.17 _(0.07)	0.65 _(0.51)	-0.48 _(0.12)
DIR	0.02 _(0.02)	0.95 _(0.06)	-0.24 _(0.07)	0.00 _(*)	-0.43 _(0.10)
PPI	0.01 _(0.01)	0.95 _(0.06)	-0.26 _(0.07)	0.01 _(*)	-0.47 _(0.13)
PCEPR1	0.04 _(0.04)	0.93 _(0.06)	0.18 _(0.07)	0.55 _(0.51)	0.90 _(0.06)
PCEPR2	0.05 _(0.00)	0.88 _(0.06)	0.30 _(0.06)	0.00 _(*)	0.94 _(0.06)
AHE	0.02 _(0.01)	0.83 _(0.05)	0.42 _(0.06)	-0.00 _(*)	0.94 _(0.08)

Variable	SC Test	NONL Test
CPI	0.275	0.020
IP	0.999	0.477
IP2	0.998	0.154
MW	0.496	0.295
PCE	0.928	0.849
PCE2	0.975	0.955
PCE3	0.488	0.323
HA	0.995	0.621
MH	0.896	0.032
I	0.691	0.003
NAPM1	0.785	0.294
NAPM2	0.987	0.018
NO	0.256	0.522
MFG	0.614	0.201
FER	0.936	0.789
IR	0.813	0.523
BY	1.000	0.452
SPR	0.440	0.790
MS2	0.993	0.111
MS3	0.922	0.215
MS	0.774	0.008
MB	0.994	0.092
DIR	0.929	0.615
PPI	0.100	0.079
PCEPR1	0.876	0.255
PCEPR2	0.395	0.000
AHE	0.886	0.002

Data Appendix

The identities of the variables considered in this paper are given below.

1. CPI: CPI-U: ALL ITEMS (82-84=100,SA)
2. IP: INDUSTRIAL PRODUCTION: MANUFACTURING (1992=100,SA)
3. IP2: INDUSTRIAL PRODUCTION: NONDURABLE MANUFACTURING (1992=100,SA)

4. MW: MERCHANT WHOLESALERS:NONDURABLE GOODS (MIL OF CHAINED 1992 DOLLARS)(SA)
5. PCE: PERSONAL CONSUMPTION EXPEND (CHAINED)-TOTAL (BIL 92\$,SAAR)
6. PCE2: PERSONAL CONSUMPTION EXPEND (CHAINED)-TOTAL DURABLES (BIL 92\$,SAAR)
7. PCE3: PERSONAL CONSUMPTION EXPEND (CHAINED)-NONDURABLES (BIL 92\$,SAAR)
8. HA: HOUSING AUTHORIZED: TOTAL NEW PRIV HOUSING UNITS (THOUS.,SAAR)
9. MH: MOBILE HOMES: MANUFACTURERS' SHIPMENTS (THOUS.OF UNITS,SAAR)
10. I: INVENTORIES, BUSINESS DURABLES (MIL OF CHAINED 1992 DOLLARS, SA)
11. NAPM1: NAPM INVENTORIES INDEX (PERCENT)
12. NAPM2: NAPM VENDOR DELIVERIES INDEX (PERCENT)
13. NO: NEW ORDERS, NONDEFENSE CAPITAL GOODS, IN 1992 DOLLARS (BCI)
14. MFG: MFG NEW ORDERS:DURABLE GOODS INDUST WITH UNFILLED ORDERS(MIL\$,SA)

15. FER: FOREIGN EXCHANGE RATE: CANADA (CANADIAN \$ PER U.S.\$)
16. IR: INTEREST RATE: U.S.TREASURY CONST MATURITIES,5-YR.(% PER ANN,NSA)
17. BY: BOND YIELD: MOODY'S BAA CORPORATE (% PER ANNUM)
18. SPR: Spread FYBAAC - FYFF
19. MS2: MONEY STOCK:M2(M1+O'NITE RPS,EURO\$,G/P&B/D MMMFS&SAV&SM
TIME DEP(BIL\$,SA)
20. MS3: MONEY STOCK: M3(M2+LG TIME DEP,TERM RP'S&INST ONLY MMMFS)(BIL\$,SA)
21. MS: MONEY SUPPLY-M2 IN 1992 DOLLARS (BCI)
22. MB: MONETARY BASE, ADJ FOR RESERVE REQUIREMENT CHANGES(MIL\$,SA)
23. DIR: DEPOSITORY INST RESERVES:NONBORROW+EXT CR,ADJ RES REQ
CGS(MIL\$,SA)
24. PPI: PRODUCER PRICE INDEX: FINISHED GOODS (82=100,SA)
25. PCEPR1: PCE,IMPL PR DEFL:PCE (1987=100)
26. PCEPE2: PCE,IMPL PR DEFL:PCE; NONDURABLES (1987=100)
27. AHE: AVG HR EARNINGS OF CONSTR WKRS: CONSTRUCTION (\$,SA)

References

- BALKE, N. S., AND T. B. FOMBY (1993): “Large shocks, Small Shocks, and Economic Fluctuations: Outliers in Macroeconomic Time Series,” *Journal of Applied Econometrics*, 9, 181–200.
- CHAN, K. S. (1993): “Consistency and Limiting Distribution of the Least Squares Estimator of a Threshold Autoregressive Model,” *The Annals of Statistics*, 21(1), 520–533.
- CLEMENTS, M., AND D. HENDRY (1999): *Forecasting Non-stationary Economic Time Series*. MIT Press.
- ENGLE, R. F., AND A. D. SMITH (1999): “Stochastic Permanent Breaks,” *Review of Economics and Statistics*, 81(4), 553–573.
- HAMILTON, J. D. (1989): “A New Approach to the Economic Analysis of Time Series,” *Econometrica*, 57(2), 357–384.
- HARVEY, A. C. (1989): *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge University Press.
- KAPETANIOS, G. (2000): “Small sample properties of the conditional least squares estimator in SETAR models,” *Economics Letters*, 69(3), 267–276.
- KAPETANIOS, G., AND E. TZAVALIS (2004): “The Impact of Large Structural Shocks on Economic Relationships: Evidence from Oil Price Shocks,” *Queen Mary, University of London Working Paper no. 524*.
- LIN, C. F., AND T. TERÄSVIRTA (1994): “Testing the Constancy of Regression Parameters Against Continuous Structural Change,” *Journal of Econometrics*, 62, 211–228.
- STOCK, J. H., AND M. W. WATSON (2002): “Macroeconomic Forecasting Using Diffusion Indices,” *Journal of Business and Economic Statistics*, 20, 147–162.
- TERÄSVIRTA, T., C. F. LIN, AND C. W. J. GRANGER (1993): “Power of the Neural Network Linearity Test,” *Journal of Time Series Analysis*, 14, 209–220.
- TZAVALIS, E., AND M. WICKENS (1996): “Forecasting inflation from the term structure,” *Journal of Empirical Finance*, 3, 103–122.

Figure 1: $T = 1000$, $r = 2.5$

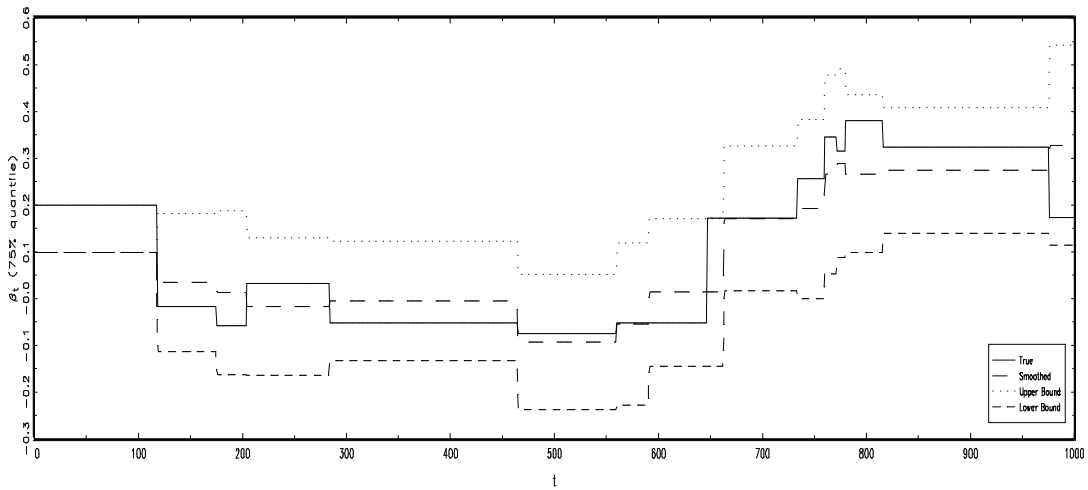
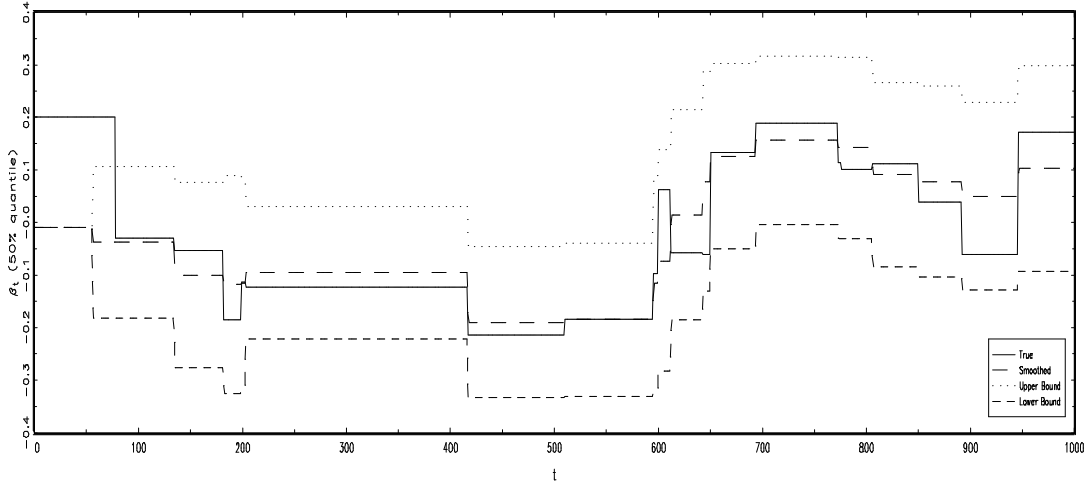
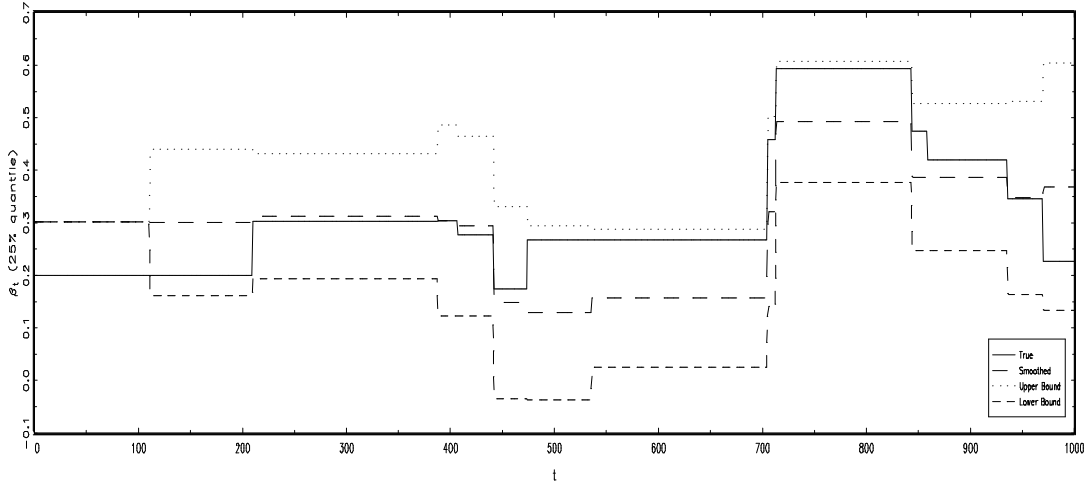


Figure 2: $T = 200, r = 2.5$

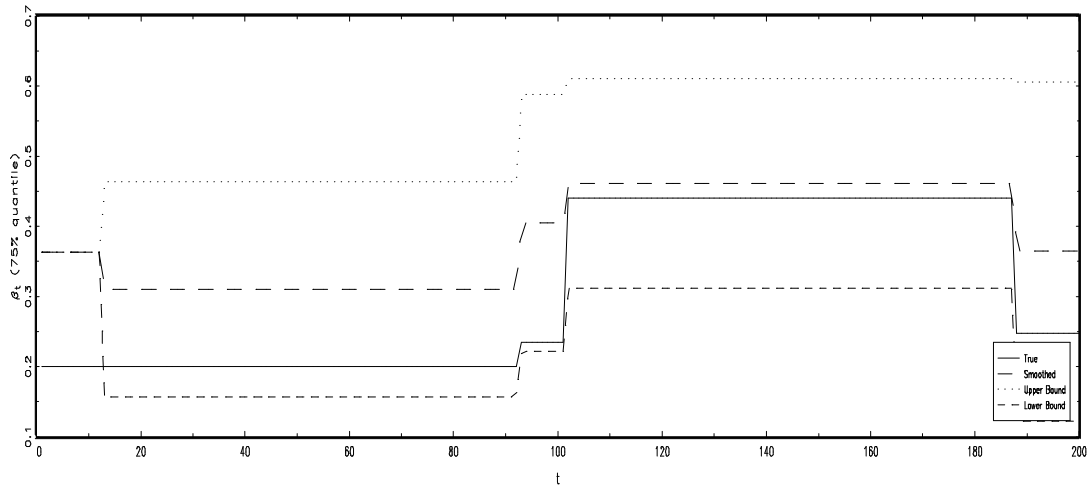
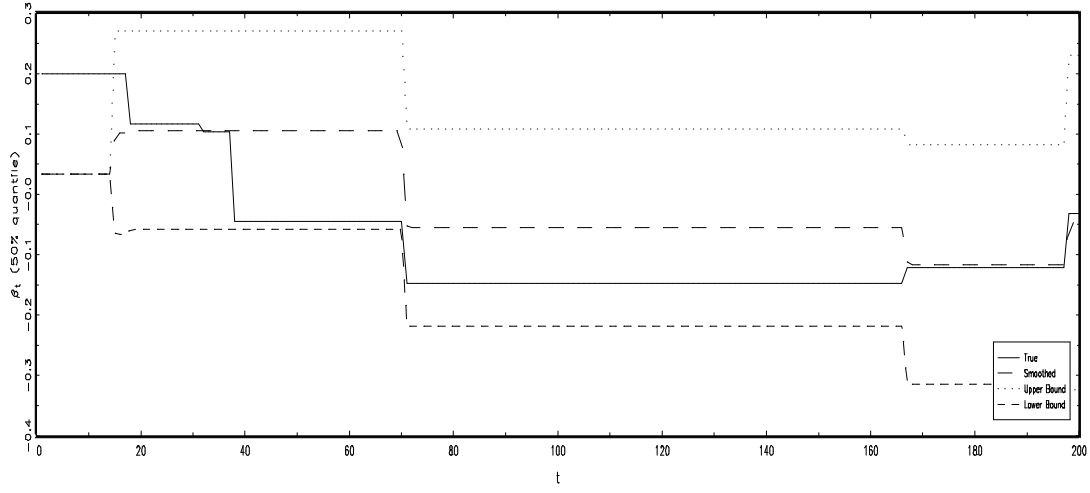
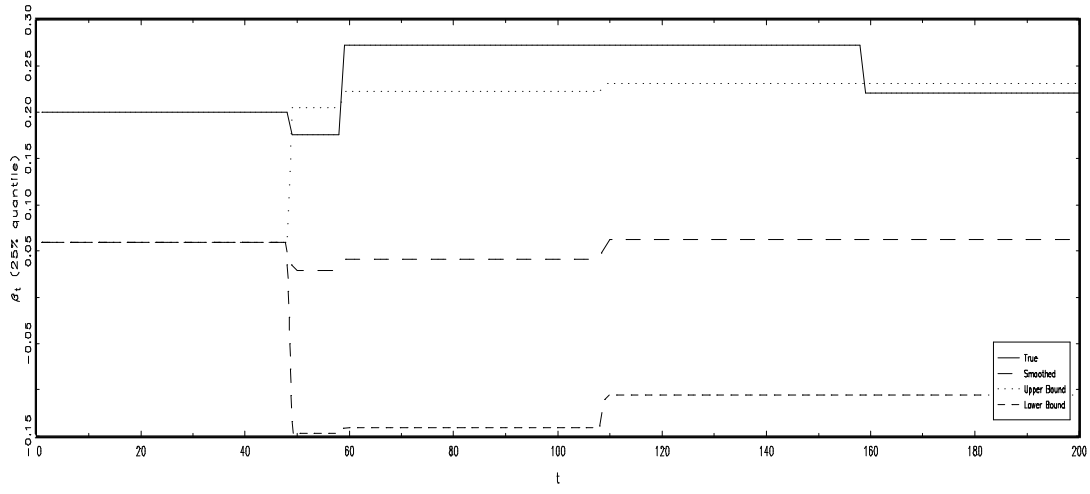


Figure 3: $T = 1000$, $r = 2.93$

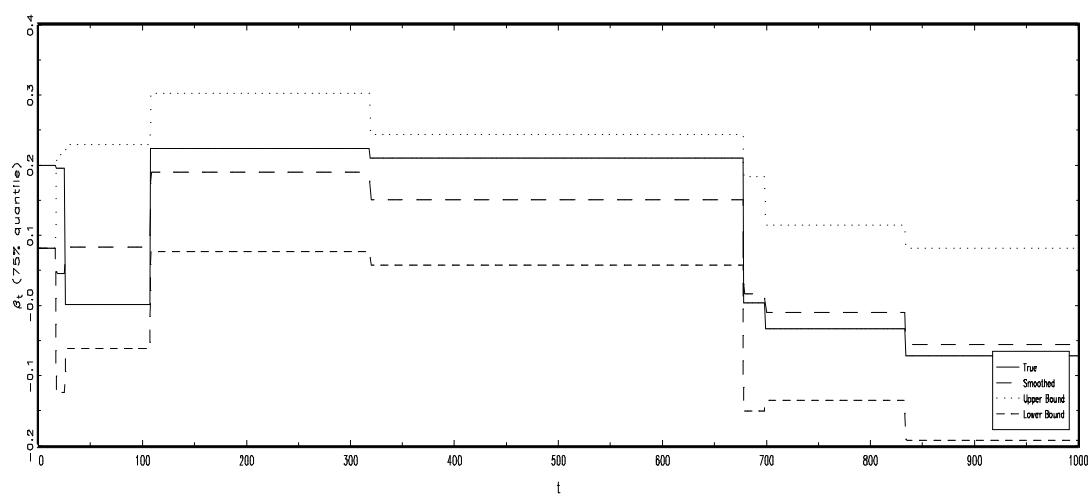
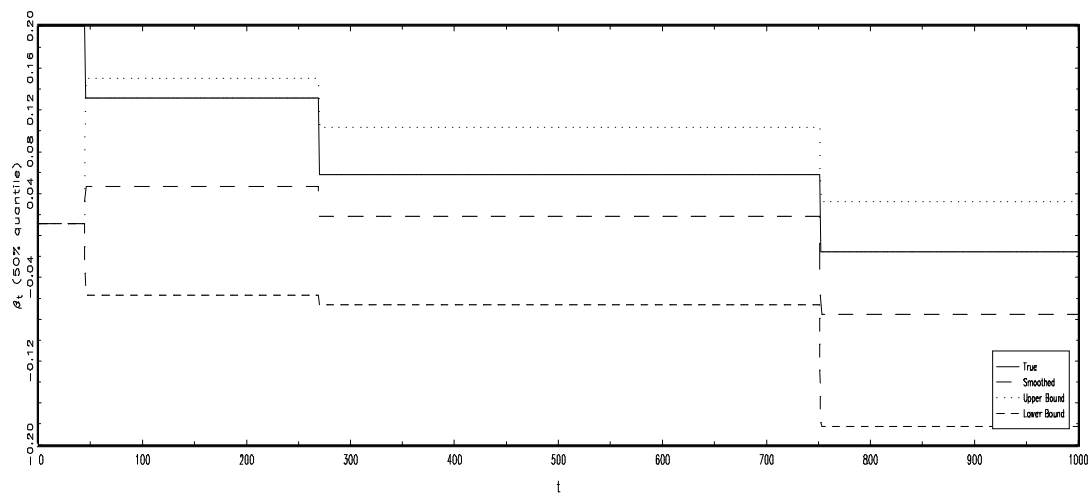
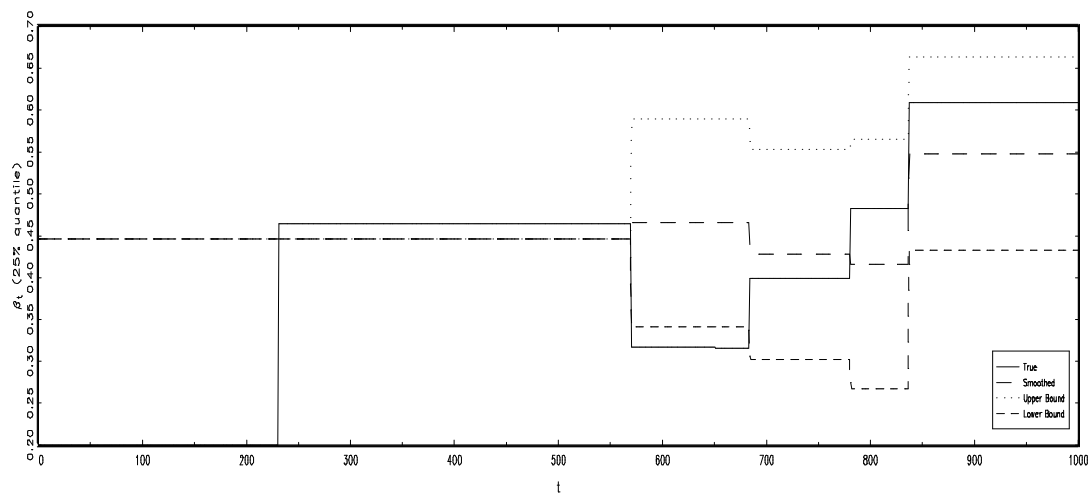


Figure 4: Series: CPI, IP, IP2, MW, PCE, PCE2, PCE3, HA, MH

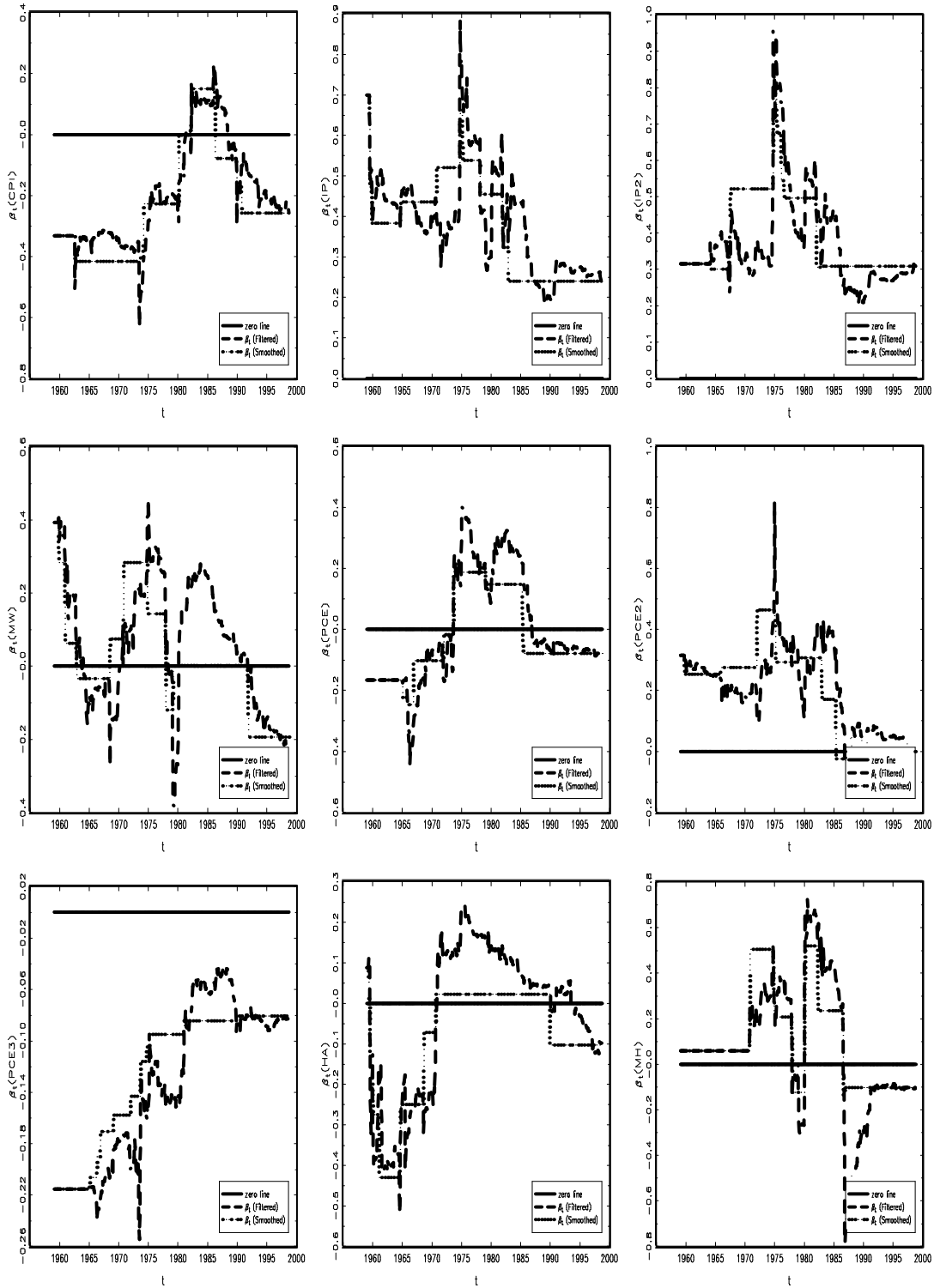


Figure 5: Series: I, NAPM1, NAPM2, NO, MFG, FER, IR, BY, SPR

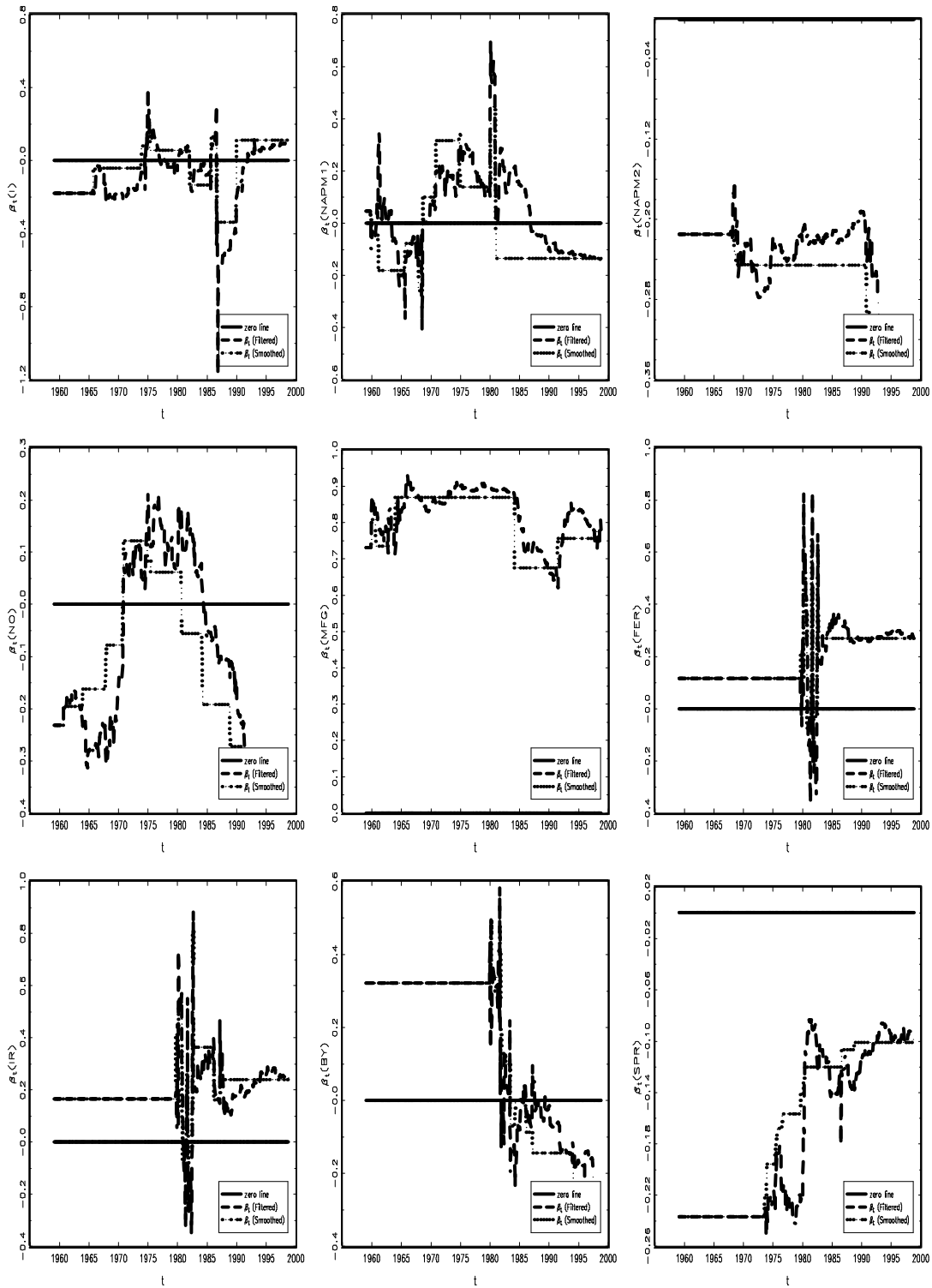
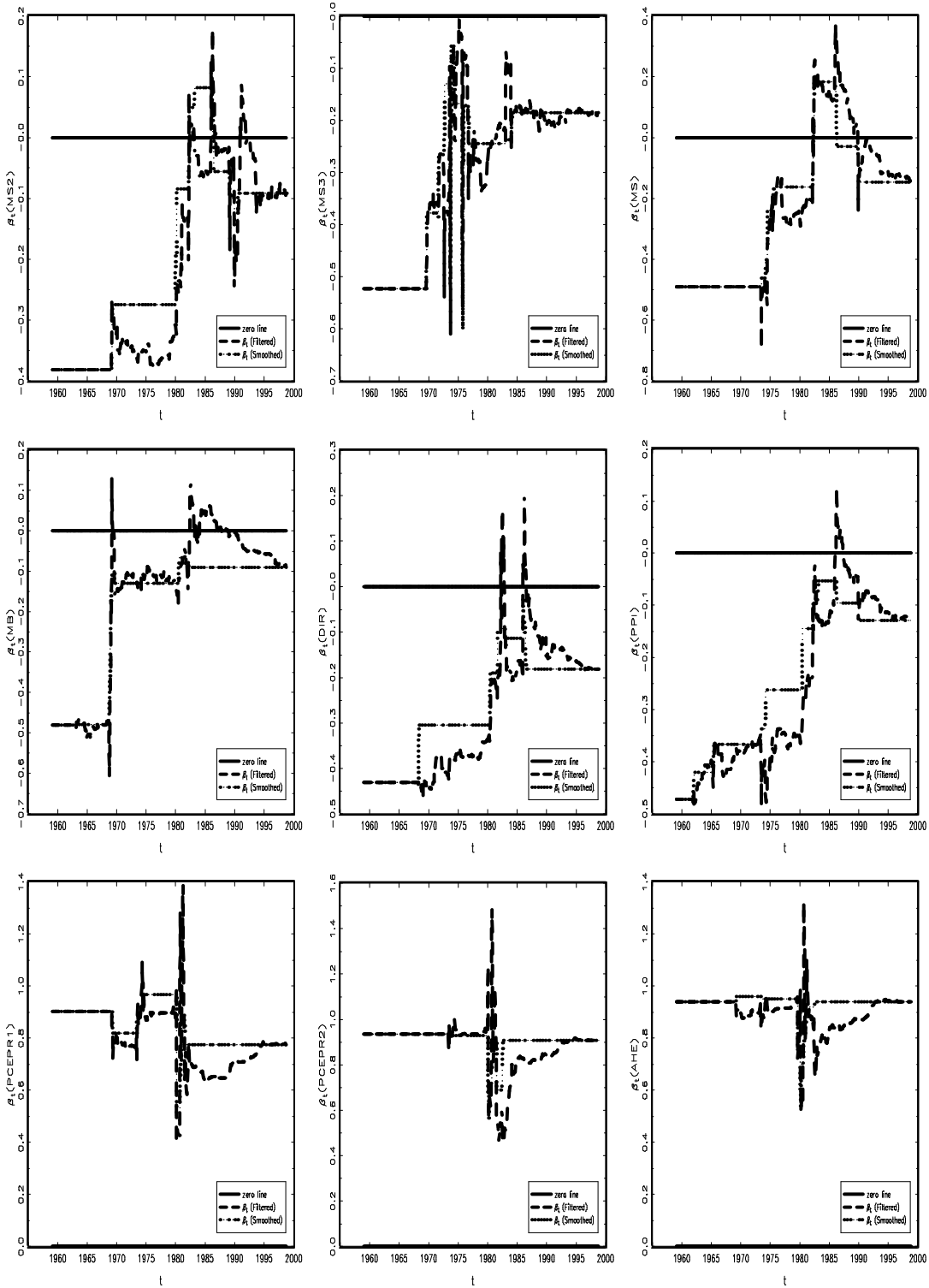


Figure 6: Series: MS2, MS3, MS, MB, DIR, PPI, PCEPR1, PCEPR2, AHE



**This working paper has been produced by
the Department of Economics at
Queen Mary, University of London**

**Copyright © 2005 George Kapetanios and Elias Tzavalis
All rights reserved**

**Department of Economics
Queen Mary, University of London
Mile End Road
London E1 4NS
Tel: +44 (0)20 7882 5096
Fax: +44 (0)20 8983 3580
Web: www.econ.qmul.ac.uk/papers/wp.htm**