TIME-VARYING COEFFICIENTS IN A GMM FRAMEWORK: ESTIMATION OF A FORWARD LOOKING TAYLOR RULE FOR THE FEDERAL RESERVE.

Harry Partouche

September 2007

NER - E # 177
TIME-VARYING COEFFICIENTS IN A GMM FRAMEWORK:
ESTIMATION OF A FORWARD LOOKING TAYLOR RULE FOR THE FEDERAL RESERVE.

Harry Partouche

September 2007

NER - E # 177
Time-Varying Coefficients in a GMM Framework: Estimation of a Forward Looking Taylor Rule for the Federal Reserve¹

Harry Partouche²

¹I would like to thank the attendants at the research seminar of the Banque de France for their helpful remarks. I would also like to thank the Banque de France for the year I spent in its research staff. The views expressed are those of the author and not necessarily those of the Banque de France.

²DGTEP, on an internship with DGEI-DIR-POMONE at the Banque de France when the article was prepared. Email: harry.partouche@dgtpe.fr
Abstract

This article deals with the estimation of a time-varying coefficients equation with endogenous regressors. A non-parametric approach is proposed, combining the Generalized Method of Moments (GMM) with the smoothing splines literature as in Hodrick and Prescott (1981).

This new method is used to analyze the evolution of a forward-looking Taylor rule for the Federal Reserve (FED) from 1960 until 2006. It suggests that monetary policy accommodated inflation during the 60s and the 70s whereas the chairmanship of P. Volcker was a turning point toward a more aggressive stance on inflation. In addition, monetary policy became more and more countercyclical.

JEL classification: E5; C14; C32

Keywords: Monetary policy rules; Generalized Method of Moments; Time-varying coefficients; Smoothing splines

Résumé

Ce papier traite de l’estimation d’une équation linéaire à coefficients variables dans le temps et dont les erreurs sont corrélées aux variables explicatives. On développe une méthode d’estimation non paramétrique qui emprunte à la fois à la Méthode des Moments Généralisée (MMG) et à celle des "smoothing splines" type Hodrick-Prescott (1981).

Dans la lignée des travaux de Taylor (1993), on applique cette méthodologie à l’étude de l’évolution d’une règle de politique monétaire forward-looking pour la FED. Au niveau des résultats, la réponse de long terme du taux d’intérêt à l’inflation est croissante au cours de la période. En particulier, on montre que la politique monétaire est accommodante jusqu’à la fin des années 1970. La réponse à l’écart de production croit elle aussi continuellement, ce qui constitue un résultat nouveau.

Codes JEL : E5; C14; C32

Mots Clés : Règle de politique monétaire; Méthode des moments généralisés; Equation à coefficients aléatoires; Smoothing splines
Non-technical summary

Taylor rules have now become very popular to describe the decision making of central bankers. They do not only succeed in terms of econometric estimation but they are also relevant for a theoretical use, for instance to ensure the uniqueness of the equilibrium in DSGE models. However, monetary policy is one of the numerous fields subject to variations across time. Central bankers change as well as the understanding of the economy or even the aims of monetary policy. The policy implications of such time variations are important. There is a debate among economists to know whether the poor economic performance in the United States during the 70s was just the result of adverse shocks or if it was policy related, debate known as bad luck vs bad policy.

Many authors have studied the possible time variation in Taylor rules. Clarida, Gali and Gertler (2000) estimate monetary policy rules on subperiods they assume to be homogenous. Kim and Nelson (2006) go further, allowing for continuous changes in the coefficients of the rule. Our stand is to estimate such time-varying coefficients. To do so, I try to impose as few priors as possible. Especially, I do not have any prior, neither on the rule at the beginning of our sample nor on the form of the evolution. Especially, we do not impose the time variation to be a structural break.

The results confirm that monetary policy accommodated inflation during the 60s and the 70s and that the appointment of P.Volcker was a turning point toward a more aggressive stance on inflation. The profile of the long term response of the interest rate to the inflation is very close to a structural break. At the same time, the rule became more and more countercyclical but in a continuous way.
Résumé non technique

L'utilisation de règles de Taylor s'est généralisée dans l'analyse de la politique monétaire. Elles sont utiles non seulement pour l'estimation de règles de politique monétaire mais jouent aussi un rôle théorique central, notamment pour assurer la détermination locale de l'équilibre dans les modèles DSGE. Cependant, la politique monétaire est sujette à des variations importantes au cours du temps. Les banquiers centraux se succèdent, la connaissance du fonctionnement de l'économie s'améliore et les objectifs assignés à la politique monétaire changent. Ces évolutions ont des conséquences importantes sur la politique économique. Ainsi, les économistes cherchent à savoir si la politique monétaire a joué un rôle dans les performances économiques décevantes des années soixante-dix aux États-Unis ou si elles étaient entièrement liées à des chocs adverses.


Les résultats confirment le caractère accommodant de la politique monétaire pendant les années soixante et soixante-dix ainsi que le tournant provoqué par la nomination de P. Volcker. En particulier, l'évolution de la réponse de long terme du taux d'intérêt à l'inflation est très proche d'une rupture structurelle. Dans le même temps, la politique monétaire est devenue de plus en plus contracyclique mais de manière continue.
1 Introduction

This paper deals with the estimation of a time-varying coefficients equation. In what follows, these coefficients are assumed to follow a random walk, allowing for permanent changes. Moreover, I suppose that the error term is correlated to the explanatory variables, which forbids the use of the Kalman filter. This endogeneity problem arises because of the presence of expected values in the regressors: under the rational expectations hypothesis, these expectations are replaced with future variables.

I use this procedure to assess the way the FED has set its interest rate over the past 40 years. Simple monetary policy rules depending on expected inflation and output gap with some degree of interest rate smoothing fit the data quite well as documented in Clarida, Gali and Gertler (2000) for instance. Nevertheless, when the time period is too long, it is difficult to rely on a constant rule. Changes may occur because of a better economic knowledge or changes in the doctrine or even structural changes in the economy. For instance, Romer and Romer (2002) see changes in the understanding of the functioning of the economy as the main reason for the changes in the way monetary policy was conducted. Many authors studied the evolution in the conduct of monetary policy, to assess whether the poor economic performance in the United States during the 70s was the result of bad luck or bad policies. Sims and Zha (2006) or Primiceri (2005) find that the changes in the volatility of the shocks were the main source of time variation. Cogley and Sargent (2005) account for both sources of time variation.

Clarida, Gali and Gertler (2000) use the GMM framework to estimate monetary policy rules on subperiods they assume to be homogenous. They find that monetary policy accommodated inflation in the 60s and the 70s and a long term response of the interest rate to future inflation close to two under P.Volcker or A.Greenspan. The choice of the break date may seem a bit arbitrary and there is a risk to carry out an estimation on too short time periods.

Kim and Nelson (2006) estimate monetary rules with a two-step procedure close to the two stages least squares (2SLS) allowing for the use of the Kalman filter. Their approach, described in appendix A.2, allows for heteroskedasticity of the monetary policy shock and thus time periods during which discretion increased. However, their results may be questioned on several grounds: they only display results of filtered estimates instead of smoothed estimates; initial conditions are estimated with the 2SLS on 40 quarters. The impact of initial conditions is decreasing when time goes on but, since their sample begins in 1970, their results, at least until 1980, are sensitive to this fragile prior. They

---

3The likelihood is computed from the prediction error decomposition based on the forward step of the Kalman filter. Once one has the parameters that maximize the likelihood, it is easy to compute filtered estimates but also smoothed estimates based on the backward recursion of the Kalman filter. Smoothed estimates are hence based on all the information available.
account for this uncertainty in the initial variance of the state vector but, since they use filtered estimates, the first years depend strongly on this prior whereas the use of smoothed estimates would have partly solved this issue. One can notice that there is another solution to estimate initial conditions: the use of generalized least squares as in Lemoine and Pelgrin (2003). Finally, they choose a particular form of correlation between the regressors and the error term and a particular form of heteroskedasticity of the error term, namely a GARCH(1,1).

The use of real-time forecasts, as Boivin (2006), is another way to get round the endogeneity problem. The model is then a standard state-space problem. Boivin (2006) uses real-time forecasts provided by the Greenbook. This real-time approach raises some questions. The unemployment gap is negative on the whole part of the sample\(^4\). As the state vector follows a random walk, Boivin (2006) chooses some particular initial conditions, a long term response of the interest rate to inflation equal to 2 and a long term response of the interest rate to the unemployment gap of 2.5 in 1970. Finally, the results seem to depend on how heteroskedasticity is handled.

I suggest a new estimation strategy which aims at minimizing the empirical counterpart of moment conditions subject to the coefficients evolution. My approach belongs to the smoothing splines litterature in line with the seminal paper by Hodrick and Prescott (1981). I get a whole class of estimates depending on the weight put on the empirical couterpart of the moment conditions and on the amount of time variation allowed in the coefficients. It is then possible to choose the estimate that minimizes the distance between the true coefficients and the estimated ones and which is locally optimal too. My method makes minimal assumptions on initial conditions or on the error terms. The results show the accommodative pattern during the 60s and the 70s and the turning point at the end of the 70s when P. Volcker was appointed chairman of the Board. As opposed to Kim and Nelson (2006), the results suggest a countercyclical feature increasing over the sample.

Section 2 describes the approach, section 3 computes the estimates and derives the optimality conditions and section 4 illustrates the implementation of this methodology on the FED’s behaviour between 1960 and 2006.

\(^4\)Orphanides (2004) notices that the output gap is below zero on the most part of the period, which means that forecasters had always anticipated an economy below its potential. Boivin (2006) uses rather the unemployment gap defined as the difference between the natural rate of unemployment and the expected unemployment rate, setting the natural rate of unemployment equal to the historical average of unemployment in the past. This choice gives data similar to Orphanides (2004). One can notice that such data does not verify the rational expectations hypothesis.
2 Econometric strategy

2.1 Equation of interest and estimation strategy

This paper focuses on estimating a forward-looking Taylor rule:

\[ i_t = a_{0,t} + a_{x,t}E_{t-1} (\pi_{t+h}) + a_{y,t}E_{t-1} (y_{t+l}) + a_{i,t}i_{t-1} + \omega_t \]

(1)

where \( i_t, \pi_t, y_t \) and \( \omega_t \) stand for the Federal Funds Rate, the inflation at an annual rate, the output gap and a monetary policy shock. The vector of coefficients \( \delta_t = (a_{0,t}, a_{x,t}, a_{y,t}, a_{i,t})' \) follows a random walk:

\[
\begin{align*}
\delta_t &= \delta_{t-1} + u_t \\
u_t & \sim \mathcal{N}(0, Q_u)
\end{align*}
\]

From the notations, (1) contains expectations \( E_{t-1} (\cdot) = E (\cdot | I_{t-1}) \) conditionnal on all the information set available at the beginning of date \( t \) which consists of the interest rate, the output gap, the inflation, and other variables until \( t-1 \). In other words, when the central bank takes its decision at the beginning of the quarter, all the current variables are unknown. Under the rational expectations hypothesis, (1) can be expressed as:

\[ i_t = a_{0,t} + a_{x,t} \pi_{t+h} + a_{y,t} y_{t+l} + a_{i,t} i_{t-1} + v_t \]

(2)

with \( v_t = -a_{x,t} (\pi_{t+h} - E_{t-1} (\pi_{t+h})) - a_{y,t} (y_{t+l} - E_{t-1} (y_{t+l})) + \omega_t \). The error term \( v_t \) in (2) is thus correlated to explanatory variables \( \pi_{t+h} \) and \( y_{t+l} \) for two possible reasons: first, forecast errors \( \pi_{t+h} - E_{t-1} (\pi_{t+h}) \) and \( y_{t+l} - E_{t-1} (y_{t+l}) \) are correlated with \( \pi_{t+h} \) and \( y_{t+l} \); second, if the monetary policy shock \( \omega_t \) affects the real economy between \( t \) and \( t + h/t + l \), it is correlated with \( \pi_{t+h} \) and \( y_{t+l} \) too.

In the following, \( x_t = i_t \) is the endogenous variable, here the interest rate. The \( K \) explanatory variables are \( z_t = (1, \pi_{t+h}, y_{t+l}, i_{t-1}) \) and the time-varying coefficients are \( \delta_t = (a_{0,t}, a_{x,t}, a_{y,t}, a_{i,t})' \). \( K \) instrumental variables \( Z_t \in I_{t-1} \) are used to deal with the endogeneity source in (2). If the coefficients were set equal to some \( \delta \), the dynamic GMM would apply:

\[ E (x_t - z_t'\delta | I_{t-1}) = 0 \Rightarrow E (x_t - z_t'\delta) Z_t = 0 \]

Properties of static GMM are valid and especially the optimality conditions. The 2SLS are used and are equivalent to the GMM when the error term is homoskedastic. To deal with the potential heteroskedasticity of the disturbance term, the Three Stages Least Squares (3SLS) are applied. The GMM framework allows to estimate a small number of coefficients, here four, with many moment conditions. The weighting matrix \( S \) arbitrates between these moment conditions:

\[
\min_{\delta} \left\{ \left( \sum_{t=1}^{T} (x_t - z_t'\delta) Z_t \right) S \left( \sum_{t=1}^{T} (x_t - z_t'\delta) Z_t \right)' \right\}
\]

(3)
This minimization problem is equivalent to the 2SLS or the 3SLS for an appropriate $S$. Kim and Nelson (2006) modify the two-step approach. On the contrary, I propose to adapt the minimization problem (3).

In my setup, one can still write moment conditions:

$$E(x_t - z_t' \delta_t | I_{t-1}) = 0 \Rightarrow E((x_t - z_t' \delta_t) Z_t) = 0$$

Since there are less moment conditions than the number of unknown coefficients, the amount of time variation of the coefficients has to be limited, which can be written as a constrained minimization problem:

$$\min_{(\delta_t)_{t \in [0,T]}} \left\{ \left( \sum_{t=1}^{T} (x_t - z_t' \delta_t) Z_t \right)' S \left( \sum_{t=1}^{T} (x_t - z_t' \delta_t) Z_t \right) \right\}$$

$$\text{u.c.} \quad \forall k \in [1, K], \sum_{t=1}^{T} (\delta_{k,t} - \delta_{k,t-1})^2 = TQ_{emp,kk}$$

where $Q_{emp}$ is the empirical covariance matrix of $\Delta \delta_t$. Program (4) may appear arbitrary. In fact, it is the discrete formulation of a more general functional problem:

$$\min_{f \in C^1([0,K])} \quad C(x(t), z(t), Z(t), f(t))$$

$$\int \left\| \frac{\partial m}{\partial x^t} \right\|^2 dt = q$$

where $m > 0$. This paper adopts a non-parametric estimation strategy, smoothing splines to be precise. The quadratic penalization of the first difference of $\delta_t$ is quite standard in the literature, see Craven and Wahba (1978) for instance. The higher the $m$, the smoother the $f$. I focus on $m = 1$, the first difference in the discrete case is equivalent to the first derivative in the continuous case. Therefore, I choose the $m$ which constrains the less the transition to be smooth. Nevertheless, this penalization of the variation in the coefficients tends to avoid one-time breaks.

The choice of $S$ has to be grounded as well as the amount of variability in the coefficients, that is $Q_{emp}$, where $Q_{emp}$ is diagonal here.

### 2.2 Economic interpretation

In the standard neo-Keynesian framework, three equations describe the economy: the New Keynesian Philips Curve (NKPC), the IS curve and the monetary policy rule, here forward looking$^5$:

$$\begin{cases} 
\pi_t = 4\kappa y_t + \beta E_t (\pi_{t+1}) + \varepsilon_t'' \\
y_t = E_t (y_{t+1}) - \sigma [i_t - E_t (\pi_{t+1})] + \varepsilon_t' \\
i_t = a_\pi E_{t-1} (\pi_{t+h}) + a_y E_{t-1} (y_{t+h}) + a_t i_{t-1} + \varepsilon_t' 
\end{cases} \quad (5)$$

$^5$ $\pi_t$ and $i_t$ are at an annual rate, which explains the presence of $4\kappa$ instead of $\kappa$. 

where $\varepsilon_t^m, \varepsilon_t^g, \varepsilon_t^i$ respectively stand for a mark-up shock, a demand shock and a monetary policy shock and where $\pi_t, y_t$ and $i_t$ are taken as deviation to their steady state values. In the present paper, I will only estimate the last equation of (5).

In the application, both the reduced form (1) and a more structural form will be commented:

$$i_t = (1 - \rho_t) i^*_t + \rho_t i_{t-1} + \omega_t$$
$$i^*_t = \beta_{0,t} + \beta_{\pi,t} E_{t-1} (\pi_{t+h}) + \beta_{y,t} E_{t-1} (y_{t+l})$$
$$= \tau_t + \pi^*_t + \beta_{\pi,t} (E_{t-1} (\pi_{t+h}) - \pi^*_t) + \beta_{y,t} E_{t-1} (y_{t+l})$$

(6)
derived from (1) with the transformation:

$$(a_0, a_\pi, a_y, a_i)^T = ((1 - \rho_t) \beta_{0,t}, (1 - \rho_t) \beta_{\pi,t}, (1 - \rho_t) \beta_{y,t}, \rho_t)^T$$

$i^*_t$ is the target interest rate which is reached gradually by the FED. $\beta_{\pi,t}$ is the elasticity of the target to expected inflation or the long-term response of the interest rate to inflation. The coefficient $\beta_{0,t}$ has no simple structural interpretation. It is linked to the equilibrium real rate $\tau_t$ and to the inflation objective $\pi^*_t$: $\beta_{0,t} = \tau_t + (1 - \beta_{\pi,t}) \pi^*_t$.

Once $(\delta_t)_{t \in [1,T]}$ is known, one can see whether and when the equilibrium is locally determined. But the answer depends on the other structural equations in (5). The estimation of such equations in a time-varying framework goes further than the present study so the coefficients $\beta, \sigma$ and $\kappa$ will be held constant$^6$.

3 Estimation of a time-varying parameter model in a Generalized Method of Moments framework

The minimization problem (4) can be written with the lagrangian:

$$\min_{(\delta_t)_{t \in [0,T]}} \left( \frac{1}{T} \sum_{t=1}^{T} (x_t - z_t' \delta_t) Z_t \right)' S \left( \frac{1}{T} \sum_{t=1}^{T} (x_t - z_t' \delta_t) Z_t \right) + \frac{1}{T} \sum_{t=1}^{T} T^2 \Delta \delta_t' R \Delta \delta_t$$

(7)

with $R$ being the Lagrange multiplier, a definite positive matrix. (4) is a special case with $R$ being diagonal.

---

$^6$We neglect structural variations in the real economy. Sims and Zha (2006) consider VAR models with switching regimes where they allow for changes both in the variance of the shocks and in the coefficients. The model with changes in the volatility only gives the best fit to the data, but among time-varying coefficients models, the best fit is the model which allows only the monetary policy rule to change.
Notice that (7) is a numerical procedure which can be solved independently of any statistical model. It is a non-parametric problem which belongs to the smoothing splines framework. The underlying statistical model is:

\[
\begin{align*}
  x_t &= z_t' \delta_t + v_t \\
  E(Z_t v_t) &= 0 \\
  \delta_t &= \delta_{t-1} + u_t \\
  u_t &\text{iid } \mathcal{N}(0, Q_u)
\end{align*}
\] (8)

A whole class of estimates can be obtained, one for each possible choice of \( R \) and \( S \). (8) will be useful to study the statistical properties of the proposed estimates.

Hodrick and Prescott’s approach belongs to this kind of methods. The HP filter extracts a trend \( \bar{y}_t(\lambda) \) from the data \( x_t \):

\[
\bar{y}_t(\lambda) = \arg \min_{(y_t)_{t=1}^T} \sum_{t=1}^T (x_t - y_t)^2 + \lambda \sum_{t=3}^T (\Delta^2 y_t)^2
\] (9)

### 3.1 Solution to the minimization problem

In this section \( (\delta_t)_{t\in[0,T]} \) stand for the estimated coefficients solution of (7) and not the true value of these coefficients. Define:

\[
\Psi(\delta_1, ..., \delta_T) = \frac{1}{T} \sum_{t=1}^T (x_t - z_t' \delta_t) Z_t
\]

\[
\chi = \frac{1}{T} \sum_{t=1}^T z_t Z_t' \text{ and } \Omega = \frac{1}{T} \sum_{t=1}^T x_t Z_t
\]

The first order conditions of (7) are:

\[
\begin{align*}
  \forall t \in [1, T-1], z_t Z_t' S \Psi + T^2 R (\delta_{t+1} - 2 \delta_t + \delta_{t-1}) &= 0 \\
  -z_T Z_T' S \Psi + T^2 R (\delta_T - \delta_{T-1}) &= 0 \\
  T^2 R (\delta_1 - \delta_0) &= 0
\end{align*}
\]
After some calculation\(^7\), let us look for \(\Psi\) and \((\delta_t)_{t\in[0,T]}\) such that:

\[
\begin{aligned}
&\forall t \in [1, T], \delta_t = \delta_0 + R^{-1} \frac{1}{T} \sum_{j=1}^T \sum_{i=j}^T z_i Z'_i S \Psi \\
&\Psi = \frac{1}{T} \sum_{t=1}^T (x_t - z'_i \delta_t) Z_t \\
&\chi S \Psi = 0
\end{aligned}
\]  

(10) involves:

\[
\begin{aligned}
\Psi &= \frac{1}{T} \sum_{t=1}^T x_t Z_t - \frac{1}{T} \sum_{t=1}^T z'_i \delta_0 - \frac{1}{T} \sum_{j=1}^T \left( \sum_{i=j}^T Z_i z'_i \right) R^{-1} \left( \sum_{i=j}^T z_i Z'_i \right) S \Psi \\
I \text{ rewrite the latest equation:}
J (R, S) \Psi &= \Omega - \chi' \delta_0 \\
J (R, S) &= J_{K_1} + \frac{1}{T} \sum_{j=1}^T \left( \sum_{i=j}^T z_i Z'_i \right)' R^{-1} \left( \sum_{i=j}^T z_i Z'_i \right) S
\end{aligned}
\]

(11)

From (11), \(\delta_0\) is sufficient to know \(\Psi\). \(\delta_0\) is given using the last equation of (10):

\[
\chi S \Psi = 0 \Rightarrow \chi S J^{-1} \chi' \delta_0 = \chi S J^{-1} \Omega \Rightarrow \delta_0 = \left( \chi S J^{-1} \chi' \right)^{-1} \chi S J^{-1} \Omega
\]  

(12) reminds us of the GMM formula with weighting matrix \(S J^{-1}\) which is definite positive as can be seen from (11).

\(^7\)\(T^2 R (\delta_T - \delta_{T-1}) = T^2 R \sum_{t=1}^{T-1} [(\delta_{t+1} - \delta_t) - (\delta_t - \delta_{t-1})] + T^2 R (\delta_1 - \delta_0)
\)

\(\Rightarrow \) \(z'_t Z'_t S \Psi = - \left( \sum_{t=1}^{T-1} z_i Z'_i \right) S \Psi \Rightarrow \left( \sum_{t=1}^{T} z_i Z'_i \right) S \Psi = 0 \Rightarrow \chi S \Psi = 0
\)

Let us compute \(\forall t \in [1, T]\) a formula of \(\Delta \delta_t\) function of \(\Psi\):

\(T^2 R (\delta_t - \delta_{t-1}) = T^2 R \sum_{t=1}^{T-1} [(\delta_{t+1} - \delta_t) - (\delta_t - \delta_{t-1})] + T^2 R (\delta_1 - \delta_0)
\)

\(\Rightarrow T^2 R (\delta_t - \delta_{t-1}) = - \left( \sum_{t=1}^{T-1} z_i Z'_i \right) S \Psi = \left( \sum_{t=1}^{T} z_i Z'_i \right) S \Psi
\)

\(\forall t \in [1, T]\), there is a formula of \(\delta_t\) function of \(\Psi\):

\(T^2 R (\delta_t - \delta_0) = \left( \sum_{t=1}^{T} z_i Z'_i \right) S \Psi = \left( \sum_{t=1}^{T} \min (i, t) z_i Z'_i \right) S \Psi
\)

\(\Rightarrow \delta_t = \delta_0 + R^{-1} \frac{1}{T^2} \left( \sum_{t=1}^{T} \min (i, t) z_i Z'_i \right) S \Psi
\)
3.2 Closed form formula

From now on, \(\hat{\delta}_t\) refers to the estimated value of the coefficients whereas \(\delta_t\) refers to the true value of the coefficients. In fine, with \(R\) and \(S\) given, the formula is:

\[
\begin{align*}
\hat{\delta}_t (R, S) &= \left\{ (\chi SJ^{-1}\chi')^{-1} \chi SJ^{-1} + \frac{1}{T^2} R^{-1} \left( \sum_{i=1}^{T} \min (i, t) z_i Z_i' \right) \right\} \times \\
& \quad SJ^{-1} \left( I_{K_t} - \chi' \left( \chi SJ^{-1}\chi' \right)^{-1} \chi SJ^{-1} \right) \Omega \\
\end{align*}
\]

(13)

Let:

\[
M_t (R, S) = \left\{ (\chi SJ^{-1}\chi')^{-1} \chi SJ^{-1} + \frac{1}{T^2} R^{-1} \left( \sum_{i=1}^{T} \min (i, t) z_i Z_i' \right) \right\} \times \\
& \quad SJ^{-1} \left( I_{K_t} - \chi' \left( \chi SJ^{-1}\chi' \right)^{-1} \chi SJ^{-1} \right) \\
\]

The time \(t\) estimate depends on all the data set. (13) gives:

\[
\hat{\delta}_t = M_t \frac{1}{T} \sum_{i=1}^{T} Z_i x_i \\
\]

Since \(M_t \chi' = I_{K_t}\), the difference between the estimate and the true value is:

\[
\hat{\delta}_t - \delta_t = M_t \frac{1}{T} \sum_{i=1}^{T} Z_i z_i' \left( \sum_{j=1}^{T} u_j - \sum_{j=1}^{t} u_j \right) + Z_i v_i \\
\]

\[
\hat{\delta}_t - \delta_t = M_t \left[ -\frac{1}{T} \sum_{j=1}^{T} \left( \sum_{i=1}^{j} Z_i z_i' \right) u_j + \frac{1}{T} \sum_{j=t+1}^{T} \left( \sum_{i=j}^{T} Z_i z_i' \right) u_j \right] \\
\]

Conditional on the data generating process (8), the mean squared error is given by:

\[
MSE_t = E \left[ \left( \hat{\delta}_t - \delta_t \right) \left( \hat{\delta}_t - \delta_t \right)' \right] \\
MSE_t = \frac{1}{T} M_t \left[ \frac{1}{T} \sum_{i=1}^{T} Z_i z_i' E \left( \psi_i^2 \right) + \frac{1}{T} \sum_{j=1}^{T} \left( \frac{1}{T} \sum_{i=1}^{j} Z_i z_i' \right) T^2 Q u \left( \frac{1}{T} \sum_{i=1}^{j} Z_i z_i' \right)' \\
+ \frac{1}{T} \sum_{j=t+1}^{T} \left( \frac{1}{T} \sum_{i=j}^{T} Z_i z_i' \right) T^2 Q u \left( \frac{1}{T} \sum_{i=j}^{T} Z_i z_i' \right)' \right] M_t' \\
\]

(14)

3.3 Choice of \((R, S)\)

Given \((R, S)\), one can solve (7), thus obtaining a whole class of estimates. In order to choose the best one, I rely on an efficiency criterion, trying to minimize

---

This formula accounts for the heteroskedasticity of \(\nu_t\). If the forward-looking horizon \(h\) or \(l\) exceeds the current quarter, there might be a source of autocorrelation of \(\nu_t\) which is not taken into account. Both Kim and Nelson (2006) and Boivin (2006) choose situations such that they can ignore autocorrelation.
the mean squared error of the estimate. One can prove that this estimate is the best possible in terms of the global distance between the curve of the true coefficients and the curve of the estimates. Since (14) depends on \( Q_u \), the first step is to know \( Q_u \), the covariance matrix of the innovation of \( \delta_t \). To achieve this, I use the Median Unbiased Estimate developed in Stock and Watson (1998)\(^9\).

### 3.3.1 Modification of Stock and Watson (1998) to deal with endogeneity

This section adapts Stock et Watson’s median unbiased estimate to allow for endogeneity as described in Sowell (1996). Following Stock and Watson (1998), assume that there is \( \mu \) such as, for every \( T \), the covariance matrix \( Q_u \) can be put in the restricted form:

\[
Q_u = \frac{\mu^2}{T^2} \overline{Q}
\]

where \( \overline{Q} \) is fixed to some particular value. The restricted form (15) can be seen in terms of frequency of the observations: for a given time period, if the number of observations is two times larger, the variance between two consecutive coefficients is two times smaller.

The state vector \( \delta_t \) is a discretization of a time continuous process\(^10\):

\[
\delta_t = \delta_0 + \frac{1}{\sqrt{T}} \mu \overline{Q}^{\frac{1}{2}} B_2 \left( \frac{t}{T} \right)
\]

where \( B_2 : \mathbb{R} \rightarrow \mathbb{R}^K \) is a standard brownian motion. Let us define for each \( s \in [0, 1] \):

\[
\xi_T (s, \delta) = \frac{1}{\sqrt{T}} \sum_{t=1}^{\lfloor sT \rfloor} (x_t - z_t' \delta) Z_t
\]

\(^9\)Here, a maximum likelihood (ML) approach can not be used without additional information on the distribution of the shocks. Moreover Stock and Watson (1998) show that when one estimates \( Q_u \) with ML, there is pile-up problem in the algorithm, which means that it converges far too often toward zero. That is why they build an alternative approach.

\(^10\) \( u_t = \frac{\mu}{\sqrt{T}} \overline{Q}^{\frac{1}{2}} \eta_t \) where \( \eta_t \approx \mathcal{N} (0, 1) \) hence \( \delta_t = \delta_0 + \frac{1}{\sqrt{T}} \mu \overline{Q}^{\frac{1}{2}} \frac{1}{\sqrt{T}} \sum_{k=1}^{\lfloor t \rfloor} \eta_k \) and from the definition of a brownian motion \( B_2 (\cdot) \), \( \frac{1}{\sqrt{T}} \sum_{k=1}^{\lfloor t \rfloor} \eta_k \Rightarrow B_2 (s) \)
Finally, define $\Sigma$, $M$, $S_T$ and $\hat{\delta}_{GMM,T}$:

$$\Sigma = \lim_{T \to \infty} E (\xi_T (1, \delta_0) \xi_T (1, \delta_0'))$$

$$M = -p \lim_{T \to \infty} \frac{\partial \xi_T}{\partial \delta} (1, \delta_0) = E (Z_t z_t')$$

$$S_T \xrightarrow{p} \Sigma^{-1}$$

$$\hat{\delta}_{GMM,T} = \arg \min_{\delta} \xi_T (1, \delta)' S_T \xi_T (1, \delta)$$

Theorem 1 in Sowell (1996) gives\(^{11}\):

$$S_T^{\frac{1}{2}} \xi_T \left( s, \hat{\delta}_{GMM,T} \right) \Rightarrow B_1 (s) - s \Sigma^{-\frac{1}{2}} M (M' \Sigma^{-1} M)^{-1} \left( \Sigma^{-\frac{1}{2}} M \right)' B_1 (1)$$

$$+ \mu \Sigma^{-\frac{1}{2}} M \mathcal{Q}^{\frac{1}{2}} \int_0^s \left( B_2 (v) - \frac{1}{0} \int B_2 (v) \right) dv$$

(16)

where $B_1 : \mathbb{R} \to \mathbb{R}^{K'}$ is another standard brownian motion. As can be seen, when there is no endogeneity problem, the same results as Stock and Watson’s first theorem can be obtained because $\Sigma \propto M$. In the more general case, the limiting distribution in (16) depends on some nuisance terms $\Sigma^{-\frac{1}{2}} M$ and $\mathcal{Q}^{\frac{1}{2}}$. $\Sigma$ and $M$ can be consistently estimated whereas $\mathcal{Q}$ cannot. I make the assumption $\mathcal{Q} = \text{Diag} \left( (M' \Sigma^{-1} M)^{-1} \right)$. It means that the variance of the innovations $u_{i,t}$ is proportional to the asymptotic covariance of the GMM estimate and that there is no cross correlation in $u_{i,\mu}^{12}$. This formula is quite natural but appendix A.4 proves the robustness toward this arbitrary choice, mainly in two directions: the diagonal form of $\mathcal{Q}$ and the measurement error on $\mathcal{Q}$’s diagonal terms. This formula is neutral with respect to any multiplicative transformation of the regressors or the instruments and it is the equivalent of the quantity proposed in Stock and Watson (1998). Moreover, the less precise the GMM estimate is for a given coefficient, the more one can think it is due to some time variation in that coefficient. Rewrite (16):

$$S_T^{\frac{1}{2}} \xi_T \left( s, \hat{\delta}_{GMM,T} \right) \Rightarrow W \left( s, \mu, D \right)$$

$$W \left( s, \mu, D \right) = B_1 (s) - s D (D' D)^{-1} D' B_1 (1) + \mu D \left( \int_0^s B_2 (v) dv - \frac{1}{0} \int B_2 \right)$$

(17)

\(^{11}\)For any definite positive matrix $A$, $A^{\frac{1}{2}}$ is the unique definite positive matrix such that $\left( A^{\frac{1}{2}} \right)^2 = A$.

\(^{12}\)corr $(u_{i,t}, u_{j,t}) = 0$ for any $i \neq j$. This is a disputable choice. Since $u_t$ is the innovation in a reduced form, there must be some cross-correlation. As there is no clue on its form, I prefer to remain neutral.
where $D = \Sigma^{-\frac{1}{2}} M Q^{-\frac{1}{2}}$ which involves:

$$L_T = \frac{1}{T} \sum_{t=1}^{T} \xi_T \left( s, \hat{\delta}_{GMM,T} \right)' S_T \xi_T \left( s, \hat{\delta}_{GMM,T} \right) = \int W(s, \mu, D)' W(s, \mu, D) \, ds$$

$$\sup_{s \in [s_0, s_1]} \xi_T \left( s, \hat{\delta}_{GMM,T} \right)' S_T \xi_T \left( s, \hat{\delta}_{GMM,T} \right) = \sup_{s \in [s_0, s_1]} W(s, \mu, D)' W(s, \mu, D)$$

To implement the estimate of $\mu$, one only need to compute $L_T$ and, for each $\mu$, the median $m(\mu)$ of the distribution $\int W(s, \mu, D)' W(s, \mu, D) \, ds$ taking $D$ equal to its estimated value. The estimate of $\mu$ is then:

$$\hat{\mu} = m^{-1}(L_T)$$

Stock and Watson’s method is implemented with:

$$\hat{S}_T^{-1} = \hat{\Sigma} = \frac{1}{T} \sum_{t=1}^{T} \left( x_t - z_t \hat{\delta}_{2SLS} \right)^2 Z_t Z_t'$$

$$\hat{M} = \frac{1}{T} \sum_{t=1}^{T} Z_t z_t'$$

$$\hat{Q} = \text{Diag} \left( \hat{M}^{-1} \hat{\Sigma}^{-1} \hat{M} \right)^{-1}$$

$$\hat{D} = \hat{\Sigma}^{-\frac{1}{2}} \hat{M}^{-\frac{1}{2}}$$

$$\hat{\delta}_{GMM,T} = \hat{\delta}_{3SLS} = \arg \min_{\delta} \xi_T (1, \delta)' S_T \xi_T (1, \delta)$$

### 3.3.2 Optimal estimate

I have now a whole set of estimates, one for each choice of $(R, S)$. One can prefer one estimate to another and there are many possible choices. The more natural choice is the estimate that is the closest to the true value of the coefficients, after giving a definition of what close means. My stand is to choose the estimate which has the smallest mean squared error. Moreover, I will show that it is the one closest to the true coefficients. (14) gives for $t = 0$:

$$\hat{MSE}_0 = \frac{1}{T} M_0 \left[ \hat{\Sigma} + \frac{1}{T} \sum_{t=1}^{T} \left( \frac{1}{T} \sum_{i=1}^{T} z_i Z_i' \right)' \hat{\mu}^2 \hat{Q} \left( \frac{1}{T} \sum_{i=1}^{T} z_i Z_i' \right) \right] M_0$$

with $M_0 = (\chi S J^{-1} \chi')^{-1}$. Following (11), define:

$$N = S J^{-1} = \left( S^{-1} + \frac{1}{T} \sum_{k=1}^{T} \left( \frac{1}{T} \sum_{i=k}^{T} z_i Z_i' \right)' R^{-1} \left( \frac{1}{T} \sum_{i=k}^{T} z_i Z_i' \right) \right)^{-1}$$

15
**Proposition 1** The optimal \((R, S)\) is the choice that minimizes \(\overline{MSE}_0\). The solution is:

\[
R^* = \frac{1}{\hat{\mu}^2 Q} \hat{\Sigma}^{-1} \\
S^* = \hat{\Sigma}^{-1}
\]  

**Proof.** The demonstration follows the one in the GMM case:

\[
\hat{V} = \hat{\Sigma} + \frac{1}{T} \sum_{t=1}^{T} \left( \frac{1}{T} \sum_{i=t}^{T} z_i Z_i' \right) \hat{\mu}^2 Q \left( \frac{1}{T} \sum_{i=t}^{T} z_i Z_i' \right)'
\]

If \(S = \hat{\Sigma}^{-1}\) and \(R = \frac{1}{\hat{\mu}^2 Q} \hat{\Sigma}^{-1}\):

\[
N^* = \left( \hat{\Sigma} + \frac{1}{T} \sum_{t=1}^{T} \left( \frac{1}{T} \sum_{i=t}^{T} z_i Z_i' \right) \hat{\mu}^2 Q \left( \frac{1}{T} \sum_{i=t}^{T} z_i Z_i' \right) \right)^{-1} = \hat{V}^{-1}
\]

\[
\overline{MSE}^*_0 = \frac{1}{T} \left( \chi \hat{V}^{-1} \chi' \right)^{-1} \chi \hat{V}^{-1} \hat{V} \chi' \left( \chi \hat{V}^{-1} \chi' \right)^{-1} = \frac{1}{T} (C'C)^{-1}
\]

For a given \(S\) and \(R\):

\[
\overline{MSE}_0 = \frac{1}{T} \left( \chi N \chi' \right)^{-1} \chi N \hat{V} \chi' \left( \chi N \chi' \right)^{-1}
\]

\[
\overline{MSE}_0 = \frac{1}{T} \left( \chi N \chi' \right)^{-1} \chi N \hat{V} \chi' \left( \chi N \chi' \right)^{-1} = \frac{1}{T} \hat{BB}'
\]

Now, \(BC = (\chi N \chi')^{-1} \chi N \hat{V} \chi' = I_K\) so that:

\[
\overline{MSE}_0 - \overline{MSE}_0 = \frac{1}{T} \hat{BB}' - \frac{1}{T} (C'C)^{-1} = \frac{1}{T} B \left( I_K - C (C'C)^{-1} C' \right) B'
\]

which is a semi definite positive matrix\(^{13}\) meaning that \(\delta_i^*\) is optimal.  

**Proposition 2** \((R^*, S^*)\) minimizes the mean squared error \(\overline{MSE}_t\) for each single \(t\). The optimal estimate is thus efficient for each \(t\).

**Proof.** Consider \(\delta_i\) as a starting point instead of \(\delta_0\) in the solution of the minimization program. All details are given in appendix A.3  

**Proposition 3** Define the loss \(L_i\) on coefficient \(i \in [1, K]\) and its expectation \(R_i\) which is called the risk by:

\[
L_i (R, S) = \frac{1}{T} \sum_{t=1}^{T} \left( \delta_{i,t} - \hat{\delta}_{i,t} (R, S) \right)^2
\]

\[
R_i (R, S) = E \left( L_i (R, S) \right) = E \left( \frac{1}{T} \sum_{t=1}^{T} \left( \delta_{i,t} - \hat{\delta}_{i,t} (R, S) \right)^2 \right)
\]

then \((R^*, S^*)\) minimizes the risk \(R_i\) on coefficient \(i\).  

\(^{13}C(C'C)^{-1} C'\) is the orthogonal projector on the space generated by the columns of matrix \(C\). Now a vector \(X\) always has a greater norm than its projection so \(X'X > X'C(C'C)^{-1} C'C(C'C)^{-1} C'X\) so \(I_K - C (C'C)^{-1} C'\) is definite positive.
Proof. \( \overline{MSE}_t \) is minimum for the choice \( (R^*, S^*) \).

As a consequence, \( E \left( \left( \delta_{i,t} - \hat{\delta}_{i,t} (R^*, S^*) \right)^2 \right) \) is minimum, being a diagonal element of the matrix \( \overline{MSE}_t \), as well as \( E \left( \frac{1}{T} \sum_{t=1}^{T} \left( \delta_{i,t} - \hat{\delta}_{i,t} (R, S) \right)^2 \right) \) \( \blacksquare \).

In line with the spline literature, the estimate minimizes the risk and, in addition, it is efficient for a given \( t \).

The optimal values (19) are the equivalent of the sound to noise ratio used in Hodrick and Prescott (1981). There is an analogy between the choice of \( (R, S) \) and the choice of the parameter \( \lambda \) of the HP filter (9). In their seminal work, Hodrick and Prescott justify the choice of \( \lambda \) by the writing of the joint likelihood of the signal and the trend. It is possible to show that the sound to noise ratio has many interesting properties. For instance, Schlicht (2006) proves that it is the unique \( \lambda \) such that \( \hat{y}_t (\lambda) = E(y|x) \). One could prove that it minimizes the mean squared error between the estimated trend and the real one.

The weighting matrix \( S^* \) is the same as the one used in the standard GMM when heteroskedasticity matters. Finally, \( \overline{Q} \) is diagonal, so is \( R^* \) and therefore cross terms are not penalized in (7).

3.4 Comments

First, the problem is equivalent to a minimization program (7) which reminds us of a filter approach such as Hodrick and Prescott (1981). The method belongs to non-parametric statistics and more precisely smoothing splines.

In contrast to Kim and Nelson (2006), the method does not require specifying initial conditions \( \delta_0 \). If \( \delta_0 \) were stationary, then one could set initial conditions to the long term value of the coefficients. Here, \( \delta_t \) is integrated and a Kalman filter approach requires initial conditions often estimated on a time invariant form of (1) on the beginning of the sample. The initial conditions are not accurate and that error source in the procedure is decreasing but persistent as time goes by. Furthermore, the 40 first observations at least should not be accounted for in the main estimation.

My method is derived from the GMM litterature and is based on moment conditions rather than sum of squares. As in the GMM case the estimate is heteroskedasticity robust. As a constrast, in a Kalman filtering approach, a special form of heteroskedasticity for the residuals has to be imposed. The random walk assumption for \( \delta_t \) allows for permanent changes. A more general autoregressive form would have led to a mean reverting process.

This method has some drawbacks. First of all, it does account for the heteroskedasticity of \( v_t \) but not for its possible autocorrelation when the forecast
horizon is greater than one period. If autocorrelation is significant, the estimate is not optimal any more. Stock and Watson (1998) propose to apply an autoregressive filter to the model (8) so that the error term $v_t$ is not autocorrelated. Then, a part of arbitrary choice remains in the definition of $Q$. Other choices shown in appendix A.4. assess the robustness to this issue: for example, if the relative variances were wrong, the estimated coefficients would adjust around the true coefficients. Then, because of the quadratic penalization of the changes in the coefficients, a smooth transition is favoured rather than one-time breaks. It is possible to penalize the absolute value of the first difference of the coefficients but this solution is analytically intractable. Another issue deals with the deep nature of $\delta_t$ and thus the properties of $u_t$. To implement Stock and Watson’s method, it is implicitly assumed that $u_t$ is independent of the other shocks in the economy. This is not likely and there must have been some interactions between the real economy and the way monetary policy was conducted. It is a substantial limitation of the time-varying approach. In the spline framework, this issue matters only for the optimality conditions because this is the only part where distribution hypotheses are made.

4 What does it tell us about the FED monetary policy?

The method can be implemented to various specifications depending on $h$ and $l$:

$$i_t = a_{0,t} + a_{\pi,t}E_{t-1}(\pi_{t+h}) + a_{y,t}E_{t-1}(y_{t+l}) + a_{i,t}i_{t-1} + \omega_t$$

which is more easily understandable as (6):

$$i_t = (1 - \rho_t) \left( \beta_{0,t} + \beta_{\pi,t}E_{t-1}(\pi_{t+h}) + \beta_{y,t}E_{t-1}(y_{t+l}) \right) + \rho_t i_{t-1} + \omega_t$$

I consider two baseline formulations: $h = l = 0$ which corresponds to the situation in which the FED reacts to current values, not available at the time it takes a decision and the case $h = 3$ et $l = 0$ when the central bank reacts to expected inflation a year ahead $\pi_{t+3} = \frac{1}{4} (\pi_t + \pi_{t+1} + \pi_{t+2} + \pi_{t+3})$ as well as to the current output gap. Kim and Nelson (2006) corresponds to $h = l = 0$ given the notations, Boivin (2005) to the second case.

4.1 Data

I use data from the FRED database (Federal Reserve Economic Data) available on the FED of St' Louis website.

The rule I consider depends only on inflation and real activity. It would have been possible to extend the set of explanatory variables to money or commodity price. However, there is no evidence that the central bank looked at such variables. Then, they could enter the instruments $Z_t$ meaning that the FED care about, say commodity price, as part of future inflation.
I use the quarterly difference of the GDP deflator (GDPDEF) annualized as a measure of inflation.

For the output gap, I follow Kim and Nelson (2006) and take the definition of the Congressional Budget Office, which is a semi-structural estimate (GDPC96-GDPPOT) very close to the extraction of a quadratic trend as shown in appendix A.4.3.

\( i_t \) is set equal to the average Federal Fund Rate in the first month of each quarter (FEDFUNDS).

The set of instruments includes four lags of quarterly inflation, output gap, interest rate, quarterly change of M2 (M2SL), commodity price inflation and the spread between one year and ten years treasury bonds (GS1 and GS10). These variables should bring information on future inflation and the output gap. Four lags are widespread in the literature. For the commodity price, I take the producer price index covering all the commodities (PPIACO in FRED database, or WPU00000000 in the Bureau of Labor Statistics database).

Other definitions are chosen in appendix A.4.3. to check for robustness.

### 4.2 Break estimates

As a first analysis, one can estimate a constant rule (1) on subperiods. I implement a GMM technique, following Clarida, Gali and Gertler (2000). Here the 2SLS allow to estimate the optimal weighting matrix in order to minimize the asymptotic covariance of the GMM estimate. Then, we compute the 3SLS, heteroskedasticity robust.

I consider the pre-Volcker era 1960Q1 – 1979Q2 and the Volcker-Greenspan era 1979Q3 – 2006Q1. A third period 1987Q3 – 2006Q1 is added for more stable results, following Jondeau, Le Bihan and Gallès (2004). The following tables display the structural form (6) in order to ease the interpretation.

<table>
<thead>
<tr>
<th>Period</th>
<th>( \beta_0 )</th>
<th>( \beta_\pi )</th>
<th>( \beta_y )</th>
<th>( \rho )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1960Q2 – 1979Q2</td>
<td>1.51</td>
<td>0.90</td>
<td>0.45</td>
<td>0.72</td>
</tr>
<tr>
<td></td>
<td>(0.37)</td>
<td>(0.09)</td>
<td>(0.11)</td>
<td>(0.06)</td>
</tr>
<tr>
<td>1979Q3 – 2006Q3</td>
<td>-0.40</td>
<td>2.45</td>
<td>0.97</td>
<td>0.79</td>
</tr>
<tr>
<td></td>
<td>(0.63)</td>
<td>(0.25)</td>
<td>(0.17)</td>
<td>(0.03)</td>
</tr>
<tr>
<td>1987Q3 – 2006Q3</td>
<td>0.95</td>
<td>1.87</td>
<td>1.45</td>
<td>0.88</td>
</tr>
<tr>
<td></td>
<td>(1.17)</td>
<td>(0.50)</td>
<td>(0.23)</td>
<td>(0.02)</td>
</tr>
</tbody>
</table>

Notes: 3SLS
The results are close to Clarida, Gali and Gertler (2000) even if the time periods or the variables are slightly different.

<p>|
| --- |
| Table 2: Baseline Estimates of Clarida, Gali &amp; Gertler (2000) |
|</p>
<table>
<thead>
<tr>
<th>$\pi^*$</th>
<th>$\beta_\pi$</th>
<th>$\beta_y$</th>
<th>$\rho$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1960Q1 – 1979Q2</td>
<td>4.24</td>
<td>0.83</td>
<td>0.27</td>
</tr>
<tr>
<td></td>
<td>(1.09)</td>
<td>(0.07)</td>
<td>(0.08)</td>
</tr>
<tr>
<td>1979Q3 – 1996Q4</td>
<td>3.58</td>
<td>2.15</td>
<td>0.93</td>
</tr>
<tr>
<td></td>
<td>(0.50)</td>
<td>(0.40)</td>
<td>(0.42)</td>
</tr>
</tbody>
</table>

Notes: Table II from Clarida, Gali and Gertler (2000)

The second specification gives:

<p>|
| --- |
| Table 3: GMM on 3 subperiods: Annual inflation, contemporaneous output gap |
|</p>
<table>
<thead>
<tr>
<th>$\beta_0$</th>
<th>$\beta_\pi$</th>
<th>$\beta_y$</th>
<th>$\rho$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1960Q2 – 1979Q2</td>
<td>1.03</td>
<td>0.98</td>
<td>0.46</td>
</tr>
<tr>
<td></td>
<td>(0.41)</td>
<td>(0.10)</td>
<td>(0.13)</td>
</tr>
<tr>
<td>1979Q3 – 2005Q4</td>
<td>-0.82</td>
<td>2.58</td>
<td>1.03</td>
</tr>
<tr>
<td></td>
<td>(1.05)</td>
<td>(0.40)</td>
<td>(0.18)</td>
</tr>
<tr>
<td>1987Q3 – 2005Q4</td>
<td>-0.43</td>
<td>2.39</td>
<td>1.41</td>
</tr>
<tr>
<td></td>
<td>(1.51)</td>
<td>(0.65)</td>
<td>(0.23)</td>
</tr>
</tbody>
</table>

Notes: 3SLS

One can notice that the rule changed between the two subperiods. Until 1979, monetary policy seemed to accommodate inflation whereas the long term response of the interest rate to inflation is above one under the chairmanship of P.Volcker and A.Greenspan. The rule seems to be more and more countercyclical.

This method suffers from two main drawbacks. First, I assume one single structural break. Second, identification is weak as GMM estimation is carried out on a small time period.
4.3 Main results

4.3.1 Median unbiased estimate

Following Stock and Watson’s approach modified to allow for endogeneity, I can compute the medians of the limiting distributions of $L_T^{14}$:

<table>
<thead>
<tr>
<th>$\mu$</th>
<th>$h = l = 0$</th>
<th>$h = 3, l = 0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>10.9</td>
<td>10.9</td>
</tr>
<tr>
<td>1</td>
<td>11.1</td>
<td>11.2</td>
</tr>
<tr>
<td>2</td>
<td>11.8</td>
<td>12.0</td>
</tr>
<tr>
<td>3</td>
<td>12.7</td>
<td>13.0</td>
</tr>
<tr>
<td>4</td>
<td>13.8</td>
<td>14.4</td>
</tr>
<tr>
<td>5</td>
<td>15.2</td>
<td>15.9</td>
</tr>
<tr>
<td>6</td>
<td>17.0</td>
<td>17.3</td>
</tr>
<tr>
<td>7</td>
<td>18.9</td>
<td>19.4</td>
</tr>
<tr>
<td>8</td>
<td>20.8</td>
<td>21.5</td>
</tr>
<tr>
<td>9</td>
<td>23.3</td>
<td>24.0</td>
</tr>
<tr>
<td>10</td>
<td>25.9</td>
<td>26.9</td>
</tr>
</tbody>
</table>

For $h = l = 0$, $L_T = 19.1$ which leads to $\hat{\mu} = 7.1$. For the second specification, $L_T = 18.1$ and $\hat{\mu} = 6.4$.

4.3.2 Results

In the first specification, the FED reacts to expected current inflation and output gap. Figure 1 shows the evolution of the coefficients of the rule (1). Figure 2 focuses on the second specification with medium term concerns. Because of the delay needed for monetary policy to be effective, it is often assumed that decisions are motivated with a medium term perspective. In contrast with Kim and Nelson (2006), the long term response of the interest rate to inflation starts from a value below one, 0.8, in the 60s and reaches a value above one, between 1.4 and 1.8, in the 80s. The transition is quite brutal and occurs when P. Volcker became chairman of the Board. The degree of interest rate smoothing raises from 0.7/0.75 to 0.8/0.85. The response to the output gap raises from 0.6 to 1.1.

Both the evolution of $(\beta_0, \beta_x, \beta_y, \rho)$ and $(a_0, a_{\pi}, a_y, a_i)$ are plotted because the first set of coefficients describes the long term response of the FED whereas the second one describes the short term dynamics.

The monetary policy shock is not strictly equal to the residual which depends also on forecast errors. Nevertheless, one can distinguish the oil shocks and the appointment of P. Volcker. During these periods, discretion increased.

\footnote{Replacing $L_T$ with the other statistic $\sup_{s \in [s_0, s_1]} \xi_T \left( s, \hat{\delta}_{GMM,T} \right) S_T \xi_T \left( s, \hat{\delta}_{GMM,T} \right)$ does not change $\hat{\mu}$ very much: in the second baseline $\hat{\mu} = 6.8$ instead of 6.4.}
Figure 1: $h = l = 0$

Notes: Optimal matrix $S^*$ and $R^*$; confidence interval $\pm \sqrt{MSE}$
Figure 2: Expected annual inflation, contemporaneous output gap

Notes: Optimal matrix $S^*$ and $R^*$; confidence interval $\pm \sqrt{MSE}$
4.3.3 Results analysis

To allow for comparison, Figure 3 superposes my results to those by Kim and Nelson (2006) and Clarida, Gali and Gertler (2000).

First, the central bank accommodated inflation during the 60s and the 70s, consistently with Clarida, Gali and Gertler (2000). It is true that the degree of uncertainty around the point estimate of $\beta_\pi$ involves that $\beta_\pi$ is never statistically different from one should not expect a high precision from a time-varying approach. It can not be ruled out that the poor economic performance in the 70s was due to a weak reaction of the central bank (ie bad policy rather than bad luck). The curve $t \rightarrow \beta_{\pi,t}$ is closer to a stepwise function rather than a continuous one. One could conclude it fully justifies the structural break approach.

The long term response of the interest rate to the output gap continuously increased from a value close to 0.6 in 1960 to 1.1 in 2000. The confidence interval seems to move away from 0 as time increases. These results are different from Kim and Nelson (2006). The central bank became more and more countercyclical while implementing its disinflation policy. This increasing profile of $\beta_y$ is robust as shown in appendix A.4 but contradicts some previous studies by Orphanides (2004) or Boivin (2006). Their results rather favor a decrease in this response. The main explanation of this huge difference relies on the type of data used, real time or ex-post.
The degree of interest rate smoothing $\rho$ is increasing. Kim and Nelson’s results are particular on this issue. They find $\rho$ to be close to 0.5 during the first oil shock, which may be due to an imperfect heteroskedasticity specification in their model.

Romer and Romer (2002) describe the changes in the conduct of monetary policy in the postwar United States. They find that the FED focused more on real activity than on inflation in the 60s and 70s. However, as soon as P.Volcker was appointed chairman of the Board, the FED vigourously struggled against inflationary pressures without giving up growth concerns:

In the 1950s, policy-makers cautiously balanced concerns over inflation and real activity; in the 1960s, they focused vigorously on increasing real activity; in the 1970s, they pursued policies ranging from rapid expansion to full-fledged contraction to grudging tolerance of inflation; in the early 1980s, they followed a policy of aggressive disinflation; and since that time, they have again cautiously balanced the pursuit of real growth with concern about the possibility of inflation. (Romer and Romer, 2002, p 70-71)

To go further in the bad luck v. bad policy debate, it is necessary to know whether and when monetary policy was able to insure the local determinacy of the equilibrium. As Woodford (2003) shows, the answer is often depending on a weighted sum of $\beta_\pi$ and $\beta_y$.

**Proposition 4** The economy is described by (5), in the case $h = l = 0$. The equilibrium is locally and uniquely determined if and only if:

$$\beta_{\pi,t} + \frac{1 - \beta}{4\kappa} \beta_{y,t} - 1 > 0$$

or, as a function of $a_{0,t}$, $a_{\pi,t}$, $a_{y,t}$ and $a_{i,t}$:

$$a_{\pi,t} + \frac{1 - \beta}{4\kappa} a_{y,t} + a_{i,t} - 1 > 0$$

(20)

**Proof.** See appendix A.5. ■

Let us calibrate $\beta = 0.99$ and $\kappa = 0.024$ according to table 5.1 in Woodford (2003). It is not only possible to check if (20) is true or not but also the uncertainty surrounding this criterion and the probability that (20) is satisfied.
\[ \tilde{a}_{\pi,t} + \frac{1-\beta}{4\kappa} \tilde{a}_{y,t} + \tilde{a}_{i,t} - 1 \quad P \left( a_{\pi,t} + \frac{1-\beta}{4\kappa} a_{y,t} + a_{i,t} > 1 \right) \]

Figure 4: Local determinacy of the equilibrium

Notes: \( h = l = 0 \)

One has to be careful with figure 4. A simple DSGE model has been chosen and \( \beta \) or \( \kappa \) may have changed during the time period. One may argue (5) has micro-foundations so the IS curve or the NKPC did not change but it is only a simplified form\(^{15}\). Moreover, we derived the criterion in a static way, as if the economic agents did not know the rule was changing. This is not a too strong hypothesis since the coefficients follow a random walk and therefore today values are the best estimates of tomorrow ones. Conditionnal on these hypotheses, it is not likely that the equilibrium was locally determined in the 60ies or in the 70ies.

5 Conclusion

This article deals with the estimation of a forward looking monetary rule with time-varying parameters. This framework faces two econometric issues, endogeneity of the regressors and changing coefficients.

This article is new mainly because of its econometric strategy. I try to answer a key question, here the way monetary policy evolved in the United States, by imposing as few hypotheses as possible. More precisely, I combine the GMM framework to the smoothing splines litterature. Using a non-parametric approach is quite appropriate here. The framework implemented does not impose any restrictions on the form of the heteroskedasticity or the way regressors are

\(^{15}\beta \) is a mesure of the subjective discount factor and hence is less likely to have changed than \( \kappa = \frac{(1-\theta)(1-\beta)}{\sigma} \left( \eta + \frac{1}{2} \right) \) which depends on \( \theta \), the probability of no price reoptimization, on \( \sigma \), the intertemporal elasticity of substition and on \( \eta \), the elasticity of labour supply. Ball (2006) favors a decrease of \( \kappa \) on the period that he links to three main reasons : the globalization which exerts a downward pressure on national prices; the central bank credibility which anchors expectations ; a moderate inflaiton which allows firms to change their price less often and thus increases \( \theta \).
correlated to the disturbance term. Progresses are made in three main directions: initial conditions are estimated with all the information; the estimation is carried out in one single step; robustness is checked.

Moreover, the results can be compared to the existing literature. As many other authors, I find an increasing long term response of the interest rate to expected inflation, especially strong during the chairmanship of P.Volcker and A.Greenspan. Concerning the output gap, I find evidence that the central bank became more and more countercyclical, which is a new result. My categorisation is consistent with the time division in Romer and Romer (2002), especially the accommodative pattern before 1980 and the stability of the period 1980-2006.
References


Appendix

A.1 Data

Figure 5 plots the data used in the econometric application.

![Graphs of Inflation, Output gap, Spread, Interest rate, Change of M2, Commodities](image)

Figure 5: Data

Then I carry out stationarity tests with a sequential approach.

For the interest rate, an augmented Dickey-Fuller test is implemented with a constant in the regressors and a number of lags according to the AIC criterion. Student statistic $\tau$ is equal to $-3.19$ and the critical values at 1%, 5%, 10% are respectively $-3.47$, $-2.88$, and $-2.57$. This test rejects the null hypothesis of a unit root at a 5% level (p-value equal to 2.2%). A KPSS stationarity test gives a value of 0.33 whereas critical values at 1%, 5% and 10% are 0.74, 0.46 et 0.35. We do not reject the null hypothesis of stationarity.

For the inflation, an ADF test with a constant is chosen. $\tau = -2.10$ and hence it does not reject the null hypothesis of a unit root at a 5% level (p-value equal to 24%). A KPSS stationarity test gives a value of 0.40. The null hypothesis of stationarity can not be rejected.

A.2 Kim and Nelson’s approach
Kim and Nelson (2006) propose a two-step procedure for the estimation of (2), time-varying equivalent of the 2SLS. They instrument (2) with lags of inflation, output gap, interest rate, quarterly change of M2 or the inflation of commodities.

First, they compute the forecast errors. They estimate the following state-space model with a MLE based on the prediction error decomposition:

\[
\begin{align*}
\pi_{t+h} &= Z_i^t \delta_t + v_t^\pi \\
\delta_t &= \delta_{t-1} + u_t
\end{align*}
\]

where \(v_t^\pi\) is assumed to follow a \(GARCH(1,1)\) in order to take into account the heteroskedasticity of the forecast error. The Kalman filter is hence slightly modified following Harvey, Ruiz & Sentana (1992). They compute the forecast errors \(v_{it-1}^\pi = \pi_{t+h} - Z_i^t \delta_{t-1} = \pi_{t+h} - E(\pi_{t+h} | Z_t)\) and the normalized forecast errors \(v_{it-1}^\pi = \pi_{t+h} - Z_i^t \delta_{t-1} / \sqrt{\sigma_{v,t}^2 + \sigma_{v,t-1}^2}\).

To implement the second step, they model the correlation between \(v_t\) and the explanatory variables as:

\[v_t = \sigma_{v,t} \left( \tau^t (v_{it-1}^\pi) + \omega_t \right)\]

They are then able to estimate (2), following Harvey, Ruiz & Sentana (1992):

\[
\begin{align*}
x_t &= z_t^t \delta_t + \sigma_{v,t} \left( \tau^t (v_{it-1}^\pi) + \omega_t \right) \\
\delta_t &= \delta_{t-1} + u_t \\
\sigma_{v,t}^2 &= \alpha_0 + \alpha_1 v_{t-1}^2 + \alpha_2 \sigma_{v,t-1}^2
\end{align*}
\]

A.3 Optimality criterion

Writing the estimator formula in a different way shows that minimizing \(MSE_0\) is equivalent to minimizing each \(MSE_{t_0}\). To get this result, the stress is put on \(\delta_{t_0}\) instead of \(\delta_0\) in the resolution of (7).

\[T^2 R(\delta_t - \delta_{t-1}) = \sum_{i=t}^{T} z_i Z_i' S \Psi = -\sum_{i=1}^{T} z_i Z_i' S \Psi\]

\[\Rightarrow T^2 R(\delta_t - \delta_{t_0}) = \begin{cases} 
\sum_{j=t_0+1}^{T} z_i Z_i' S \Psi & \text{if } t > t_0 \\
0 & \text{if } t = t_0 \\
-\sum_{j=t+1}^{T} z_i Z_i' S \Psi = \sum_{j=t+1}^{T} z_i Z_i' S \Psi & \text{if } t < t_0
\end{cases}\]

\[\Psi = \frac{1}{T} \sum_{t=1}^{T} (x_t - z_t^t \delta_t) Z_t = \Omega - \chi^T \delta_{t_0} - \frac{1}{T} \left[ \sum_{i=1}^{T} Z_i^T R_1 \sum_{j=t+1}^{T} z_i Z_i' + \sum_{t=t_0+1}^{T} Z_t^T R_1 \sum_{j=t_0+1}^{T} z_i Z_i' \right] S \Psi\]

31
\[ \Psi = \Omega - \chi' \delta_{t_0} - \frac{1}{T^3} \left[ \sum_{j=2}^{T} \left( \sum_{i=1}^{T} Z_{i} \delta_{t_0} \right) R^{-1} \left( \sum_{i=1}^{T} z_i Z'_i \right) + \sum_{j=t_0+1}^{T} \left( \sum_{i=1}^{T} z_i Z'_i \right) R^{-1} \left( \sum_{i=1}^{T} z_i Z'_i \right) \right] S \Psi \]

Define:

\[ J(t_0) = I_{K^*} + \frac{1}{T^3} \left[ \sum_{j=1}^{T} \left( \sum_{i=1}^{T} Z_{i} \delta_{t_0} \right) R^{-1} \left( \sum_{i=1}^{T} z_i Z'_i \right) + \sum_{j=t_0+1}^{T} \left( \sum_{i=1}^{T} z_i Z'_i \right) R^{-1} \left( \sum_{i=1}^{T} z_i Z'_i \right) \right] S \]

Hence:

\[ J(t_0) \Psi = \Omega - \chi' \delta_{t_0} \]

Since \( \chi S \Psi = 0 \):

\[ \hat{\delta}_{t_0} = \left( \chi S J(t_0)^{-1} \chi' \right)^{-1} \chi S J(t_0)^{-1} \Omega \]

The mean squared error is:

\[ \hat{MSE}_{t_0} = \frac{1}{T} \left( \chi S J(t_0)^{-1} \chi' \right)^{-1} \chi S J(t_0)^{-1} \left\{ \hat{\Sigma} + \frac{1}{T} \sum_{j=1}^{t_0} \left( \sum_{i=1}^{T} Z_{i} \delta_{t_0} \right) \hat{\mu}^2 \hat{Q} \left( \sum_{i=1}^{T} z_i Z'_i \right) \right\} \left( \chi S J(t_0)^{-1} \chi' \right)^{-1} \chi S J(t_0)^{-1} \]

which is minimized for \( R^* = \frac{1}{\hat{\mu}^2} \hat{Q}^{-1}, S^* = \hat{\Sigma}^{-1} \). From here, the proof follows the one detailed for \( t_0 = 0 \).

**A.4 Robustness analysis**

**A.4.1 Choice of \( Q \)**

I focus on the second baseline with inflation one year ahead and contemporaneous output gap. I have to impose an arbitrary value for \( Q \) following Stock and Watson (1998). When the dimension of the hidden variable is greater than one, it is equivalent to fix the relative variances of the innovations of the coefficients. It is possible to test for robustness with respect to this choice. More precisely, I change one diagonal element of \( \hat{Q} \) from a ten factor, which is huge. Figure 6 illustrates the robustness of the results.
Figure 6 (1): Robustness with respect to $\overline{Q}$

Notes: in blue $\overline{Q}' = \overline{Q}$; in green $\overline{Q}'_{11} = \overline{Q}_{11}/10$; in red $\overline{Q}'_{22} = \overline{Q}_{22}/10$; in cyan $\overline{Q}'_{33} = \overline{Q}_{33}/10$; in purple $\overline{Q}'_{44} = \overline{Q}_{44}/10$

If $\overline{Q}$ is not constrained to be diagonal but simply $\overline{Q} = (M'\Sigma^{-1}M)^{-1}$, one
can notice the robustness, especially on $\beta_y$:  

![Figure 6 (2): Robustness with respect to $Q$](image)

Notes: in blue $Q$ diagonal; in red $Q$ non restricted

A.4.2 Uncertainty surrounding $\mu$

Stock & Watson’s method makes it possible to build a confidence interval for $\mu$. Let us define $q_\alpha (\mu)$ as the $\alpha$-quantile of the limiting distribution $\int_0^1 W (s, \mu, D)' W (s, \mu, D) ds$, then:

$$P (\mu < q_\alpha^{-1} (L_T)) = P (q_\alpha (\mu) < L_T) = 1 - \alpha$$

hence, the confidence interval is:

$$I_{1-2\alpha} = [q_{1-\alpha}^{-1} (L_T), q_{\alpha}^{-1} (L_T)] = [\mu, \bar{\mu}]$$

I take $\alpha = 0.16$, which is one standard error in the gaussian case, and I compute the curves obtained for $\mu$ and $\bar{\mu}$. This modification allows to see how the results would have been modified if a mistake was made on $\mu$. Figure 7 stresses the robustness on $\beta_y$.  

34
A.4.3 Variables definition

In this section, I investigate the robustness to various changes in the data set.

Kim & Nelson (2006) do not include the spread between the short term and the log term bonds, as opposed to Clarida, Gali & Gertler (2000). I chose to include it in the instruments because of its informationnal content about future inflation.
Figure 8: No spread in the instruments

Notes: inflation one year ahead, contemporaneous output gap

One can choose the trend of the GDP to be quadratic. The time serie is close to the CBO definition.
Figure 9: Quadratic trend for GDP

Notes: in the upper part, CBO’s definition in full line
quadratic trend in dashes; inflation one year ahead,
contemporaneous output gap

A.5 Determinacy of the equilibrium

Let us rewrite (5) in the case $h = l = 0$ with the shocks being omitted:

\[
\begin{align*}
\pi_t &= 4\kappa y_t + \beta E_t (\pi_{t+1}) \\
y_t &= E_t (y_{t+1}) - \sigma [i_t - E_t (\pi_{t+1})] \\
i_t &= a_x E_{t-1} (\pi_t) + a_y E_{t-1} (y_t) + a_i i_{t-1}
\end{align*}
\]
To check whether the equilibrium is locally determined, I follow Woodford (2003), ie check Blanchard & Kahn (1980) conditions. The model is slightly different from those detailed in Woodford (2003), that is why the whole proof is given. Define \( P_t = (\pi_t, y_t)' \) the non predetermined variables:

\[
\begin{pmatrix} 1 & -4\kappa \\ 0 & 1 \end{pmatrix} P_t = \begin{pmatrix} \beta & 0 \\ \sigma & 1 \end{pmatrix} E_t P_{t+1} \begin{pmatrix} 0 \\ 0 \end{pmatrix} i_t
\]

\[
E_t P_{t+1} = \begin{pmatrix} \beta & 0 \\ \sigma & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 & -4\kappa \\ 0 & 1 \end{pmatrix} P_t + \begin{pmatrix} \beta & 0 \\ \sigma & 1 \end{pmatrix}^{-1} \begin{pmatrix} 0 \\ 0 \end{pmatrix} i_t
\]

\[
E_t P_{t+1} = \begin{pmatrix} \frac{1}{\beta} & -\frac{4\kappa}{\beta} \\ -\frac{\sigma}{\beta} & 1 + \frac{4\kappa\sigma}{\beta} \end{pmatrix} P_t + \begin{pmatrix} 0 \\ 0 \end{pmatrix} i_t
\]

Now:

\[
i_{t+1} = a_x E_t (\pi_{t+1}) + a_y E_t (y_{t+1}) + a_i i_t \Rightarrow (-a_x, -a_y, 1) \left( E_t P_{t+1} \right) = a_i i_t
\]

Hence:

\[
\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -a_x & -a_y & 1 \end{pmatrix} \begin{pmatrix} E_t P_{t+1} \\ i_{t+1} \end{pmatrix} = \begin{pmatrix} \frac{1}{\beta} & -\frac{4\kappa}{\beta} \\ -\frac{\sigma}{\beta} & 1 + \frac{4\kappa\sigma}{\beta} \end{pmatrix} \begin{pmatrix} P_t \\ i_t \end{pmatrix}
\]

\[
\begin{pmatrix} E_t P_{t+1} \\ i_{t+1} \end{pmatrix} = \begin{pmatrix} \frac{1}{\beta} & 0 \\ 0 & \frac{1}{\beta} \end{pmatrix} \begin{pmatrix} P_t \\ i_t \end{pmatrix}
\]

\[
\begin{pmatrix} E_t P_{t+1} \\ i_{t+1} \end{pmatrix} = \begin{pmatrix} \frac{1}{\beta} & 0 \\ \frac{-\sigma}{\beta} & 1 + \frac{4\kappa\sigma}{\beta} \end{pmatrix} \begin{pmatrix} P_t \\ i_t \end{pmatrix}
\]

\[
\begin{pmatrix} E_t P_{t+1} \\ i_{t+1} \end{pmatrix} = \begin{pmatrix} \frac{1}{\beta} & 0 \\ \frac{-\sigma a_y}{\beta} & 1 + \frac{4\kappa a_y}{\beta} \end{pmatrix} \begin{pmatrix} a_x - a_y \\ 1 + \frac{4\kappa}{\beta} \end{pmatrix} \begin{pmatrix} P_t \\ i_t \end{pmatrix}
\]

I only consider parameters sets such that \( a_y > 0 \). The equilibrium is uniquely and locally determined if and only if \( A \) has one root inside the unit circle and two outside. I compute the characteristic polynomial of matrix \( A \):

\[-P_A (X) = X^3 + A_2 X^2 + A_1 X + A_0\]

with \( A_2 = -1 - \frac{1}{\beta} - a_x - \frac{4\kappa a_y}{\beta} - \sigma a_y < 0, A_1 = \frac{4\kappa a_y}{\beta} a_x + \frac{a_y}{\beta} a_x + \frac{4\kappa}{\beta} a_x + a_x + a_y + \frac{4\kappa a_y}{\beta} \frac{1}{\beta} + \frac{1}{\beta} > 0 \) and \( A_0 = -\frac{a_y}{\beta} < 0 \). I look for values of the parameters such that one of the three cases of proposition C.2 in Woodford (2003) is verified:

I 1 + A_2 + A_1 + A_0 < 0 and \(-1 + A_2 - A_1 + A_0 > 0\)

II 1 + A_2 + A_1 + A_0 > 0 and \(-1 + A_2 - A_1 + A < 0\) and \(A_0^2 - A_0 A_2 - A_1 - 1 > 0\)

III 1 + A_2 + A_1 + A_0 > 0 and \(-1 + A_2 - A_1 + A < 0\) and \(A_0^2 - A_0 A_2 - A_1 - 1 < 0\) and \(|A_2| > 3\)

Given the signs of \( A_0, A_1 \) and \( A_2, -1 + A_2 - A_1 + A_0 < 0 \), case I is excluded. Then, for case II or III to be verified, I must have:

38
\[1 + A_2 + A_1 + A_0 > 0 \iff a_x + \frac{1 - \beta}{4\kappa} a_y + a_i - 1 > 0 \quad (21)\]

Let us show that if (21) is verified then either case II or III is, which would achieve the proof.

\[A_0^2 - A_0 A_2 + A_1 - 1 = -\frac{4\sigma}{\beta^2} (1 - \beta) a_i + \frac{\sigma}{\beta} (1 - a_i) a_y + \frac{4\sigma}{\beta} a_x + \frac{1}{\beta^2} (\beta - a_i) (1 - a_i) (1 - \beta)\]

If (21) is true:

\[A_0^2 - A_0 A_2 + A_1 - 1 > -\frac{4\sigma}{\beta^2} (1 - \beta) a_i + \frac{\sigma}{\beta} (1 - a_i) a_y + \frac{1}{\beta^2} (\beta - a_i) (1 - a_i) (1 - \beta) + \frac{4\sigma}{\beta} \left(1 - a_i - \frac{1 - \beta}{4\kappa} a_y\right)\]

\[A_0^2 - A_0 A_2 + A_1 - 1 > -\frac{4\sigma}{\beta^2} (1 - \beta) a_i + \frac{4\sigma}{\beta} \beta (1 - a_i) + \frac{\sigma}{\beta} (1 - a_i) a_y - \frac{\sigma}{\beta} (1 - \beta) a_y + \frac{1}{\beta^2} (\beta - a_i) (1 - a_i) (1 - \beta)\]

\[A_0^2 - A_0 A_2 + A_1 - 1 > \frac{4\sigma}{\beta^2} (\beta - a_i) + \frac{\sigma}{\beta} (\beta - a_i) a_y + \frac{1}{\beta^2} (\beta - a_i) (1 - a_i) (1 - \beta)\]

\[A_0^2 - A_0 A_2 + A_1 - 1 > (\beta - a_i) \left[\frac{4\sigma}{\beta^2} + \frac{\sigma}{\beta} a_y + \frac{1}{\beta^2} (1 - a_i) (1 - \beta)\right]\]

If \(\beta > a_i\), \(A_0^2 - A_0 A_2 + A_1 - 1 > 0\) and case II is verified.

If \(a_i > \beta, \frac{1}{\beta} + a_i + \frac{4\sigma}{\beta} + \sigma a_y > \frac{1}{\beta} + \beta > 2\) because \(x \rightarrow x + \frac{1}{2}\) is decreasing on \([0, 1]\). Hence \(|A_2| = 1 + \frac{1}{\beta} + a_i + \frac{4\sigma}{\beta} + \sigma a_y > 3\), and therefore either case II, or case III is verified.

Finally, the equilibrium is locally determined if and only if:

\[a_x + \frac{1 - \beta}{4\kappa} a_y + a_i - 1 > 0 \iff \beta_x + \frac{1 - \beta}{4\kappa} \beta_y > 1\]
<table>
<thead>
<tr>
<th>Notes d'Études et de Recherche</th>
</tr>
</thead>
</table>


73. F. Chesnay and E. Jondeau, “Does correlation between stock returns really increase during turbulent period?,” April 2000.


Pour tous commentaires ou demandes sur les Notes d’Études et de Recherche, contacter la bibliothèque de la direction de la recherche à l’adresse suivante :

For any comment or enquiries on the Working Papers, contact the library of the Research Directorate at the following address :

BANQUE DE FRANCE
41-1404 Labolog
75049 Paris Cedex 01
tél : 0033 (0)1 42 92 49 55 ou 62 65
fax :0033 (0)1 42 92 62 92
email : thierry.demoulin@banque-france.fr
jeannine.agoutin@banque-france.fr