67
DISCUSSION PAPER

Originally published by Ipea in March 1997 as
number 463 of the series Texto para Discussão.

THE VARIANCE OF INFLATION AND THE STABILITY OF THE DEMAND FOR MONEY IN BRAZIL: A BAYESIAN APPROACH

Elcyon Caiado Rocha Lima
Ricardo Sandes Ehlers

ipea

67
DISCUSSION PAPER
Originally published by Ipea in March 1997 as
number 463 of the series Texto para Discussão.
Brasília, January 2015

THE VARIANCE OF INFLATION AND THE STABILITY OF THE DEMAND FOR MONEY IN BRAZIL: A BAYESIAN APPROACH

Elcyon Caiado Rocha Lima[1]
Ricardo Sandes Ehlers[2]

1. Ipea and Santa Úrsula University.
2. Department of Statistics, Fluminense Federal University.

DISCUSSION PAPER

A publication to disseminate the findings of research directly or indirectly conducted by the Institute for Applied Economic Research (Ipea). Due to their relevance, they provide information to specialists and encourage contributions.

Discussion paper / Institute for Applied Economic

Research.- Brasília : Rio de Janeiro : Ipea, 1990-

ISSN 1415-4765

1. Brazil. 2. Economic Aspects. 3. Social Aspects. I. Institute for Applied Economic Research.

CDD 330.908

SUMÁRIO

RESUMO

Em trabalho recente Tourinho (1995) sugeriu que, em processos de inflação elevada, deve ser considerada não somente a esperança da taxa de inflação mas também a variância esperada da taxa de inflação

O modelo apresentado em Lima & Ehlers (1993) é aqui estendido para lidar com a incerteza produzida pela variabilidade da taxa de inflação : um termo proporcional à esperança do erro quadrático médio, na previsão da taxa de inflação, é incluído na equação de demanda por moeda. O problema de que estimativa utilizar, para a variância esperada da taxa de inflação, é resolvido através de um procedimento de estimação Bayesiano. É permitida a alteração dos parâmetros do modelo ao longo do tempo e são adotados procedimentos de monitoramento e intervenção Bayesianos para detectar-se mudanças estruturais.

O modelo foi estimado com dados trimestrais entre o primeiro trimestre de 1973 e o quarto trimestre de 1995, e portanto considerando-se os diversos planos recentes de estabilização da economia Brasileira. Nós concluímos que a presença da variância esperada da taxa de inflação, na equação de demanda por moeda, é importante por duas razões principais: **a**) ela impede que o monitor sinalize em 1990 após uma intervenção no modelo em 1986 e **b**) o seu efeito se torna significante depois de 1986 quando diversos planos de estabilização contribuíram para aumentar a incerteza a respeito da taxa de inflação.

ABSTRACT

When analyzing the demand for money in high inflation processes it has been suggested [Tourinho (1995)] that we should consider not only the effects of changes in the expected inflation rate but also changes in the expected variability of inflation. The model in Lima & Ehlers (1993) is extended here to deal more accurately with the uncertainty produced by the variability of inflation: a term proportional to the expected quadratic error in forecasting inflation is included in the demand for money equation. The problem of what estimate to use for the expected variance of inflation, is addressed by a Bayesian estimation procedure. Model parameters are allowed to vary slowly over time and Bayesian monitoring and intervention procedures are then used to cater for structural changes. We estimate the model with data ranging from first quarter of 1973 to fourth quarter of 1995, thus taking into account many stabilization plans for the Brazilian economy. We find that the presence of variance of inflation in our money demand equation is important in two ways: **a**) it prevents the monitor from signaling again in 1990 after an intervention period in 1986 and **b**) its effect turns out to be significant after 1986 when many stabilization plans contributed to increase uncertainty.

1 - INTRODUCTION

The stability of money demand has become of great concern in empirical studies since the work of Goldfeld (1976). The Brazilian economy has faced high inflation rates and an unstable economic environment in its recent history. The sudden changes in economic policy and in the liquidity of government bonds and other financial assets make it unlikely for money demand in Brazil to show structural stability in its parameters.

The behavior of money demand is relevant due to the possible implications of its stability (or not) for monetary policy. The stability of the equation's parameters is of great concern since the sample period covers many stabilization plans. Furthermore, detecting periods of money demand instability is important to explain high inflation episodes. Finally, the correct specification of the money demand equation increases the predictive performance of the model.

Money demand stability in Brazil has been investigated recently in a number of works [Rossi (1988, 1990), Pereira (1989), Pastore (1991) and Tourinho (1995)]. These works follow the prior assumption of constancy, over time, of the parameters that describe structural relations among economic variables besides the classical assumption that these parameters are unknown fixed (non-stochastic) quantities.

This work differs from the previous ones by specifying a statistical model for the money demand which allows for continuous or abrupt stochastic changes of the parameters over time and assumes that they are random variables with an associated probability distribution.

Estimation is sequential, testing the model in use by comparing it with an alternative one, over the whole sample period. Furthermore, the model is open, in the sense of allowing for the incorporation of external subjective information relevant to the estimation process. These are the basic ideas underlying the monitoring and intervention procedures described in a later section. Certainly, the great advantage of the Bayesian approach is in the case in which one combines subjective information with data information.

It has been suggested [Tourinho (1995)] that we should consider not only the effects of changes in the expected inflation rate but also changes in the expected variability of inflation when analyzing the demand for money in high inflation processes. The model in Lima & Ehlers (1993) is extended here to deal with the uncertainty produced by the variability of inflation. A term proportional to the expected quadratic error in forecasting inflation is included in the demand for money equation.

The main results are: (**a**) there is evidence of a continuous increase, over time, in the absolute value of the short-run elasticity of the money demand with respect to the interest rate; (**b**) there is evidence of an abrupt increase (in absolute value) of this elasticity in the second quarter of 1986.

We used the logarithm of the following macroeconomics variables: the real money stock (M1), the monthly inflation rate detected by the price index of Fundação Getúlio Vargas, the quarterly GNP in real terms and the real interest rate. The model was estimated using quarterly data from the first quarter of 1973 to the fourth quarter of 1995.

The paper is organized as follows: in Section 2 we discuss the Bayesian estimation procedure and the monitoring method adopted to detect structural changes in model parameters; in Section 3 we present the probabilistic model, the adopted parameterization of priors and more details about the estimation method; in Section 4 we present the results for the search procedure and the use of the monitor. In Section 5 there are some conclusions.

2 - METHODOLOGY

2.1 - Bayesian Estimation Procedure

Conventional methods of developing parsimonious probabilistic models for econometric time series are, in general, quite arbitrary and do not give probabilistic treatment to the uncertainty yielded by inexact knowledge of the true model specification. For instance, the classical procedure of reducing the number of variables in the model or the number of lags is equivalent to setting the coefficients of these variables to zero. From the Bayesian point of view this indicates complete uncertainty about the behavior of the coefficients up to lag k and complete prior certainty about the value zero of the coefficients in longer lags. The estimation procedure applied in this work is designed to treat this uncertainty adequately.

Let $\boldsymbol{y}$ be the vector with sample information, $\boldsymbol{\theta}$ a vector of parameters and $p(\boldsymbol{y}|\boldsymbol{\theta})$ the model for the data, represented by a density function for $\boldsymbol{y}$ conditioned by $\boldsymbol{\theta}$. In the usual procedure, a subspace of the possible values of $\boldsymbol{\theta}$ is selected as the parameter space, for example omitting variables or reducing the number of lags of certain variables on the right side of the equation, which corresponds to setting to zero some elements of the vector $\boldsymbol{\theta}$. In this work, the parameter space indexed by $\boldsymbol{\theta}$ is reduced to a smaller one indexed by a vector $\boldsymbol{\lambda}$ of prior parameters. The priors are parameterized so as to represent ignorance about the true data generating model and to allow convergence to different models as the prior parameters converge to certain regions of their space of possible values.

In the standard Bayesian approach, a model is specified to data in the form of a density function to $\boldsymbol{Y}$ conditioned by $\boldsymbol{\theta}$, $p(\boldsymbol{y}|\boldsymbol{\theta})$, a prior density function, $q(\boldsymbol{\theta})$, and, once $\boldsymbol{y}$ is observed, the distribution of θ is updated via Bayes theorem as $f(\theta \mid \boldsymbol{y}) \propto p(\boldsymbol{y}|\theta)\, q(\theta)$.

An additional level of parameters is introduced here conditioning the prior distribution of θ by a new set of parameters λ, specifying a density function $g(\theta \mid \lambda)$. The uncertainty with respect to the true values of λ is represented by a prior distribution of $h(\lambda)$. Thus, the prior density of θ is obtained as:

$$q(\theta) = \int g(\theta \mid \lambda)\, h(\lambda)\, d\lambda.$$

However, rather than evaluate this integral, an alternative method will be used assuming that $h(\lambda)$ is flat in the relevant region.

Note that $p(\boldsymbol{y}|\theta)g(\theta \mid \lambda)$ is the joint density function of $\boldsymbol{y}$ and θ conditioned by λ, which can be integrated with respect to θ leading to a marginal distribution of $\boldsymbol{y}$ given λ, $m(\boldsymbol{y} \mid \lambda)$. Once $\boldsymbol{y}$ is observed, $m(\boldsymbol{y} \mid \lambda)$ plays the formal role of a likelihood function and the posterior distribution of λ can be obtained via Bayes theorem as $n(\lambda \mid \boldsymbol{y}) \propto m(\boldsymbol{y} \mid \lambda)\, h(\lambda)$. If $h(\lambda)$ is flat in the region where $m(\boldsymbol{y} \mid \lambda)$ assumes its highest values then $n(\lambda \mid \boldsymbol{y})$ is proportional to $m(\boldsymbol{y} \mid \lambda)$ and it is not necessary to specify $h(\lambda)$ more precisely. In this case, to make inferences about the more likely values of λ is equivalent to the classical procedure of estimating parameters from the sample data.

After calculating the posterior distribution of θ given λ as $k(\theta \mid \boldsymbol{y},\lambda) = p(\boldsymbol{y}|\theta)\, g(\theta \mid \lambda) / m(\boldsymbol{y} \mid \lambda)$ the posterior marginal distribution of θ is finally obtained via integration,

$$f(\theta \mid \boldsymbol{y}) = \int k(\theta \mid \boldsymbol{y},\lambda)\, n(\lambda \mid \boldsymbol{y})\, d\lambda,$$

i.e., as a weighted average of the posterior distribution of θ fixing different values of λ with weights $n(\lambda \mid \boldsymbol{y})$ which are proportional to $m(\boldsymbol{y} \mid \lambda)$ if $h(\lambda)$ is flat.

In fact, the practical procedure is to estimate the vector θ given $\boldsymbol{y}$ via the mode of the posterior distribution of θ conditioned by the value of λ $[k(\theta \mid \boldsymbol{y},\lambda)]$ that maximizes $m(\boldsymbol{y}|\lambda)$, rather than to evaluate this integral. This is the procedure adopted in this work.

2.2 - Automatic Monitoring

Forecasting time series models must be sufficiently flexible in the sense of adequately treating discontinuities and structural changes frequently encountered in practice. Such events imply an increased uncertainty about the model and its forecasts. One or more observations that are far from the forecasts can indicate for example a series break down with the model's form, which in this case must be modified or adapted, or they can be outliers that have no relation with time evolution of the series and should not be used as information in the estimation the model parameters.

The method presented in this paper, developed in West (1986) and generalized in West & Harrison (1986), involves the use of a standard model with a monitor, used sequentially, that measures predictive performance of the model over time, from the observed values of predictive densities, comparing it with an alternative model. Essentially, the alternative model is expected to provide a predictive density that is consistent with those observations judged discrepant by the standard model.

At time t-1, the predictive distribution of Y_t given D_{t-1}, the data up to that point, has a density function of the form

$$p(Y_t \mid D_{t-1}) = \int p(Y_t \mid \theta_t)\, p(\theta_t \mid D_{t-1})\, d\theta_t$$

the observed value $p(y_t \mid D_{t-1})$ being a measure of predictive performance for the standard model, that is the likelihood of the model.

Let $p_A(y_t \mid D_{t-1})$ be the predictive density of the alternative model. Then after observing $Y_t = y_t$ the Bayes factor based on this observation is defined as :

$$H_t = p(y_t \mid D_{t-1}) / p_A(y_t \mid D_{t-1})$$

with small values of H_t indicating low predictive performance of the standard model with respect to the alternative one. For a sequence of observations the joint predictive density is $p(y_1,...,y_t \mid D_0) = p(y_t|D_{t-1})\, p(y_1,...,y_{t-1}|D_0)$ leading to an overall Bayes factor defined as:

$$H_t(t) = \frac{p(y_1,\cdots,y_t|D_0)}{p_A(y_1,\cdots,y_t|D_0)} = \frac{p(y_t|D_{t-1})}{p_A(y_t|D_{t-1})}\frac{p(y_1,\cdots,y_{t-1}|D_0)}{p_A(y_1,\cdots,y_{t-1}|D_0)} = H_t H_{t-1}(t-1)$$

However, in a context of monitoring one must take into account the local behavior of the model. The cumulative Bayes factor, based on the k most recent observations, is defined as:

$$H_t(k) = p(y_t, \cdots, y_{t-k+1} | D_{t-k}) / p_A(y_t, \cdots, y_{t-k+1} | D_{t-k})$$

and defining $H_t(0) = 1$ it follows that $H_t(k) = H_t H_{t-1}(k-1)$, $1 \leq k \leq t$. This way, the evidence in favor or against the standard model accumulates multiplicatively as the data are observed.

When observations are increasingly far from the forecasts then the individual Bayes factors, H_t, are not sufficiently small and need to be accumulated to indicate some evidence against the standard model. In this case, the monitor identifies the most discrepant group of consecutive observations calculating V_t and l_t such that:

$$V_t = \min_{1 \leq k \leq t} H_t(k) = H_t(l_t)$$

being sequentially calculated through the following recursions:

$$V_t = H_t \min\{1, L_{t-1}\} \qquad \text{and} \qquad l_t = \begin{cases} l_{t-1} + 1 & \text{if } L_{t-1} < 1 \\ 1 & \text{if } L_{t-1} \geq 1 \end{cases}$$

as shown in West (1986).

The standard model is accepted as satisfactory until the occurrence of a value L_t less than a prespecified value $\tau < 1$ (the lower limit on acceptability of L_t) when the occurrence of a discontinuity in the series is then signaled. If $l_t = 1$ then a single discrepant observation is identified as the most likely cause of failure, although the onset of a change is also a possibility $l_t > 1$ indicates that a change started to occur l_t periods later in $t - l_t + 1$. Furthermore, if a slow structural change is occurring in the series then the most recent observations will indicate evidence against the standard model which will not be sufficient to make $L_t < \tau$. So, to increase the sensitivity of the monitor to these changes a discontinuity should be signaled if $l_t > 3$ or 4.

In order to specify the alternative model consider that the predictive densities are normal with common mean f_t and variances Q_t and Q_t / ρ where $0 < \rho < 1$, so that the Bayes factor is:

$$H_t = \sqrt{\frac{1}{\rho}} \exp\left\{-\frac{1}{2} \frac{(y_t - f_t)^2}{Q_t} (1 - \rho)\right\} = \sqrt{\frac{1}{\rho}} \exp\left\{-\frac{1}{2}(1 - \rho) e_t^{\,2}\right\}$$

where e_t is the standardized forecast error.

The choice of ρ may be guided by rewriting the Bayes factor as $H_t = \exp(-0.5 \log \rho + (1-\rho)\, e_t^2)$. Clearly $H_t = 1$ or equivalently $e_t^2 = -(\log \rho) / (1-\rho)$ indicates no evidence to discriminate between the models. The value of ρ may be chosen so as to provide the maximum value of $|e_t|$ indicating no evidence against the standard model. For instance, $\rho \in (0.1, 0.3)$ implies that evidence against the standard model must be accumulated for $1.3 < |e_t| < 1.6$ which are roughly the 0.90 and 0.95 percentiles of the standard normal distribution.

Of course, for fixed ρ, the evidence against the standard model increases with $|e_t|$. West & Harrison (1989) illustrate that the choice of ρ matters little when the error becomes extreme relative to the alternative model. The alternative model may be viewed as a general one in the sense that it caters for various types of changes in addition to outlying observations. Essentially, this procedure may be viewed as an exploratory method yielding information on the type and more likely period of structural change.

3 - THE MODEL, THE CHOICE OF PRIOR PARAMETERS AND THE ESTIMATION METHOD

3.1 - The probabilistic model

The postulated model for the money demand series is represented by the following observation equation:

$$X_{1t} = \mu_t + \alpha_t + \delta_t + c_{2t}(0)X_{2t} + c_{3t}(0)X_{3t} + c_{4t}(0)X_{4t} + \phi_t V_t + \sum_{j=1}^{4}\sum_{k=1}^{3} c_{jt}(k)X_{j,t-k} + \nu_t \qquad (1)$$

while the evolution equations for trend, seasonal and regression components are given by:

$$\begin{aligned}
\mu_t &= \mu_{t-1} + \omega_{1t} \\
\alpha_t &= \beta_{t-1} + \omega_{2t} \\
\beta_t &= -\alpha_{t-1} + \omega_{3t} \\
\delta_t &= -\delta_{t-1} + \omega_{4t} \\
\phi_t &= \phi_{t-1} + \omega_{5t} \\
c_{jt}(k) &= c_{j,t-1}(k) + \sigma_{jt}(k)
\end{aligned} \qquad (2)$$

for $j = 1,2,3,4$ and $k = 0,1,2,3$, with $c_{1t}(0) = 0$. Equations (2) allow the modeling of time variation of the parameters. α, β and δ are the coefficients in the harmonics of the Fourier representation of seasonal component, the contribution of these harmonics to the seasonal factor being $\alpha + \delta$. The variables X_{it} are the logarithms of the real money stock, the inflation rate, the real GNP and the real interest rate, **i.e.**,

$$X_{1t} = \log (M1/P)_t\, , X_{2t} = \log (1+ \pi)_t\, , X_{3t} = \log y_t,\ X_{4t} = \log (1+ r)_t$$

where r and π are in percent. Finally, V_t is an estimate of the variance of inflation constructed from the past squared deviations of the actual logged inflation rate from its expected value,

$$V_t = \sum_{j=1}^{k} b_j\, (\pi*_{t+1\,-j} - \log(1+ \pi_{t+1\,-\,j}))^2 \tag{3}$$

where

$$\pi*_t = \sum_{j=1}^{k} b_j \log(1+ \pi_{t\,-\,j}). \tag{4}$$

The number k of lags in equations (3) and (4) is the same as for the other variables in the model, **i.e.**, $k = 3$.

Using a matrix notation these equations can be written as:

$$X_{1t} = \boldsymbol{F}'_t\, \boldsymbol{\theta}_t + \nu_t$$
$$\boldsymbol{\theta}_t = \boldsymbol{G}\boldsymbol{\theta}_{t\text{-}1} + \boldsymbol{\omega}_t$$

where: $\boldsymbol{\theta}$ is the vector of parameters to be estimated, $\boldsymbol{F}'_t$ is the known regression vector at time t given by $\boldsymbol{F}'_t = (1,1,0,1,V_t,\ X_{2t},\ X_{3t},\ X_{4t},\ X_{1,t\,\text{-}1},\dots,\ X_{4,t\,\text{-}1},\dots,\ X_{1,t\,\text{-}3},\dots,\ X_{4,t\,\text{-}3})$, and $\boldsymbol{\omega}_t$ is the vector of evolution errors.

Furthermore, $\boldsymbol{G}$ is the block diagonal evolution matrix, given by $\boldsymbol{G}$ = block diagonal $(G_1\, ,\boldsymbol{G}_2,\boldsymbol{G}_3)$,

where:

$G_1 = 1 \qquad \boldsymbol{G}_2 = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} \qquad \boldsymbol{G}_3 = \boldsymbol{I}$.

Essentially, $\boldsymbol{G\theta}_{t\text{-}1}$ describes the deterministic part of the time evolution of $\boldsymbol{\theta}$ with the variance matrix of the errors ω_t increasing the uncertainty about this evolution. This model representation is known as the state space representation and $\boldsymbol{\theta}_t$ is the state vector.

The error sequences $\{\nu_t\}$ and $\{\omega_t\}$ are subject to the usual assumptions of normality and no serial correlation. They are uncorrelated with each other at any period t and uncorrelated with $(\boldsymbol{\theta}_t \mid D_0)$. Furthermore, they have zero means and unknown variances Σ and $\boldsymbol{W}$. The posterior distribution of $\boldsymbol{\theta}$ at time t-1 is normal with mean $\boldsymbol{m}_{t\text{-}1}$ and variance matrix $\boldsymbol{C}_{t\text{-}1}$, i.e., $(\boldsymbol{\theta}_{t\text{-}1} \mid D_{t\text{-}1}) \sim N(\boldsymbol{m}_{t\text{-}1}, \boldsymbol{C}_{t\text{-}1})$, where $D_{t\text{-}1}$ denotes the use of information up to t-1. Furthermore, given the distributions of the error sequence, from the observation and evolution equations, we have that:

$$(\boldsymbol{\theta}_t \mid \boldsymbol{\theta}_{t\text{-}1}, D_{t\text{-}1}) \sim N(\boldsymbol{G\theta}_{t\text{-}1}, \boldsymbol{W})$$

$$(X_{1t} \mid \boldsymbol{\theta}_t, D_{t\text{-}1}) \sim N(\boldsymbol{F}'_t\boldsymbol{\theta}_t, \Sigma)$$

and one obtains a joint normal distribution for X_{1t}, $\boldsymbol{\theta}_t$ and $\boldsymbol{\theta}_{t\text{-}1}$ conditioned by data up to period t-1. The Kalman filter is a set of recursions which allows obtaining this joint normal distribution for each t. Given this joint distribution it is possible to obtain the marginal distribution of $\boldsymbol{\theta}_t \mid D_t$ and $X_{1t} \mid \boldsymbol{\theta}_t, D_{t\text{-}1}$. To use the Kalman filter it is necessary to specify the prior distribution of $\boldsymbol{\theta}_0$ as an initial information which sufficiently summarizes the previous knowledge about the vector of parameters. Assuming that this prior information is appropriately summarized by a normal distribution with mean $\boldsymbol{m}_0$ and variance matrix $\boldsymbol{C}_0$ and that one of them or both depend on a set of unknown parameters λ (as mentioned in section 2.1), we have:

$$(\boldsymbol{\theta}_0 \mid D_0, \lambda) \sim N(\boldsymbol{m}_0, \boldsymbol{C}_0),$$

and one obtains the following conditional distributions:

$$(\boldsymbol{\theta}_1 \mid \boldsymbol{\theta}_0, D_0, \lambda) \sim N(\boldsymbol{G\theta}_0, \boldsymbol{W})$$

$$(X_{11} \mid \boldsymbol{\theta}_1, D_0, \lambda) \sim N(\boldsymbol{F}_1'\boldsymbol{\theta}_1, \Sigma)$$

Note that these two last conditional distributions become completely specified when the observational error variance Σ, the variance matrix $\boldsymbol{W}$ of the errors in the evolution equations and λ are known.

The variance matrix $\boldsymbol{W}$ of the evolution errors in Equation (2) is specified for each t using the discount approach [Ameen & Harrison (1985)]. In this approach the prior variance matrix of the parameters is inflated to reflect the amount of information lost between two consecutive periods. If the model is viewed as one single component model with just one global discount factor β ($0 < \beta \leq 1$), which is the interpretation given in this article, then $\boldsymbol{W}_t = \boldsymbol{G}\boldsymbol{C}_{t\text{-}1}\,\boldsymbol{G}'(1/\beta\text{-}1)$. The analysis of predictive performance, as a function of β, allows the investigation of the improvements that can be obtained by relaxing the usual hypothesis of fixed coefficients. The β hyper-parameter controls the amount of time variation of model's coefficients (note that $\beta = 1$ implies $\boldsymbol{W}_t = \boldsymbol{0}$ so that $\boldsymbol{\theta}_t = \boldsymbol{G}\boldsymbol{\theta}_{t\text{-}1}$ with probability 1).

3.2 The Estimation Procedure

To simplify the notation we include β, the global discount factor, in the set of parameters λ defined in 3.1. Than, once the probabilistic model is defined, for a fixed value of the vector λ, one obtains via Kalman filter the joint distribution of Y_t, $\boldsymbol{\theta}_t$ and $\boldsymbol{\theta}_{t\text{-}1}$ conditioned on $D_{t\text{-}1}$, λ and Σ. The marginal distribution of Y_t follows directly from:

$$(Y_t \mid \lambda,\Sigma, D_{t\text{-}1}) \sim N\,(\boldsymbol{F}'_t\boldsymbol{G}\boldsymbol{m}_{t\text{-}1}, Q_t)$$

where: $Q_t = \boldsymbol{F}'_t\,[\boldsymbol{G}\boldsymbol{C}_{t\text{-}1}\boldsymbol{G}' + \boldsymbol{W}_t]\,\boldsymbol{F}_t + \Sigma$, is the variance of the one step ahead forecast error (e_t), $e_t = y_t - \boldsymbol{F}'_t\boldsymbol{G}\boldsymbol{m}_{t\text{-}1}$.

Now, $\boldsymbol{m}_{t\text{-}1}$ and $\boldsymbol{C}_{t\text{-}1}$ depend on λ and will typically have different values at each point of the λ grid. The density function of the above marginal distribution written in terms of e_t is then given by $Q_t^{-1/2}\ \exp\{-e_t^2\ /\ 2Q_t\ \}$.The log of the likelihood, for the whole sample, is given by the summation across t of the log of the above expression and can be concentrated with respect to Σ given what we call m($\boldsymbol{y}/\lambda$) — the likelihood for λ — in Section 2.1. It is a measure of model's predictive ability which should be maximized with respect to the vector λ of hyper-parameters.

Another measure of predictive performance of the model which does not depend on the probability distribution of the observation error is the Theil's U statistic. It is to be minimized with respect to the vector λ and is given by the following ratio of root mean square forecast errors,

$$\left(\sum_{t=1}^{T} e_t^{\ 2} \Bigg/ \sum_{t=1}^{T} n_t^{\ 2} \right)^{1/2}$$

where n_t is the one step ahead forecast error in the naive model in which the dependent variable follows a random walk. The estimated vector $\boldsymbol{\theta}$ will be conditioned by the chosen value of λ.

3.3 - The Choice or Estimation of Prior's and Monitoring Parameters

The Bayesian approach requires that we specify our prior beliefs to construct the model. In the first step we estimated the variance of inflation that enters Equation (1) or, equivalently, the b_{js} of Equations (3) and (4). This was accomplished indirectly by estimating Equation (1) in a slightly different way. By developing Equations (3) and (4) one substitutes ϕV_t (if ϕ is not time dependent) for a sum of terms of the type $f_{k,m}(\boldsymbol{b},\phi)\ \pi_{t-k}\ \pi_{t-m}$ for $k,m = 0,\ldots,6$. Then, Equation (1) — with the common discount factor fixed to one ($\beta = 1$, as defined in Section 3.1) and a specification where $\phi\ V_t$ is substituted for the sum mentioned above — was estimated.

We have used normal and independent priors for all coefficients of Equation (1). The prior standard deviations for the constant term, seasonal component coefficients and contemporaneous variables coefficients were specified as diffuse priors with zero mean. The prior means for the $f_{k,m}$ coefficients were obtained fixing all b_j to zero, with the exception of b_1 considered equal to 1 [Equation (4) is a random walk] and $\phi = 1$. All priors for the $f_{k,m}$ coefficients were modeled as having the same standard deviation which was estimated maximizing predictive likelihood [Doan et al. (1984)]. The priors for the lagged variable coefficients were specified with zero mean and standard deviation obtained from a tightness parameter, as suggested by Litterman.

For the coefficients of lagged variables, $X_{j,t-k}$, the Litterman prior simplifies the choice of prior standard deviations through the following function:

$$\delta(j,k) = (\sigma_1/\sigma_j)\ \gamma\ g(k) \qquad (5)$$

for $j = 1,2,3,4$ and $k = 1,2,3$ where σ_j is the estimated standard deviation of residuals of an ordinary least squares regression of the variable j on a constant and 3 of its own lagged values (Litterman, 1986). The function $g(k)$ describes the tightness on the $k^{\underline{th}}$ lag with respect to the first one and γ is the overall tightness parameter. Since the variables in the model have different magnitudes the ratio of standard errors $(\sigma_1/\sigma_j\ )$ is a correction for the scale that adjusts for the units in which the data are measured. Any choice of the function $g(k)$ is expected to reflect a prior certainty about the mean of the lagged variables coefficients that increases with the number of lags. In this paper we used the function $g(k) = k^{-1/2}$, as suggested in Doan (1990) thus tightening the prior distribution with a harmonic series for the standard deviation as the lag increases.

In the first step, the vector λ described in the last section, is composed of two parameters : the prior standard deviation for the $f_{k,m}$ coefficients and tightness for the lagged variables coefficients. Using the maximum likelihood method described in last section, we estimated the following values : 29.08 for the prior standard deviation and 10.84 for the tightness. The implied prior standard deviation is simply $10.84k^{-1/2}$ for the coefficients of lagged $M1$ and $10.84k^{-1/2}$ scaled by (σ_1/σ_j) for the coefficients of the other lagged variables, for j = 2,3,4. By fixing these two parameters at their estimated values we were able, using the Kalman filter, to construct an estimate for ϕV_t .

In the second step, we investigate possible values for the global discount factor (β).The discount factor could be specified using the notion of loss of information over time, setting $\beta = (3n-1)/(3n+1)$, where n is the number of periods that it takes to lose half of the information [Harrison & Johnston (1984)]. In this paper however this approach was not used and we actually estimate β.This was accomplished by substituting ϕV_t, in Equation (1), for its estimated value times a new parameter (that we hope will have an estimated value of 1). The prior for the new parameter is diffuse with zero mean. All other prior's parameters where fixed at the same values of the first step.

In the second step $\lambda = \beta$. We used the likelihood function $m(\boldsymbol{y}/\lambda)$, to estimated β (β=0.9998), searching for the likelihood maximizing value. Although the usual hypothesis of constant coefficients may be relaxed, the amount time variation detected in the coefficients is very small.

Finally, we used the monitoring techniques described in Section 2.2, to investigate the occurrence of structural changes in the money demand series. We used ρ = 0.01 which implies accumulation of evidence against the standard model whenever $|e_t| \geq 2.16$. This is the standard normal distribution's 0.985 order percentile and we expect to accumulate evidence against the standard model 3% of the time. For this chosen value of ρ any standardized forecast errors greater than 2.80, in absolute value, should be signaled as discrepant (relative to the standard model) or, equivalently, that τ = 0.21. It means that we expect the monitor to signal less than 1% of the time.

Once the monitor indicates a change, the answer is to increase model uncertainty in that period. The adopted procedure increases the prior variance of model parameters, by reducing the global discount factor. This implies putting more weight on more recent observations when estimating those parameters. In this case we say that an intervention was made in the process.

4 - RESULTS

In this section we present the results for two types of estimation methods: Bayesian and Classical. The Bayesian method use the prior structure and the two steps procedure mentioned in Section 3.2. The Classical method use diffuse priors for all model parameters and for the two steps discussed in Section 3.2. In both methods, Equation (1) — with the term $\phi\ V_t$ substituted for its estimated value times a new parameter (see Section 3.2 for a detailed description) — is estimated.

4.1 - Results for the Bayesian Method

At first, we did not make any intervention (using $\beta = 0.9998$) , thus allowing the periods with model break down to be observed. This is illustrated in Figure 1(a), which shows the logged cumulative Bayes factor; the evidence against (or in favor) of the standard model is being accumulated additively. Later, we made an intervention sequentially when the monitor signals and Figure 1(b) shows the resulting logged cumulative Bayes factor.

Our approach was to try different interventions and choose, sequentially, the ones that increased the predictive likelihood. We could not find such an intervention for the third quarter of 1980 but we found an increase in the likelihood when dividing the global discount factor by 4.5 in the second quarter of 1986. This prevented the monitor from signaling again in the second quarters of 1990 and 1995.

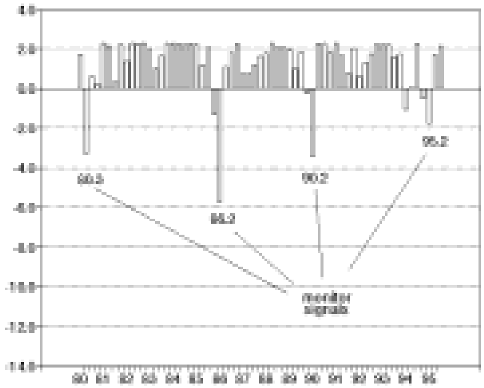
Figures 1(c) and 1(d) show the same exercise but now dropping the variance of inflation from the model (keeping, nevertheless, β=0.9998). As before the monitor signals in the third quarter of 1980 but we could not find a suitable intervention. The next period of model break down is the first quarter of 1986. A successful intervention dividing the global discount factor by 10 increased the likelihood, even so another break is detected in the fourth quarter of 1989, when another intervention — dividing the global discount factor by 2,5 — is called for.

Figures 2 and 3 show the effects of the intervention, in the model with the expected variance of inflation, on the constant term coefficient, short-run elasticities of money demand with respect to inflation rate, GNP and real interest rate in terms of both nonsmoothed and smoothed trajectories respectively. The trajectories of the short-run elasticities of $M1$ with respect to inflation and real interest rates were the most affected ones. Long-run elasticities are shown in Figures 4 and 5 (nonsmoothed and smoothed respectively).

Figure 1
Sequential logged Bayes factors for the models with and without variance of inflation. Bayesian model, discount factor is 0.9998 except in intervention periods.

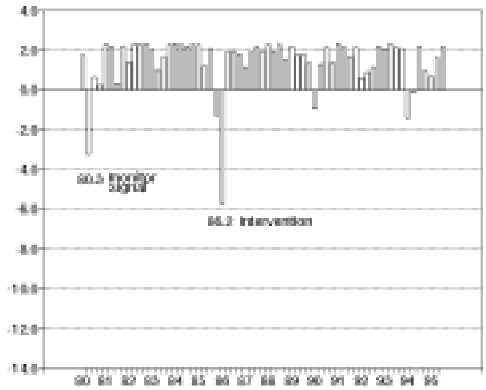
WITH VARIANCE OF INFLATION

(a)



no intervention

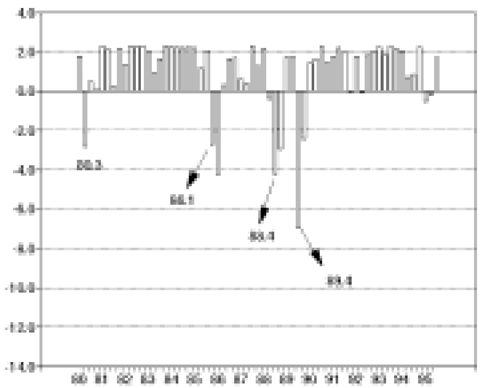
(b)



with intervention

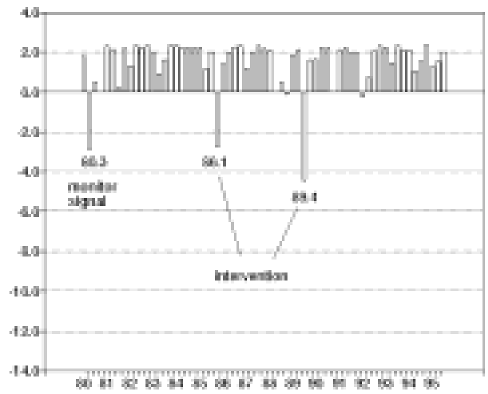
WITHOUT VARIANCE OF INFLATION

(c)



no intervention

(d)

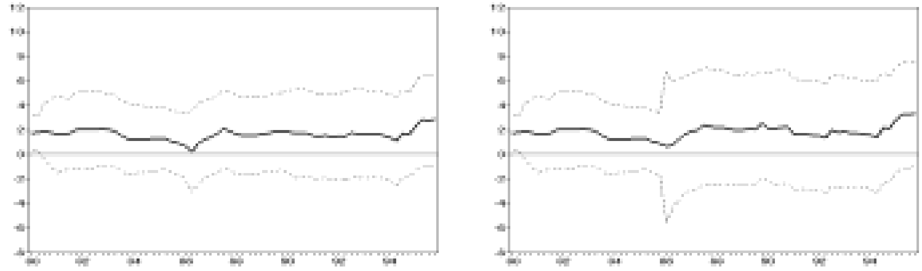


with intervention

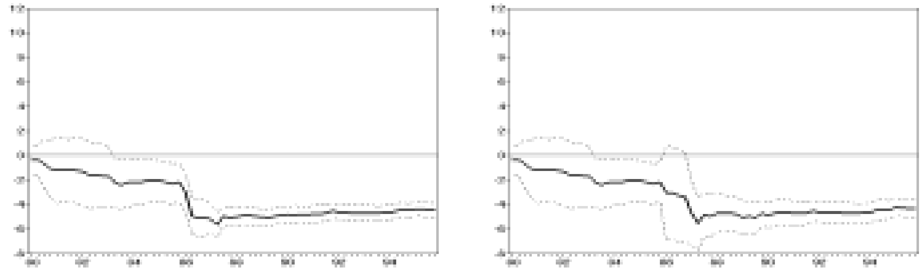
Figure 2
Nonsmoothed trajectories of the constant term coefficient, short-run elasticities and the variance of inflation coefficient. Without and with intervention in 86.2.

Bayesian Model

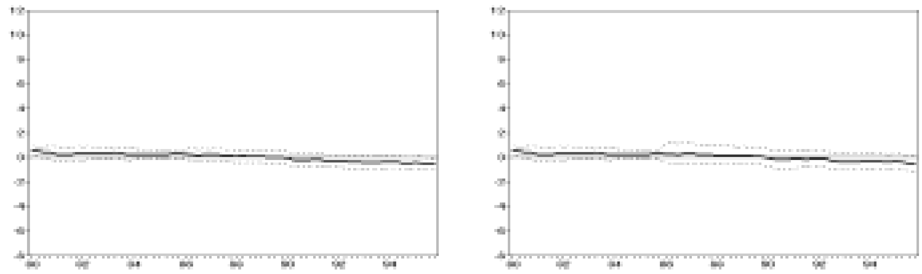
constant term



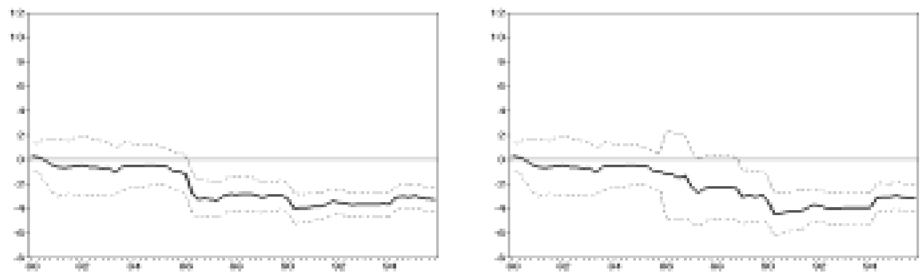
Short-Run Elasticity - Inflation



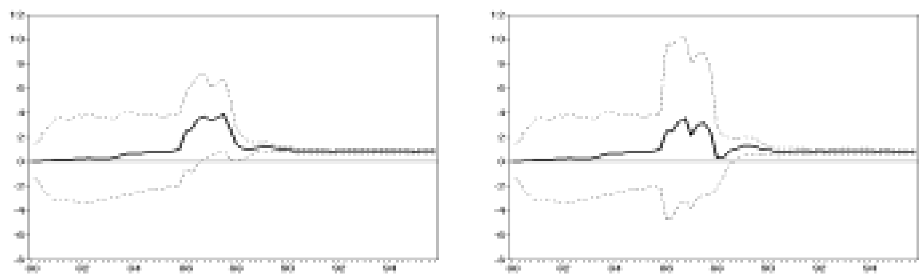
Short-Run Elasticity - GNP



Short-Run Elasticity - real interest rate



variance of inflation

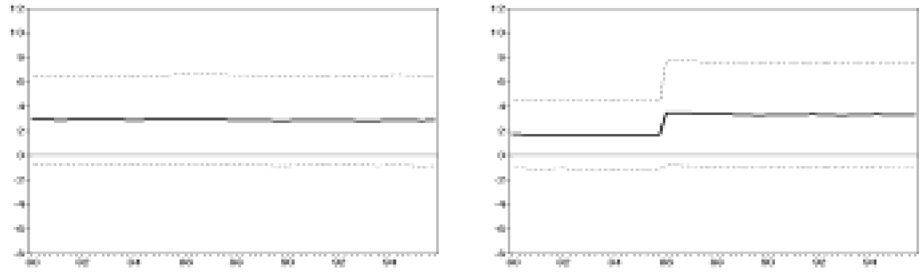


no intervention intervention

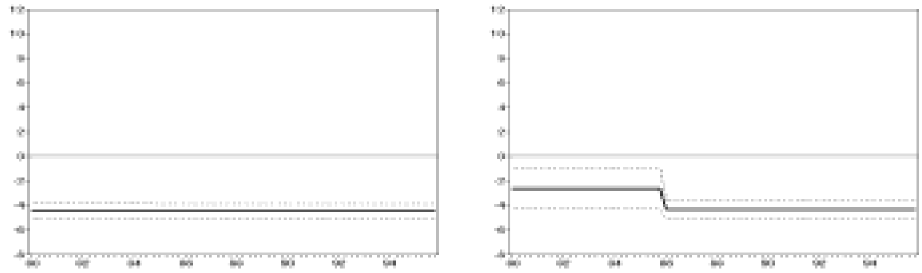
Figure 3
Smoothed trajectories of the constant term coefficient, short-run elasticities and the variance of inflation coefficient. Without and with intervention in 86.2.

Bayesian Model

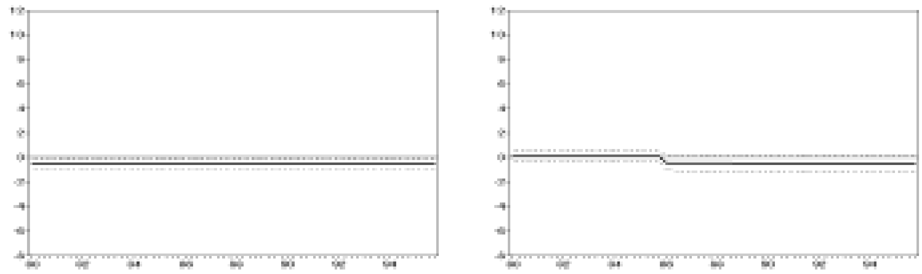
constant term



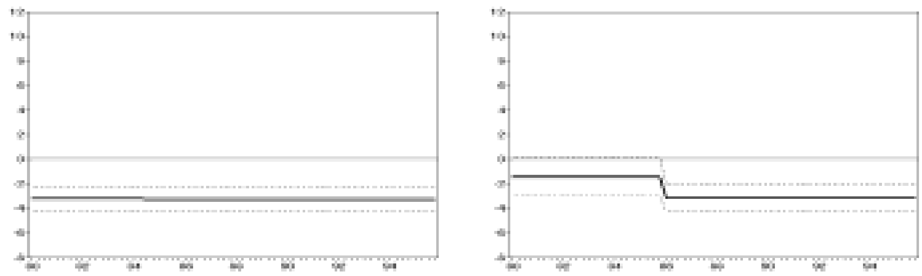
Short-Run Elasticity - Inflation



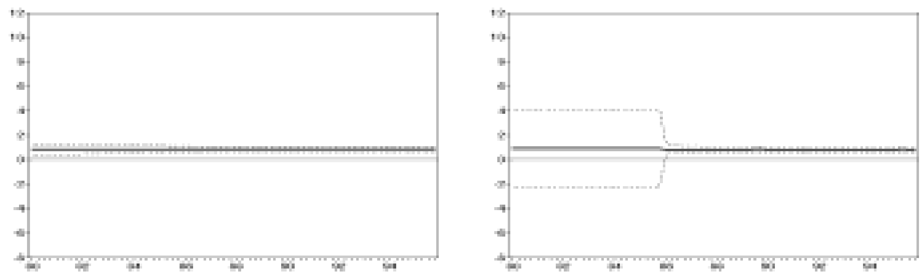
Short-Run Elasticity - GNP



Short-Run Elasticity - Interest



Variance of Inflation



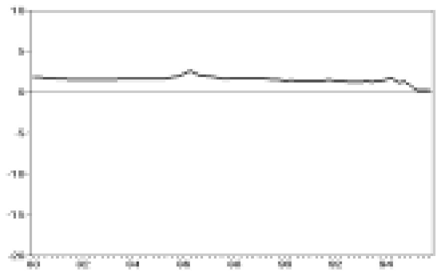
no intervention intervention

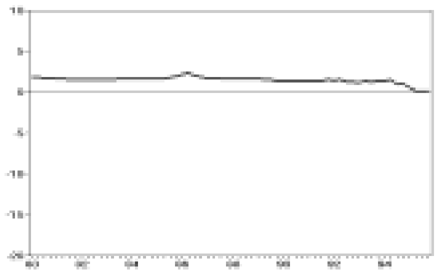
Figure 4
Nonsmoothed trajectories of long-run elasticities. Without and with intervention in 86.2.

Bayesian Model

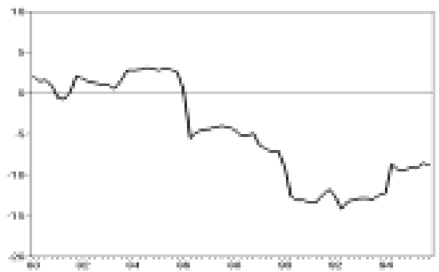
Long-Run Elasticity - Inflation

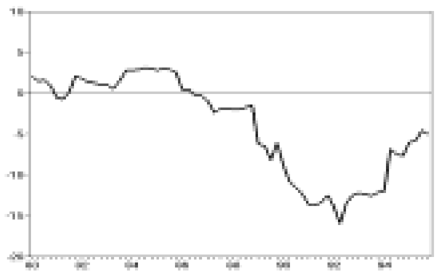
Long-Run Elasticity - GNP



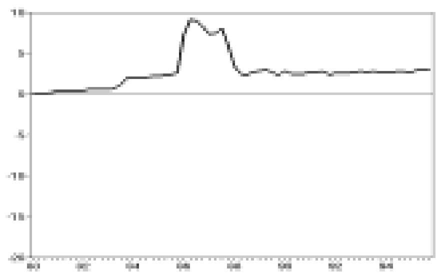


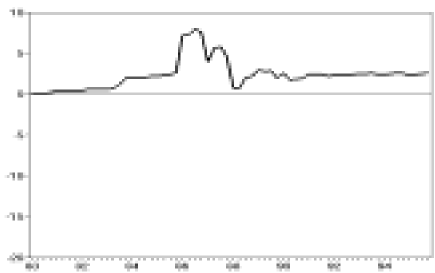
Long-Run Elasticity - Real interest rate





Variance of Inflation



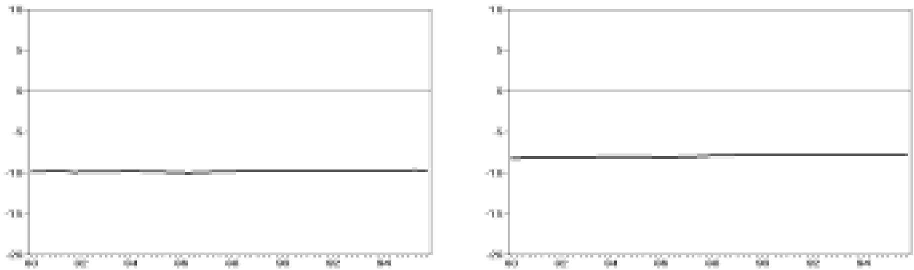


no intervention intervention

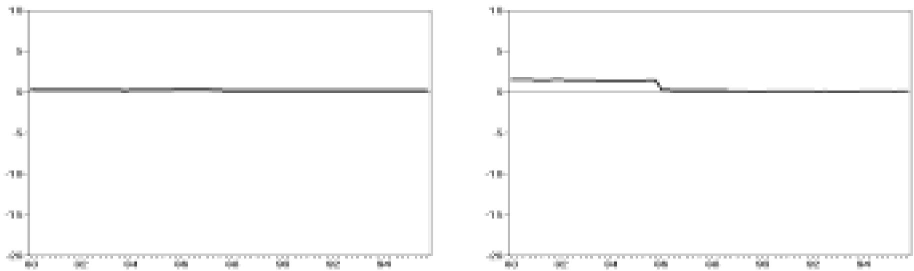
Figure 5
Smoothed trajectories of long-run elasticities. Without and with intervention in 86.2.

Bayesian Model

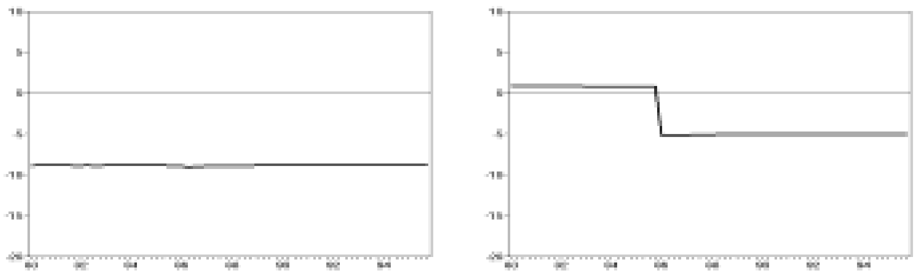
Long-Run Elasticity - Inflation



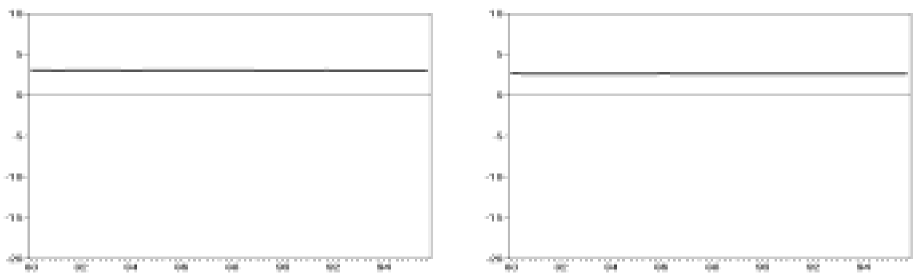
Long-Run Elasticity - GNP



Long-Run Elasticity - Interest



Long-Run Elasticity - Variance of Infation



no intervention intervention

4.2 - Results for the Classical Method

At first we fixed the global discount factor to 1 (β = 1) and, as in the Bayesian case, we did not make any intervention, thus allowing the periods with model break down to be observed. This is illustrated in Figure 6(a), which shows the logged cumulative Bayes factor, and the evidence against (or in favor) of the standard model is being accumulated additively. Later, we made interventions sequentially when the monitor signals and Figure 6(b) shows the resulting logged cumulative Bayes factor.

We chose the interventions that increased the predictive likelihood. We could not find such an intervention for the third quarter of 1980 but we found an increase in the likelihood when dividing the global discount factor (β = 1) by 1.1 in the third quarter of 1986 and by 7 in the first quarter of 1990.

Figures 6(c) and 6(d) show the same exercise but now dropping the variance of inflation from our model. As before the monitor signals in the third quarter of 1980 but we could not find a suitable intervention. We could only find interventions that increased the predictive likelihood in 86,3 , 89,4 and 92,1. For these years the global discount factor (β = 1) was divided, respectively, by 2 , 10 and 1.1

Figures 7 and 8 show the effects of the intervention, for the model with the variance of inflation, on the constant term coefficient, short-run elasticities of money demand with respect to inflation rate, GNP and real interest rate in terms of both non-smoothed and smoothed trajectories respectively. Again, the trajectories of the short-run elasticities of *M*1 with respect to inflation and real interest rates were the most affected ones. Long-run elasticities are shown in Figures 9 and 10 (nonsmoothed and smoothed respectively).

5 - CONCLUSIONS

Our main conclusions are :

- the inclusion of the expected variance of inflation in the money equation increases the stability of the demand for money after 1986 (using Classical or Bayesian estimation procedures);

- even after the inclusion of the expected variance of inflation the money demand shows a structural and permanent brake in 1986;

- there is evidence of a permanent increase in money demand elasticities, with respect to inflation and real interest rate, after 1986;

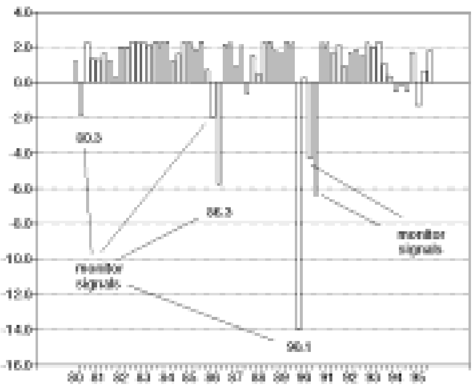
- the increase in money demand elasticities is not altered by the Real Plan till the end of 1995, the period covered by our analysis.

Figure 6

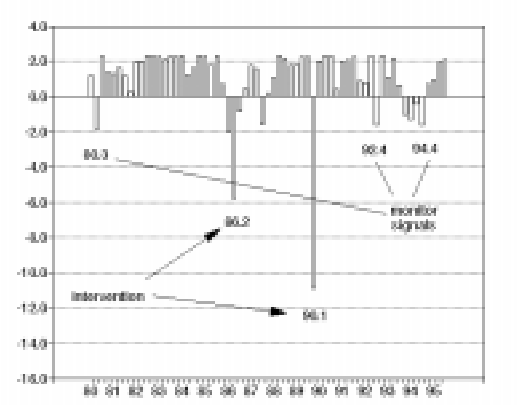
Sequential logged Bayes factors for the models with and without variance of inflation. Classical model, discount factor is 1 except in intervention periods.

WITH VARIANCE OF INFLATION

(a)

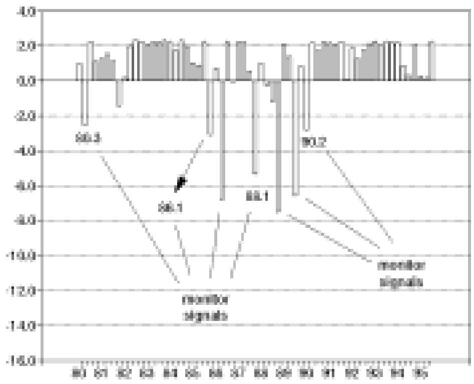


(b)



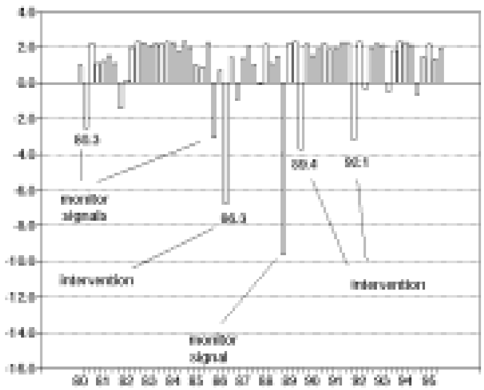
WITHOUT VARIANCE OF INFLATION

(a)



no intervention

(b)

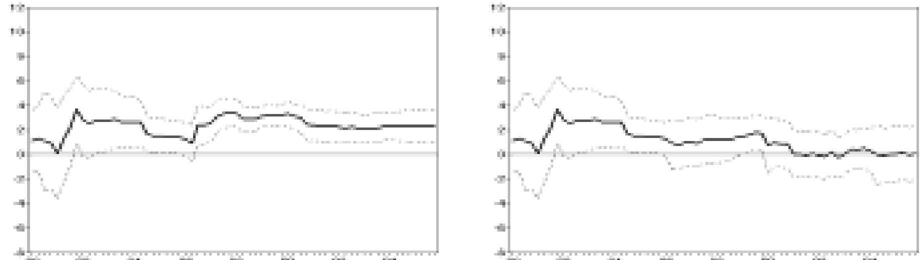


with intervention

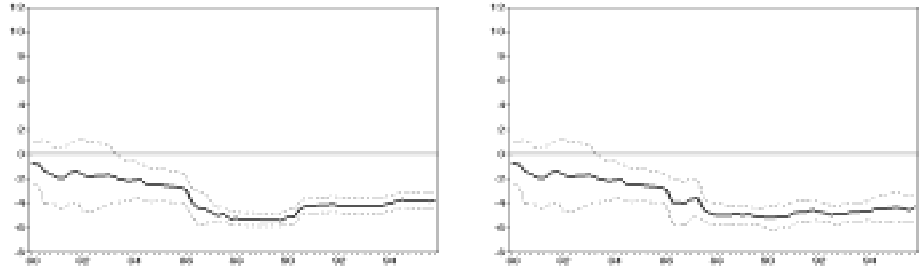
Figure 7
Nonsmoothed trajectories of the constant term coefficient, short-run elasticities and the variance of inflation coefficient. Without and with intervention (classical model).

Classical Model

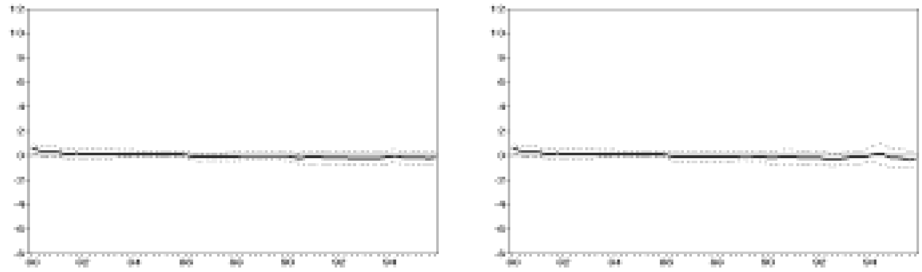
constant term



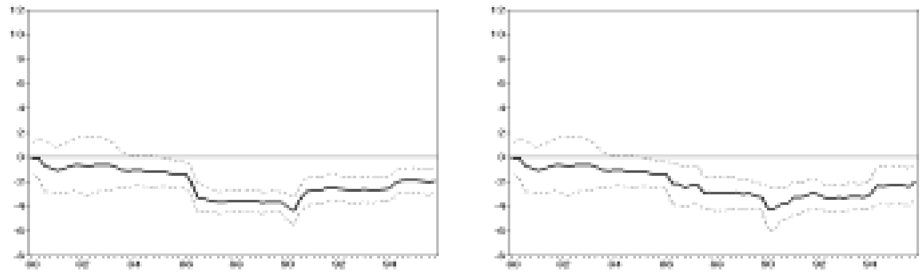
Short-Run Elasticity - Inflation



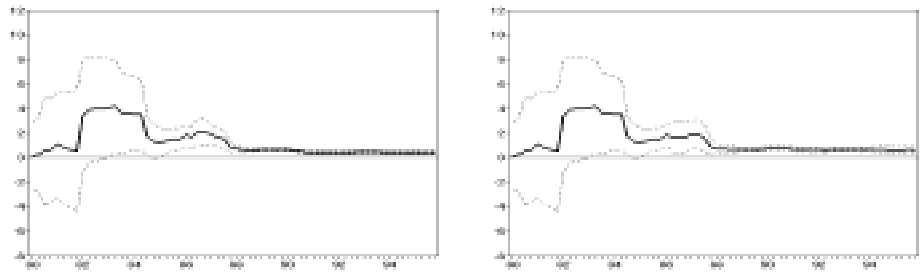
Short-Run Elasticity - GNP



Short-Run Elasticity - real interest rate



variance of inflation



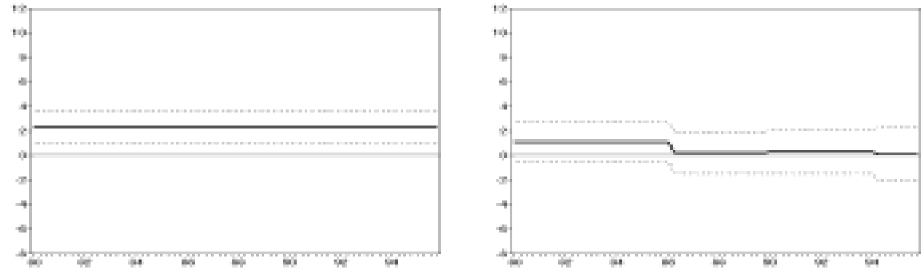
no intervention intervention

Figure 8

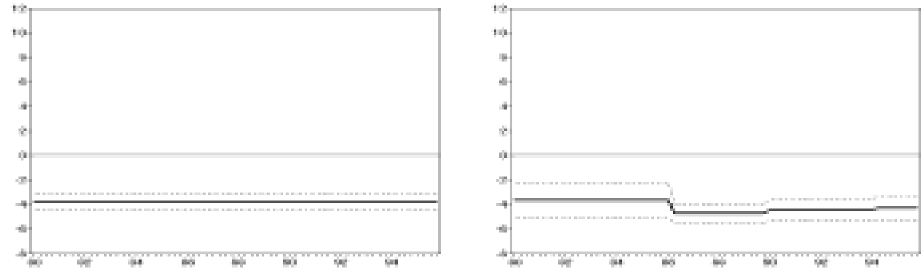
Smoothed trajectories of the constant term coefficient, short-run elasticities and the variance of inflation coefficient. Without and with intervention.

Classical Model

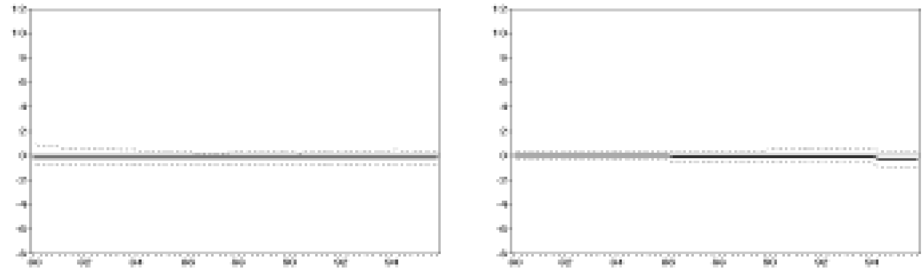
constant term



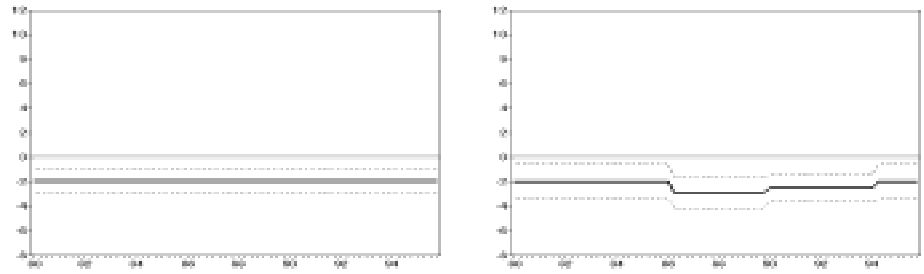
Short-Run Elasticity - Inflation



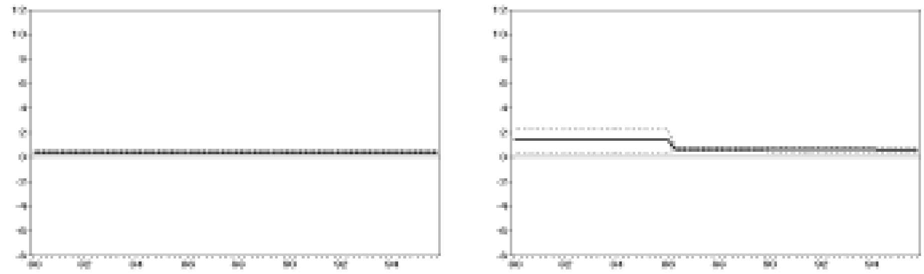
Short-Run Elasticity - GNP



Short-Run Elasticity - Interest



Variance of Inflation

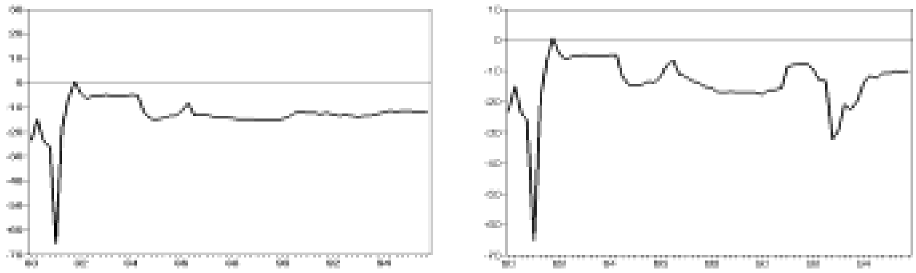


no intervention intervention

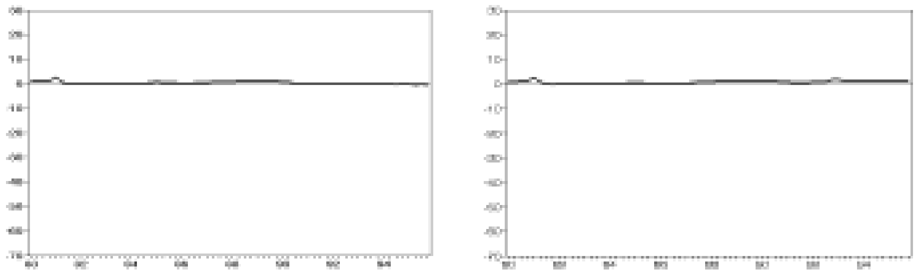
Figure 9
Nonsmoothed trajectories of long-run elasticities. Without and with intervention.

Classical Model

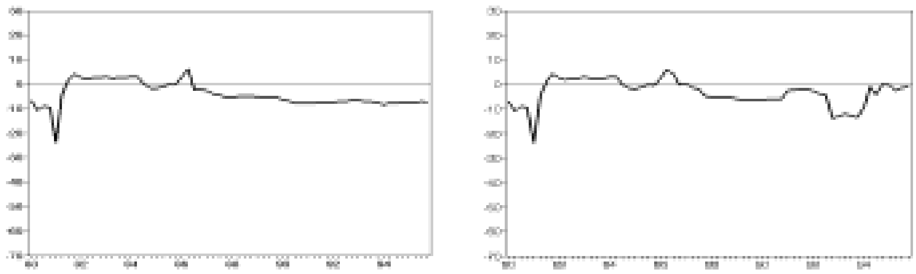
Long-Run Elasticity - Inflation



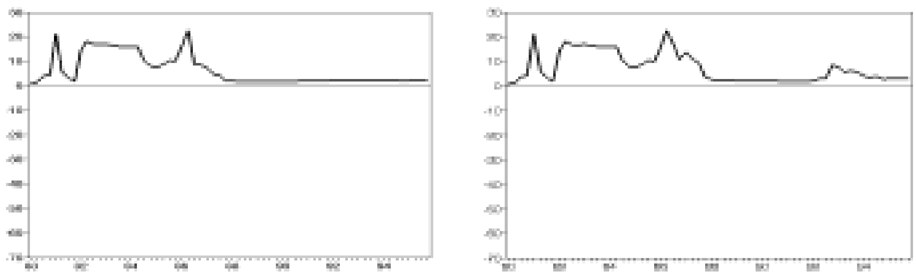
Long-Run Elasticity - GNP



Long-Run Elasticity - Real interest rate



Variance of Inflation

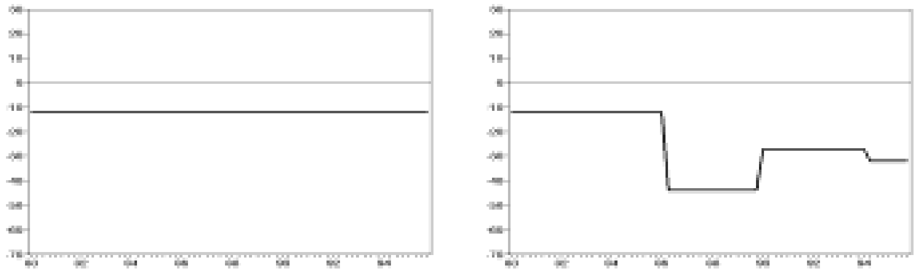


no intervention intervention

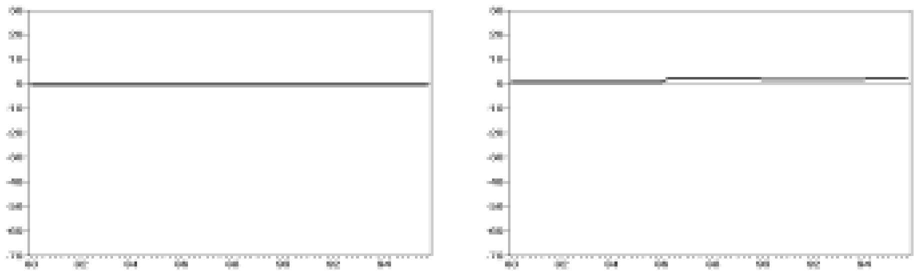
Figure 10
Smoothed trajectories of long-run elasticities. Without and with intervention.

Classical Model

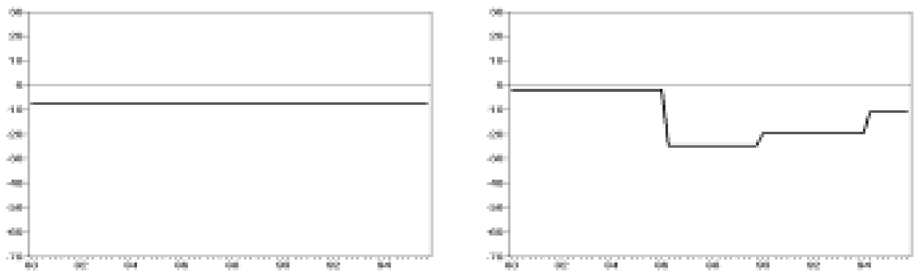
Long-Run Elasticity - Inflation



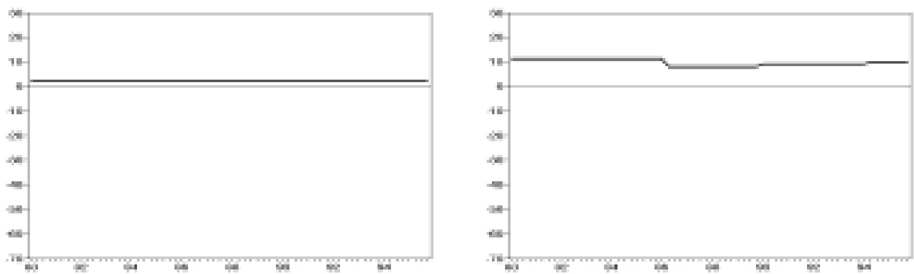
Long-Run Elasticity - GNP



Long-Run Elasticity - Interest



Long-Run Elasticity - Variance of Inflation



no intervention intervention

BIBLIOGRAPHY

AMEEN, J. R. M. and HARRISON, P. J. Normal Discount Bayesian Models. In: BERNARDO, J. M. **et alii. Bayesian Statistics 2.** Valencia University Press, 1985.

CALOMIRIS and DOMOWITZ. Asset Substitutions, Money Demand and the Inflation Process in Brazil. **Journal of Money, Credit and Banking**, 1989.

DOAN, T., LITTERMAN, R. and SIMS, C. Forescasting and Conditional Projections Using Realistic Prior Distributions. **Econometric Review**, 3 (1): 1-100, 1984.

GOLDFELD S. M. The Case of the Missing Money. **Brookings Papers on Economic Activity**, n. 3: 588-635, 1976.

HARRISON, J. and JOHNSTON, F.R. Discount Weighted Regression. **Journal of the Operational Research Society**, 35(10): 923-932, 1984.

LIMA, E. C. R. and EHLERS, R. S. A Estabilidade da Demanda por Moeda no Brasil: Uma Abordagem Bayesiana. **Anais do XV Encontro Brasileiro de Econometria**, 1993.

LITTERMAN, R. B. Forecasting Bayesian Vector Autoregressions - Five Years of Experience. **Journal of Business and Economic Statistics**, 4: 25-38, 1986.

NAKANE, M.I. Testes de Exogeneidade Fraca e de Superexogeneidade para a Demanda por Moeda no Brasil. **16º Prêmio BNDES de Economia**, Rio de Janeiro, 1994.

PASTORE, A. C. Déficit Público e Inflação- Uma Resenha. **Unpublished Mimeo,** 1991.

PEREIRA, P. L. V. Empirical Analysis of Brazilian Money Demand (1966-1987) an Application of Cointegration Methods. **Anais do XI Encontro Brasileiro de Econometria**, 317-333, 1989.

ROSSI, J. W. A Demanda por Moeda no Brasil - O que Ocorreu a Partir de 1980?. **Pesquisa e Planejamento Econômico**, 18(1): 37-54, 1988.

ROSSI J. W. Comportamento de Agregados e Multiplicadores Monetários no Brasil. **Revista Brasileira de Economia**, Rio de Janeiro, 44(2): 217-236, 1990.

TOURINHO, O. A. F. 1995. A Demanda por Moeda em Processos de Inflação Elevada, **Pesquisa e Planejamento Econômico**, abril de 1995 (forthcoming).

TRICHES, D. Demanda por Moeda no Brasil e a Causalidade entre as Variáveis Monetárias e a Taxa de Inflação:1972/87. **16º Prêmio BNDES de Economia**, Rio de Janeiro, 1992.

WEST, M. Bayesian Model Monitoring. **Journal of the Royal Statistical Society** *B* 48: 70-78, 1986.

WEST M. and HARRISON J. **Bayesian Forecasting and Dynamic Models**, New York, Springer-Verlag, NY, 1989.

__________. Monitoring and Adapting in Bayesian Forcasting Models. **Journal of the American Statistical Association** 81: 741-750, 1986.

Ipea – Institute for Applied Economic Research

PUBLISHING DEPARTMENT

Coordination
Cláudio Passos de Oliveira

Supervision
Everson da Silva Moura
Reginaldo da Silva Domingos

Typesetting
Bernar José Vieira
Cristiano Ferreira de Araújo
Daniella Silva Nogueira
Danilo Leite de Macedo Tavares
Diego André Souza Santos
Jeovah Herculano Szervinsk Junior
Leonardo Hideki Higa

Cover design
Luís Cláudio Cardoso da Silva

Graphic design
Renato Rodrigues Buenos

The manuscripts in languages other than Portuguese published herein have not been proofread.

Ipea Bookstore

SBS – Quadra 1 – Bloco J – Ed. BNDES, Térreo
70076-900 – Brasília – DF
Brazil
Tel.: + 55 (61) 3315 5336
E-mail: livraria@ipea.gov.br

Composed in Adobe Garamond 11/13.2 (text)
Frutiger 47 (headings, graphs and tables)
Brasília – DF – Brazil

Ipea's mission
Enhance public policies that are essential to Brazilian development by producing and disseminating knowledge and by advising the state in its strategic decisions.

ISSN 1415-4765

ipea Institute for Applied Economic Research

Secretariat of Strategic Affairs

BRAZILIAN GOVERNMENT BRASIL