Does Inflation Walk on Unstable Paths?

Guido Ascari
University of Pavia
guido.ascari@unipv.it

Paolo Bonomolo
University of Pavia
paolo.bonomolo@eco.unipv.it

Hedibert F. Lopes
University of Chicago Booth School of Business
hlopes@chicagobooth.edu

May 22, 2012

Abstract

PRELIMINARY AND INCOMPLETE

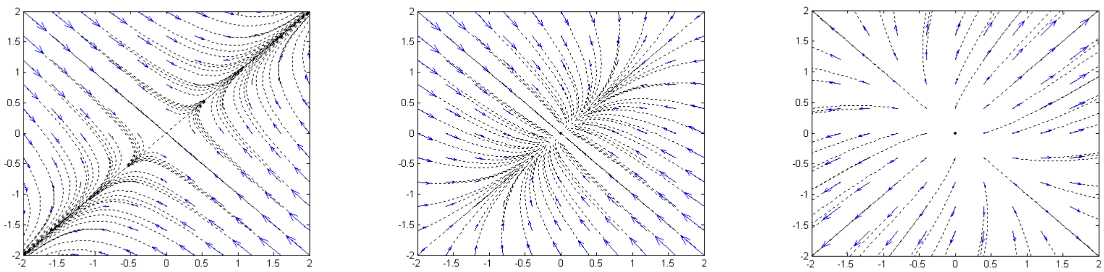
1 Introduction

Is there any evidence that the U.S. inflation is described, at least for a while, by unstable equilibrium paths? In this paper we explore this possibility estimating a basic New Keynesian model under different assumptions about the set of valid equilibria.

The possibility of unstable solutions was previously excluded a priori, because explosive paths would violate terminal conditions, and they will not be optimal. The literature refers to the rational expectations hypothesis as a sufficient condition to ensure stability: the economic agents have a complete knowledge of the environment, so they are able to select the path that respects optimality. The literature strived to find sufficient conditions also for the uniqueness of the stable solution: in that case the problem is said to be "determinate". By contrary, when there are many stable solutions, the economy can jump from one equilibrium path to another, because of self fulfilling believes. This situation of "indeterminacy" would create an additional source of inflation variations, and it is considered, by the New Keynesian literature (Clarida Galì and Gertler, 2000, and Lubik and Schorfheide, 2004, among others), as the cause of price instability in the U.S., during the Seventies (look at Figure 1, in Section 5). Then, the dichotomy between determinacy / indeterminacy is considered as the starting point to give policy prescriptions: the central banks should implement policies apt to avoid situations with infinite stable solutions.

This reasoning has at least two weak points. First, this dichotomy is based on the assumption that inflation can not explode, but hyperinflations happen, unfortunately. How much is reliable such an assumption? And, more important, is it proper to describe a situation in which inflation oscillates from 3% to 15%?

Moreover, note the paradoxical result of a stable system that generate instability, as opposed to an unstable system to ensure stability. Consider the observed series of inflation during the Seventies, in the so called "Great Inflation" period. Suppose we want to describe this pattern with a dynamic model. In the figure below there are the dynamics of three possible bivariate systems around an equilibrium point.



The first equilibrium is a saddle: the variables will explode unless they are on the saddle path that brings them in the equilibrium point. The second is a sink: no matter where the economy starts, it will reach the (stable) equilibrium asymptotically. The third case is a source: the equilibrium is unstable, and the variables will explode somewhere unless the starting point is exactly the steady state. Imagine to ask someone, who is not an expert in economics, the following question: "which system will you choose to describe inflation in the Great Inflation period?" The answer would probably be: "the third one". If we also ask to that person: "if you can, will you exclude a priori one of these systems?" He will say: "Yes, I exclude the second system: the stable one." It is exactly the opposite of what economists use to do. The usual practice is to use a system with multiple stable equilibrium paths to describe the unstable behavior of macroeconomic variables. Explosive paths, on the contrary, are a priori excluded by hypothesis.

This practice is at least counter-intuitive, and we want to test it. We suppose the economy is described by the New Keynesian model, and we compare the fit of this model under different assumptions on the set of valid solutions. We consider the case in which unstable paths are not a priori excluded, and we verify if it can help in describing the U.S. inflation.

We first present a theoretical framework in which we clarify the role of unstable solutions in models with rational expectations. We refer to the rational sunspots approach in which "temporary" unstable paths are not in contrast with the hypothesis of rational expectations.

We proceed estimating the model's parameters and the latent states using the particle learning approach of Carvalho, Johannes, Lopes and Polson (2010). This method relays on the assumption that the posterior distribution of the parameters depends on a set of sufficient statistics that are recursively updated. When we can not use this assumption, we approximate the posterior distribution of the parameters using mixtures of Normals, as in Liu and West (2001).

To compare the different models we use the sequential Bayes factor presented in West (1986). We provide evidence that the high inflation during the Seventies is better explained by unstable dynamics.

2 The Rational Sunspots Approach

2.1 Motivation

We want to verify if unstable paths can better explain the inflation dynamics. Then, a theoretical point raises: stability is implied by rational expectations, should we abandon this hypothesis? Cochrane (2011) asserts that "*Transversality conditions can rule out real explosions, but not nominal explosions*". In other words rational expectations leaves the possibility of hyperinflations. Alas, the New Keynesian model we analyze in the next section has both nominal and real variables that will eventually explode together with inflation.

However, the possibility of a temporary walk on unstable paths is not at all in contrast with rational expectations: a short lived deviation from the stable solution would not violate any transversality condition. This is the approach followed by Ascari and Bonomolo (2012), in which temporary walks on unstable path can be justified by temporary changes in the expectations formation process.

2.2 A simple example

Let formalize this idea with a simple example. Consider the following model inspired by Cochrane (2011), including the Fisher equation (1) and the Taylor rule (2):

$$i_t = r + E_t \pi_{t+1} \tag{1}$$

$$i_t = r + \phi \pi_t + \varepsilon_t \qquad \varepsilon_t \sim N(0, \sigma_\varepsilon^2) \tag{2}$$

i_t is the nominal interest rate at time t, r is the real interest rate (assumed constant for simplicity), π_t is inflation and ε_t is a white noise exogenous shock. Finally, $E_t \pi_{t+1} = E(\pi_{t+1}|I_t)$ that is the expected value of inflation at $t+1$, conditional on the information set available at time t. Coherently with the rational expectations hypothesis (in the strong form), we assume that the set I_t contains all the relevant information: all the present and past values of the endogenous and exogenous variables, and the structure of the model with its parameters.

Let concentrate on inflation: the two equations above implies the following model:

$$\pi_t = \frac{1}{\phi} E_t \pi_{t+1} - \frac{1}{\phi} \varepsilon_t \tag{3}$$

Equation (3) has an infinite number of solutions, because one can find an infinite number of couples $(\pi_t, E_t \pi_{t+1})$ that clear the equation. However, remember that $E_t \pi_{t+1}$ is a function of π_t because it is an expected value conditional on an information set that contains also inflation at time t. Then, we can impose some additional restrictions to limit the set of allowed solutions. In this spirit, we use the original method by Muth (1961), supposing that inflation at time t is a linear function of only present, past, and expected future values of the exogenous shock. Using this assumption, we derive the set of solutions parametrized by $b \in (-\infty, +\infty)$ (see the Appendix):

$$\pi_t = \phi \pi_{t-1} + \varepsilon_{t-1} - \frac{b}{\phi} \varepsilon_t \tag{4}$$

Equation (4) represents all the solutions of equation (3), each one corresponding to a particular value of b. As an example, we write two important cases, often considered in the literature: following the terminology used by Blanchard (1979), we have the pure forward looking solution corresponding to $b = 1$,

$$\pi_t^F = -\frac{1}{\phi} \varepsilon_t \tag{5}$$

and the pure backward looking solution, corresponding to $b = 0$,

$$\pi_t^B = \phi \pi_{t-1}^B + \varepsilon_{t-1} \tag{6}$$

We can understand how the central bank can influence the dynamics of inflation. The Taylor rule (2) describes the monetary policy implemented. The Taylor principle states that the central bank conducts an "active" policy if it moves the nominal interest rate more than proportionally with respect to inflation's variations, that is when $|\phi| > 1$. Otherwise the policy is "passive". From equation (4) it is clear that when the Taylor principle is not respected, for every $b \in (-\infty, +\infty)$ the implied dynamics are stable. By contrary when the central bank conducts an active monetary policy all the solutions are unstable but the forward looking one: equation (5). Now, fostering on common sense, a central bank should not respect the Taylor principle, to be sure the inflation is described by stable dynamics. Posing $\phi > 1$ can be highly risky, because the probability of being on the unique stable path (among infinite unstable) is practically zero! Not if we are under the rational expectations hypothesis. In that case, when ϕ is inside the unit circle, the economy can choose randomly the solution, and change it at every period. This creates an additional source of inflation's variation that enhances its volatility. Moreover, this translates in uncertainty in the conduction of monetary policy, because the response of inflation to a monetary policy shock is not predictable. This regrettable situation can be avoided posing ϕ greater than one in absolute value. In that case only the forward looking solution is valid: any other solution will violate a terminal condition.

The economy can select equation (5) by mean of expectations. At a particular time, each solution corresponds to a particular expected value, and the Taylor principle forces the agents to coordinate their expectations on the unique value that corresponds to the forward looking solution, that in this simple example is $E_t\pi_{t+1} = 0$ $\forall t$. This correspondence suggests an economic interpretation for b.

2.3 An interpretation for b

Following Ascari and Bonomolo (2012) we can interpret b as the way agents form their expectations, under the rational expectations hypothesis. To understand this point, go back to the original spirit of Muth (1961). One of the purpose of that paper is to write the expectation at time t as an exponentially weighted average of past observations. The previous paper by Muth (1960), in fact, demonstrates that, under some assumptions, this is an optimal estimator. In the simple case of equation (3) we obtain the following expression:

$$E_t\pi_{t+1} = (b-1)\sum_{i=1}^{\infty}\left(\frac{\phi}{b}\right)^i \pi_{t+1-i} \tag{7}$$

Then, b determine how the agents consider past observations in making forecasts. $E_t\pi_{t+1}$ is the product of two terms: $(b-1)$, that tells how much past is important in the expectations formation, and the weighted average in which the weights are function of b. Intuitively, the agents can form expectations looking relatively more to the past (when b is far from one), or looking to the steady state value (when b is exactly equal to one): in the last limiting case, represented by the forward looking solution, we have $E_t\pi_{t+1} = 0$.

2.4 The Rational Sunspots

Start considering the case of a passive monetary policy. When $|\phi| < 1$ the agents are free to choose any solutions, and the economy can jump among equilibrium paths just because of self fulfilling believes. This behavior can be modeled with sunspots. Ascari and Bonomolo (2012) construct sunspots randomizing among the infinite rational expectations equilibria. Because they are parametrized by b, let randomize over this parameter. More precisely, assume

$$b_t = b_{t-1} + \zeta_t \qquad \zeta_t \sim N(0, \sigma_\zeta^2) \tag{8}$$

where ζ_t is a Normally distributed sunspot shock. This last condition, in terms of equation (7), can be interpreted as changes in the expectations formation process. Suppose that in some periods the agents form their expectations trusting a lot the past, while in other periods they expect inflation to be more or less around its steady state. This mechanism is labeled "rational sunspots" because we are randomizing over the set of rational expectations solutions: what pushes the economy in changing the expected value, is not an external element, but it is something related to a degree of freedom we have in making forecasts, coherently with the rational expectations hypothesis.

Is this possibility only restricted to the case of passive policy? When the Taylor principle is respected any unstable solution is excluded because it is not optimal: the transversality condition will be violated. But it happens for sure if b is constant. On the other hand, if b is time varying, you can imagine temporary deviations from the stable solution, so that the terminal condition will not be violated. These deviations, in terms of equation (7) can be interpreted as temporary changes in the expectations formation process. The opinion of the present paper, is that in the empirical analysis, we can not exclude this situation. Hence, in the following sections, we propose a test to verify the empirical validity of these temporary unstable paths.

When rational sunspots affect the solution of the model, supposing that b_t follows equation (8), and assuming that expectations have the form of equation (7) in every period, we have that the complete set of solutions are represented by equation (9)

$$\begin{aligned} \pi_t &= \vartheta_t \pi_{t-1} + \varepsilon_{t-1} - \frac{b_t}{\phi}\varepsilon_t \ , \\ \vartheta_t &= \phi \frac{(1-b_t)}{(1-b_{t-1})} \end{aligned} \tag{9}$$

that has the same form of equation (4), but with time varying parameters. In other words, rational sunspots can be considered as an economic explanation of drifting parameters and stochastic volatility.

2.5 The General Solution

We consider the class of models that can be written in the form of Blanchard and Kahn (1980):

$$\begin{bmatrix} X_{t+1} \\ E_t P_{t+1} \end{bmatrix} = A \begin{bmatrix} X_t \\ P_t \end{bmatrix} + \boldsymbol{\gamma} Z_t$$

where X_t is a $(n \times 1)$ vector of predetermined variables, and P_t is a $(m \times 1)$ vector of non-predetermined variables. The exogenous disturbances are collected in the $(\kappa \times 1)$ vector Z_t, that has a multivariate normal distribution: $Z_t \sim i.i.d.\ N(\mathbf{0}, \Sigma)$. The exogenous shocks in Z_t are called fundamental errors. Finally, A and $\boldsymbol{\gamma}$ are matrices with the parameters of the model.

The matrix A can be rewritten using the Jordan decomposition

$$A = C^{-1} J C$$

and we define the following set of block matrices:

$$C^{-1} = \begin{bmatrix} B_{11} & B_{12} \\ (n \times n) & (n \times m) \\ B_{21} & B_{22} \\ (m \times n) & (m \times m) \end{bmatrix}, \quad C = \begin{bmatrix} C_{11} & C_{12} \\ (n \times n) & (n \times m) \\ C_{21} & C_{22} \\ (m \times n) & (m \times m) \end{bmatrix},$$

$$J = \begin{bmatrix} J_1 & \mathbf{0} \\ (n \times n) & (n \times m) \\ \mathbf{0} & J_2 \\ (m \times n) & (m \times m) \end{bmatrix}, \quad \boldsymbol{\gamma} = \begin{bmatrix} \gamma_1 \\ (n \times \kappa) \\ \gamma_2 \\ (n \times \kappa) \end{bmatrix}$$

In the Appendix we show that the rational sunspots solutions take the form of the following system:

$$X_t = (B_{11}J_1C_{11} + B_{12}J_2C_{21})\,X_{t-1} + (B_{11}J_1C_{12} + B_{12}J_2C_{22})\,P_{t-1} + \gamma_1 Z_{t-1} \tag{10}$$

$$\begin{aligned} C_{21}X_t + C_{22}P_t &= J_2H_t\left(C_{21}X_{t-1} + C_{22}P_{t-1}\right) + \\ &\quad + H_t(C_{21}\gamma_1 + C_{22}\gamma_2)Z_{t-1} + \mathbf{b}_t J_2^{-1}(C_{21}\gamma_1 + C_{22}\gamma_2)Z_t \end{aligned} \tag{11}$$

where $\mathbf{b}_t$ is a $(m \times m)$ diagonal matrix:

$$\mathbf{b}_t = \begin{bmatrix} b_{1,t} & 0 & \cdots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & b_{m,t} \end{bmatrix}$$

in which

$$b_{i,t} = b_{i,t-1} + \zeta_{i,t} \qquad \zeta_{i,t} \sim N(0, \sigma^2_{\zeta_i}) \quad \forall i$$

and $H_t = (I + \mathbf{b}_t)(I + \mathbf{b}_{t-1})^{-1}$. The matrix $\mathbf{b}_t$ plays the same role of the coefficient b introduced above, parametrizing the solutions. In general, if we have m non predetermined variables, the cardinality of the set of solutions is infinite to the power of m. However, as in the simple example here presented, when the eigenvalues of the model are outside the unit circle, we can put some restrictions on the elements in

$\mathbf{b}_t$, reducing the set of solution to a smaller one. In practice, consider the following *stability criterion*:

> *stability criterion*: for $i = 1...m$, if $|J_{2,i}| > 1$ put $b_{i,t} = -1\ \forall t$,
>
> where $J_{2,i}$ is the i^{th} element in the main diagonal of J_2, and $b_{i,t}$ is the i^{th} element in the main diagonal of $\mathbf{b}_t$.

The criterion reduces the degrees of freedom in the matrix $\mathbf{b}_t$, and then, it downsizes the set of valid solutions. If, for example, there are $r \leq m$ number of eigenvalues outside the unit circle, the number of solutions, applying the criterion, is $\infty^{(m-r)}$. The limiting case is when the Blanchard-Kahn condition is satisfied, that is when the number of eigenvalues outside the unit circle is equal to the number of non predetermined variables: the criterion forces all the elements in the main diagonal of $\mathbf{b}$ to be equal to -1, and this is the unique stable solution. If the criterion is not satisfied, the dynamics of the variables will be unstable.

3 The New Keynesian Model

We want to test the validity of the stability criterion, when the New Keynesian model is called to explain the U.S. inflation dynamics. Let introduce a basic version of the model, described by the following three equations:

$$x_t = E_t x_{t+1} - \frac{1}{\sigma}\left(i_t - E_t \pi_{t+1}\right) + e_t^x \tag{NKIS}$$

$$\pi_t = \beta E_t \pi_{t+1} + k x_t + e_t^\pi \tag{NKPC}$$

$$i_t = \rho_i i_{t-1} + (1-\rho_i)\left[\phi_x x_t + \phi_\pi \pi_t\right] + \varepsilon_t^i \qquad \varepsilon_t^i \sim N(0, \sigma_i^2)\ . \tag{TR}$$

The first equation is the New Keynesian IS curve, that relates the output gap x_t to the real interest rate. The dynamics of the inflation rate π_t are described by the second equation, the New Keynesian Phillips curve. The (NKIS) and the (NKPC) come from the maximization problem of the households and the firms, and they are found loglinearizing, around the steady state, the respective first order conditions. A standard Taylor rule (TR) closes the model. It describes how the central bank conducts the monetary policy, moving the nominal interest rate i_t, in response to the deviations of inflation and output gap from their steady state.

We also suppose that the shocks in the NKIS and in the NKPC are autocorrelated, that is

$$e_t^x = \rho_x e_{t-1}^x + \varepsilon_t^x \qquad \varepsilon_t^x \sim N(0, \sigma_x^2) \tag{12}$$

$$e_t^\pi = \rho_\pi e_{t-1}^\pi + \varepsilon_t^\pi \qquad \varepsilon_t^\pi \sim N(0, \sigma_\pi^2) \tag{13}$$

Even if the last two equations do not come from the microfounded framework, this is a standard hypothesis: its aim is to capture the empirical persistence of the data, that the model seems to ignore. However, this is a crucial hypothesis only when the model is described by a particular solution: the forward looking one. In all the other cases expectations are formed taking explicitly into account the past history of the variables, and the model is able per se to display persistence.

The model has five variables, three predetermined and two non predetermined. Then, the matrix $\mathbf{b}_t$ has dimension two. We also know that among the five eigenvalues of the matrix A, in the Blanchard - Kahn form, three of them are inside the unit circle (because ρ_x, ρ_π, and ρ_i are less than one in absolute value), and one is always outside the unit circle (see Bullard and Mitra, 2002). The remaining eigenvalue can be inside or outside the unit circle, depending on the conduction of monetary policy. It is straightforward to verify that, when the following condition holds,

$$\phi_\pi > 1 - \frac{1-\beta}{k}\phi_x \tag{14}$$

the model has two eigenvalues greater than one in absolute value. In that case, the Blanchard - Kahn condition holds (that is the number of eigenvalues outside the unit circle is equal to the number of non predetermined variables), and there is a unique stable solution: the forward looking. If the condition (14) is respected we are under "determinacy", and the monetary policy is said to be "active". Vice versa, the policy is "passive".

We can test the validity of the stability criterion in a particular sample comparing the relative performance of the New Keynesian model here presented, under different hypotheses on the set of valid solutions. In this spirit, we compare two assumptions: one in which the stability criterion is imposed, and one in which we take solutions excluded by the same criterion.

Model M_S: the subset of stable solutions

Many researcher impose condition (14) before estimating the New Keynesian model. This assumption can be too strong (see, for example, Lubik and Schorfheide, 2004). We allow for a passive monetary policy to be implemented, and we only exclude unstable solutions. We label this case as model M_S, and the matrix $\mathbf{b}_t$ is:

$$\mathbf{b}_t = \begin{bmatrix} b_{1,t} & 0 \\ 0 & 1 \end{bmatrix}$$

$$b_{1,t} = \begin{cases} 1 & \text{if } \phi_\pi > 1 - \frac{1-\beta}{k}\phi_x \\ b_{1,t-1} + \zeta_t & \zeta_t \sim N(0, \sigma_\zeta^2) \quad \text{otherwise.} \end{cases}$$

The south east element in $\mathbf{b}_t$ is imposed to be -1 because, in the matrix A of the Blanchard Kahn canonical form, there is always one "explosive" eigenvalue. The first element, on the other hand, is $b_{1,t}$, and it follows a random walk if we have an infinite number of stable solutions. Otherwise, if there is also another eigenvalue outside the unit circle, it is automatically posed equal to minus one.

Model M_U: a subset of unstable solutions

The assumption, here, is that the stability criterion, in general, does not hold: we define the matrix $\mathbf{b}_t$ as:

$$\mathbf{b}_t = b_{1,t} I$$

$$b_{1,t} = b_{1,t-1} + \zeta_t \qquad \zeta_t \sim N(0, \sigma_\zeta^2)$$

The set of solutions considered does not contain the stable set allowed in M_S: the intersection of the two is the forward looking solution, that is the unique possibility for the stability criterion to hold in this case.

The next section explains the method used to compare the two assumptions just presented.

4 Econometric strategy

We use an econometric strategy that is thought to deal with the following peculiarities: i) the model has stochastic volatility, then the likelihood distribution is not Gaussian; ii) we are interested in tracking the behavior of b_t, that can be considered as a stochastic latent process; iii) we would like to study the fit of different models, and eventually compare them, during different periods. Then, the econometric strategy is based on Bayesian methods, in particular on *Particle filtering*, and on *Sequential model monitoring.*

4.1 Particle filtering

Start with notation: indicate with $y_{1:t} = \{y_1, y_2, ...y_t\}$ the observed data (containing series for inflation, output gap and nominal interest rate); ϑ_t is the vector of latent processes at time t, but the variable b_t; ψ is the vector with the variances of exogenous shocks; finally, ω is the vector with the other parameters. We are interested in approximating the posterior distribution:

$$f\left(\vartheta_{0:T}, b_{0:T}, \omega, \psi | y_{1:T}\right)$$

To estimate latent processes and parameters we use the Particle Learning approach by Carvalho, Johannes, Lopes and Polson (2010). We make a sequential inference on the parameters using their full conditional distributions, when it is possible to derive them analytically. An analytical expression is typically available for the variances of the exogenous shocks (the vector ψ), using appropriate conjugate distributions. For the other parameters, collected in ω, we implement the Liu and West (2001) approach, approximating the posterior distribution with mixtures of Normals.

In what follows we derive the algorithm we use in making inference, starting from the basic particle filter.

The basic particle filter

Consider, for the moment, the parameters as known, and indicate the latent processes as $l_t = (\vartheta_t, b_t)$. l_t is a Markovian process, and we are interested in the expected value:

$$E_f\left[l_t | l_{t-1}, y_{1:t}\right] = \int l_t f\left(l_t | l_{t-1}, y_{1:t}\right) dl_t$$

where the subscript f in the expectation operator indicates that the expected value is computed under the density distribution f. We can write

$$E_f\left[l_t | l_{t-1}, y_{1:t}\right] = \int l_t \frac{f\left(l_t | l_{t-1}, y_{1:t}\right)}{q\left(l_t\right)} q\left(l_t\right) dl_t$$

where $q\left(l_t\right)$ is a proposal distribution for l_t. Then, posing

$$w_t = \frac{f\left(l_t | l_{t-1}, y_{1:t}\right)}{q\left(l_t\right)} \tag{15}$$

we have

$$E_f\left[l_t | l_{t-1}, y_{1:t}\right] = \int l_t w_t q\left(l_t\right) dl_t = E_q\left[l_t w_t | l_{t-1}, y_{1:t}\right]$$

We can draw a large number of particles N from $q\left(l_t\right)$, evaluate and normalize the weights w_t, and approximate $E_f\left[l_t | l_{t-1}, y_{1:t}\right]$ using

$$E_f\left[l_t | l_{t-1}, y_{1:t}\right] \approx \frac{1}{N} \sum_{i=1}^{N} l_t^{(i)} w_t^{(i)}$$

The superscript (i) indicates the i^{th} particle. The proposal distribution $q\left(l_t\right)$, also called importance density, plays a crucial role. As an example, consider $q\left(l_t\right) = q\left(l_t | l_{t-1}\right)$, that is the prior distribution. Then, the weights are simply proportional to the likelihood, and the algorithm, known as the *bootstrap filter*, is:

`For every` t `:`

`1- Propagate:draw` $l_t^{(i)}$ `from` $q\left(l_t | l_{t-1}^{(i)}\right)$

`2- Resample:  compute` $w_t^{(i)} \propto f(y_t | l_t^{(i)})$ `and resample` $l_t^{(i)}$ `according to` $w_t^{(i)}$

The second step is optional. Using the prior as the proposal density is not very useful. The optimal distribution, that is the one who minimizes the variance of the estimator, is represented by the filtering distribution of the Kalman filter: $q\left(l_t|l_{t-1}, y_t\right)$. Moreover, the *bootstrap filter* can be improved reversing the steps: first resample, and then propagate. These two improvements lead to the Particle Learning approach.

Particle learning

We want to make inference also on the parameters. Using Particle Learning we can do it sequentially, together with the latent variables, using their full conditional distributions. Consider, in the set of unknowns, also the vector ψ, containing the variances. For these parameters we can use sufficient statistics $s_t = S(s_{t-1}, l_t, y_t)$ to represent the posteriors. The sufficient statistics can be recursively updated, and they are random variables because they depend on l_t, so they can be part of the latent vector $z_t = (x_t, s_t)$.

To improve efficiency with respect to the basic particle filter, use $q\left(l_t|l_{t-1}, y_t\right)$ as the importance distribution. In that case the weights are proportional to the predictive likelihood. Moreover, use a Resample-Propagate framework. The *Particle Learning* filter is:

For every t :
1- Resample: compute $w_t^{(i)} \propto f(y_{t+1}|z_t^{(i)})$ and resample $z_t^{(i)}$ according to $w_t^{(i)}$
2- Propagate: draw $x_t^{(i)}$ from $q\left(x_t|x_{t-1}^{(i)}, y_t\right)$ and $s_t^{(i)}$ from $S(s_{t-1}^{(i)}, x_t^{(i)}, y_t)$
3- Draw $\psi^{(i)}$ from $f\left(\psi|s_t^{(i)}\right)$

Our algorithm

There are two difficulties in implementing Particle Learning in our framework: i) the optimal proposal distribution is not available in our case, because the likelihood distribution is not Gaussian; ii) we can not use sufficient statistics to represent the posterior of the parameters in ω.

The first problem is addressed noting that we have a conditional linear model: considering $b_{0:T}$ as given, the posterior $f\left(\vartheta_{0:T}|b_{0:T},\omega,\psi,y_{1:T}\right)$ is Gaussian, then we can use the "optimal" distribution to approximate it. In fact we can write:

$$f\left(\vartheta_{0:T},b_{0:T}|\omega,\psi,y_{1:T}\right)=\underbrace{f\left(\vartheta_{0:T}|b_{0:T},\omega,\psi,y_{1:T}\right)}_{\text{approximated with optimal proposal}}\underbrace{f\left(b_{0:T}|\omega,\psi,y_{1:T}\right)}_{\text{approximated with blind proposal}} \tag{16}$$

and we use a blind proposal to make inference only on b_t. This practice, also called "Rao-Blackwellization", use the Rao-Blackwell theorem to reduce the variance of our estimator. Then, we use the prior as the proposal distribution for b_t, that is $q(b_t|b_{t-1},\psi)$, and the optimal proposal $q\left(\vartheta_t|\vartheta_{t-1},b_t,\omega,\psi,y_t\right)$ for the other latent variables.

Also the predictive likelihood, used in the Particle Learning filter to compute weights, is not available. We can use, instead, a predictive density that is conditional on a good guess for b_t (we use its expected value).

Finally, we want to make inference on the other parameters collected in ω. We can approximate the posterior of ω using mixtures of Normals as in Liu and West (2001):

$$f(\omega|y_{1:t})=\sum_{i=1}^{N}N\left(m^{(i)};h^2\Sigma\right)$$

where

$$\begin{aligned}m^{(i)}&=a\omega^{(i)}+(1-a)\bar{\omega}\\ \Sigma&=Var(\omega)\end{aligned}$$

$\bar{\omega}$ is the sample mean of $\left\{\omega^{(i)}\right\}_{i=1}^{N}$, while a and h are parameters that govern the shrinkage and the degree of overdispersion of the mixture.

Then, we use the following algorithm:

`For` $t=1...T$`:`

`0` `Compute` $\bar{\omega}=E(\omega)$ `and` $\Sigma=Var(\omega)$`. For` $i=1...N$ `put`

$$\begin{aligned}m^{(i)}&=a\omega^{(i)}+(1-a)\bar{\omega}\\ g(b_{t-1}^{(i)})&=E(b_t|b_{t-1}=b_{t-1}^{(i)})\end{aligned}$$

For $i = 1...N$

1 Compute weights: $\tilde{w}_t^{(i)} \propto w_{t-1}^{(i)} q\left(y_t | \vartheta_{t-1}^{(i)}, g(b_{t-1}^{(i)}), m^{(i)}, \psi^{(i)}\right)$

2 Resample $\left\{\tilde{\vartheta}_{t-1}^{(i)}\right\}_{i=1}^{N} \left\{\tilde{b}_{t-1}^{(i)}\right\}_{i=1}^{N} \left\{\tilde{s}_{t-1}^{(i)}\right\}_{i=1}^{N} \left\{\tilde{m}^{(i)}\right\}_{i=1}^{N} \left\{\tilde{\psi}^{(i)}\right\}_{i=1}^{N}$ according to $\tilde{w}_t^{(i)}$

3 Propagate:

(i) draw $\tilde{\omega}^{(i)}$ from $N(\omega; \tilde{m}^{(i)}, h^2\Sigma)$

(ii) draw $\tilde{b}_t^{(i)}$ from $q\left(b_t | \tilde{b}_{t-1}^{(i)}, \tilde{\psi}^{(i)}\right)$

(iii) draw $\tilde{\vartheta}_t^{(i)}$ from $q\left(\vartheta_t | \tilde{\vartheta}_{t-1}^{(i)}, \tilde{b}_t^{(i)}, \tilde{\omega}^{(i)}, \tilde{\psi}^{(i)}, y_t\right)$

4 Compute new weights: $w_t^{(i)} = \dfrac{f(y_t | \tilde{\vartheta}_{t-1}^{(i)}, \tilde{b}_t^{(i)}, \tilde{\omega}^{(i)}, \tilde{\psi}^{(i)})}{q(y_t | \tilde{\vartheta}_{t-1}^{(i)}, g(\tilde{b}_{t-1}^{(i)}), \tilde{m}^{(i)}, \tilde{\psi}^{(i)})}$

5 Update sufficient statistics $\tilde{s}_t^{(i)} = S(\tilde{s}_{t-1}^{(i)}, \tilde{\vartheta}_t^{(i)}, y_t)$

6 Draw $\tilde{\psi}^{(i)}$ from $f\left(\tilde{\psi}^{(i)} | \tilde{s}_t^i\right)$

7 Final draws: $\left\{\vartheta_t^{(i)}\right\}_{i=1}^{N} \left\{b_t^{(i)}\right\}_{i=1}^{N} \left\{s_t^{(i)}\right\}_{i=1}^{N} \left\{\omega^{(i)}\right\}_{i=1}^{N} \left\{\tilde{\psi}^{(i)}\right\}_{i=1}^{N}$ using $w_t^{(i)}$

The weights at step 4 are computed as indicated in equation (15):

$$w_t^{(i)} = \frac{f\left(\tilde{\vartheta}_t^{(i)}, \tilde{b}_t^{(i)}, \tilde{\omega}^{(i)}, \tilde{\psi}^{(i)} | \tilde{\vartheta}_{t-1}^{(i)}, \tilde{b}_{t-1}^{(i)}, y_t\right)}{q\left(\tilde{\vartheta}_t^{(i)}, \tilde{b}_t^{(i)}, \tilde{\omega}^{(i)}, \tilde{\psi}^{(i)}\right)}$$

where $q\left(\tilde{\vartheta}_t^{(i)}, \tilde{b}_t^{(i)}, \tilde{\omega}^{(i)}, \tilde{\psi}^{(i)}\right)$ is the proposal density, equal to

$$q\left(y_t | \tilde{\vartheta}_{t-1}^{(i)}, g(\tilde{b}_{t-1}^{(i)}), \tilde{m}^{(i)}, \tilde{\psi}^{(i)}\right) q\left(\vartheta_t | \tilde{\vartheta}_{t-1}^{(i)}, \tilde{b}_t^{(i)}, \tilde{\omega}^{(i)}, \tilde{\psi}^{(i)}, y_t\right) q\left(b_t | \tilde{b}_{t-1}^{(i)}, \tilde{\psi}^{(i)}\right) f\left(\tilde{\psi}^{(i)} | \tilde{s}_t^i\right) N(\omega; \tilde{m}^{(i)}, h^2\Sigma)\ .$$

Then, the weights simplify to

$$w_t^{(i)} = \frac{f\left(y_t | \tilde{\vartheta}_t^{(i)}, \tilde{b}_t^{(i)}, \tilde{\omega}^{(i)}, \tilde{\psi}^{(i)}\right) f\left(\tilde{\vartheta}_t^{(i)} | \tilde{\vartheta}_{t-1}^{(i)}, \tilde{b}_t^{(i)}, \tilde{\omega}^{(i)}, \tilde{\psi}^{(i)}\right)}{q\left(y_t | \tilde{\vartheta}_{t-1}^{(i)}, g(\tilde{b}_{t-1}^{(i)}), \tilde{m}^{(i)}, \tilde{\psi}^{(i)}\right) q\left(\tilde{\vartheta}_t^{(i)} | \tilde{\vartheta}_{t-1}^{(i)}, \tilde{b}_t^{(i)}, \tilde{\omega}^{(i)}, \tilde{\psi}^{(i)}, y_t\right)}\ . \qquad (17)$$

Consider the density $q\left(\tilde{\vartheta}_t^{(i)}|\tilde{\vartheta}_{t-1}^{(i)},\tilde{b}_t^{(i)},\tilde{\omega}^{(i)},\tilde{\psi}^{(i)},y_t\right)$ in the denominator: it can be rewritten as

$$q\left(\tilde{\vartheta}_t^{(i)}|\tilde{\vartheta}_{t-1}^{(i)},\tilde{b}_t^{(i)},\tilde{\omega}^{(i)},\tilde{\psi}^{(i)},y_t\right)=\frac{f\left(y_t|\tilde{\vartheta}_t^{(i)},\tilde{b}_t^{(i)},\tilde{\omega}^{(i)},\tilde{\psi}^{(i)}\right)f\left(\tilde{\vartheta}_t^{(i)}|\tilde{\vartheta}_{t-1}^{(i)},\tilde{b}_t^{(i)},\tilde{\omega}^{(i)},\tilde{\psi}^{(i)}\right)}{f\left(y_t|\tilde{\vartheta}_{t-1}^{(i)},\tilde{b}_t^{(i)},\tilde{\omega}^{(i)},\tilde{\psi}^{(i)}\right)}$$

that substituted in (17) gives the expression at step 4 in the algorithm.

4.2 Sequential model monitoring

The use of a Sequential Monte Carlo has the advantage of comparing different models in different periods of time. To reach this goal we use the sequential Bayes factor by West (1986), as suggested in Carvalho, Johannes, Lopes and Polson (2010).

Suppose you want to compare two models: M_S and M_A. You can implement the following method:

`For` $t=1...T$

`1    Compute the predictive likelihood:` $f(y_t|y_{0:t-1},M_i)$ $i=S,A$

`2    Compute the likelihood ratio`

$$H_t=\frac{f(y_t|y_{0:t-1},M_S)}{f(y_t|y_{0:t-1},M_A)}$$

`3    Compute` $W_t(\kappa)=H_tH_{t-1}...H_{t-\kappa+1}$

$W_t(\kappa)$ is called the cumulative Bayes factor and it assesses the fit of the most recent κ observations.

With the sequential model monitoring we can track the relative performance of different models during time, and distinguish periods where a model is much better than another.

5 Empirical Results

5.1 Data and subsamples

We estimate the New Keynesian model of Section 3 using quarterly data for inflation, output gap and nominal interest rate. All the series are from the FRED database. In particular, inflation is computed as the first difference in the logarithm of the price level (Consumer Price Index) between two subsequent quarters; the output gap is obtained detrending the logarithm of the real output using the Hodrick-Prescott filter; finally, about the nominal interest rate, we suppose the central bank's instrument is the Federal Funds Rate.

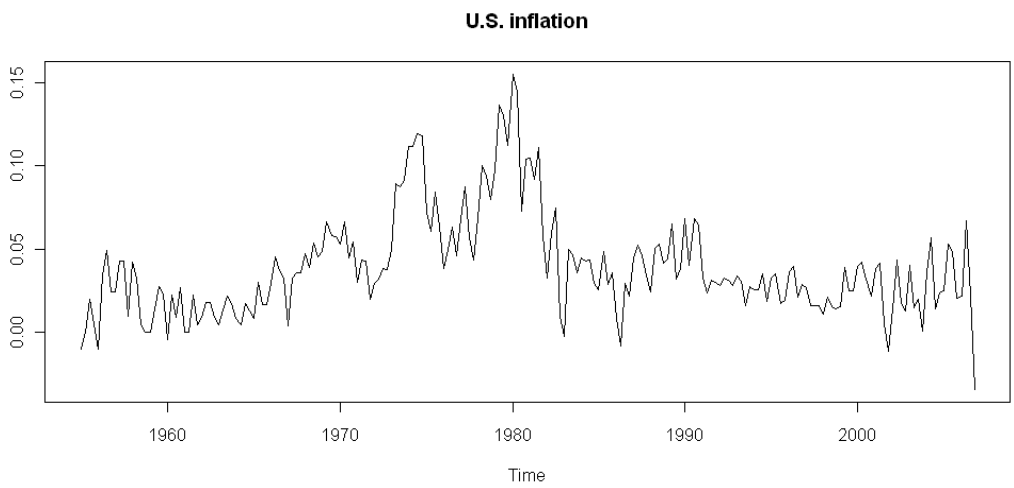


Figure 1: CPI inflation, quarterly data. Sample: 1955Q1 - 2006Q4

Figure 1 plots the inflation series. As it is clear, from the mid Sixties until the end of Seventies, the U.S. experienced a period of price instability, also labeled as "Great Inflation". Then, since the first Eighties, when the Fed was under the chair of Paul Volcker, prices came back under control: inflation became low, as low became the volatility of prices and of other macroeconomic variables. By contrast to the previous period, these times are known as the "Great Moderation". The New Keynesian literature explains the shift from the Great Inflation to the Great Moderation with the shift from a passive to an active monetary policy, in the terms we referred above.

As we underlined previously, this interpretation excludes a priori unstable paths, despite inflation exceeded 15%. This consideration raises the following question: would you explain the Great Inflation with a stable system, as in the New Keynesian literature, or with unstable dynamics? To have a better comparison of our results with the pre-existing works (Clarida Galì Gertler, 2000, Lubik and Schorfheide, 2004, among others), we consider two subsample: the pre-Volcker period, from 1960q1 to 1979q3, and the Volcker-Greenspan period, from 1979q4 to 1997q4.

5.2 The Evidence

Table 1 collects the priors for the parameters. Two parameters, the intertemporal elasticity of substitution and the subjective discount factor, are calibrated to conventional values. We use conjugate priors for the variance of the shocks (see the previous section). The distributions for the other parameters are in accordance with usual restrictions about the signs, and with the estimates found in the literature.

Table 1

Priors and calibrations		
Parameter	Distribution	Calibrated
β		0.99
σ		1
ρ_i	$U(0,1)$	
ρ_x	$U(0,1)$	
ρ_π	$U(0,1)$	
ϕ_x	$G(0.5,1)$	
ϕ_π	$G(2,2)$	
k	$G(0.5,1)$	
σ_i^2	$IG(2,0.001)$	
σ_x^2	$IG(2,0.001)$	
σ_π^2	$IG(2,0.001)$	
σ_ζ^2	$IG(2,0.01)$	

Table 2 reports the estimates of the parameters, for the two models in the two subsamples analyzed. The convergence to these values through time (we are using an on line estimator) is shown in the Appendix. Start considering the first subsample. Under stability (model M_S) the forward looking solution is selected: the estimate of ϕ_π is greater than one, suggesting that the Fed respected the Taylor principle. Because M_S excludes unstable equilibrium paths, the matrix $\mathbf{b}_t$ is equal to the identity by mean of the stability criterion. Then, if we restrict the solutions to the stable ones, in a certain sense, data select "the most unstable system", instead of describing the behavior of the variables through an infinite number of stable equilibrium dynamics.

Table 2

Estimates				
	1960Q1 - 1979Q3		1979Q4 - 1997Q4	
Parameter	M_S	M_U	M_S	M_U
ρ_i	0.57 [0.45 0.68]	0.63 [0.45 0.79]	0.74 [0.69 0.79]	0.66 [0.55 0.75]
ρ_x	0.5 [0.12 0.78]	0.12 [0.01 0.3]	0.23 [0.11 0.39]	0.58 [0.47 0.67]
ρ_π	0.04 [0.01 0.09]	0.03 [0.01 0.22]	0.03 [0.01 0.06]	0.06 [0.01 0.16]
k	0.04 [0.02 0.07]	0.02 [0.004 0.06]	0.01 [0.001 0.04]	0.21 [0.08 0.42]
ϕ_x	0.28 [0.25 0.32]	0.032 [0.029 0.037]	0.3 [0.29 0.34]	0.33 [0.31 0.35]
ϕ_π	1.047 [1.037 1.05]	0.58 [0.56 0.61]	1.36 [1.36 1.38]	1.37 [1.26 1.46]
σ_i	0.0058 [0.0051 0.0067]	0.0057 [0.005 0.0067]	0.0063 [0.0056 0.0073]	0.0072 [0.0062 0.0085]
σ_x	0.01 [0.009 0.013]	0.011 [0.01 0.014]	0.0089 [0.0076 0.011]	0.008 [0.007 0.01]
σ_π	0.0065 [0.0057 0.0075]	0.007 [0.006 0.009]	0.0068 [0.006 0.008]	0.0074 [0.0063 0.0088]
σ_ς		0.18 [0.15 0.22]		0.3 [0.25 0.38]

90% credibility interval in brackets

The estimates of the parameters in the first subsample are very similar in the two models, except for the policy parameters of the Taylor rule. The inference on those parameters changes because M_U gives a completely different interpretation about the instability of that period. Independently from the Fed policy, the instability comes from the model itself: it is not the monetary policy to destabilize the system (ϕ_π is less than one), but the latter is unstable because of the presence of an eigenvalue outside the unit circle. Even if there is the possibility of selecting a stable solution, the forward looking one chosen by M_S, the model with unstable paths discards this possibility, and it happens exactly when inflation starts growing away from the steady state.

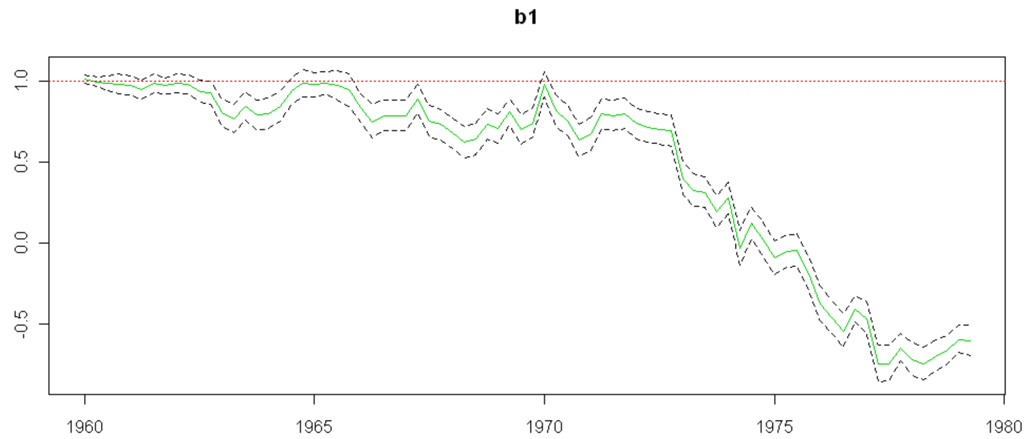


Figure 2: $b_{1,t}$ during the Great Inflation

This intuition is clear looking at Figure 2 that plots the filtered estimate of $b_{1,t}$ during the Great Inflation, under model M_U. The latent process fluctuates close to the forward looking value (that represents the unique stable solution) until the first Seventies. Then, it walks away from one, selecting unambiguously unstable paths, exactly when inflation starts fluctuating between 3% and 15%.

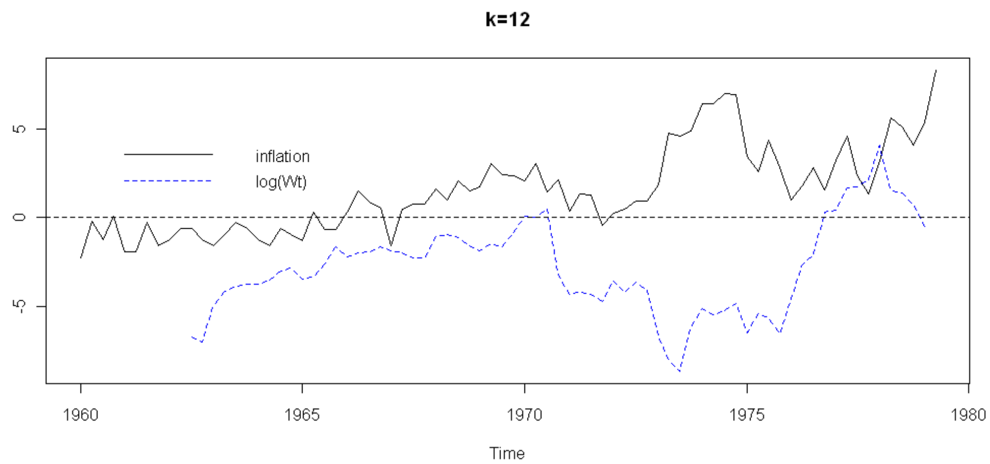


Figure 3: Comparing M_S - M_U. Great Inflation

We compare the relative fit of the two models computing the Sequential Bayes factor as in West (1986). The results for the first subsample are reported in Figure

3, where we plot together the logarithm of the Sequential Bayes factor and inflation. Twelve quarters are compared at each time. Model M_S is at the numerator and model M_U is at the denominator, and we take the logarithm so that when the Bayes factor is zero it means that the two models have the same performance in terms of predictive likelihood, when it is positive it means that M_S is preferred, and vice versa, we prefer M_U when it is negative. The advantage of the Sequential Bayes factor, with respect to the conventional measures in Bayesian Econometrics, is that we can compare two models through time, and verify the sub-periods in which a model is better than another. In our specific case, as expected, the unstable model is much preferred when inflation reaches high values.

The inference about the two models in the second subsample is very similar: in both cases they select the forward looking solution. The Taylor principle is respected, so that in M_S the process for $b_{1,t}$ degenerates to the value of one. In the case of model M_S, on the other hand, $b_{1,t}$ remains close to one, and in the second half of the period considered it is not statistically different from one.

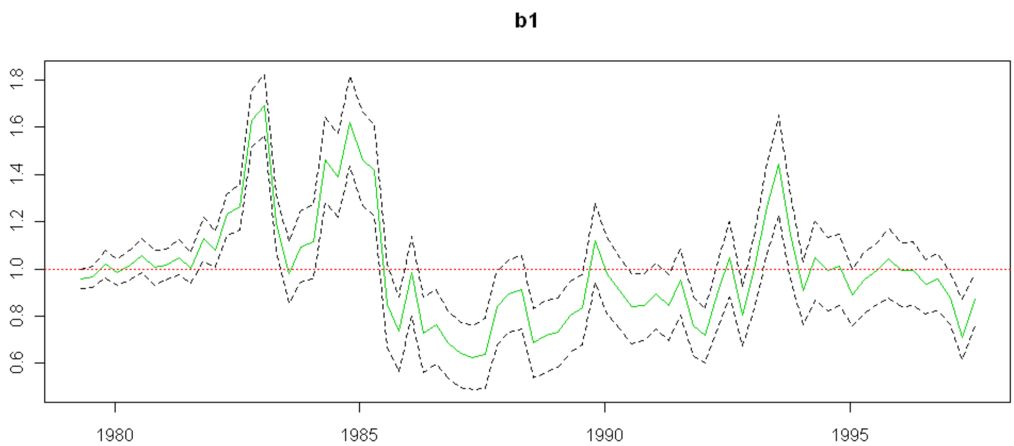


Figure 4: $b_{1,t}$ during the Great Moderation

Now, comparing the two models as in the previous case, we find that the two models have the same explanatory power, except in the first part, when the stabilization period is better interpreted by M_S.

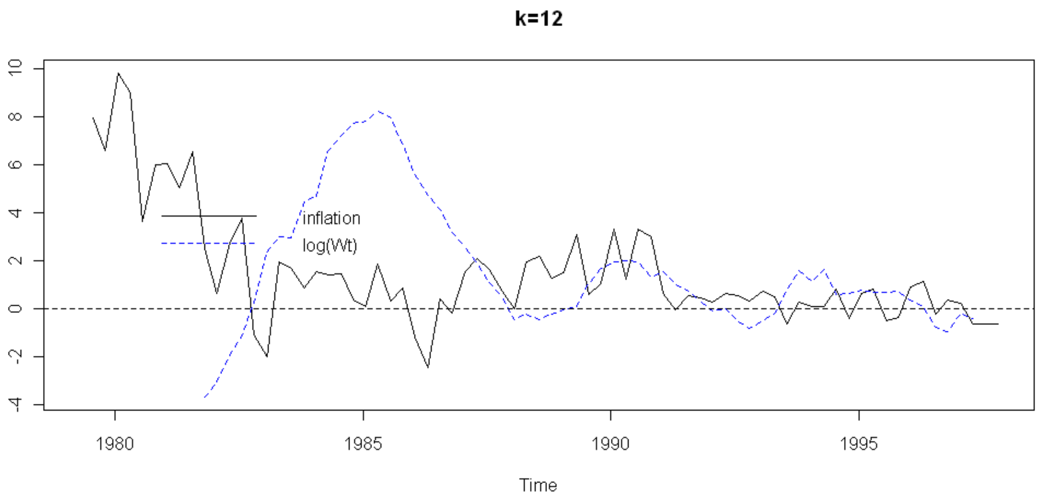


Figure 5: Comparing M_S - M_U. Great Moderation

6 Conclusion

The empirical evidence we show in the present paper suggests that the Great Inflation in the U.S. can be explained by temporary unstable paths, as in the rational sunspots interpretation. We show that the usual practice of excluding a priori unstable solutions is not supported by the data. Obviously the model we use plays a role, because it has an intrinsic source of instability (one eigenvalue is always greater than one). It is precisely using that model that we try to confute the New Keynesian interpretation about the Great Inflation period: the main cause was not the presence of an infinite number of stable solutions, generated by a passive monetary policy. The results go on the opposite direction with respect to the conclusions of popular papers like the one of Lubik and Schorfheide (2004): interpreting instability with a stable system is not only counter-intuitive, but also not very reliable under an empirical point of view.

Appendix A

The solution for the simple model

Consider equation (3) in the paper:

$$\pi_t = \frac{1}{\phi}E_t\pi_{t+1} + e_t \tag{18}$$

$$e_t = -\frac{1}{\phi}\varepsilon_t \quad \varepsilon_t \sim i.i.d.N(0, \sigma_\varepsilon^2) \tag{19}$$

In this appendix we treat the general case with a time varying b_t (then, the case with b_t constant is simply obtained). Suppose that

$$b_t = b_t(\zeta_t) \tag{20}$$

where ζ_t is a random variable, called sunspot shock, orthogonal to the fundamental shocks e_s $(s = 1, 2, ...)$ and such that $E_t\zeta_t = 0 \ \forall t$.

Following Muth (1961) and Blanchard (1979) we guess the solution for model (18):

$$\pi_t = \sum_{j=1}^{\infty} u_{j,t}e_{t-j} + b_t e_t + \sum_{j=1}^{\infty} c_{j,t}E_t e_{t+j} \tag{21}$$

where $u_{j,t}$, b_t and $c_{j,t}$ are coefficients to be determined. Hence verify using undetermined coefficients:

$$\begin{aligned} \pi_t &= \frac{1}{\phi}E_t\pi_{t+1} + e_t \\ \sum_{j=1}^{\infty} u_{j,t}e_{t-j} + b_t e_t + \sum_{j=1}^{\infty} c_{j,t}E_t e_{t+j} &= \frac{1}{\phi}E_t\left(\sum_{j=1}^{\infty} u_{j,t+1}e_{t+1-j} + b_{t+1}e_{t+1} + \sum_{j=1}^{\infty} c_{j,t+1}E_t e_{t+1+j}\right) + e_t \end{aligned}$$

that is:

$$\begin{aligned} & u_{1,t}e_{t-1} + u_{2,t}e_{t-2} + u_{3,t}e_{t-3} + ... + b_t e_t + c_{1,t}E_t e_{t+1} + c_{2,t}E_t e_{t+2} + ... \\ = \ & \frac{1}{\phi}E_t\left(u_{1,t+1}e_t + u_{2,t+1}e_{t-1} + u_{3,t+1}e_{t-2} + ... + b_{t+1}e_{t+1} + c_{1,t+1}e_{t+2} + c_{2,t+1}e_{t+3} + ...\right) + e_t \end{aligned}$$

equal coefficients to find an expression for the $u's$:

$$\begin{aligned}
e_t &: & b_t = \frac{1}{\phi}E_t u_{1,t+1} + 1 \Rightarrow E_t u_{1,t+1} = \phi(b_t - 1) \\
e_{t-1} &: & u_{1,t} = \frac{1}{\phi}E_t u_{2,t+1} \Rightarrow E_t u_{2,t+1} = \phi u_{1,t} \\
&\vdots & \\
e_{t-j} &: & u_{j,t} = \frac{1}{\phi}E_t u_{j+1,t+1} \Rightarrow E_t u_{j+1,t+1} = \phi u_{j,t}
\end{aligned}$$

and for the $c's$:

$$\begin{aligned}
e_{t+1} &: & c_{1,t} = \frac{1}{\phi}E_t b_{t+1} \\
e_{t+2} &: & c_{2,t} = \frac{1}{\phi}E_t c_{1,t+1} \\
&\vdots & \\
e_{t+j+1} &: & c_{j+1,t} = \frac{1}{\phi}E_t c_{j,t+1}
\end{aligned}$$

These equations need an assumption on the stochastic process governing b_t to be satisfied. Otherwise, in general the system can not be solved.

Random walk process for b_t

Assume that b_t is following a random walk process as $b_t = b_{t-1} + \zeta_t$, with $\zeta_t \sim i.i.d.N(0, \sigma_\zeta^2)$. Then $E_{t+1}b_{t+1} = b_t$. Hence:

$$r_t: \qquad b_t = \frac{1}{\phi}E_t u_{1,t+1} + 1 \Rightarrow E_t u_{1,t+1} = \phi(b_t - 1)$$

However, given $E_t u_{1,t+1} = \phi(b_t - 1)$ what can we say about $u_{1,t+1}$? Assuming that $u_{1,t+1} = F(b_{t+1})$, the problem then is to find the function F such that $E_t u_{1,t+1} = \phi(b_t - 1)$, given the stochastic process for b_t. Assuming that F is linear then we are looking for a linear function such that $E_t(a_1 b_{t+1} + a_0) = \phi(b_t - 1)$, that is: $a_1 E_t b_{t+1} + a_0 = \phi b_t - \phi \Rightarrow a_1 b_t + a_0 = \phi b_t - \phi \Rightarrow$

$$\begin{aligned}
a_1 &= \phi \\
a_0 &= -\phi
\end{aligned}$$

so

$$u_{1,t+1} = \phi b_{t+1} - \phi \tag{22}$$

Equal coefficients that multiply e_{t-1}:

$$E_t u_{2,t+1} = \phi u_{1,t}$$

Then, $u_{1,t} = \phi b_t - \phi$ needs to be equal to $\frac{1}{\phi} E_t u_{2,t+1}$. Following the same reasoning, assuming $u_{2,t+1}$ is a linear function of b_{t+1}, we need to solve for

$$E_t \left(a_1 b_{t+1} + a_0 \right) = \phi u_{1,t} = \phi^2 b_t - \phi^2.$$

Then, it must be

$$\begin{aligned} a_1 &= \phi^2 \\ a_0 &= -\phi^2 \end{aligned}$$

so that:

$$u_{2,t+1} = \phi^2 b_{t+1} - \phi^2 \tag{23}$$

generally

$$u_{j,t} = \phi^j b_t - \phi^j \tag{24}$$

Having solved for the $u's$ let's solve for the $c's$. This is easy since $E_{t+1} b_{t+1} = b_t$:

$$e_{t+1}: \qquad c_{1,t} = \frac{1}{\phi} E_t b_{t+1} = \frac{1}{\phi} b_t \tag{25}$$

Following the method implemented above we obtain, in general:

$$c_{j,t} = \frac{1}{\phi^j} b_t \tag{26}$$

Equations (24) and (26) are the coefficients of equation (21), written as function of b_t. Equation (21) is a solution for model (18) only if it satisfies these restrictions.

In our case, because the exogenous shocks are *i.i.d.* with zero mean, the sum $\sum_{j=1}^{\infty} c_{j,t} E_t e_{t+j}$ in equation (21) is zero. Then, substituting equation (24) we have:

$$\pi_t = (b_t - 1) \sum_{j=1}^{\infty} \phi^j e_{t-j} + b_t e_t \tag{27}$$

so that we have the pure forward looking solution when $b_t = 1$ (equation, 5 in the paper):

$$\pi_t^F = e_t = -\frac{1}{\phi} \varepsilon_t$$

and the pure backward looking solution when $b_t = 0$:

$$\begin{aligned} \pi_t^B &= -\sum_{j=1}^{\infty} \phi^j e_{t-j} = -\phi e_{t-1} - \phi \sum_{j=1}^{\infty} \phi^j e_{t-j-1} \\ \pi_t^B &= \phi \left(\pi_{t-1}^B - \pi_{t-1}^F \right) = \phi \pi_{t-1}^B + \varepsilon_t \end{aligned}$$

that corresponds to equation (6). Note that equation (27) can be rewritten as:

$$\pi_t = (1 - b_t)\, \pi_t^B + b_t \pi_t^F \tag{28}$$

that is, each particular solution depends on b_t, and it can be written as a linear combination of the backward and the forward one.

The recursive formulation

We first report the important equations:

$$\begin{aligned} \pi_t &= (1 - b_t)\pi_t^B + b_t \pi_t^F & (29) \\ \pi_t^B &= \phi \pi_{t-1}^B - \phi \pi_{t-1}^F & (30) \\ \pi_t^F &= e_t & (31) \end{aligned}$$

substituting π_t^B and π_t^F in the first equation we obtain

$$\pi_t = \phi(1 - b_t)\pi_{t-1}^B - \phi(1 - b_t) e_{t-1} + b_t e_t \tag{32}$$

Multiply for $(1 - b_t)$ equation (30) and substitute in the last equation to find π_t^B:

$$\begin{aligned} (1 - b_t)\pi_t^B &= \phi(1 - b_t)\pi_{t-1}^B - \phi(1 - b_t) e_{t-1} \\ (1 - b_t)\pi_t^B &= \pi_t - b_t e_t \\ \pi_t^B &= \frac{\pi_t - b_t e_t}{(1 - b_t)} \end{aligned}$$

Use this expression, lagged, in (32) to derive the complete set of solutions for model (18), when $b_{t-1} \neq 1$:

$$\pi_t = \alpha_t \pi_{t-1} - \alpha_t e_{t-1} + b_t e_t \qquad (33)$$

with $\alpha_t = \phi \dfrac{(1-b_t)}{(1-b_{t-1})}$. The particular case of $b_t = b$ constant is obtainable offsetting the sunspot shocks, that is imposing $\sigma_\zeta^2 = 0$. The coefficient α_t becomes:

$$\alpha_t = \phi \frac{1-b_{t-1}}{1-b_{t-1}} = \phi$$

and π_t is described by equation (4):

$$\pi_t = \phi \pi_{t-1} + \varepsilon_{t-1} - \frac{b}{\phi} \varepsilon_t$$

Expectations as a weighted average of past observations

Under the rational expectations hypothesis, the expected value in model (18), can be written as a weighted average of the past observations (see Muth, 1961):

$$\begin{aligned} E_t \pi_{t+1} &= \sum_{i=1}^{\infty} V_{i,t} \pi_{t+1-i} = \\ &= V_{1,t} \pi_t + V_{2,t} \pi_{t-1} + V_{3,t} \pi_{t-2} + ... \end{aligned} \qquad (34)$$

where we need to determine the coefficients $V_{i,t}$. Using equation (28) we have:

$$\begin{aligned} E_t \pi_{t+1} &= V_{1,t} \left[(b_t - 1) \sum_{j=1}^{\infty} \phi^j e_{t-j} + b_t e_t \right] + \\ &\quad + V_{2,t} \left[(b_{t-1} - 1) \sum_{j=1}^{\infty} \phi^j e_{t-j-1} + b_{t-1} e_{t-1} \right] + \\ &\quad + V_{3,t} \left[(b_{t-2} - 1) \sum_{j=1}^{\infty} \phi^j e_{t-j-2} + b_{t-2} e_{t-2} \right] + ... \end{aligned}$$

Rearrange:

$$\begin{aligned} E_t \pi_{t+1} &= V_{1,t} b_t e_t + \\ &\quad + \left[V_{1,t} (b_t - 1) \phi + V_{2,t} b_{t-1} \right] e_{t-1} + \\ &\quad + \left[V_{1,t} (b_t - 1) \phi^2 + V_{2,t} (b_{t-1} - 1) \phi + V_{3,t} b_{t-2} \right] e_{t-2} + \\ &\quad + ... \end{aligned}$$

Then, bring equation (28) one step ahead,

$$\begin{aligned} E_t\pi_{t+1} &= (b_t-1)\sum_{j=1}^{\infty}\phi^j e_{t-j+1} = \\ &= (b_t-1)\left[\phi e_t+\phi^2 e_{t-1}+\phi^3 e_{t-2}+...\right] \end{aligned}$$

and compare coefficients:

$$e_t: \qquad (b_t-1)\,\phi = V_{1,t}b_t$$
$$V_{1,t} = \frac{(b_t-1)}{b_t}\phi$$

$$e_{t-1}: \qquad (b_t-1)\,\phi^2 = \left[V_1\,(b_t-1)\,\phi + V_2 b_{t-1}\right]$$
$$V_2 = \frac{(b_t-1)}{b_t b_{t-1}}\phi^2$$

$$e_{t-2}: \qquad (b_t-1)\,\phi^3 = \left[V_{1,t}\,(b_t-1)\,\phi^2 + V_{2,t}\,(b_{t-1}-1)\,\phi + V_{3,t}b_{t-2}\right]$$
$$V_{3,t} = \frac{(b_t-1)}{b_t b_{t-1} b_{t-2}}\phi^3$$

in general:

$$V_{i,t} = \frac{(b_t-1)}{\prod_{0}^{i-j=1} b_{t-i}}\phi^i$$

and when $b_t = b$ constant:

$$V_i = \frac{(b-1)}{b^i}\phi^i$$

The multivariate case

b constant

We show how to compute the complete set of solutions of a system with rational expectations.

Consider a system with rational expectations written in the form of Blanchard and Kahn (1980):

$$\begin{bmatrix} X_{t+1} \\ E_t P_{t+1} \end{bmatrix} = A \begin{bmatrix} X_t \\ P_t \end{bmatrix} + \boldsymbol{\gamma} Z_t \tag{35}$$

X_t is a $(n \times 1)$ vector of predetermined variables and P_t is a $(m \times 1)$ vector of non predetermined variables. $Z_t \sim i.i.d.\ N(\mathbf{0}, \Sigma)$ is a $(\kappa \times 1)$ vector of exogenous random variables.

Use the Jordan form to rewrite A

$$A = C^{-1} J C.$$

In the main diagonal of J there are the eigenvalues of A, ordered by increasing absolute value. We decompose the matrices C^{-1}, J, C and $\boldsymbol{\gamma}$ as follows:

$$C^{-1} = \begin{bmatrix} \underset{(n \times n)}{B_{11}} & \underset{(n \times m)}{B_{12}} \\ \underset{(m \times n)}{B_{21}} & \underset{(m \times m)}{B_{22}} \end{bmatrix}, \quad C = \begin{bmatrix} \underset{(n \times n)}{C_{11}} & \underset{(n \times m)}{C_{12}} \\ \underset{(m \times n)}{C_{21}} & \underset{(m \times m)}{C_{22}} \end{bmatrix},$$

$$J = \begin{bmatrix} \underset{(n \times n)}{J_1} & \underset{(n \times m)}{\mathbf{0}} \\ \underset{(m \times n)}{\mathbf{0}} & \underset{(m \times m)}{J_2} \end{bmatrix}, \quad \boldsymbol{\gamma} = \begin{bmatrix} \underset{(n \times \kappa)}{\gamma_1} \\ \underset{(n \times \kappa)}{\gamma_2} \end{bmatrix}.$$

Define

$$\begin{bmatrix} Y_t \\ Q_t \end{bmatrix} = C \begin{bmatrix} X_t \\ P_t \end{bmatrix},$$

and rewrite equation (35) in terms of $\begin{bmatrix} Y_t \\ Q_t \end{bmatrix}$:

$$\begin{bmatrix} E_t Y_{t+1} \\ E_t Q_{t+1} \end{bmatrix} = \begin{bmatrix} J_1 & \mathbf{0} \\ \mathbf{0} & J_2 \end{bmatrix} \begin{bmatrix} Y_t \\ Q_t \end{bmatrix} + \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix} \begin{bmatrix} \gamma_1 \\ \gamma_2 \end{bmatrix} Z_t\ . \tag{36}$$

Now consider the second block of equation (36),

$$Q_t = J_2^{-1} E_t Q_{t+1} - \Omega_t \tag{37}$$

where $\Omega_t = J_2^{-1}(C_{21}\gamma_1 + C_{22}\gamma_2) Z_t$. The system (37) has m disjoined equations, and each of them admits an infinite number of solutions because of the presence of an expected value. Defining $q_{i,t}$ as the i^{th} element of Q_t, and $\omega_{i,t}$ the corresponding disturbance, we write all the solutions of the generic row of equation (37) as

$$q_{i,t} = \sum_{j=1}^{\infty} u_{i,j}\omega_{i,t-j} + b_i\omega_{i,t} + \sum_{j=1}^{\infty} c_{i,j} E_t \omega_{i,t+j} \ . \tag{38}$$

Using matrices instead of scalars the solutions can be rewritten as

$$Q_t = \sum_{j=1}^{\infty} \mathbf{u}_j \Omega_{t-j} + \mathbf{b}\Omega_t + \sum_{j=1}^{\infty} \mathbf{c}_j E_t \Omega_{t+j} \tag{39}$$

where $\mathbf{u}_j$, $\mathbf{b}$ and $\mathbf{c}_j$ are diagonal matrices of coefficients to be determined. Bring equation (39) one step ahead

$$E_t Q_{t+1} = \sum_{j=1}^{\infty} \mathbf{u}_j \Omega_{t+1-j} + E_t \mathbf{b}\Omega_{t+1} + \sum_{j=1}^{\infty} \mathbf{c}_j E_t \Omega_{t+1+j}$$

and substitute in equation (37)

$$Q_t = J_2^{-1} \sum_{j=2}^{\infty} \mathbf{u}_j \Omega_{t+1-j} + J_2^{-1}\mathbf{u}_1 \Omega_t - \Omega_t + J_2^{-1} E_t \mathbf{b}\Omega_{t+1} + J_2^{-1} \sum_{j=1}^{\infty} \mathbf{c}_j E_t \Omega_{t+1+j} \ . \tag{40}$$

We find the coefficients comparing the matrices of equation (39) to the ones of equation (40):

$$\begin{aligned}
&\mathbf{b} = J_2^{-1}\mathbf{u}_1 - I \Longrightarrow \boxed{\mathbf{u}_1 = J_2\mathbf{b} + J_2} \\
&\mathbf{u}_1 = J_2^{-1}\mathbf{u}_2 \Longrightarrow \boxed{\mathbf{u}_{j+1} = J_2\mathbf{u}_j} \quad j = 1...\infty \\
&\boxed{\mathbf{c}_1 = J_2^{-1}\mathbf{b}} \\
&\mathbf{c}_2 = J_2^{-1}\mathbf{c}_1 \Longrightarrow \boxed{\mathbf{c}_{j+1} = J_2^{-1}\mathbf{c}_j} \quad j = 1...\infty
\end{aligned}$$

The matrices $\mathbf{u}_j$ and $\mathbf{c}_j$ are functions of $\mathbf{b}$ and J_2, and since J_2 is given, the complete set of solutions is parametrized by $\mathbf{b}$. There are two particular cases: the pure backward looking solution, corresponding to $\mathbf{b} = \mathbf{0}$, that implies $\mathbf{c}_j = \mathbf{0}$ and $\mathbf{u}_j = J_2^j$, $j = 1...\infty$; the pure forward looking solution corresponding to $\mathbf{b} = -I$, that implies $\mathbf{u}_j = \mathbf{0}$ and $\mathbf{c}_j = -J_2^{-j}$, $j = 1...\infty$. The backward looking solution can be written as follows:

$$Q_t^B = \sum_{j=1}^{\infty} \mathbf{u}_j \Omega_{t-j} \tag{41}$$

$$Q_t^B = \sum_{j=1}^{\infty} J_2^j \Omega_{t-j} = J_2\Omega_{t-1} + J_2^2\Omega_{t-2} + J_2^3\Omega_{t-3} + ...$$

$$Q_t^B = J_2\Omega_{t-1} + J_2\left[J_2\Omega_{t-2} + J_2^2\Omega_{t-3} + J_2^3\Omega_{t-4} + ...\right]$$

$$Q_t^B = J_2 Q_{t-1}^B + J_2\Omega_{t-1} \tag{A7}$$

The forward looking solution is

$$Q_t^F = b\Omega_t + \sum_{j=1}^{\infty} \mathbf{c}_j E_t\Omega_{t+j} = -I\Omega_t - J_2^{-1}E_t\Omega_{t+1} - J_2^{-2}E_t\Omega_{t+2} - ...$$

and since $E_t\Omega_{t+j} = \mathbf{0} \ \ \forall j \geq 1$, we obtain

$$Q_t^F = -\Omega_t \ . \tag{42}$$

Following Blanchard (1979) we write any other solution as a linear combination of the backward and the forward looking solutions. In compact form

$$Q_t = \boldsymbol{\lambda} Q_t^B + (I - \boldsymbol{\lambda})Q_t^F \tag{43}$$

where $\boldsymbol{\lambda} = I + \mathbf{b}$ is a diagonal matrix. The elements in the main diagonal of $\mathbf{b}$ are such that $\mathbf{b} = \mathbf{0} \Rightarrow Q_t = Q_t^B$, and $\mathbf{b} = -I \Rightarrow Q_t = Q_t^F$.

Substitute the equations (A7) and (42) in equation (43)

$$\begin{aligned} Q_t &= \boldsymbol{\lambda}\left(J_2 Q_{t-1}^B + J_2\Omega_{t-1}\right) - (I - \boldsymbol{\lambda})\Omega_t \\ &= \boldsymbol{\lambda} J_2 Q_{t-1}^B - \boldsymbol{\lambda} J_2 Q_{t-1}^F + J_2 Q_{t-1}^F - J_2 Q_{t-1}^F - (I - \boldsymbol{\lambda})\Omega_t \ . \end{aligned}$$

In the last passage we have added and subtracted $J_2 Q_{t-1}^F$. Since both J_2 and $\boldsymbol{\lambda}$ are diagonal matrices the commutative property holds and we can write

$$Q_t = J_2\left(\boldsymbol{\lambda} Q_{t-1}^B + (I - \boldsymbol{\lambda})Q_{t-1}^F\right) + J_2\Omega_{t-1} - (I - \boldsymbol{\lambda})\Omega_t$$

$$Q_t = J_2 Q_{t-1} + J_2\Omega_{t-1} + b\Omega_t \tag{44}$$

Equation (44) represents the infinite number of solutions for Q_t parametrized by $\mathbf{b}$. The complete set of solutions for model (35) is found using the definition of Q_t and the first n rows of the model written with the Jordan matrices:

$$X_t = \left(B_{11}J_1C_{11} + B_{12}J_2C_{21}\right)X_{t-1} + \left(B_{11}J_1C_{12} + B_{12}J_2C_{22}\right)P_{t-1} + \gamma_1 Z_{t-1} \tag{45}$$

$$C_{21}X_t + C_{22}P_t = J_2(C_{21}X_{t-1} + C_{22}P_{t-1}) \quad +(C_{21}\gamma_1 + C_{22}\gamma_2)Z_{t-1} + $$
$$+\mathbf{b}J_2^{-1}(C_{21}\gamma_1 + C_{22}\gamma_2)Z_t \qquad (46)$$

In the paper we focus on the case in which the matrix A has at least n eigenvalues inside the unit circle. This means that the model admits at least one stable solution. If this condition is not satisfied the equations (45) and (46) continue to represent the complete set of solutions that are all unstable.

Adding sunspots

Add the hypothesis that each element in the main diagonal of **b** is described by the following stochastic process:

$$b_{i,t} = b_{i,t-1} + \zeta_{i,t}$$

with $\zeta_{i,t} \sim i.i.d.N(0, \sigma^2_{\zeta_i})$, $i = 1, 2, ...m$. With this hypothesis equation (38) becomes

$$q_{i,t} = \sum_{j=1}^{\infty} u_{i,t}^{(j)} \omega_{i,t-j} + b_{i,t}\omega_{i,t}$$

and its solution is:

$$\begin{aligned} q_{i,t} &= \alpha_{i,t} q_{i,t} + \alpha_{i,t}\omega_{i,t-1} + b_{i,t}\omega_{i,t} \\ \alpha_{i,t} &= J_{2,i}\frac{(1 - b_{i,t})}{(1 - b_{i,t-1})} \end{aligned}$$

where $J_{2,i}$ is the i^{th} eigenvalue in the main diagonal of J_2. Putting in matrix form the system with these m disjoined equations, we obtain the following system, analogous to equation (44):

$$Q_t = J_2 \left(I + \mathbf{b}_t\right) \left(I + \mathbf{b}_{t-1}\right)^{-1} Q_{t-1} + J_2 \left(I + \mathbf{b}_t\right) \left(I + \mathbf{b}_{t-1}\right)^{-1} \Omega_{t-1} + \mathbf{b}_t \Omega_t$$

withFinally, the solution is represented by the following system:

$$X_t = \left(B_{11}J_1C_{11} + B_{12}J_2C_{21}\right) X_{t-1} + \left(B_{11}J_1C_{12} + B_{12}J_2C_{22}\right) P_{t-1} + \gamma_1 Z_{t-1}$$

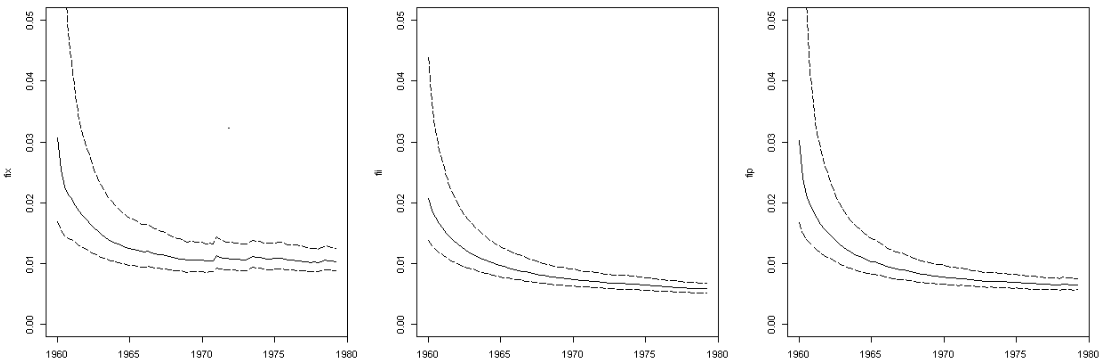
$$\begin{aligned} C_{21}X_t + C_{22}P_t \;=\; & J_2 \left(I + \mathbf{b}_t\right) \left(I + \mathbf{b}_{t-1}\right)^{-1} \left(C_{21}X_{t-1} + C_{22}P_{t-1}\right) + \\ & + \left(I + \mathbf{b}_t\right) \left(I + \mathbf{b}_{t-1}\right)^{-1} \left(C_{21}\gamma_1 + C_{22}\gamma_2\right)Z_{t-1} + \mathbf{b}_t J_2^{-1}(C_{21}\gamma_1 + C_{22}\gamma_2)Z_t \end{aligned}$$

Appendix B

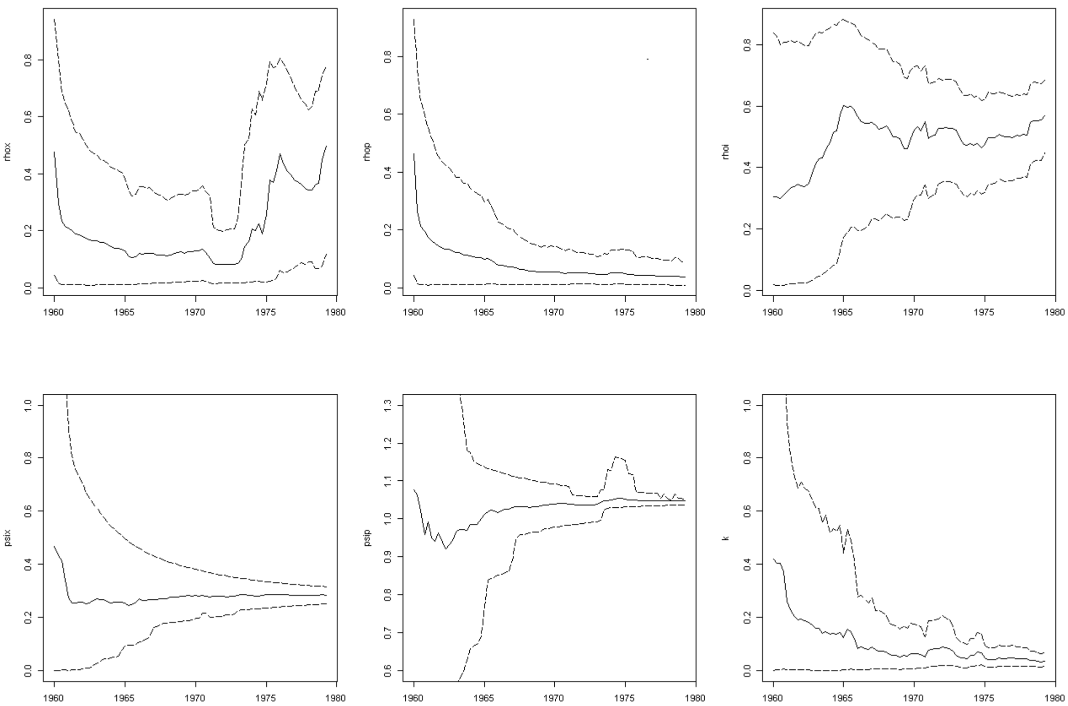
I report the sequential inference on the parameters for the four cases studied.

M_S **Great Inflation**

The variances

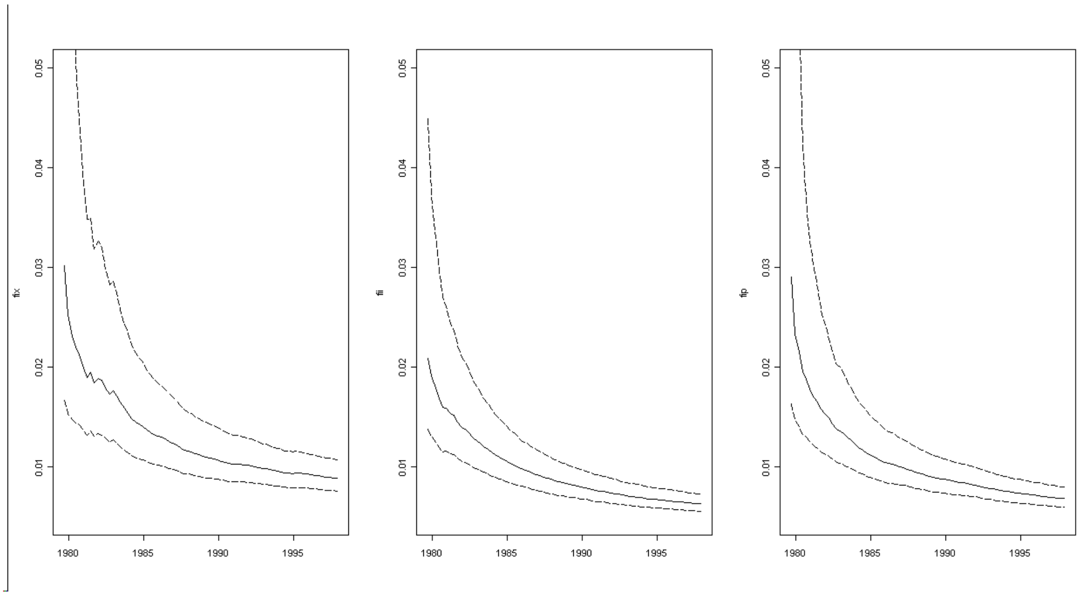


and the other parameters:

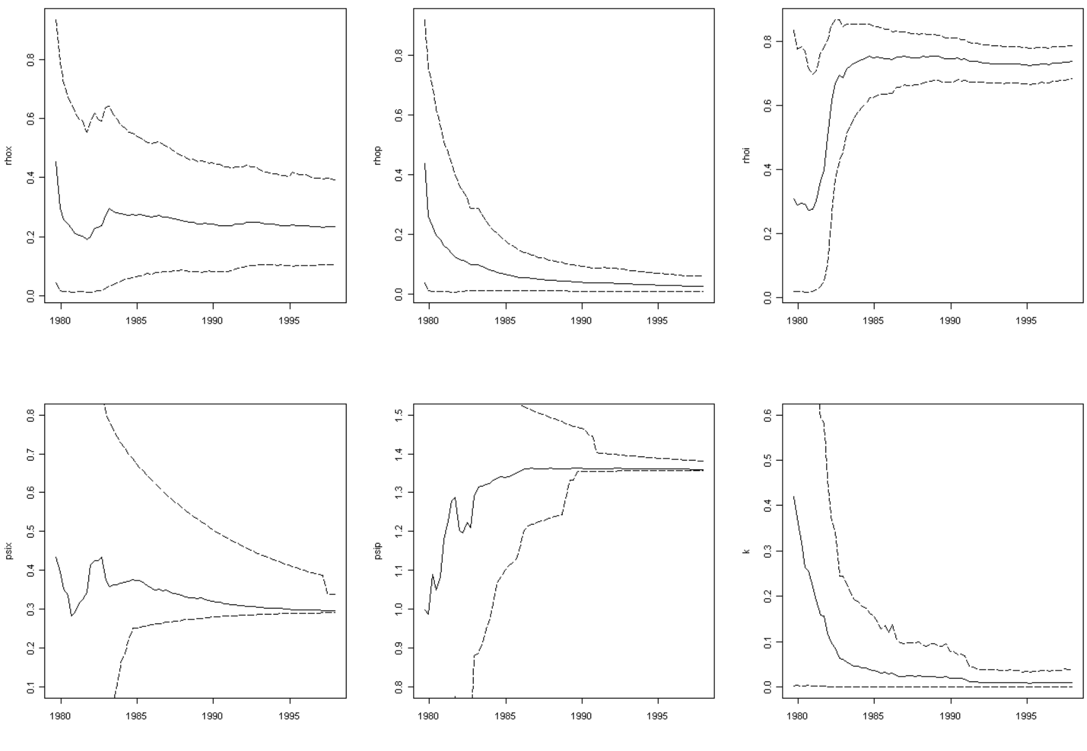


M_S **Great Moderation**

The variances:

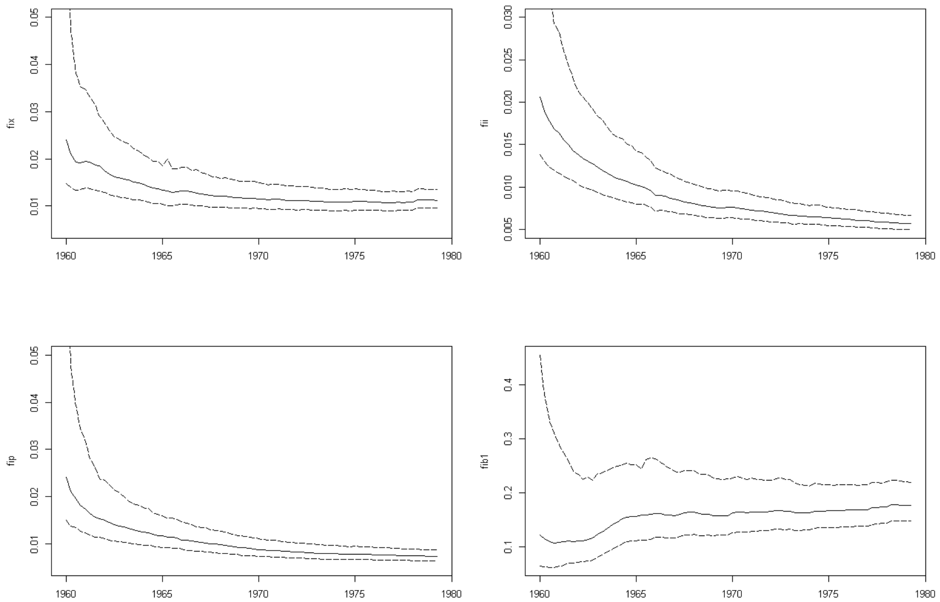


and the other parameters:

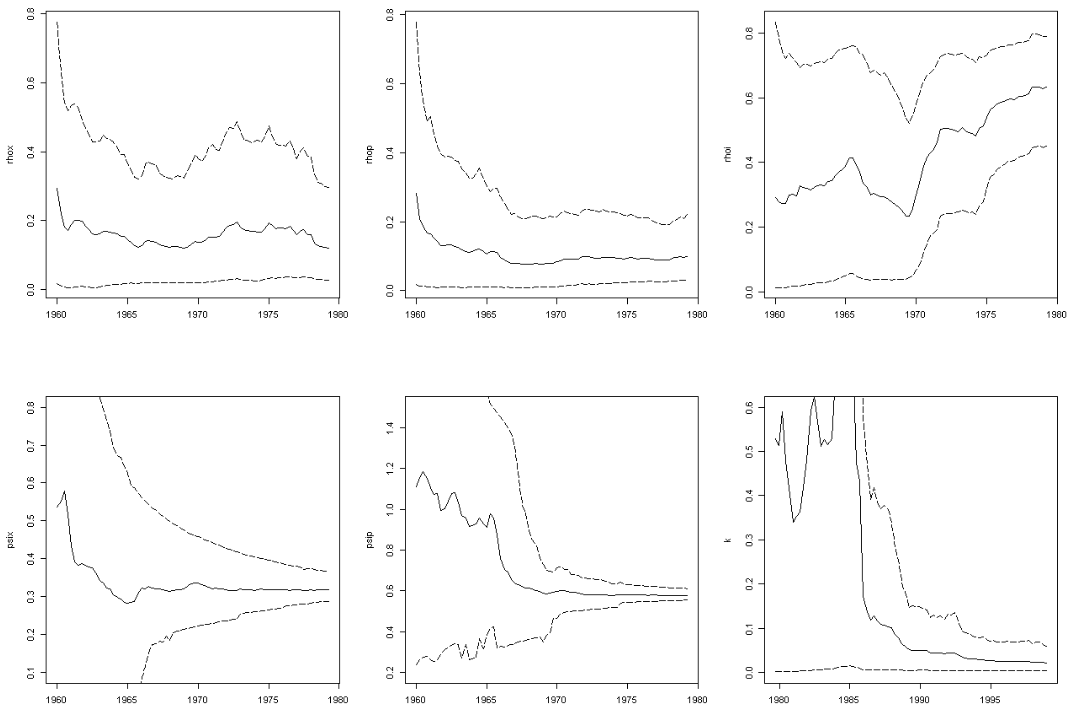


M_U **Great Inflation**

The variances:

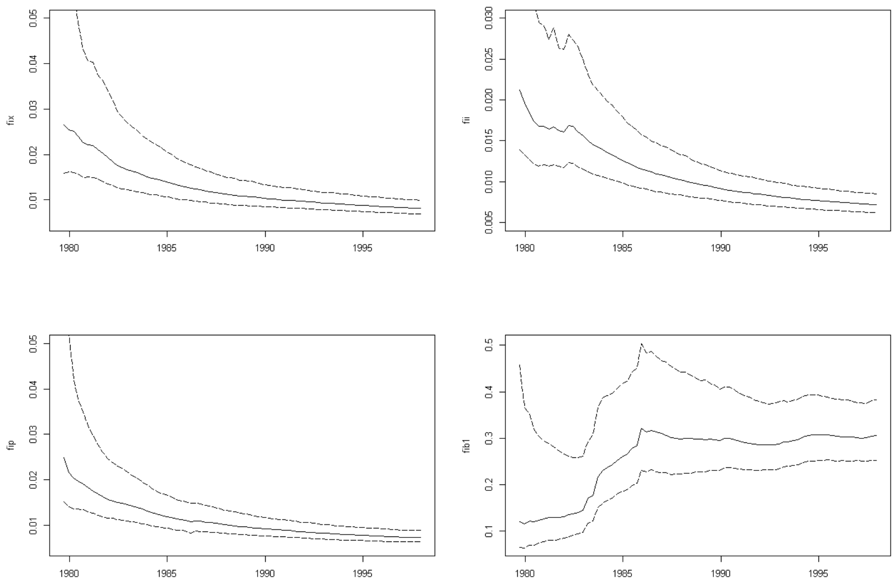


and the other parameters:

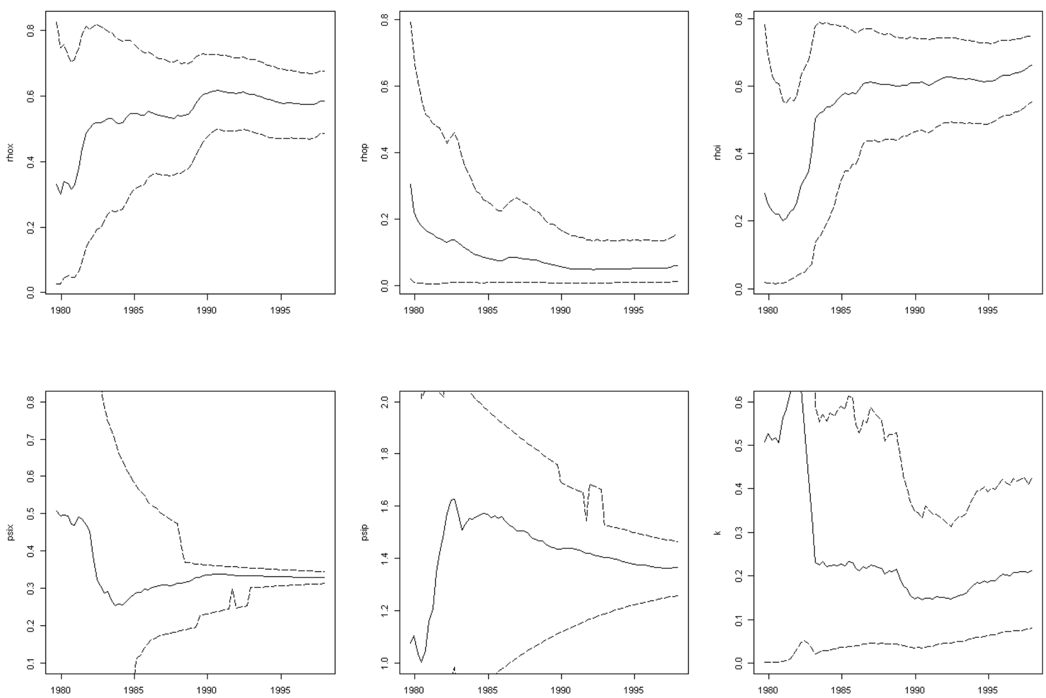


M_U **Great Moderation**

The variances:



and the other parameters:



References

ASCARI, G. AND P. BONOMOLO (2012): "Rational Sunspots and Drifting Parameters," Mimeo, Università di Pavia.

BLANCHARD, O. J. (1979): "Backward and Forward Solutions for Economies with Rational Expectations," *American Economic Review*, American Economic Association, vol. 69(2), pages 114-18, May.

BLANCHARD, O. J. AND C. M. KAHN (1980): "The Solution of Linear Difference Models under Rational Expectations," *Econometrica*, Econometric Society, vol. 48(5), pages 1305-11, July.

BULLARD, J. AND K. MITRA (2002): "Learning about monetary policy rules," *Journal of Monetary Economics*, Elsevier, vol. 49(6), pages 1105-1129, September.

CARVALHO, C. M., M. S. JOHANNES, H. F. LOPES AND N. G. POLSON (2010), "Particle learning and smoothing," *Statistical Science*, Institute of Mathematical Statistics, vol. 25(1), pages 88-106, February.

CLARIDA, R., J. GALI AND M. GERTLER (2000): "Monetary Policy Rules And Macroeconomic Stability: Evidence And Some Theory," *The Quarterly Journal of Economics*, MIT Press, vol. 115(1), pages 147-180, February.

COCHRANE, J. H. (2011): "Determinacy and identification with Taylor rules," *Journal of Political Economy*, The University of Chicago Press, vol. 119(3), pages 565-615, June.

LIU, J. AND M. WEST (2001): "Combined parameters and state estimation in simulation-based filtering," in A. Doucet, N. de Freitas and N. Gordon (ed.), *Sequential Monte Carlo Methods in Practice,* Springer.

LUBIK, T. A. AND F. SCHORFHEIDE (2004): "Testing for Indeterminacy: An Application to U.S. Monetary Policy," *American Economic Review*, American Economic Association, vol. 94(1), pages 190-217, March.

MUTH J. F. (1960): "Optimal Properties of Exponentially Weighted Forecasts", *Journal of the American Statistical Association*, American Statistical Association, vol. 55(290), pages 299-306, June.

MUTH J. F. (1961): "Rational Expectations and the Theory of Price Movements," *Econometrica*, Econometric Society, vol. 29(3), pages 315-335, July.

WEST M. (1986): "Bayesian Model Monitoring," *Journal of the Royal Statistical Society. Series B (Methodological)*, Blackwell Publishing, vol. 48(1), pages 70-78