MODELLING THE DISCRETE AND INFREQUENT OFFICIAL INTEREST RATE CHANGE IN THE UK[1]

Juan de Dios Tena and Edoardo Otranto[2]

Abstract

This paper is an empirical analysis of the manner in which official interest rates are determined by the Bank of England. We use a nonlinear framework that allow for the separate study of factors affecting the magnitude of positive and negative interest rate changes as well as their probabilities. Using this approach, new kinds of monetary shocks are defined and used to evaluate their impact on the UK economy. Among them, unanticipated negative interest rate changes are especially important. The model generalizes previous approaches in the literature and provides a rich methodology to understand central banks' decisions and their consequences.

Keywords: Markov Switching models, monetary shocks, impulse-response functions, monetary policy.

JEL Codes: C22, E32.

[2] Tena, Universidad de Concepcion, Departamento de Economia, Victoria 471 - Oficina 242 - Concepcion, Chile, email: juande@udec.cl and Universidad Carlos III, Departamento de Estadística, C/Madrid 126. 28903 Getafe (Madrid), Spain, e-mail: jtena@est-econ.uc3m.es; Otranto, Universita di Sassari, Dipartimento di Economia, Impresa e Regolamentazione, Via Torre Tonda, 34, 07100 Sassari, Italy, e-mail: eotranto@uniss.it.

1 Introduction

This paper proposes a new methodology for the empirical analysis of the manner in which official interest rates are determined by The Bank of England. In particular, the framework presented here allows for the separate study of factors affecting the magnitude of positive and negative interest rate changes as well as their probabilities. Recently, empirical literature has shown an increasing interest in modeling the discrete and infrequent changes of central bank rates. Here, we extend previous research allowing for a different characterization of the probability of positive and negative movements in a single model[1].

The proposed approach is particularly appealing as it captures the most relevant stylized facts inherent in the behaviour of official interest rates in the UK. Figure 1 shows the official interest rate series in monthly basis for the period 1974:01-2004:12. Although the period of analysis is comprised of different monetary regimes, at least, three stylized facts are always present. Firstly, this series does not seem to be influenced by the effect of stochastic noise but it is more the outcome of discrete and infrequent decisions by the Monetary Policy Committee. This is corroborated by the 220 occasions in which the level of the series did not change with respect to the previous month. Secondly, two consecutive interest rate movements in opposite direction is a very rare event. This only happens in 13 occasions. The final point refers to the different nature of positive and negative movements. We find that, in the considered period , there have been 99 negative compared to 53 positive changes. However, the average magnitude of negative changes have been of 61 basis points whereas the average magnitude of positive changes has been of 98 basis points.

[INSERT FIGURE 1 HERE]

Economic literature provides some explanations for these facts. For example, regarding the first point, Orphanides and Wilcox (2002) suggest that, when inflation is moderate, central banks do not take any action to reduce inflation. Instead, they wait for external shocks to obtain the desired reduction. For the second fact, Lowe and Ellis (1997) propose a psychological explanation suggesting that policy makers are likely to be embarrassed by reversals in the direction of interest rate changes as this can be interpreted by the public as a repudiation of previous actions. Regarding the third point, economic literature suggests the presence of asymmetric preferences with respect to output or (and) inflation by policy makers; see, for example, Cukierman and Gerlach (2003) and Ruge-Murcia (2003). This theory could be used to explain the different nature of positive and negative interest rate movements.

At the empirical level, a traditional methodology for modeling the dy-

---

[1]Gauss codes used in this analysis can be obtained from the authors upon request.

namics of official interest rate series is the use of conventional ordered logit and probit models; Eichengreen et al (1985) and Davutyan and Parke (1995) are two relevant examples. In these models, the magnitude of interest rate changes is conditioned to a set of fundamental economic variables. Hamilton and Jorda (2002) propose a framework denoted as autoregressive conditional hazard (ACH) models to separately analyze two separate decisions by policy makers: 1) whether or not to change interest rates; and 2) by how much. They specify a hazard rate associated to an interest rate change, for a given interval, as a function of the length of time between two previous interest rate changes and a set of economic variables. For the analysis of factors affecting the magnitude of change they use an ordered probit model.

In our particular context, two main drawbacks of this approach can be mentioned. Firstly, the ACH model does not differentiate between the probability law governing positive and negative changes. This is especially relevant for the UK case because, as we mentioned above, negative movements are more likely and of less magnitude than positive movements. A second drawback is due to the fact that the two questions related to the magnitude and probability of change are considered in the context of two different models that are specified and estimated separately. This is an implicit hypothesis of independence of changes; the use of a single model includes the structure of dependence among the changes.

Here, we propose a model that explains interest rate changes as a function of three fundamental economic variables: output gap and movements in inflation and exchange rate. The parameters of the function are allowed to change depending on a latent three state variable that drives negative, no movement and positive interest rate interventions respectively. In turn, the probability of being in each state is also allowed to depend on interest rate decisions in the previous period as well as in a set of economic variables. We denote this framework as the general probabilistic and magnitude (GPM henceforth) model for interest rate changes. The GPM model allows us for the analysis of the different factors affecting four fundamental decisions about interest rate movements: the magnitude of (1) positive and (2) negative interest rate changes; and the probability of (3) positive and (4) negative changes.

An additional advantage of this methodology is that specification and estimation of the GPM model is particularly simple as it shares some similar features with the Markov Switching model (MS hereafter) introduced by Hamilton (1989) and later extended by Filardo (1994) and Diebold et al. (1994) for the case where the transition probability matrix can change along the time.

A separate contribution of the paper is the use of GPM models to define 4 different types of monetary shocks. The first two types relates to the situation where the monetary authority increases and decreases official interest rates when no change was expected. The third and forth shocks are when

central bankers do not change official interest rates even though a positive and a negative change had been expected. We compute numerically the impact of these shocks on the UK economy and compare our results with other procedures in the literature to estimate responses to monetary shocks such as the standard VAR approach and the generalized impulse response function advocated by Koop et al (1996). Our methodology is proved to be useful as results indicate substantial differences in the impact of the four different shocks.

This paper is structured as follows. Next section provides a brief description of the different approaches in the literature to estimate the determinants of changes in official interest rate set by central bankers within a reaction function context. In the same section we present the GPM model. Estimation of the model for the UK and analyses of the main results are shown in Section 3. Section 4 explains the use of GPM to define different types of monetary shocks and indicates how to estimate impulse-response functions to these shocks. Section 5 analyzes results from this estimation. This section also compares the reactions proposed here with those obtained in more standard approaches. Conclusions are drawn in Section 6.

2 A GPM Model for Interest Rate Decisions

This section provides a brief description of the different approaches in the literature to estimate the determinants of changes in official interest rate set by central bankers within a reaction function context. Then, we present the GPM model that allows for asymmetries in both the magnitude and the probability of changes in interest rate.

Two big groups of methodologies can be mentioned for the estimation of reaction functions of central bankers. The first of them was motivated by the seminal paper by Taylor(1993). He showed that a simple reaction function that uses short-term interest rate as a policy instrument responding to movements in fundamental variables (inflation and output gap) follows closely the observed path of the US. Federal Funds Rate in the late 1980s and early 1990s. Following this paper, it has been a main concern in the more recent literature to specify and estimate simple policy rules that are also able to capture the smoothing nature of official interest rates and can be used for policy recommendation; see Clarida et al (2000) and Orphanides (2001) for some relevant examples. They suggest that short-term interest rate converge to the desired rate through a partial adjustment mechanism such as:

$$i_t = \rho i_{t-1} + (1-\rho)\left\{\alpha + \beta\pi_{t+1} + \gamma y_{t+1}\right\} \tag{1}$$

where i_t is the interest rate controlled by the monetary authority; and π_t and y_t are respectively the rate of inflation and cyclical output in period t.

A second approach relates to the characterization of the discrete and infrequent changes in bank rates. A typical example is based on the use of logit and probit models in which the dependent variable is referred to the magnitude of change of overnight interest rate (instead of interest rate levels) set by the monetary authority and assume that all the relevant conditioning variables are included in the study; see for example Eichengreen et al (1985) and Davutyan and Parke (1995). A clear advantage of this approach is that it is well suited for the nature of changes in interest rates set by central bankers.

However, as pointed by Hamilton and Jorda (2002), an important drawback of this methodology is the presence of a potentially significant serial correlation in the latent residuals. A solution to this problem can be found in the autoregressive conditional duration (ACD henceforth) model of Engle and Russel (1997) and Engle (2000). In the ACD model, previous length of time between events are taken to forecast future durations.

Starting from this consideration, Hamilton and Jorda (2002) proposed their autoregressive conditional hazard (ACH) model. In this case, their interest is not on the length of time between events but on the probability of an interest change tomorrow given the information up today. In their framework they study separately two types of decisions by central bankers: 1) the decision on whether or not to change interest rate; and 2) the magnitude of change. For the first one, they specify and estimate an ACH model in which the hazard rate for interest rate change in a given period depends on a vector of fundamental variables. The second decision, on the other hand, is evaluated using an ordered probit model for the magnitude of change. However, an important problem to apply the ACH approach to the analysis of interest rate changes in the UK arises: while the observation of changes seems to indicate that the probability of positive is clearly different to negative movements, ACH models make no distinction between the probability law goberning positive and negative interest rate movements.

Here, we propose a simple model in which probability of positive and negative interest rate movements are treated differently. Moreover, in contrast to the ACH model that studies the probability and magnitude of change in two separate models, the GPM model integrates these two decisions into a single framework.

For clarity of exposition, given the nature of interest rate movements, a good description of the series is given by the following equation

$$
i_t = \begin{cases}
i_{t-1} + \alpha_1 + \gamma_1^1 \Delta \pi_t + \gamma_2^1 y_t + \gamma_3^1 \Delta e_t + \sigma_1 u_t & if \quad s_t = 1 \\
i_{t-1} & if \quad s_t = 2 \\
i_{t-1} + \alpha_2 + \gamma_1^2 \Delta \pi_t + \gamma_2^2 y_t + \gamma_3^2 \Delta e_t + \sigma_3 u_t & if \quad s_t = 3
\end{cases}, \quad (2)
$$

where Δ indicates the difference operator; e_t denotes the foreign exchange rate included to account for open market considerations in the interest rate

rule[2]; u_t is the error component; and s_t is a variable that takes values 1,2, and 3 when there is a negative change, no change, and a positive change in interest rate at time t respectively.

Two points must be mentioned at this stage. First, the inclusion of $\Delta\pi_t$ instead of π_t must be empirically justified on econometric grounds and indicates that the monetary authority cares about the stabilization and not the level of inflation. Second, model (2) can be easily generalized to allow for lagged explanatory variables. These two points will be more specifically outlined in the following section.

Notice that a main feature of model (2) is that positive and negative interest rate changes are affected differently by fundamental variables. However, although specification (2) is a plausible representation of the data, if the interest of the analyst is to study what factors affect interest rate movements, then it is necessary to define the probability law governing these changes. Therefore, for our purpose it is useful to consider the process to be influenced by an unobserved latent variable, s_t^*, that drives interest rate movements. This variable has two possible interpretations. First, it can be an indicator on how prone central bankers are to perform positive or negative interest rate changes. Also, it can be used to evaluate the probability of positive and negative interest rate changes at period t given the information at $t-1$. In other words, in our model we are supposing that we know the variable s_t^* at time $t-1$, but not at time t. In other words the variable s_t and s_t^* are referred to the same phenomenon, but their interpretation is different: s_t is observable, whereas s_t^* is latent. In practice, predicting s_t^* we try to forecast the state s_t before that it was observable.

Then, the GPM model can be defined as:

$$
i_t = \begin{cases} i_{t-1} + c_1 + \beta_1^1 \Delta\pi_t + \beta_2^1 y_t + \beta_3^1 \Delta e_t + \sigma_1 \varepsilon_t & \text{if } s_t^* = 1 \\ i_{t-1} & \text{if } s_t^* = 2 \\ i_{t-1} + c_2 + \beta_1^2 \Delta\pi_t + \beta_2^2 y_t + \beta_3^2 \Delta e_t + \sigma_3 \varepsilon_t & \text{if } s_t^* = 3 \end{cases} , \qquad (3)
$$

where ε_t is a standard Normal disturbance and the rest of components has been previously defined.[3]

To define the probability of being in one of the three states, let $\{p_{ij}\}_{i,j=1,2,3}$ denotes the transition probability $P(s_t^* = j / s_{t-1}^* = i)$. We assume that s_t^* follows a three-state Markov change with transition matrix:

---

[2]In this case e_t is defined as the number of US dollars to one British pound. Hence, positive values of Δe_t means a bigger value of British pound with respect to US dollar.

[3]The exposition of this model is analogous to the model with changes in regime presented by Hamilton (1994), Chapter 22. He observes a different evolution of the volume of dollar-denominated accounts held in Mexican banks before and after 1982. Then, he argues that although it is possible to estimate a different model for these two periods, the change in regime should not be considered as the outcome of a perfectly foreseeable deterministic event but a random variable.

$$P = \begin{bmatrix} p_{11} & p_{21} & p_{31} \\ p_{12} & p_{22} & p_{32} \\ p_{13} & p_{23} & p_{33} \end{bmatrix} \tag{4}$$

where $p_{i3} = 1 - \sum_{j=1}^{2} p_{ij}$.

The described framework in (3)-(4) is a standard MS model where the only peculiarity is the absence of a stochastic error term in state $s_t^* = 2$.[4] MS models were initially introduced by Hamilton (1989) and later extended by Filardo (1994) and Diebold et al. (1994) for the case where the transition probability matrix can change along the time, depending on some observed variables $\mathbf{z}_t$. In this case, we need some specification for the probabilities $p_{ij,t}$. Filardo (1994) uses the following logistic functions for the 2-states case:

$$p_{ii,t} = \frac{\exp(\phi_i + \mathbf{z}_t \vartheta_i)}{1 + \exp(\phi_i + \mathbf{z}_t \vartheta_i)} \tag{5}$$

where ϕ_i and ϑ_i $(i = 1, 2)$ are unknown parameters. Of course, the choice of $\mathbf{z}_t$ is crucial and implies computational efforts and complications in the likelihood function. Filardo (1998) indicates the condition to select $\mathbf{z}_t$ to avoid estimation problems; in general a sufficient condition to justify the use of Hamilton filter (Hamilton, 1990) in a time varying transition probability Markov Switching (TVTP-MS hereafter) model to develop the maximum likelihood estimation, is that the elements in $\mathbf{z}_t$ are conditionally uncorrelated with s_t. This is a plausible assumption according to our empirical analysis in the following section.

Here, we also extend the specification of Filardo to a 3-states case, using a multinomial logit:

$$p_{ij,t} = \frac{\exp(\phi_{ij} + \mathbf{z}_t \vartheta_{ij})}{1 + \sum_{h=1}^{2} exp(\phi_{ih} + \mathbf{z}_t \vartheta_{ih})} \tag{6}$$

Details about estimation of the model are confined to Appendix 2.

The described framework allows for the separate analysis of four different decisions by central bankers: the probability of (1) an increase and (2) a decrease in interest rate; the magnitude of (3) a positive and (4) a negative interest rate change. The remaining of this paper shows the estimation of the proposed model for the UK and, based on this, we study the transmission of monetary shocks to fundamental interest rate shocks.

---

[4]Notice that an intuitive methodology to estimate the transition probabilities could be just counting the proportion of events in which the observable state, s_t, is i at time $t - 1$ and j at time t. However, in our particular case, probabilities obtained in this way are very similar to those obtained using the approach advocated by Hamilton (1989).

3   Estimation of a GPM Model for the UK

For the interest rate estimation, we use UK data on inflation, the output gap, nominal exchange rates and the overnight interest rate set by the Bank of England. Here, inflation is measured by the seasonal difference of the price of consumer goods and services in logs. Output is measured by the seasonally adjusted Industrial Production Index (IPI)[5]. The natural output level is the Hodrick-Prescott (HP) trend of the logged IPI. Then, the output gap is computed as the difference between the logged IPI and its HP trend. Nominal exchange rate is measured as first differences of the log of US/UK foreign exchange rate All series are in monthly frequencies and they are defined for the sample period 1974:01 to 2004:12.

For the sake of illustration, we first show the estimation of a Taylor rule in the spirit of Clarida et al (2000) and Orphanides (2001) similar to the one in (1). Given our set of data, we estimate this equation by Hansen's (1982) Generalized Method of Moments (GMM) using as instruments twelve lags of the policy instrument and the policy targets (output and inflation). The following expression reports the main results (standard errors are between brackets):

$$i_t = \underset{(0.01)}{0.97} i_{t-1} + (1 - \underset{(0.01)}{0.97}) \left\{ \underset{(1.10)}{3.49} + \underset{(13.39)}{59.88} \pi_{t+1} + \underset{(49.69)}{76.52} y_{t+1} \right\} \qquad (7)$$

with $R^2 = 0.970$.

It is important to notice that the estimation of the autoregressive parameter in this expression is very close to 1. In fact, this could be considered an unbalanced regression as i_t is clearly integrated of order one ($I(1)$ henceforth); π_t is on the borderline $I(1)/I(0)$; and y_t is $I(0)$[6]. Important problems related to the interpretation and statistical inference in this type of regressions can be found in Banerjee et al (1993).

Therefore, it seems reasonable to estimate an interest rate equation where the dependent variable is the magnitude of change instead of levels of interest rates. Initially, we estimate this simple linear model

$$\Delta i_t = \underset{(0.030)}{-0.022} + \underset{(5.375)}{5.552} \Delta \pi_t + \underset{(1.634)}{3.199} y_t - \underset{(1.249)}{5.163} e_t + 0.584 u_t \qquad (8)$$

with a poor fitting ($R^2 = 0.047$).

---

[5]There is considerable evidence in the literature that, in the UK, short-term movements in manufacturing output are strongly correlated with movements in other output components; see Salazar et al. (1997) for a discussion.

[6]The integration order can be inferred from the observation of figures and correlograms of the series. Also, we developed simple ADF tests for these series finding that the value of the test is -2.93 for i_t; -6.56 for y_t; and -3.22 for π_t. MacKinnon (1991) critical values for rejection of hypothesis of a unit root are respectively -3.99, -3.42 and -3.13 at the 1%, 5% and 10% significant level.

Three points must be mentioned at this stage. First, although, in equation (8), the parameter associated to $\Delta\pi_t$ is not significantly different from zero at the 0.05 level, we include this variable to be consistent with the common practice of considering the rate of inflation to estimate interest rate equations in the previous literature; see, inter alia, Eichengreen et al (1985) and Davutyan and Parke (1995). The equation is specified with $\Delta\pi_t$ instead of π_t because the estimated coefficient associated to π_t is clearly less significant than the coefficient associated to $\Delta\pi_t$. Besides, variable $\Delta\pi_t$ will play a fundamental role in the estimation of the GPM model as we will prove later. Thus, from specification (8), we interpret that central bankers move interest rates to stabilize inflation and not to control its level. Second, nominal exchange rate are included to account for open market consideration in the interest rate rule. In fact, the estimated parameter associated to this variable is significantly different from zero at the conventional levels. The third point relates to adding lagged components in expression (8). We considered this possibility but, in our case, the BIC criterion (Schwarz, 1978) suggests to adopt a model without lagged variables.

In general, the sign of the estimated parameters in this equation is consistent with economic insight. More specifically, positive changes in cyclical output and first differences in inflation lead an increase of i_t and a bigger value of the currency pushes i_t down. A potential drawback in this interpretation is that one may argue that dependent and independent variables in (8) are determined simultaneously. To deal with this issue, we test the null hypothesis of exogeneity with a Hausman (1978) test accepting the null hypothesis for all the explanatory variables at the conventional significance levels. This result is consistent with economic insight as one can assume that central bankers have contemporaneous monthly information on exchange rate, inflation and output gap; see, for example, Bernanke and Blinder (1992) and Wright (2002).

The presence of regimes in equation (8) is formally tested by using the nonparametric Bayesian approach of Otranto and Gallo (2002)[7]. In this procedure, we have used three different priors for the number of states, represented in Table 1 by the parameter A; higher A corresponds higher prior probability on a larger number of regimes. We can note that all the cases favor the presence of 3 regimes, according with the hypothesis of model (2); in addition the cases relative to 1 and 2 regimes are excluded.

[INSERT TABLE 1]

The first column in Table 2 reports the estimates of model (3) with fixed transition probabilities. To obtain the final specification we eliminate variables whose estimated parameters are not significantly different from zero at the 5% significant level in a step-wise procedure. Note that the magnitude of a contractionary monetary policy is more affected by concerns

---

[7]A more formal description of this procedure is confined to Appendix 2.

about the state of cyclical output. However, the magnitude of expansionary policies are more affected by movements in inflation. Also, the smoothed probabilities, obtained by the Hamilton filter, show a good inference on the regime as they generally assign an high probability to the correct regime for each time.

Results of this estimation are consistent with economic literature and with the comments in Section 1. In particular, the low values of p_{13} and p_{31} indicate that two interest rate movements in opposite direction are very unlikely. Secondly, negative movements are more likely than positive ones. From the estimated probabilities, the expected permanence in regime 1 is 1.9 months, in regime 2 is 3 months, in regime 3 is 1.6 months[8]. Finally, the most likely action after a positive movement is no movement whereas after a negative movement, the most likely action is another negative movement.

[INSERT TABLE 2]

Results for the estimation of model (3) with TVTP-MS are shown in the second column of Table 2. We run a number of different experiments, not explicitly reported here, checking if variations in the estimated transition probabilities can be explained by different sets of fundamental variables. The only case in which probabilities are significantly affected by, at least, one of the fundamental variables was the third row of the transition probability matrix. Hence, we estimate the model with the following transition probability matrix:

$$\mathbf{P}_t = \begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31,t} & p_{32,t} & p_{33,t} \end{bmatrix}$$

where the last row is parameterized as:

$$
\begin{aligned}
p_{31,t} &= \frac{\exp(\phi_{13})}{1 + \exp(\phi_{13}) + \exp(\phi_{23} + \vartheta_{23}\pi_{t-1})} \\
p_{32,t} &= \frac{\exp(\phi_{23} + \vartheta_{23}\pi_{t-1})}{1 + \exp(\phi_{13}) + \exp(\phi_{23} + \vartheta_{23}\pi_{t-1})} \\
p_{33,t} &= 1 - p_{31,t} - p_{32,t} = \frac{1}{1 + \exp(\phi_{13}) + \exp(\phi_{23} + \vartheta_{23}\pi_{t-1})}
\end{aligned}
$$

In this case, the Likelihood Ratio test favors the TVTP model with respect to the MS model with fixed transition probabilities (the corresponding p-value is equal to 0.005).

Interestingly, although the level of inflation does not have a significant impact in the magnitude of interest rate changes, it clearly affects the probability of changes. In particular, after a positive interest rate movement the probability of no movement in the next period increases when inflation rate decreases. Hence, we can only observe two contractionary movements

---

[8]The expected permanence in the regime j is given by $1/(1 - p_{jj})$.

in a row when inflation is high. This result is consistent with the insight in Orphanides and Wilcox (2002). They suggest that the monetary authority only moves interest rates to counteract incipient increases of inflation and when inflation is moderate but still above the long run objective central bankers do not take any action to reduce inflation. Instead , they wait for external shocks to obtain the desired reduction of inflation.

[INSERT FIGURE 2]

For robustness, two kind of experiments were run and not explicitly reported to save space. Firstly, given that one can argue that central bankers should consider real instead of nominal exchange rates, we estimate the GPM model including real exchange rate without any significant change in the results. The second type of experiments regards to the presence of different monetary regimes for the whole period of analysis. A particularly important one is known as the "hard" Exchange Rate Mechanism from October 1990 to September 1992. During this regime, The Bank of England sacrificed domestic control to fix exchange rates with respect to the Deutsche Mark. To evaluate the influence of this period in our results, we eliminate observations belonging to this period from our sample and estimated the GPM model again. However, results were qualitatively similar to those reported in Table 2. We also run the estimation considering only the most recent observations from January 1993 to December 2004. In this case, results about changes in the magnitude of interest rate remain similar and the only significant change is that transition probabilities do not depend on any variable. This can be explained for the low and stable inflation level in the most recent period. In any case, main results in the subsequent analysis about the effect of interest rate shocks are not affected when only this period is considered.

4 Monetary Policy Shocks in a GPM Model

Consider a linear vector autoregression (VAR) for a n-dimensional vector, $\mathbf{Y}_t$

$$\mathbf{Y}_t = \mathbf{C} + \mathbf{B}_1 \mathbf{Y}_{t-1} + ... + \mathbf{B}_p \mathbf{Y}_{t-p} + \mathbf{a}_t, \quad E\mathbf{a}_t \mathbf{a}_t' = \mathbf{\Sigma}, \tag{9}$$

where: p is a positive integer; $\mathbf{Y}_t$ is a $(n\text{x}1)$ vector of jointly determined variables and includes, at least, a monetary instrument; $\mathbf{B}_i$ is a $(n\text{x}n)$ matrix of parameters; and $\mathbf{a}_t$ is a vector of zero mean, serially uncorrelated disturbances whose symmetric variance-covariance matrix has typical element σ_{ij}. Efficient estimates of the parameters of such a system with common degrees can be obtained by running OLS equation by equation (see, for example, Lütkepohl, 1993).

From model (9) it is not possible to compute the dynamic response function of $\mathbf{Y}_t$ to the fundamental shocks in the economy. This because the elements of $\mathbf{a}_t$ are, in general, contemporaneously correlated and one cannot presume that they correspond solely to a particular (single) economic

shock. To deal with this issue, a structural model is typically considered for economic analysis. Such a model is defined by

$$\mathbf{A}_0 \mathbf{Y}_t = \mathbf{\Lambda} + \mathbf{A}_1 \mathbf{Y}_{t-1} + ... + \mathbf{A}_p \mathbf{Y}_{t-p} + \varepsilon_t, \qquad (10)$$

where $E\varepsilon_t\varepsilon_t' = \mathbf{A}_0 \mathbf{\Sigma} \mathbf{A}_0' = \mathbf{I}$, an nth order identity matrix. The parameter matrices and errors in (9) and (10) are linked by: $\mathbf{B}_i = \mathbf{A}_0^{-1} \mathbf{A}_i$; $\mathbf{C} = \mathbf{A}_0^{-1} \mathbf{\Lambda}$; and $\mathbf{a}_t = \mathbf{A}_0^{-1} \varepsilon_t$ with ε_t being a $(n\mathrm{x}1)$ vector of orthogonal and standardized structural disturbances.

Once consistent estimators of the $\mathbf{B}_i$'s in (9) are obtained, one can estimate $\mathbf{\Sigma}$ from the fitted residuals. All the information about the matrix $\mathbf{A}_0$ is contained in the relationship $\mathbf{\Sigma} = \mathbf{A}_0^{-1} \left(\mathbf{A}_0^{-1}\right)'$. However, $\mathbf{A}_0$ has n^2 parameters while the symmetric matrix, $\mathbf{\Sigma}$, has at most $n(n+1)/2$ distinct elements. Christiano et al. (1999) provide a detailed discussion of this identification problem. In order to identify the structural model one usually imposes a set of linear restrictions across the elements of the individual rows of $\mathbf{A}_0$. The concomitant order condition, (i), specifies that it is necessary to specify at least $n(n-1)/2$ restrictions, to get a sufficient condition for identification. Together with this condition, it is necessary, (ii), that the diagonal elements of $\mathbf{A}_0$ be positive. However, these conditions are still not sufficient for identification. It is also necessary to ensure, (iii), that a neighborhood of $\mathbf{A}_0$ cannot contain other matrices that fulfil the aforementioned conditions. This is ensured by imposing the additional restriction that the matrix derivative with respect to $\mathbf{A}_0$ of the equations defining $\mathbf{\Sigma} = \mathbf{A}_0^{-1} \left(\mathbf{A}_0^{-1}\right)'$ is of full rank. By doing this, local identification is established. For global identification it is necessary to impose that there be no other matrices that fulfil the three restrictions (i)-(iii) in a neighborhood of $\mathbf{A}_0$. However, if the identification problem only involves systems of linear equations, local identification obtains if, and only if, global identification obtains.

When the model is identified, assuming $\mathbf{Y}_t$ to be covariance-stationary, one can use model (10) to compute analytically the responses of variables in $\mathbf{Y}_t$ to fundamental monetary shocks in different periods. Moreover, these reactions are symmetric and independent of the history of the process. However, if, at least, one of the equations in a system is non linear, impulse response functions can only be estimated by using numerically intensive methods; and the impact of fundamental shocks may depend on the sign and size of the shock as well as on the history of the process. Koop et al (1996) provide a discussion on this issue and propose a methodology based on Monte Carlo simulation of an equation system.

Hamilton and Jorda (2002) propose an alternative methodology, based on their ACH model, to evaluate the effect of monetary shocks. To describe this procedure, notice that a monetary shock can be defined as:

$$\varepsilon_{i,t} = i_t - E(i_t/\Omega_t) \qquad (11)$$

where $\varepsilon_{i,t}$ is the fundamental shock related to the equation for interest rate and Ω_t refers to all the available information to the monetary authority up to period t.

The shock in expression (11) can be written as:

$$\varepsilon_{i,t} = i_t - i_{t-1} - \left(E(i_t/\Omega_t) - i_{t-1} \right). \tag{12}$$

Expression (12) indicates that there are two types of monetary shocks that cannot be distinguished in standard VAR models. The first one relates to the situation where the monetary authority changes official interest rates $(i_t - i_{t-1} \neq 0)$ when no change was expected $(E(i_t/\Omega_t) - i_{t-1} = 0)$. The second one is when central bankers fail to change official interest rates $(i_t - i_{t-1} = 0)$ even though a change $(E(i_t/\Omega_t) - i_{t-1} \neq 0)$ had been expected.

However, a key difference of this model with the one developed by Hamilton and Jorda (2002) is that our model is entirely asymmetric and treats differently the probability of expansionary and contractionary monetary policy. Then, we can extend the previous definition allowing for the differentiation of 4 different types of monetary shocks in the following way

$$\varepsilon_{i,t} = \underbrace{(i_t - i_{t-1})^+}_{\text{Shock 1}} + \underbrace{(i_t - i_{t-1})^-}_{\text{Shock 2}} - \underbrace{(E(i_t/\Omega_t) - i_{t-1})^+}_{\text{Shock 3}} - \underbrace{(E(i_t/\Omega_t) - i_{t-1})^-}_{\text{Shock 4}},$$
$$\tag{13}$$

where $(x_t)^+$ is a function that takes value x_t if $x_t > 0$ and zero otherwise; similarly $(x_t)^-$ takes value x_t when $x_t < 0$ and zero otherwise.

Hence, using our GPM model, we estimate the effect of four different types of monetary shocks. The first two represent respectively an increase and a decrease in the official interest rates when no change is expected. Shocks 3 and 4 represent no change in interest rate when a positive and negative movements are respectively expected. In GPM models, these four events are not forced to have the same effect. An explanation of how to estimate the effect of these shocks follows.

Suppose that we are in period T and we want to estimate the effect on the economy of a certain value of i_T at different time horizons. Variables in $\mathbf{Y}_T$ can be split into two different groups: those that react with one lag delay to i_T, denoted by $\mathbf{Y}_{1,T}$; and those that react contemporaneously to i_T; denoted by $\mathbf{Y}_{2,T}$. The first thing to do is to estimate the value of the variables that react contemporaneously to i_T, $\widehat{\mathbf{Y}}_{2,T}(i_T)$. Using this information we use linear equations to predict the value of $\mathbf{Y}_{1,T+1}(i_T, \Omega_T)$. The value $\widehat{i}_{T+1}$ is then obtained using the GPM model described in Section 2. Iterating in this manner, we can compute

$$\widehat{\mathbf{Y}}_{T+k}(i_T) = \left(\widehat{\mathbf{Y}}_{1,T+k}(i_T, \Omega_{T+k}), \widehat{i}_{T+k}(i_T), \widehat{\mathbf{Y}}_{2,T+k}(i_T, \Omega_{T+k}) \right)'. \tag{14}$$

13

Then, we can easily extend the procedure in Hamilton and Jorda (2002) to answer the questions: what difference does it make if the Bank of England raises (decreases) the official interest rate by 25 basis points during month T compared to the case in which it had kept it constant? To measure these two questions we compute the following expressions:

$$(0.25)^{-1} \left[\widehat{\mathbf{Y}}_{T+k}(i_T) \mid_{i_T=i_{T-1}+0.25} - \widehat{\mathbf{Y}}_{T+k}(i_T) \mid_{i_T=i_{T-1}} \right], \quad (15)$$

$$(0.25)^{-1} \left[\widehat{\mathbf{Y}}_{T+k}(i_T) \mid_{i_T=i_{T-1}-0.25} - \widehat{\mathbf{Y}}_{T+k}(i_T) \mid_{i_T=i_{T-1}} \right]. \quad (16)$$

We can also compute what would happen if we predicted a positive and negative change in official interest rates but no change occurred from the following expression:

$$(w_T)^{-1} \left[\widehat{y}_{T+k/T}(i_T) \mid_{i_T=i_{T-1}} - \widehat{y}_{T+k/T}(i_T) \mid_{i_T=\left(\widehat{i}_{T/T-1}\right)^+} \right], \quad (17)$$

$$(w_T)^{-1} \left[\widehat{y}_{T+k/T}(i_T) \mid_{i_T=i_{T-1}} - \widehat{y}_{T+k/T}(i_T) \mid_{i_T=\left(\widehat{i}_{T/T-1}\right)^-} \right], \quad (18)$$

where

$$w_T = \begin{cases} \left(i_{T-1} - \widehat{i}_{T/T-1}\right)^{-1} & \text{if } \left| i_{T-1} - \widehat{i}_{T/T-1} \right| > 0.05 \\ 0 & \text{otherwise} \end{cases}.$$

As suggested in Hamilton and Jorda (2002), the effect of the weight w_T in expressions (17) and (18) is to ignore observations for which no change was expected and to rescale forecast errors into units comparable to (15) and (16).

5 Impulse-Response Analysis for the UK.

This section presents and analyses the transmission of interest rate shocks to the variables of interest. More specifically, we study the effect of the four types of shocks defined in the previous section and we compare them with the analysis of monetary transmission derived from more standard techniques, such as the use of linear VAR models as proposed by Sims (1980) and the Generalized Impulse Response functions advocated by Koop et al (1996).

We start with the specification of a standard linear VAR model similar to (9) as a benchmark case. Structural system is identified by imposing the recursiveness assumption. We choose this scheme because it is simple and its widespread use makes results comparable to previous studies. In our case, the variables considered in $\mathbf{Y}_t$ are (in this order) the annual rate of inflation, π_t; output gap, y_t; the US/UK foreign exchange rate in first differences, Δe_t;

first differences of interest rates in the long run, ΔR_t; first differences of the overnight interest rate set by the Bank of England, Δi_t; and first differences of the monetary aggregate M0, ΔM_t. This order assumes that overnight interest rate do not have contemporaneous impact on macroeconomic variables such as π_t, y_t, e_t and R_t. It also assume that these variables affect the behaviour of monetary policy makers contemporaneously. For example Bernanke and Blinder (1992) and Wright (2002) both hypothesize recursive structures with similar forms. The inclusion of R_t and M_t is to have a more complete description of the monetary transmission process. R_t is also important because it includes expectations about inflation; see Dale and Haldane (1995).

Reactions to one standard deviation interest rate shock for π_t, y_t, e_t and i_t using a simple VAR model with two lags are exhibited in Figure 3. First thing to notice is that reactions are consistent with the empirical literature on monetary transmission for the UK; see for example Dale and Haldane (1995). More concretely, inflation increases substantially even in the medium and long term after a positive shock in interest rate; cyclical output decreases after the first month and the value of the currency increases after a positive movement in interest rate.

[INSERT FIGURE 3]

The estimated inflation reaction after a positive monetary shock on the official interest rates is a controversial result. This outcome is known in the literature as the price puzzle and has been found by Dale and Haldane (1995) and Wright (1998) amongst others. One possible explanation for this is that the monetary authority sets policy by using private information that is not shared by the rest of the economy and cannot be captured in a VAR model. Hence, when The Bank of England increases interest rate, economic agents interpret this as a signal for future inflation pressures. In order to test this hypothesis, we included the inflation forecast provided by the Bank of England since 1993 in a linear VAR model estimated for the period 1993:01-2004:12. However, price reactions estimated from this model are very similar to the one shown in Figure 3. An alternative explanation is suggested by Dale and Haldane (1995) and Tillmann (2006) among others. In an oligopolistic market, a tight monetary policy affects the capital market raising firms' marginal costs and, henceforth, pushing inflation up. This is known in the literature as the cost channel of monetary transmission.

We developed a number of different experiments to check for robustness of the estimated reactions. For example, 1) the number of lags was changed to 4, 6 and 8; 2) the model was estimated for different subperiods; 3) the case without the inclusion of R_t, M_t and both (R_t and M_t) in the VAR model were also considered; 4) different combinations to order the variables under the recursiveness assumption were tried; and 5) two alternative VAR models using levels of price and first differences of inflation instead of inflation were estimated. However, results were qualitatively similar in all cases.

15

For our second benchmark case, we considered a similar VAR system in which all the equations are linear except the equation for the official interest rate set by the Bank of England. More specifically, our equation for the official interest rate was substituted by our estimated GPM model. Then, we used the methodology developed by Koop et al (1996) to estimate the effect of shocks of different sign and size. A detailed description of the use of this methodology for our particular case can be found in Appendix 3. Figure 4 shows reactions to interest rate shocks of different sign and size under this approach. The size of the shock , (u), is measured in terms of the standard deviation of the structural residual. A visual inspection of the figure reveals that reactions are qualitatively similar to those in the linear case. Moreover, we do not find significant evidence of asymmetric responses to shocks of different size and sign.

Two points must be mentioned about responses obtained in this way. First, although the procedure in Koop et al (1996) allows for the distinction of the effect of shocks of different size and sign, however this approach is not as general as the one described in the previous section that allows for the identification of 4 different types of shocks depending on the previous expectations about interest rate movements. Second, the standard approach to estimate generalized impulse response functions in nonlinear models advocated by Koop et al (1996) is not a very accurate procedure in our particular context. The reason is that it ignores the nature of the official interest rates assuming in the simulation that this variable is affected by stochastic random shocks in all the periods.

[INSERT FIGURE 4]

Figure 5 shows reactions to a positive and a negative movement in interest rate by 25 basis points when no movement is expected (shocks 1 and 2 in Section 4). The effect of responses to negative movements are multiplied by -1 to make it comparable to responses to positive movements. On the other hand, Figure 6 exhibits reactions to no movement in interest rate when a positive and negative movement is expected (shocks 3 and 4 in Section 4). As we did before, the effect of reactions to shock 4 are multiplied by -1 to make them comparable to the effect of shock 3. A detailed description of how these reactions are obtained can be found in Appendix 3.

[INSERT FIGURE 5]

[INSERT FIGURE 6]

Two important features must be highlighted. First, the effect of shocks 3 and 4 are negligible when they are compared to the effect of shocks 1 and 2. These results are consistent with those in Hamilton and Jorda (2002). Second, negative interest rate movements have a bigger effect compared to positive movements. These two features provide valuable economic intuitions. The first one can be used to explain why interest rate remains at the same level for long periods of time even when changes are expected by economic agents. The reason for this is simply that the effect of this kind

16

of shock on the economy is negligible. The second one present a new explanation for the asymmetric effect of positive and negative monetary shocks. The explanation of this asymmetry is straightforward. In our framework a negative movement in interest rate increase the expectation about subsequent negative movements in a near future, however this effect is not so strong for positive movements.

6 Concluding Remarks

This paper presents a methodology for the statistical analysis of the discrete and infrequent changes in the official interest rates set by the Bank of England. The clear advantage of our methodology is that it allows us to differentiate the probabilistic law of positive and negative interest rate movements. Our estimation results are consistent with the observed stylized facts of the official UK rates. More specifically, it shows that negative changes are more likely than positive ones and that two consecutive movements in opposite direction is a very unlikely event. Moreover, when transition probabilities are allowed to depend on fundamental economic variables we find that the probability of two consecutive positive changes in interest rate is an increasing function of inflation rate.

The model is used to define four different types of economic shocks and estimate their separate effects on inflation, output gap, nominal exchange rates and overnight interest rates. Two main results were found. Firstly, changing interest rates when no change is expected have a much bigger economic effect than failing to move interest rates. A similar result was found by Hamilton and Jorda (2002). Secondly, negative changes almost double the economic impact of positive change. An explanation for this asymmetric effect is that after a negative interest rate movement the probability of subsequent negative movements is higher than the probability of positive changes after a positive movement.

References

Abramowitz M. and Stegun I. A., 1972, Handbook of Mathematical Functions, New York: Dover Publications.

Antoniak C. E., 1974, "Mixtures of Dirichlet Processes with Applications to Bayesian Nonparametric Problems", The Annals of Statistics, 2, 1152-1174.

Banerjee, A., J. Dolado, J. Galbraith, and D. Hendry, 1993, Cointegration, Error Correction and the Econometric Analysis of Nonstationary Series, Oxford University Press.

Bernanke, B.S. and A.S. Blinder, 1992, The Federal Funds Rate and the Channels of Monetary Transmission, American Economic Review, 82, 901-921.

Christiano, L.J., M. Eichenbaum and C. Evans, 1999, Monetary Shocks: What We Have Learned and to What End, in: Handbook of Macroeconomics, eds., J.B.Taylor, M.Woodford, Volume A (North-Holland) 65-148.

Clarida, R., J. Galí and M.Gertler, 2000, "Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory", Quarterly Journal of Economics, 115, 145-80.

Cukierman, A. and S. Gerlach, 2003, "The Inflation Bias Result Revisited: Theory and Some International Evidence", CEPR DP#3761.

Dale, S. and A.G. Haldane, 1995, Interest Rates and the Channels of Monetary Transmission: Some Sectoral Estimates, European Economic Review, 39, 1611-26.

Davutyan, N. and W.R. Parke, 1995, "The Operations of the Bank of England, 1890-1908: A Dynamic Probit Approach", Journal of Money, Credit and Banking, 27, 1099-1112.

Diebold F.X., Lee J.H., Weinbach G.C. (1994): Regime Switching with Time-Varying Transition Probabilities, in Nonstationary Time Series Analysis and Cointegration (Hargreaves P. *ed.*), Oxford University Press, 283-302.

Eichengreen, B., M.W. Watson and R.S. Grossman, 1985, "Bank Rate Policy under the Interwar Gold Standard: A Dynamic Probit Model", Economic Journal, 95, 725-745.

Engle, R.F., 2000, "The Econometrics of Ultra-High-Frequency Data", Econometrica, 68, 1-22.

Engle, R.F. and J.R. Russell, 1997, "Forecasting the Frequency of Changes in Quoted Foreign Exchange Prices with the ACD Model", Journal of Empirical Finance, 12, 187-212.

Escobar M. D. and West M. (1995): "Bayesian Density Estimation and Inference Using Mixtures". Journal of the American Statistical Association 90, 577-588.

Ferguson T. S., 1973, "A Bayesian Analysis of Some Nonparametric Problems". The Annals of Statistics 1, 209-230.

Filardo A. J., 1994, "Business-Cycle Phases and their Transitional Dynamics", Journal of Business and Economic Statistics, 12, 299-308.

Filardo A. J., 1998, "Choosing Information Variables for Transition Probabilities in a Time-Varying Transition Probability Markov Switching Model", Federal Reserve Bank of Kansas City, RWP 98-09.

Garcia R., 1995, "Asymptotic Null Distribution of the Likelihood Ratio Test in Markov Switching Models", International Economic Review, 39, 763-88.

Hamilton J. D., 1989, "A New Approach to the Economic Analysis of Nonstationary time Series and the Business Cycle", Econometrica, 57, 357-384.

Hamilton J. D., 1990, "Analysis of Time Series Subject to Change in Regime", Journal of Econometrics, 45, 39-70.

Hamilton J.D., 1994, Time Series Analysis, Princeton: Princeton University Press.

Hamilton, J.D., and Jorda, O., 2002, A Model of the Federal Funds Rate Target, Journal of Political Economy, 110, 1135-1167.

Hansen, L.P., 1982, "Large Sample Properties of Generalized Method of Moments Estimators", Econometrica, 50, 1029-54.

Hansen B. E., 1992, "The Likelihood Ratio Test under Nonstandard Conditions: testing the Markov Switching Model of GNP", Journal of Applied Econometrics 7, S61-S82.

Hausman, J. A., 1978, "Specification Tests in Econometrics", Econometrica, 46, 1251-1272.

Kim C. J., 1994, "Dynamic Linear Models with Markov Switching" , Journal of Econometrics, 60, 1:22.

Koop, G., Pesaran, M.H. and Potter, S.M., 1996, "Impulse Response Analysis in Nonlinear Multivariate Models", Journal of Econometrics, 74, 119-147.

Lowe, P. and L. Ellis, 1997, "The Smoothing of Official Interest Rates" in Lowe, P.(ed.) Monetary Policy and Inflation Targeting Proceedings of a Conference, Reserve Bank of Australia.

Lütkepohl H., 1993, Introduction to Multiple Time Series Analysis, Springer-Verlag, Berlin.

MacKinnon, J.G., 1991, "Critical Values for Cointegration Tests, Chapter 13 in Long-run Economic Relationships: Readings in Cointegration", edited by R.F.Engle and C.W.J. Granger, Oxford University Press.

Orphanides, A., 2001, "Monetary Policy Rules Based on Real Time Data", American Economic Review, 91, 964-985.

Orphanides, A. and D.W. Wilcox, 2002, "The Opportunistic Approach to Disinflation", International Finance, 5, 47-71.

Otranto E., Gallo G. M., 2002, "A Nonparametric Bayesian Approach to Detect the Number of Regimes in Markov Switching Models", Econometric Reviews, 21, 477-496.

Ruge-Murcia, F.J., 2003, "Does the Barro-Gordon Model Explain the Behavior of US Inflation? A Reexamination of the Empirical Evidence", Journal of Monetary Economics, 50, 1375-1390.

Salazar, E., R. Smith, M.Weale and S.H. Wright, 1997, A monthly indicator of GDP, National Institute Economic Review, 161, 84-90.

Schwarz, G., 1978, "Estimating the dimension of a model", The Annals of Statistics 6, 461–464.

Sims, C.A., 1980, "Macroeconomics and Reality", Econometrica 48, 1-48.

Taylor, J.B., 1993, "Discretion versus Policy Rules in Practice", Carnegie-Rochester Conference Series on Public Policy, 39, 195-214.

Tillmann, P., 2006, Inflation Dynamics and the Cost Channel of Monetary Transmission: Empirical Evidence, Royal Economic Society Annual

Conference 2006, Royal Economic Society.

Wright, S., 2002, Monetary policy, nominal interest rate, a long-horizon inflation uncertainty, Scottish Journal of Political Economy, 40, 61-90.

Appendix 1

This appendix describes the time series considered.

- Economic activity: the industrial output of production industries (IOP): all production industries. Seasonally adjusted Period: 1974:01-2004:12. Source: National Statistics (www.statistics.gov.uk).

- Price: long term indicator of price of consumer goods and services. Not seasonally adjusted. Period: 1974:01-2004:12. Source: National Statistics (www.statistics.gov.uk).

- Exchange rate: US/UK exchange rate. Not seasonally adjusted. Period: 1974:01-2004:12. Source: Federal Reserve of St. Louis (research.stlouisfed.org).

- Official Interest Rates at the last day of each month. Not seasonally adjusted. Period: 1974:01-2004:12. Source: Bank of England (www.bankofengland.co.uk).

- Monetary Aggregate: M0 wide monetary base (end period). Seasonally Adjusted. Period: 1974:01-2004:12. Source: National Statistics (www.statistics.gov.uk).

- Long term interest rates: 20 years rates. Par yield - per cent per annum. Not seasonally adjusted. Period: 1974:01-2004:12. Source: National Statistics (www.statistics.gov.uk).

- Inflation forecast from the Bank of England. Not seasonally adjusted. Period: 1974:01-2004:12. Source: Bank of England (www.bankofengland.co.uk).

All series are in monthly basis.

Appendix 2

This appendix describes the procedures to estimate a MS model via the Hamilton (1990) and Kim (1994) filter (Section A) and to apply the nonparametric Bayesian approach of Otranto and Gallo (2002) to detect the number of regimes.

A: Estimation of a Markov Switching Model

The estimations of the unknown parameters contained in the MS and TVTP-MS models can be obtained by the maximum likelihood method. The way to obtain the likelihood function is described in Hamilton (1990); we recall its formulation, reminding to Hamilton (1990) and Kim (1994) for details.

Let us recall our GPM model:

$$i_t = \begin{cases} i_{t-1} + c_1 + \beta_1^1 \Delta\pi_t + \beta_2^1 y_t + \beta_3^1 \Delta e_t + \sigma_1 \varepsilon_t & \text{if } s_t^* = 1 \\ i_{t-1} & \text{if } s_t^* = 2 \\ i_{t-1} + c_2 + \beta_1^2 \Delta\pi_t + \beta_2^2 y_t + \beta_3^2 \Delta e_t + \sigma_3 \varepsilon_t & \text{if } s_t^* = 3 \end{cases}, \quad t = 1,...,T \tag{19}$$

where ε_t is a standard Normal disturbance, whereas the switch from a state to another is driven by a Markov Chain with transition probability matrix:

$$P = \begin{bmatrix} p_{11} & p_{21} & p_{31} \\ p_{12} & p_{22} & p_{32} \\ p_{13} & p_{23} & p_{33} \end{bmatrix}$$

in the case of MS model, or:

$$P_t = \begin{bmatrix} p_{11,t} & p_{21,t} & p_{31,t} \\ p_{12,t} & p_{22,t} & p_{32,t} \\ p_{13,t} & p_{23,t} & p_{33,t} \end{bmatrix}$$

in the case of TVTP-MS model.

The likelihood function is given by:

$$L = \prod_{t=1}^{T} f(i_t | \boldsymbol{\Psi}_{t-1})$$

where $\boldsymbol{\Psi}_t$ represents the information available at time t. The density function $f(i_t | \boldsymbol{\Psi}_{t-1})$ is a mixture of Normal densities; in fact:

$$\begin{aligned} f(i_t | \boldsymbol{\Psi}_{t-1}) &= \sum_{j=1}^{3}\sum_{i=1}^{3} f(i_t, s_t^* = j, s_{t-1}^* = i | \boldsymbol{\Psi}_{t-1}) = \tag{20} \\ &= \sum_{j=1}^{3}\sum_{i=1}^{3} f(i_t | s_t^* = j, s_{t-1}^* = i, \boldsymbol{\Psi}_{t-1}) \Pr\left(s_t^* = j, s_{t-1}^* = i | \boldsymbol{\Psi}_{t-1}\right) = \\ &= \sum_{j=1}^{3}\sum_{i=1}^{3} f(i_t | s_t^* = j, s_{t-1}^* = i, \boldsymbol{\Psi}_{t-1}) \Pr\left(s_t^* = j | s_{t-1}^* = i\right) \Pr\left(s_{t-1}^* = i | \boldsymbol{\Psi}_{t-1}\right) \end{aligned}$$

In the final equation the three elements of the products are easily obtained. The first one, $f(i_t | s_t^* = j, s_{t-1}^* = i, \boldsymbol{\Psi}_{t-1})$, is the Normal density:

$$\frac{1}{\sqrt{2\pi\sigma_j^2}} \exp\left[-\frac{\left(i_t - i_{t-1} - c_j - \beta_1^j \Delta\pi_t - \beta_2^j y_t - \beta_3^j \Delta e_t\right)^2}{2\sigma_j^2} \right].$$

Note that in (19) the equation referred to the second state is deterministic. To provide maximum likelihood estimates we have supposed that, when

$s_t^* = 2$, the variable $(i_t - i_{t-1})$ is Normal with mean 0 and variance 10^{-9}; in this way we can have a stochastic equation to estimate the transition probabilities relative to state 2, but with a variance practically equal to zero.

The second element of the product in (20), $\Pr\left(s_t^* = j | s_{t-1}^* = i\right)$, is the parameter p_{ij} of the transition probability matrix for the MS model, or the element $p_{ij,t}$ of the time varying transition probability matrix in the TVTP-MS model.

Finally, $\Pr\left(s_{t-1}^* = i | \boldsymbol{\Psi}_{t-1}\right)$ is the so-called filtered probability of the state at time $t - 1$, obtained via the Hamilton (1990) filter. Given a starting value $\Pr(s_0^* = 0)$ each element is obtained by the recursive formula:

$$
\begin{aligned}
\Pr\left(s_t^* = j | \boldsymbol{\Psi}_t\right) &= \sum_{i=1}^{3} \Pr(s_t^* = j, s_{t-1}^* = i | \boldsymbol{\Psi}_t) = \\
&= \sum_{i=1}^{3} \frac{f(i_t | s_t^* = j, s_{t-1}^* = i, \boldsymbol{\Psi}_{t-1}) \Pr\left(s_t^* = j | s_{t-1}^* = i\right) \Pr\left(s_{t-1}^* = i | \boldsymbol{\Psi}_{t-1}\right)}{f(i_t | \boldsymbol{\Psi}_{t-1})}
\end{aligned}
$$

Furthermore, the filtered probabilities can be used to make inference on the state for each time t, assigning to the regime j the observation at time t for which $\Pr\left(s_t^* = j | \boldsymbol{\Psi}_t\right)$ is the highest one. A similar function is covered by the smoothed probabilities $\Pr\left(s_t^* = j | \boldsymbol{\Psi}_T\right)$, where the probability at each time is conditional on the full information available. This is a sort of ex post inference with respect to the filtered probabilities, which represent a simultaneous inference on the regime. For details on the algorithm to obtain the smoothed probabilities, see Kim (1994).

B: The Nonparametric Bayesian Method to Detect the Number of Regimes

In this section we describe the nonparametric Bayesian approach of Otranto and Gallo (2002) to detect the number of regimes in a time series. The advantage of this procedure with respect to other specific approaches for MS models (such as Hansen, 1992 or Garcia, 1995) is that it is not a formal test (requiring an estimation procedure), but a sort of identification procedure based on the data set available made at the beginning of the study. In addition, it can include the eventual a priori knowledge of the researcher on the data set and does not require strong hypotheses.

This approach is based on the idea that the data $\{x_1, ..., x_T\}$ are generated from an unknown number k of Normal densities; let us denote with $\theta_t = \left(\mu_t, \sigma_t^2\right)$ the parameter vector containing the mean and the variance of the density referred to x_t. Let us assume that θ_t follows an unknown distribution G belonging to a class of distributions $\mathcal{F}$. Following a nonparametric Bayesian approach (Ferguson, 1973), we can put a class of priors on $\mathcal{F}$ which should cover every kind of prior for $\mathcal{F}$ and be analytically manageable (Antoniak, 1974). The Dirichlet process, introduced by Ferguson (1973), is the

instrument chosen for this purpose. In other terms, we assume that θ_t is generated from an unknown distribution G, following a Dirichlet process $\mathcal{D}(AG_0)$, where A is a hyperparameter which regulates the prior probabilities on the number of regimes k, and G_0 is a bivariate distribution; in particular we suppose that:

$$\sigma_t^{-2} \sim \mathcal{G}(a/2, b/2),$$

$$\mu_t | \sigma_t^2 \sim \mathcal{N}(m, \sigma_t^2 \tau).$$

where $\mathcal{G}$ is a Gamma distribution and $\mathcal{N}$ is a Normal distribution; a, b, and m are hyperparameters to be chosen; the precision parameter τ follows an Inverse Gamma distribution:

$$\tau^{-1} \sim \mathcal{G}\left(\frac{w}{2}, \frac{W}{2}\right)$$

The choice of A is fundamental because the prior distribution for the number k of regimes depends on it; this prior is expressed by (Antoniak, 1974):

$$_n a_k A^k / A^{(n)},$$

where $_n a_k$ is the first type Stirling number in absolute value (tabulated in Abramowitz and Stegun, 1972, p. 833) and $A^{(n)} = A(A+1)...(A+n-1)$. In addition, the expected value of k depends only on A and n; in fact:

$$E(k_n) = \sum_{i=1}^{n} A/(A+i-1) \approx A\left[\log\left(\frac{n+A}{A}\right)\right].$$

Therefore, if we have some expectation about the number of regimes, we can choose the A that fits this expectation. In addition, from it we can obtain the conditional distribution $p(\theta_t | \boldsymbol{\Theta}_{[-t]})$, where $\boldsymbol{\Theta}_{[-t]} = \{\theta_1, ... \theta_{t-1}, \theta_{t+1} ..., \theta_T\}$:

$$p(\theta_t | \boldsymbol{\Theta}_{[-t]}) = \frac{A}{A+T-1} G_0(\theta_t) + \frac{1}{A+T-1} \sum_{j=1, j \neq t}^{T} \delta_{\theta_j}(\theta_t).$$

There is a probability $\frac{A}{A+T-1}$ that θ_t is different from the other terms in $\boldsymbol{\Theta}_T$ and a probability $\frac{1}{A+T-1}$ that θ_t is equal to the j-th term in the matrix $\boldsymbol{\Theta}_T \equiv (\theta_1 \theta_2 ... \theta_T)$. Of course, if there are n_j terms of the sample equal to θ_j, the probability that $\theta_t = \theta_j$ would be $\frac{n_j}{A+T-1}$.

Otranto and Gallo (2002) use the Gibbs sampling approach of Escobar and West (1995) to estimate the empirical posterior distribution of k. This procedure provides the estimation of the posterior distribution $p(\theta_t | \boldsymbol{\Theta}_{[-t]}, \mathbf{Y}_T,)$, where $\boldsymbol{\Theta}_{[-t]}$ denotes $\{\theta_1, ... \theta_{t-1}, \theta_{t+1} ..., \theta_T\}$ and $\mathbf{X}_T$ represents the available observations; the mode of the distribution of the distinct values of the parameters can be considered as an estimate of the number

of regimes k. In the rest of this Appendix we list the steps of the Gibbs sampling procedure, the posterior distributions used and the values of the fixed hyperparameters.

Gibbs sampling steps:

1) choose a starting value for $\mathbf{\Theta}_T$, call it $\mathbf{\Theta}_T^0$ drawing every θ_t from $G_t\left(\theta_t^0\right)$;

2) sample θ_1 from $(\theta_1|\mathbf{\Theta}_{[-1]}, \mathbf{X}_T)$, θ_2 from $(\theta_2|\mathbf{\Theta}_{[-2]}, \mathbf{X}_T)$, ..., θ_T from $(\theta_T|\mathbf{\Theta}_{[-T]}, \mathbf{X}_T)$, obtaining a new $\mathbf{\Theta}_T$. The last θ_t sampled is inserted immediately in $\mathbf{\Theta}_{[-(t+1)]}$ for the subsequent draw;

3) iterate step 2 until convergence, obtaining the first element $\mathbf{\Theta}_T(1)$ for the first replication;

4) repeat step 2) N times, obtaining $\mathbf{\Theta}_T(2)$, ..., $\mathbf{\Theta}_T(N)$;

5) enumerate the k distinct values in $\mathbf{\Theta}_T(i)$ and construct the empirical posterior distribution $p(k|\mathbf{X}_T)$;

6) the mode of $p(k|\mathbf{X}_T)$ is the number of regimes for the switching model.

Posterior Distributions

$$p\left(\theta_t|\mathbf{\Theta}_{[-t]}, \mathbf{X}_T\right) = q_0 G_t\left(\theta_t\right) + \sum_{j=1, j \neq t}^{T} q_j \delta_{\theta_j}\left(\theta_t\right),$$

where $G_t\left(\theta_t\right)$ is the bivariate Normal-Inverse Gamma, with components:

$$\sigma_t^{-2} \sim \mathcal{G}\left[(a+1)/2, \beta_t/2\right],$$

$$\mu_t|\sigma_t^2 \sim N\left(z_t, Z\sigma_t^2\right).$$

and:

$$q_0 \propto A \frac{\Gamma\left[(1+a)/2\right]}{\Gamma\left(a/2\right) a^{1/2}} \left\{1 + (x_t - m)^2/\left[(1+\tau)b\right]\right\}^{-(1+a)/2} \left[(1+\tau)b/a\right]^{-1/2},$$

$$q_j \propto \exp\left[-(x_t - \mu_j)^2/\left(2\sigma_j^2\right)\right]\left(2\sigma_j^2\right)^{-1/2}, \qquad j = 1, \ldots, T$$

with $q_0 + \ldots + q_{t-1} + q_{t+1} + \ldots + q_T = 1$;

$$\beta_t = b + (x_t - m)^2/(1+\tau),$$

$$z_t = (m + \tau x_t)/(1+\tau),$$

$$Z = \tau/(1+\tau).$$

$$\left(\tau^{-1}|\mathbf{\Theta}_T\right) \sim \mathcal{G}\left(\frac{w+k}{2}, \frac{W+K}{2}\right),$$

where k is the number of different components of $\mathbf{\Theta}_T$ and $K = \sum_{i=1}^{k} \frac{(\mu_i - m)^2}{\sigma_i^2}$.

Hyperparameters

The hyperparameters are the same used in Otranto and Gallo (2002), who assign large probabilities to a large set of parameters. They are:

$m = 0$; $a = 4$; $b = 50$; $w = 4$; $W = 1$.

Appendix 3

This appendix describes the procedure to estimate impulse-response functions to monetary shocks. The structure of this appendix is as follows. Section A describes the necessary steps to compute generalized impulse-response functions along the lines of Koop et al (1996). Section B explains the steps followed to compute the economic effect of positive and negative movements in interest rates when no movement was expected. Section C explain the computation of the effect of the decision of not to change official interest rates when positive and negative change was expected.

A. Estimation of the Impulse Response Function

This appendix describes the procedure to estimate the impulse-response function along the lines of Koop et al (1996). Here, this procedure is considered for the analysis of reactions to monetary shocks for the variables of interest. The steps followed in order to obtain the generalized impulse-response function are described next.

- The first step is to estimate a linear VAR system by OLS in reduced form for all the equations except for the overnight interest rate that is estimated from the GPM model.

- We picked a vector of starting values x_{t-1} for the simulation of the system.

- We picked a sample of 6-dimensional shocks. This is done by using the inverse of a Cholesky factorization of the estimated covariance matrix. This transforms the residuals of the model in contemporaneous independent shocks (ε_t). That is, $\widehat{\varepsilon}_t = P^{-1}\widehat{a}_t$, where $\widehat{a}_t$ are the residuals of the different equations and P is the lower triangular Cholesky decomposition of the residuals. Then we drew R unordered collections of these shocks randomly and independently (with replacement) which is denoted by $\left\{\widehat{\varepsilon}_1^{(j)}, \widehat{\varepsilon}_2^{(j)}, ..., \widehat{\varepsilon}_k^{(j)}\right\}$, where $j = 1, ..., r$. The residuals thus obtained were recovered by $\widehat{a}_t^{(j)} = P\widehat{\varepsilon}_t^{(j)}$. We also considered the same sample of R shocks, except that a shock of standard error size s was imposed on the fifth element of $\widehat{\varepsilon}_1^{(j)}$. The reason for this is that we need to analyze the effects of a shock in Δi_t. The sample of residuals recovered was denoted by $\widehat{a}_t^{(j)*}$.

- We simulated the evolution of Y_{t+k} using x_{t-1} and one sample of residuals $\widehat{a}_{t+k}^{(j)}$. The values thus obtained were denoted by $Y_{t+k}^j(x_{t-1}, \widehat{a}_{t+k}^{(j)})$, $k = 1, ..., K$.

- We simulated the evolution of Y_{t+k} using x_{t-1} and one sample of residuals $\widehat{a}_{t+k}^{(j)*}$. The values thus obtained were denoted by $Y_{t+k}^j(x_{t-1}, \widehat{a}_{t+k}^{(j)*})$, $k = 1, ..., K$.

- The last two steps were repeated R times for each of the samples to form an average of each individual component.

$$\overline{Y}_{t+k}(x_{t-1}, a_{t+k}) = \frac{1}{r} \sum_{j=1}^{r} Y_{t+k}^j(x_{t-1}, \widehat{a}_{t+k}^{(j)}), \ \ k = 1, 2, ..., K \quad (21)$$

$$\overline{Y}_{t+k}(x_{t-1}, a_{t+k}^*) = \frac{1}{r} \sum_{j=1}^{r} Y_{t+k}^j(x_{t-1}, \widehat{a}_{t+k}^{*(j)}), \ \ k = 1, 2, ..., K \quad (22)$$

- We took the difference of the two averages to form a Monte Carlo estimate of the reaction function to a monetary shock.

- This process was repeated B times and the estimate reaction is an average of these. Here K is set at 36, B at 500 and R at 500.

B. Estimation of Responses to Shocks 1 and 2.

This part describes a procedure to estimate economic reactions to monetary shocks 1 and 2 in expressions (13). The following steps have been followed:

- The first step is to estimate a linear VAR system by OLS in reduced form for all the equations except for the overnight interest rate equation that is estimated from the GPM model.

- We picked a vector of starting values, x_T, for the simulation of the system. x_T is a vector that includes π_T, y_T, Δe_T, ΔR_T, Δi_T, ΔM_T. In the fifth element of this vector we impose the restriction $i_T = i_{T-1}$ and the new vector is denoted by x_T^*.

- We impose the restriction $i_T = i_{T-1} \pm 0.25$ in the fifth column of vector x_T and it is denoted by x_T^{**}.

- Based on the recursiveness assumption, we simulated the evolution of Y_{t+k} using x_T^*. The values thus obtained were denoted by $Y_{t+k}^{j*}(x_T^*)$, $k = 1, ..., K$.

- Based on the recursiveness assumption, we simulated the evolution of Y_{t+k} using x_T^{**}. The values thus obtained were denoted by $Y_{t+k}^{j**}(x_T^{**})$, $k = 1, ..., K$.

- The last two steps were repeated B times for each of the samples to form an average of each individual component.

$$\overline{Y}_{t+k}^*(x_T^*) \;\; = \;\; \frac{1}{B} \sum_{j=1}^{B} Y_{t+k}^{j*}(x_T^*), \;\; k = 1, 2, ..., K \qquad (23)$$

$$\overline{Y}_{t+k}^{**}(x_T^{**}) \;\; = \;\; \frac{1}{B} \sum_{j=1}^{B} Y_{t+k}^{j**}(x_T^{**}), \;\; k = 1, 2, ..., K \qquad (24)$$

- The normalized response to shocks 1 and 2 are computed as

$$(0.25)^{-1} \left[\overline{Y}_{t+k}^{**}(x_T^{**}) - \overline{Y}_{t+k}^*(x_T^*) \right] \qquad (25)$$

- Here K is set at 36 and B at $10,000$.

C. Estimation of Responses to Shocks 3 and 4.

This part describes a procedure to estimate economic reactions to monetary shocks 2 and 3 in expressions (13). The following steps have been followed:

- The first step is to estimate a linear VAR system by OLS in reduced form for all the equations except for the overnight interest rate equation that is estimated from the GPM model.

- We picked a vector of starting values x_{T-1} for the simulation of the system. x_{T-1} is a vector that includes π_{T-1}, y_{T-1}, Δe_{T-1}, ΔR_{T-1}, Δi_{T-1}, ΔM_{T-1}. We also pick the values x_T that corresponds to the selected x_{T-1}.

- Using x_{T-1}, we compute $E[i_T/\Omega_{T-1}] = \widehat{i}_{T/T-1}$ from the GPM model. Then, we differentiate two cases depending on whether $\widehat{i}_{T/T-1} - i_T$ is positive or negative. We denote these two cases as shocks 3 and 4. In both cases we computed the associated weight w_k^j as

$$
w_T^j = \begin{cases} \left(i_{T-1} - \widehat{i}_{T/T-1}\right)^{-1} & \text{if } \left|i_{T-1} - \widehat{i}_{T/T-1}\right| > 0.05 \\ 0 & \text{otherwise} \end{cases}
$$

- We impose the restriction $i_T = \widehat{i}_{T/T-1}$ in the fifth column of vector x_T and it is denoted by x_T^{**}. In the fifth element of this vector we impose the restriction $i_T = i_{T-1}$ and the new vector is denoted by x_T^*.

- Based on the recursiveness assumption, we simulate the evolution of Y_{t+k} using x_T^{**}. The values thus obtained were denoted by $Y_{t+k}^{j*}(x_T^{**})$, $k = 1, ..., K$.

- Based on the recursiveness assumption, we simulated the evolution of Y_{t+k} using x_T^*. The values thus obtained were denoted by $Y_{t+k}^{j*}(x_T^*)$, $k = 1, ..., K$.

- The last two steps were repeated B times. Then, we classify separately the values $Y_{t+k}^{j*}(x_{T-1}^*)$, $Y_{t+k}^{j**}(x_{T-1}^{**})$ and w_T^j belonging to the types of shocks 3 and 4. We denote by B_1 the number type 3 shocks and by B_2 the number of type 4 shocks, $(B_1 + B_2 = B)$. Then, we compute

$$
\sum_{j=1}^{B_1} \left(w_T^j\right)^{-1} \left[Y_{t+k}^*(x_{T-1}^*) - Y_{t+k}^{**}(x_{T-1}^{**})\right] \tag{26}
$$

$$
\sum_{j=1}^{B_2} \left(w_T^j\right)^{-1} \left[Y_{t+k}^*(x_{T-1}^*) - Y_{t+k}^{**}(x_{T-1}^{**})\right] \tag{27}
$$

- Here K is set at 36 and B at $20,000$.

Tables and Figures
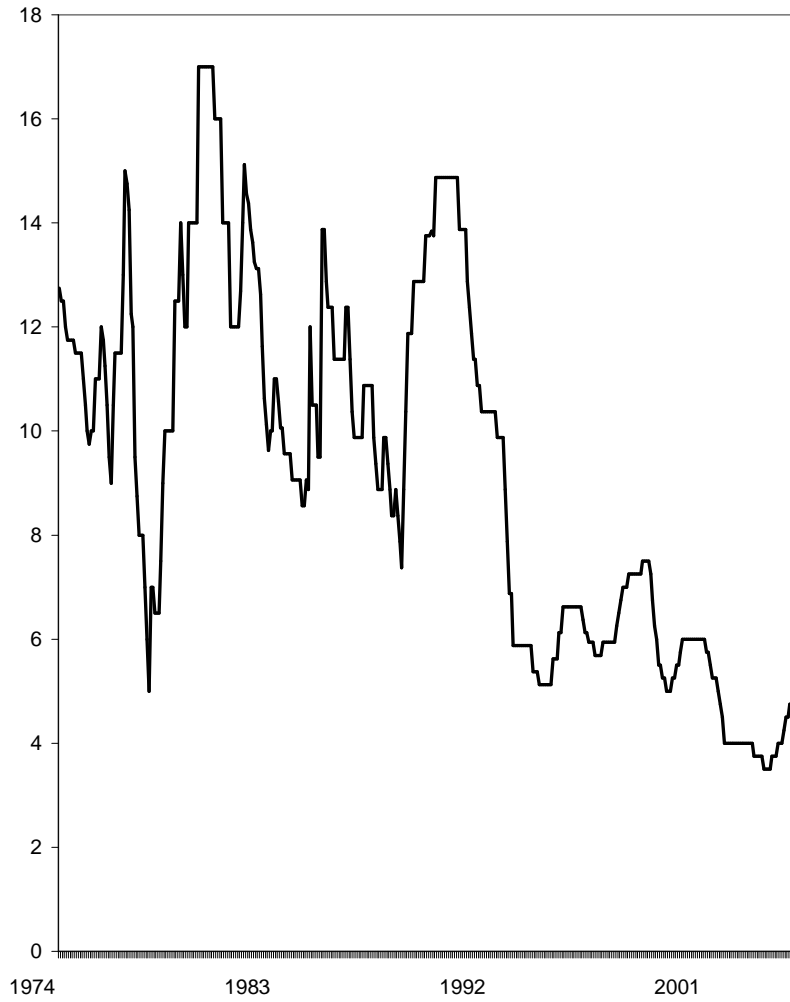
Table 1: Posterior distributions of the number of regimes of ∇i_t corresponding to three different priors; 500 Gibbs sampling replications

	Number of Regimes					
A	1	2	3	4	5	6
0.150	0.000	0.000	0.672	0.274	0.046	0.008
0.220	0.000	0.000	0.440	0.404	0.134	0.022
0.465	0.000	0.000	0.760	0.204	0.032	0.004

Table 2: Estimates of MS and TVTP-MS models (standard errors in parentheses)

Parameters	MS	TVTP-MS
c_1	-0.380 (0.021)	-0.379 (0.022)
c_3	0.078 (0.121)	0.070 (0.122)
$\beta_{1,1}$	13.675 (3.416)	14.087 (3.236)
$\beta_{2,3}$	9.483 (6.077)	9.818 (5.900)
$\beta_{3,1}$	-2.603 (0.883)	-2.641 (0.896)
$\beta_{3,3}$	-12.733 (4.149)	-12.651 (4.099)
σ_1^2	0.018 (0.004)	0.018 (0.004)
σ_3^2	1.249 (0.187)	1.228 (0.182)
p_{11}	0.476 (0.078)	0.485 (0.080)
p_{12}	0.406 (0.078)	0.408 (0.082)
p_{21}	0.083 (0.025)	0.079 (0.025)
p_{22}	0.669 (0.032)	0.670 (0.032)
p_{31}	0.119 (0.044)	
p_{32}	0.524 (0.056)	
ϕ_{31}		-1.217 (0.544)
ϕ_{32}		1.254 (0.428)
ϑ_{32}		-12.018 (4.630)
Log-Likelihood	1923.45	1927.348

Figure 1: Official Interest Rates in the UK

Source: Bank of England

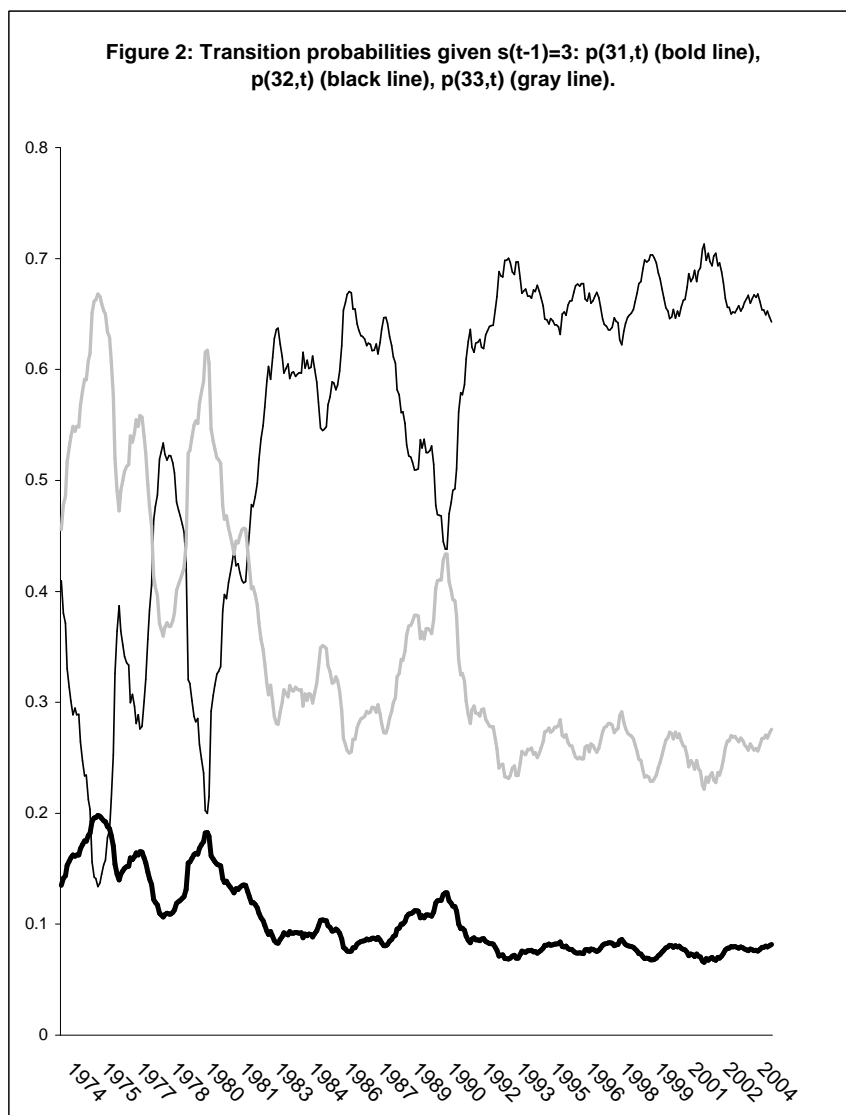**Figure 2: Transition probabilities given s(t-1)=3: p(31,t) (bold line), p(32,t) (black line), p(33,t) (gray line).**

Figure 3. Reactions to Monetary Shocks in a Linear VAR Model

Effect of one S.D. Monetary Shock on Inflation

Effect of one S.D. Monetary Shock on Cyclical Output

Effect of one S.D. Monetary Shock on Exchange Rate

Effect of one S.D. Monetary Shock on Interest Rates
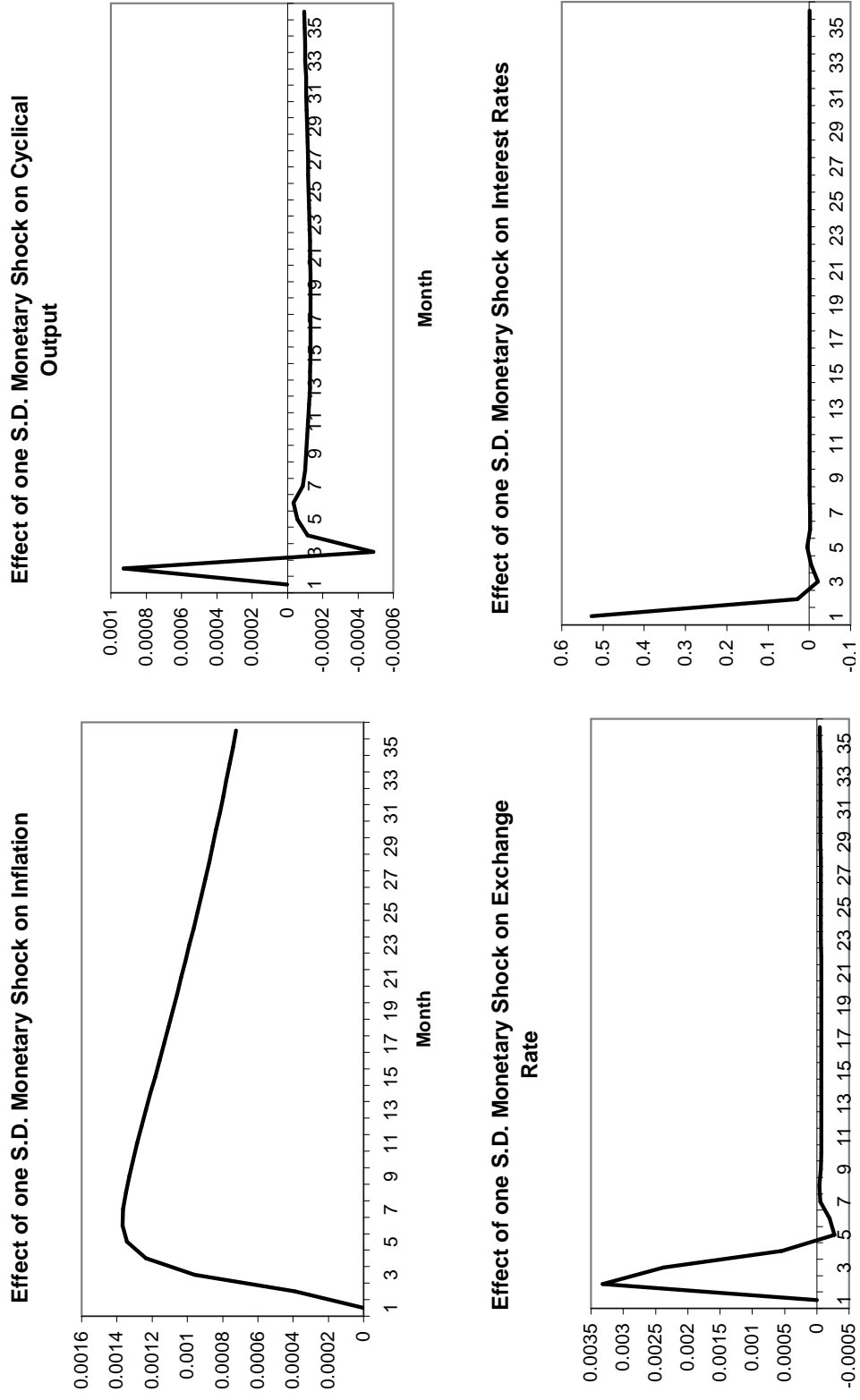
33

Figure 4. Comparable Reactions to Monetary Shocks
using the Approach in Koop et al (1996) (*)

Comparable Measures of Inflation Reactions to Monetary Shocks

Comparable Measures of Inflation Reactions to Monetary Shocks

Comparable Measures of Output Reactions to Monetary Shocks

Comparable Measures of Output Reactions to Monetary Shocks

Comparable Measures of Exchange Rate Reactions to Monetary Shocks

Comparable Measures of Exchange Rate Reactions to Monetary Shocks

Comparable Measures of Official Interest Rate Reactions to Monetary Shocks

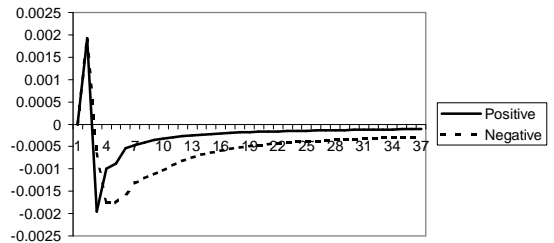**Comparable Measures of Official Interest Rates Reactions to Monetary Shocks**

(*) u refers to the size of the shock in terms of the stardard deviation of the structural residuals.
 The first column shows the effect of two shocks of u=1 and u=2 on official interest rate. To make results comparable, reactions to shocks of size u=1 are multiplied by 2.
The second column shows the effect of two shocks of size u=2 and u=-2 on official interest rate. To make results comparable, reactions to shocks of size u=2 are multiplied by -1.

Figure 5. Comparable Reactions to a Positive and Negative
Interest Rate Movement by 25 Basis Points when no Movement is Expected (*)
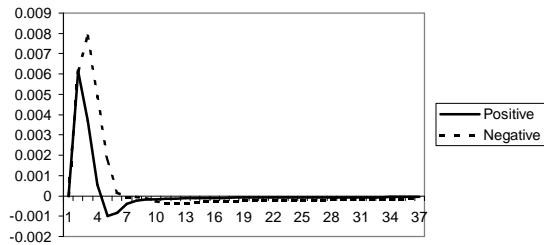
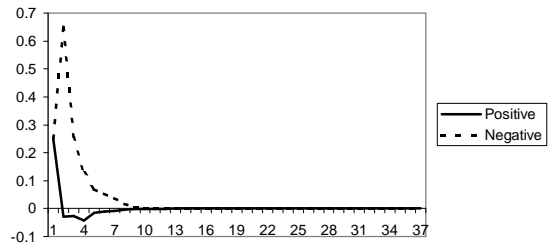**Comparable Measures of Inflation Reactions to a Non Expected Interest Rate Movement.**
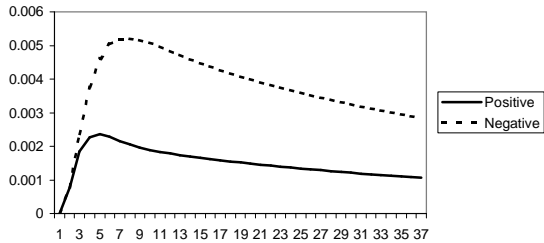
Comparable Measures of Output Reactions to a Non Expected Interest Rate Movement.

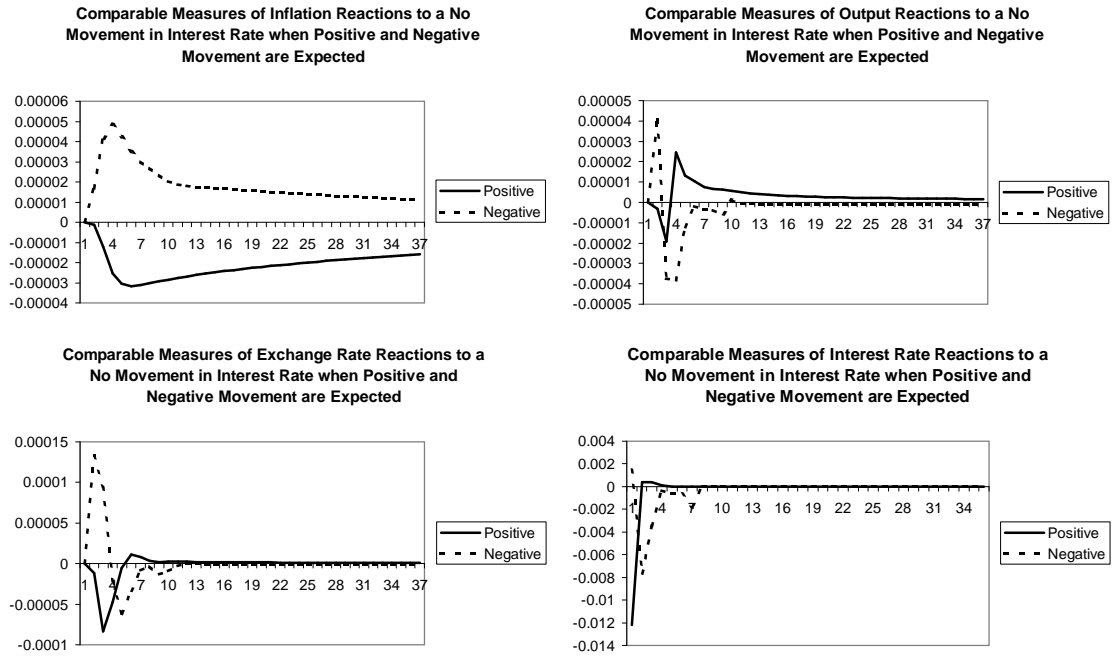**Comparable Measures of Exchange Rate Reactions to a Non Expected Interest Rate Movement.**

Comparable Measures of Interest Rate Reactions to a Non Expected Interest Rate Movement.

(*) Reactions to negative interest rate movements are multiplied by -1.

35

Figure 6. Comparable Reactions to no Change in Interest Rate when Positive and Negative Movements are Expected (*)

Comparable Measures of Inflation Reactions to a No Movement in Interest Rate when Positive and Negative Movement are Expected

Comparable Measures of Output Reactions to a No Movement in Interest Rate when Positive and Negative Movement are Expected

Comparable Measures of Exchange Rate Reactions to a No Movement in Interest Rate when Positive and Negative Movement are Expected

Comparable Measures of Interest Rate Reactions to a No Movement in Interest Rate when Positive and Negative Movement are Expected

(*) Reactions to no movement in interest rates when a negative movement is expected is multiplied by -1.