DNB Working Paper

No. 597 / May 2018

Walk on the wild side: Multiplicative sunspots and temporarily unstable paths

Guido Ascari, Paolo Bonomolo and Hedibert Lopes

DeNederlandscheBank

EUROSYSTEEM

Walk on the wild side: Multiplicative sunspots and temporarily unstable paths

Guido Ascari, Paolo Bonomolo and Hedibert Lopes *

* Views expressed are those of the authors and do not necessarily reflect official positions of De Nederlandsche Bank.

Working Paper No. 597

May 2018

De Nederlandsche Bank NV
P.O. Box 98
1000 AB AMSTERDAM
The Netherlands

Walk on the wild side: Multiplicative sunspots and temporarily unstable paths*

Guido Ascari[a], Paolo Bonomolo[b] and Hedibert Lopes[c]

[a] *University of Oxford, University of Pavia and Bank of Finland, United Kingdom, Email: guido.ascari@economics.ox.ac.uk*
[b] *De Nederlandsche Bank. The Netherlands, Email: p.bonomolo@dnb.nl.*
[c] *INSPER, Brasil, Email: hedibertfl@insper.edu*

30 May 2018

Abstract

We propose a generalization of the rational expectations framework to allow for multiplicative sunspot shocks and temporarily unstable paths. Then, we provide an econometric strategy to estimate this generalized model on the data. Our approach yields drifting parameters and stochastic volatility. The methodology allows the data to choose between different possible alternatives: determinacy, indeterminacy and temporary instability. We apply our methodology to US inflation dynamics in the '70s through the lens of a simple New Keynesian model. When temporarily unstable paths are allowed, the data unambiguously select them to explain the stagflation period in the '70s.

Keywords*:* Rational Expectations, Sunspots, Instability, Indeterminacy, Inflation, Monetary Policy.
JEL classifications*:* E31, E52.

* We thank Philippe Andrade, Paul Beaudry, Fabio Canova, Efrem Castelnuovo, John Cochrane, Andrea Colciago, Jakob de Haan, Martin Ellison, Stefano Eusepi, Luca Fanelli, Jesus Fernandez-Villaverde, Andrea Ferrero, Gaetano Gaballo, Valentina Gavazza, Tom Holden, Alejandro Justiniano, Jesper Lindé, Albert Marcet, Michael McMahon, Leonardo Melosi, Davide Raggi, Neil Rankin, Lorenza Rossi, Luca Sala, Thomas Sargent, Paolo Surico, conference participants at the SED 2012 in Limassol, the CEF 2016, the NBER Summer Institute 2016, the Cleveland Feds Inflation Conference 2016, the Learning Conference Expectations in Dynamic Macroeconomic Models 2016 in Amsterdam, the Workshop on Non-linear Models in Macroeconomics and Finance for an Unstable World in Oslo 2018, and participants in many seminars for helpful comments on earlier versions, as well as two anonymous referees. Bonomolo benefited from visiting the Research division of Sveriges Riksbank. This work acknowledges funding from the European Union's Seventh Framework Programme grant agreement number 612796, project MACFINROBODS "Integrated macro-financial modelling for robust policy design". Preliminary versions of this paper circulated under the titles "Does inflation walk on unstable paths?" and "Rational Sunspots". Views expressed are those of the authors and do not necessarily reflect official positions of De Nederlandsche Bank.

1 Introduction

The vast majority of modern dynamic macroeconomics has relied on models of rational expectations (RE henceforth) with a unique stable equilibrium, following the methodology in Blanchard and Kahn (1980). Such models are somewhat limited in terms of their ability to analyze unstable behavior in the data, which, especially now after the Great Financial Crisis, is an important issue in macroeconomics. One option is to make RE models more flexible to allow temporarily explosive paths. This work provides a novel framework to carry out this option, by considering a broader class of solutions and taking this to the data.

The rational expectations assumption generally implies multiple equilibria, that is, an infinite number of RE trajectories. Depending on the properties of the dynamic system at hand, these trajectories could be either explosive or stable. After Muth's (1961) seminal contribution, the literature faced the problem of how to select an equilibrium out of many possible ones.[1] The stability criterion was accepted as a general consistency requirement to impose on a infinite horizon RE agents model. In modern dynamic macroeconomics, the linear RE system often approximates the first-order conditions of underlying dynamic optimization problems. Explosive paths would generally violate the transversality conditions associated with such problems. Ruling out the possibility of unstable equilibria led saddle path dynamic systems to become the new standard in macroeconomics. Among the infinite RE equilibria in saddle path dynamics, only one is stable, thus the stability criterion is enough to pin down a unique admissible RE path. Blanchard and Kahn (1980) formalized this idea and conceptualized the solution algorithm on which dynamic macroeconomics is based.

The stability criterion, however, is not a sufficient selection device when the RE system admits multiple stable equilibria, as investigated by the literature on indeterminacy

[1] As Blanchard and Watson (1982, footnote 1, p. 27) put it: *"This indeterminacy arises [...] in all models in which expectations of future variables affect current decisions. It is the subject of much discussion currently in macroeconomics, under the label of 'non uniqueness'."* Sargent and Wallace (1973), Brock (1974), Phelps and Taylor (1977), Taylor (1977), Blanchard (1979), Blanchard and Kahn (1980) and Flood and Garber (1980) are some examples of this compelling debate in the literature that followed Muth's contribution. See also the discussion in Burmeister et al. (1983).

and sunspots. Moreover, given the stability criterion, RE solutions have a hard time in explaining unstable behaviour in the data, such as hyperinflations or boom and bust episodes in asset markets. The stability criterion, however, is an asymptotic condition so that one can not rule out a priori the possibility of explosive, but temporary, trajectories. In other words, the stability criterion might not be violated by paths that are temporarily on explosive RE trajectories, but whose time-path coincides with the time-path of the stationary solution after some date. Therefore, from a theoretical perspective, the possibility of temporarily unstable equilibria might not be excluded, and hence in the empirical analysis, it is appropriate to consider this possibility. We provide a novel framework to do so. Our contribution is both theoretical and empirical.

From a theoretical perspective, our framework generalizes RE solutions in two ways. First, it allows for a time-varying parameter solution, by considering *multiplicative sunspots*. Generally, any solution whose expectation error has a zero conditional mean could be considered a RE solution. The eigenvalues describe the nature of these solutions (i.e., the number of stable or unstable trajectories). In the case of *indeterminacy*, the system admits an infinite number of stable trajectories. The sunspot literature (e.g., Benhabib and Farmer, 1999; Lubik and Schorfheide, 2004) randomizes over all these infinite stable solutions, because for any given solution, it is always possible to construct another solution by adding a sunspot shock with zero conditional mean. Our framework parameterizes all these admissible solutions in a different way, through a free parameter whose value selects a particular solution among the infinitely many admissible ones.[2] This parameter has an appealing interpretation: it shows how the infinite solutions differ in the way agents form their expectations, or more precisely, in the way agents weight past data to calculate their RE. We then assume that this parameter follows a stochastic process, driven by a non-fundamental (sunspot) shock. In our interpretation, the economy randomly switches among the infinite RE solutions, because agents change the way they are forming their (rational) expectations. Our approach could also be seen as a different way to introduce

[2]This is similar to the original insight in Blanchard (1979), who showed that all these solutions could be expressed as a linear combination of the forward-looking solution and the backward-looking one, where the parameter of this linear combination corresponds to our free parameter.

sunspots randomizing over this parameter. Since this parameter enters non-linearly in the solution, our sunspots are multiplicative, instead of additive, as so far considered by the literature. In our approach, sunspot disturbances interact with the fundamental ones, and can be effective only when a fundamental error hits the economy. Given that our sunspots are multiplicative, the solution exhibits drifting parameters and stochastic volatility.

Second, once time-variation in the solution is allowed, our approach can also accommodate temporarily unstable paths. Under *determinacy*, there is an infinite number of unstable solutions and a unique stable RE solution, which satisfies the stability criterion. In other words, there is a unique admissible value of this free parameter: the one that pins down the unique (saddle-path) stable RE equilibrium. However, the stability criterion should be satisfied in the long run, so that time variation in the solution could allow temporary walks on these unstable RE trajectories, provided that the system converges to the unique stable one in the long run. Appropriate restrictions on the stochastic process describing the evolution of this parameter generate paths that could temporarily randomize over all the possible RE unstable trajectories, provided that this parameter converges to that unique admissible value in the limit. We thus propose a class of solutions where RE paths are temporarily unstable, but stable in the long run. Transversality conditions, however, usually require that at any given time the economy is expected to converge back to equilibrium in the long run, which is a somewhat more stringent requirement than the economy is actually converging in the limit. In order not to rule out a priori the possibility that the system could temporarily be on an unstable path in the estimations, we show how a minimal relaxation of the RE constraint allows us to consider unstable paths and also comply with the transversality conditions.

In our theoretical framework, therefore, the case of multiple solutions is the natural case in RE: (i) the values of the eigenvalues describe the nature of the dynamic system (i.e., the number of stable or unstable trajectories); (ii) in each period a parameter selects a particular solution among the infinitely many; (iii) this parameter follows a stochastic process and a restriction on the time-variation in this parameter imposes that the economy is expected to eventually converge back to equilibrium at the limit. The main insight is that while unstable paths are usually ruled out by imposing the stability criterion to select equilibria, the time-variation in the solution opens up the possibility of temporarily unstable paths, which are not necessarily in contrast with RE and stability.

From an empirical perspective, we develop an econometric strategy suited to our framework. Given that our sunspots are multiplicative and imply stochastic volatility, the likelihood is not Normal and we cannot use Gaussian methods. We thus proceed by estimating the model parameters and the latent states using a Bayesian approach based on sequential Monte Carlo methods. In particular, we build an econometric strategy for parameter learning that combines the approach of Carvalho et al. (2010), and the particle filter of Liu and West (2001).[3] Finally, we use the sequential Bayes factor presented in West (1986) to compare the different models. The econometric strategy allows for the cases of determinacy, indeterminacy or explosiveness, without imposing them a priori. We then propose a methodology to let the data choose the preferred equilibria among all the possible ones, and thus to test the empirical validity of temporarily unstable paths. By the same token, our approach could be seen as checking the validity of the stability criterion as usually imposed on the RE solutions.

To show the potential of our methodology, we apply our approach to explain the US inflation dynamics in the post-war sample. The Great Inflation of the '70s, and the subsequent Volcker disinflation, is among the most studied episodes of US monetary history. In an extremely influential article, Clarida et al. (2000) estimate an interest rate equation for the US and suggest that the change in the response of monetary policy to inflation could explain the different inflation behaviour between the Great Inflation period of the '70s and the so-called Great Moderation period of the late '80s and '90s. A simple New Keynesian model would predict that if monetary policy does not sufficiently react to inflation (i.e. the Taylor principle is not satisfied), there exists an infinite number of stable RE equilibrium paths. Such indeterminacy of equilibria could explain the aggregate insta-

[3] Fernández-Villaverde and Rubio-Ramírez (2007) present pioneering work on the estimation of non linear or non Gaussian DSGE models, based on particle filtering within a Markov chain Monte Carlo (MCMC) scheme. The use of sequential Monte Carlo methods is less common in the literature. Exceptions are Creal (2007), Chen et al. (2010) and Herbst and Schorfheide (2014). Our approach differs somewhat from the latter as explained in Section 3.

bility of the '70s through shifts in self-fulfilling agents' beliefs due to sunspot shocks. In a seminal contribution about the econometrics of indeterminate RE equilibria, Lubik and Schorfheide (2004) (LS henceforth) estimate a standard three-equations New-Keynesian model under both determinacy and indeterminacy. Their results provide support to the original Clarida et al.'s (2000) result in a multivariate context. Subsequently, other papers in the literature confirmed this narrative that identifies loose monetary policy as the cause of the Great Inflation period (e.g., Boivin and Giannoni, 2006; Benati and Surico, 2009; Mavroeidis, 2010; Castelnuovo et al., 2014; Castelnuovo and Fanelli, 2015; Lubik and Matthes, 2016).[4]

The New Keynesian literature, therefore, appeals to indeterminacy, induced by a dovish monetary policy, to explain the apparently explosive behaviour of inflation during the Great Inflation period, and to a hawkish one to explain the great Moderation. However, this has the rather paradoxical implication of appealing to a stable system to generate instability, as well as to an unstable system to ensure stability. From a theoretical perspective, a saddle path describes an unstable dynamic system, because there are infinite unstable trajectories while only one, that thus has measure zero, is stable. On the contrary, indeterminacy has an infinite number of stable trajectories, so it is a stable dynamic system. Indeterminacy, however, opens up the possibility of rationalizing an explosive behaviour by randomizing among all these possible stable trajectories thanks to a sunspot shock. Nonetheless, a central bank that does not respect the Taylor principle is sure that the economy is on stable dynamics, though subject to self-fulfilling beliefs, while on the contrary satisfying the Taylor principle is potentially highly risky, because the probability of being on the unique stable path (among infinitely many unstable ones) is practically zero. Macroeconomists generally assume agents are able to select this unique stable solution.

It seems to us it would be more natural to associate the unstable behaviour of inflation in the data to an unstable trajectory in the model. We thus apply our framework to ask

[4] Alternative possible explanations for the Great Inflation period put forward in the literature are stochastic volatility of the shocks (e.g., Justiniano and Primiceri, 2008; Fernández-Villaverde et al., 2010) or escape dynamics (e.g., Sargent, 1999; Cho et al., 2002; Sargent et al., 2006; Carboni and Ellison, 2009).

the following question: is there any evidence that inflation is described by temporarily unstable equilibria in the '70s?

The seminal paper of LS is the natural benchmark against which to compare our results, so we will use both their econometric model and their data. If we impose the stability criterion on the estimation, that is, allowing just for determinacy or indeterminacy while ruling out temporary instability, our econometric strategy recovers results that are practically identical to the one in LS. Our main result, however, is to provide evidence that the high inflation during the '70s is better explained by temporarily unstable dynamics: the data seem to favour a temporarily unstable equilibrium path to explain the Great Inflation period, rather than a randomization over stable trajectories, as suggested by the indeterminacy literature.

Inflation in the '70s increased quite rapidly. Intuitively, to explain the data a standard indeterminacy model needs to rely on persistent and successive sunspot shocks in the same direction. The data assign a low likelihood to such a sequence of shocks and favour a model that presents inherent temporarily explosive dynamics. However, the model also features stochastic volatility, so that, one might think that allowing for the possibility of large shocks, rather than for temporarily explosive dynamics, is what makes the model outperform the indeterminacy model. In section 6, we compare our framework to one with stochastic volatility but a unique stable trajectories (i.e., determinacy), as in Justiniano and Primiceri (2008). Again the estimation favours a model with intrinsic temporarily unstable dynamics, rather than a model that would need a series of consecutive large shocks in the same direction to capture the unstable behavior of the data during the Great Inflation period.

The paper proceeds as follows. Section 2 explains our approach by means of a simple model. Section 3 explains our econometric strategy. Section 4 presents our application, that is, how we apply our approach to the New Keynesian model in LS. Section 5 shows and comments on the empirical results. Section 6 presents a comparison with a stochastic volatility model. Section 7 concludes.

2 Multiplicative Sunspots and Unstable Paths

We use a simple example to illustrate our approach. We proceed in two steps. First, we introduce multiplicative sunspots by allowing agents to switch between all the possible fundamental solutions under stability. Second, we discuss asymptotic stability and we examine the possibility of temporarily unstable paths. Finally, we generalize our simple example to a multivariate model.

2.1 A simple example

Consider the following expectational difference equation (as in LS's Section II):

$$y_t = \frac{1}{\theta} E_t y_{t+1} + \varepsilon_t, \tag{1}$$

where ε_t is a i.i.d shock $\sim N(0, \sigma_\varepsilon^2)$ and $\theta \in \Theta = [0, 2]$. $E_t y_{t+1} = E(y_{t+1}|I_t)$ is the expected value of y at $t+1$ conditional on the information set available at time t.[5] It is important to stress from the outset the logic of a forward-looking equation as (1): the expectations regarding the value of y in the following period determines the equilibrium value of y at t (and not viceversa).[6] Here lies a fundamental degree of freedom: the way agents form their expectations about future values of y pins down the equilibrium value today. Equation (1) naturally has an infinite number of solutions, because one can find an infinite number of pairs $(y_t, E_t y_{t+1})$ that satisfy it (see footnote 1).

Muth's (1961) RE seminal idea restricts the way agents form their expectation to be coherent with the economic system, so that the expected forecast error should be zero (i.e., the error in expectation should not be correlated with anything in the available information set). Defining the forecast error as $\eta_t = y_t - E_{t-1} y_t$, thus: $E_{t-1}(\eta_t) = 0$. The RE requirement, however, is generally not enough to pin down a unique solution: any process η_t such that $E_{t-1}(\eta_t) = 0$ defines a different solution to (1). The constraint that the expectation error has zero mean simply implies that the solution is characterized up

[5] 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.

[6] See the discussion in Section 2.4 in Woodford (2003), especially pages 127-128.

to an arbitrary martingale process.[7] Let ζ_t be a mean zero non-fundamental disturbance, uncorrelated with the fundamental one (i.e., what the literature calls a sunspot shock), then any forecast error of the form:

$$\eta_t = (1 + M)\varepsilon_t + \zeta_t \tag{2}$$

yields a RE solution.[8]

The early literature on RE (see footnote 1) agreed on considering only stable RE solutions as a general consistency requirement to impose on a model of infinite horizons RE agents, because of the set of transversality conditions associated with the agents' dynamic optimization problems in the underlying model. However, whether or not the stability criterion is sufficient to select a unique solution depends on the stability properties of the expectational difference equation (1), and that in turn hinge on the value of the parameter θ.

To see it, simply introduce conditional expectations by defining $\xi_t = E_t(y_{t+1})$, so that (1) can be written as:

$$\xi_t = \theta\xi_{t-1} - \theta\varepsilon_t + \theta\eta_t, \tag{3}$$

which corresponds to the way Sims (2002) writes and solves linear rational expectations models. If $\theta > 1$, deviations of ξ_t from 0 explode with time, thus stability requires $\xi_t = 0, \forall t$. Hence, the stability criterion imposes a restriction on the forecast error (2): $\eta_t = \varepsilon_t$ and $M = \zeta_t = 0, \forall t$. If $\theta > 1$, thus, this restriction pins down the only one solution to (3), among the infinitely many RE ones, that does not violate the stability criterion. Blanchard and Kahn (1980) generalized this idea to a multivariate RE linear system with backward and forward looking variables, and conceptualized the well-known solution algorithm that is a cornerstone of dynamic macroeconomics.

If $\theta \leq 1$, however, the model is indeterminate, because all the infinite RE solutions of

[7] If $\mathcal{M}_t$ is a martingale, then $\Delta\mathcal{M}_t = \mathcal{M}_t - \mathcal{M}_{t-1}$ is a martingale difference process, and $E_{t-1}(\Delta\mathcal{M}_t) = E_{t-1}(\mathcal{M}_t - \mathcal{M}_{t-1}) = 0$. So one could interpret the error of expectations η_t as a martingale difference process, and the requirement of a zero expected error simply implies that the solution is characterized up to an arbitrary martingale (see Pesaran, 1987).

[8] Plugging (2) into (1) yields: $y_t = \theta y_{t-1} - \theta\varepsilon_{t-1} + (1 + M)\varepsilon_t + \zeta_t$, which is a way of writing all the possible fundamental solutions to (1) parameterized by M plus the sunspot shock.

(3) are stable. In other words, any deviation of ξ_t from 0 will not lead ξ_t to explode over time. Hence, the stability criterion imposes no restriction on the forecast error (2), and thus it does not solve the problem of selecting a unique equilibrium. The indeterminacy literature then assumes that the economy will choose randomly among these infinite stable solutions.[9] This randomization is usually done (e.g., LS) by adding an exogenous sunspot (i.e. non fundamental) shock, ζ_t, in (2) for a given value for M (on which the system dynamics put no restrictions).

2.2 Indeterminacy and Multiplicative Sunspots: A generalized time-varying solution

Consider the case of indeterminacy, $\theta \leq 1$. (2) suggests another possible source of multiplicity: M. M derives from the intrinsic multiplicity of the RE solutions, because it parameterizes all the possible fundamental solutions, where the expectation error is just a function of the structural shock (i.e., no additive sunspots), so that: $\eta_t = (1+M)\varepsilon_t$. Thinking along the lines of the Benhabib and Farmer's (1999) quotation in footnote 9 suggests a different way to introduce sunspot shocks by randomizing over the fundamental solutions, i.e., randomizing over M, rather than adding ζ_t. This approach introduces a multiplicative sunspot shock, rather than an additive one.

To illustrate our approach, we first show that M parameterizes all the fundamental RE solutions, and we provide an economic interpretation for M. Second, we introduce time-variation in M. We then show that the RE condition restricts the type of admissible time variation processes, and we analyze the nature of these solutions.

Parameterization. Consider only fundamental solutions where $\zeta_t = 0, \forall t$ in (2). Substituting in (3) and iterating backward yields:

$$\xi_t = M\theta \sum_{i=0}^{t-1} \theta^i \varepsilon_{t-i} = M \sum_{i=1}^{t} \theta^i \varepsilon_{t+1-i}, \tag{4}$$

[9] *"Sunspot equilibria can often be constructed by randomizing over multiple equilibria of a general equilibrium model, and models with indeterminacy are excellent candidates for the existence of sunspot equilibria since there are many equilibria over which to randomize"* Benhabib and Farmer (1999, p.390).

assuming that exists a period zero where everything starts (i.e., that the economy in a whatever distant past was in a steady state): $\varepsilon_{-i} = \xi_{-i} = 0, \forall i \geq 0$. All the possible fundamental solutions are thus parameterized by $M \in (-\infty, +\infty)$, because a particular value of M defines a particular solution.[10] Among the infinitely many, two important solutions are often considered in the literature: (i) the pure forward looking solution corresponding to $M = 0 : \eta_t = \varepsilon_t,\ \ \xi_t^F = 0,\ \ y_t^F = \varepsilon_t$; (ii) the pure backward looking solution, corresponding to $M = -1 : \eta_t = 0,\ \ \xi_t^B = -\theta \sum_{i=0}^{t-1} \theta^i \varepsilon_{t-i},\ \ y_t^B = \xi_{t-1}^B = -\sum_{i=1}^{t-1} \theta^i \varepsilon_{t-i}$.[11]

M has a very natural interpretation: it defines the way agents form their expectations. More precisely, it defines if and how agents are going to use past observations in forming their expectations. One of the purposes of Muth's (1961) original paper is to write the expectation at time t as an exponentially weighted average of past observations, because a previous paper, i.e., Muth (1960), demonstrated that this is the optimal estimator under some assumptions. In the simple case of equation (1), the expectation (when $M \neq -1$) is given by:

$$\xi_t \equiv E_t y_{t+1} = M \sum_{i=0}^{t} \left(\frac{\theta}{1+M} \right)^i y_{t+1-i}, \tag{5}$$

$E_t y_{t+1}$ is the product of two terms. First, M measures how much the past is important in forming expectations in absolute terms: if $M = 0$, then past data do not matter. This is the forward-looking solution. Second, the weights $\left(\frac{\theta}{1+M}\right)^i$ tell us how much agents relatively weight the past data. The higher is M, the less past terms are important in setting expectations. Then, M determines how the agents combine past observations in making their forecasts both in *absolute* terms (M versus 0), and in *relative* terms.

Time variation. Following Muth's RE original formulation, we just argued that M can be interpret as pinning down the infinite number of ways agents could combine past data to form their expectations. Our proposed class of solutions simply generalize the

[10] From (4), it is again evident that, if $\theta > 1$ (i.e., determinacy), ξ_t does not explode only if $M = 0 \Rightarrow \eta_t = \varepsilon_t, \forall t$, while, if $\theta < 1$ (i.e., indeterminacy), stability imposes no restrictions on M.

[11] It is easy to rewrite (4) as a linear combination of the forward and backward looking solutions as: $\xi_t = \theta M \varepsilon_t + M\theta \sum_{i=1}^{t-1} \theta^i \varepsilon_{t-i} = \theta M (y_t^F - y_t^B)$. It follows that one can write also the solution for y as a weighted average of the backward and the forward looking solution (see Blanchard, 1979).

standard one, (4), by letting M be a random variable that can change over time, so that:

$$\xi_t = M_t \theta \sum_{i=0}^{t-1} \theta^i \varepsilon_{t-i} \equiv -M_t \xi_t^B. \tag{6}$$

RE implies $E_{t-1}(\eta_t) = 0$. Plugging the proposed solution (6) into the original equation (3) and solving for η_t yields:

$$\eta_t = (1 + M_t)\varepsilon_t + (M_t - M_{t-1}) \left(\sum_{i=1}^{t-1} \theta^i \varepsilon_{t-i} \right), \tag{7}$$

which gives the forecast error implied by our proposed solution. For it to be a RE solution it must be: $E_{t-1}(\eta_t) = 0$. Thus, M_t must satisfy the following two conditions: 1) $E_{t-1}(M_t) = M_{t-1}, \forall t$, that is, M_t must be a martingale process:; 2) $E_{t-1}\left[(1 + M_t)\varepsilon_t\right] = 0$, that is, M_t must be uncorrelated with ε_t.

Implications and discussion. Our approach has a number of implications. First, from the point of view of the economic interpretation, our approach simply allows for agents to change over time the weights they assign to past shocks or past data in forming their expectations. The solution has the same form as (5) but with M being time-varying. Under indeterminacy, a given M_t selects one of the infinite stable RE path, and then agents randomly shift from one to another. In some periods agents form their expectations with great trust in the past, while in some other periods they expect ξ_t to be more or less around its steady state (i.e., $M_t = 0$, the forward looking solution in this simple case).

Second, by introducing a multiplicative sunspot shock, rather than an additive one, our solution features time-varying parameters and stochastic volatility. From (3), (6) and (7), we can write our solution as:

$$\begin{cases} y_t = \alpha_t y_{t-1} - \alpha_t \varepsilon_{t-1} + (1 + M_t)\varepsilon_t & iff \quad M_{t-1} \neq 0 \\ y_{t-1} = \varepsilon_{t-1} & if \quad M_{t-1} = 0 \end{cases} \tag{8}$$

with $\alpha_t = \theta \dfrac{M_t}{M_{t-1}}$.[12] The random variation of M_t causes both a different structural

[12] Given the two conditions above (i.e., $E_{t-1}(M_t) = M_{t-1}$ and $E_{t-1}[(1 + M_t)\varepsilon_t] = 0$), it follows that (8)

dependence of ξ_t (or y_t) from its lagged value and a different reaction of the system to the current shock. Drifting parameters naturally arise because agents change how they form their expectation formation process each period, so changing the intrinsic dynamics of the model (i.e., α_t). Stochastic volatility arises because the way the system reacts to the current fundamental shock depends on the current realization of M_t. Hence, the stochastic process for M_t interacts with the structural shock through the term $(1+M_t)\varepsilon_t$, possibly amplifying the effects of ε_t on the economy. Our approach has the potential for an economic explanation of drifting parameters and stochastic volatility, without departing from the RE hypothesis. The empirical research (Cogley and Sargent, 2005; Primiceri, 2005; Justiniano and Primiceri, 2008, and related literature) considers these as important features in explaining the dynamics of macroeconomic variables.

This important property of our solution is evident in the expression for the forecast error (7) which is the sum of two terms. The first term is the interaction term between the innovation in M_t and the structural shock. The second term is due to the fact that the change in M_t leads agents to respond differently to past shocks, putting the system on a different RE path. We can re-write the forecasts error as:[13]

$$\eta_t = (1+M_t)\varepsilon_t + (\alpha_t - \theta)(y_{t-1} - \varepsilon_{t-1}) = (1+M_t)\varepsilon_t + \frac{(M_t - E_{t-1}M_t)}{E_{t-1}M_t} E_{t-1}y_t, \qquad (9)$$

where the second term highlights how M_t determines the structural dynamics of the solution. This term derives from the time-varying coefficient α in (8), because it depends on $\alpha_t - E_{t-1}(\alpha_t) = \alpha_t - \theta$, and captures the fact that an innovation in M_t changes the equilibrium trajectory of the system and its structural dynamics. The endogenous emergence of stochastic volatility and drifting parameters within the RE framework is the direct consequence of assuming a multiplicative sunspot that makes the likelihood non Gaussian.

Last but not least, our approach allows to recover the minimum state variable solution

satisfies the original equation (1). Moreover, it can also be written as a dynamic formulation of Blanchard (1979): $y_t = -M_t y_t^B + (1+M_t)y_t^F$.

[13]Since from (1) $E_{t-1}y_t = \theta(y_{t-1} - \varepsilon_{t-1})$, then using (8): $\eta_t = y_t - E_{t-1}y_t = (\alpha_t - \theta)(y_{t-1} - \varepsilon_{t-1}) + (1+M_t)\varepsilon_t$ and given that $\alpha_t = \theta M_t / M_{t-1}$, it follows: $\eta_t = y_t - E_{t-1}y_t = \left(\frac{M_t - M_{t-1}}{M_{t-1}}\right)\theta(y_{t-1} - \varepsilon_{t-1})$.

extremely easily, simply by putting $M_t = 0$. This remains true in a more general model (see Section 2.4) and hence in the empirical implementation, where the data could choose the usual minimum state variable solution by supporting the estimate of $M_t = 0$.

2.3 Modelling Temporarily Unstable Paths

When $\theta > 1$, the solution (6) is unstable. If explosiveness is allowed in the model, as it could be the case of a model with only nominal variables or of asset pricing, then the only restriction comes form the RE requirement that constraints M_t to be: 1) a martingale; 2) uncorrelated with the fundamental shock ε_t.

In the more general case where the model needs to satisfy the stability criterion, the only stable solution when M is a constant is the forward-looking solution. The stability criterion, however, relates to the asymptotic behaviour of the solution. Hence, it does not rule out "bubbly", but temporary, trajectories, featuring unstable dynamics that are temporarily explosive, but stable in the long run following the bursting of the bubble. Such stationary paths would actually exhibit the same asymptotic behavior as the one selected by the stability criterion. In this Section, we show that our approach could consider this broader class of solutions (where equilibrium paths are temporarily unstable, but asymptotically stable).

To appreciate how our approach could allow temporarily unstable paths, note that in each period t, the solution (6) depends only on the current realization of M_t and not on its past values. We impose this restriction on our class of solution. Hence, we are not considering all possible solutions under time-variation.[14] The stability condition requires that M_t converges to zero faster than $\sum_{i=0}^{t-1} \theta^i$ goes to infinity. Hence, for example, any process such that $M_t = 0$, with probability 1 in finite time would satisfy the stability criterion. More generally, Gourieroux et al. (1982) shows that, given a stationary solution y_t^0 to (1), it is always possible to find non-stationary processes whose time path eventually coincides with the time path of the stationary solution after some random date, that depends on the

[14]For example, we do not assume: $\eta_t = (1 + M_t)\varepsilon_t$, which would yield a different solution with respect to ours: $\xi_t = \theta \sum_{i=0}^{\infty} \theta^i M_{t-i}\varepsilon_{t-i}$. In this case a change in M_t affects only the weight period t, while in our framework it affects all the weights in (6).

time path. They call these solutions *asymptotically equal to a stationary process.*[15] They are equilibrium path that are temporarily unstable, but asymptotically stable, since they will eventually coincide with the unique stationary solution in finite time. As an example, Gourieroux et al. (1982) shows that the process $M_t = \begin{cases} M_{t-1}/p & \text{with probability } p \\ 0 & \text{with probability } 1-p \end{cases}$ leads to an AES solution for (1).[16] While our solution is explosive if $\theta > 1$ and $M_t \neq 0$, time-variation in M_t makes possible to consider a martingale processes for M_t that will randomly converge to the unique stationary solution, such that $M_t = 0, \forall t > T$. This would satisfy a stability criterion that requires the economy to eventually converge so that: $\lim_{i->\infty} = \xi_{t+i} = \lim_{i->\infty} E_{t+i}(y_{t+1+i}) = 0$.

The stability criterion usually induced by transversality conditions in optimization problems, however, imposes a different restriction that relates to the *current expectation* of the asymptotic behaviour of the solution, requiring that $\lim_{i->\infty} E_t(y_{t+i}) = 0$. The RE requirement, however, restricts the admissible time-varying processes for M_t to be a martingale and uncorrelated with ε_t. The martingale requirement implies that if $\theta > 1$ and $M_t \neq 0$ the economy is expected to remain forever on the unstable path selected by M_t, so that the transversality condition would be violated, even if the instability is only temporary.[17] To allow for temporarily unstable paths in this case, we need to relax the martingale assumption, and thus the RE assumption. This deviation, however, could be minimal without practical implications when the model is taken to the data. In our empirical analysis in Section 5 we will assume that $M_{1,t}$ follows:

$$M_{1,t} = N_t \Im_{\{\|y^B_{t-1}\| < \bar{U}\}} \tag{10}$$

[15] According to Gourieroux et al. (1982), a solution y_t is said to be asymptotically equal to a stationary solution (AES) y^0_t, if, for any time-path w, there exists $T(w)$ such that $y^1_t(w) = y^0_t(w), \forall t > T(w)$. It is important to note that the date $T(w)$ depends on the time-path w and, therefore, we cannot generally find two solutions such that they are identical at a date $t < T$. If it is, then it must be that they are the same solution.

[16] This process has been indeed used in the somewhat related rational bubble literature (see Blanchard and Watson, 1982; West, 1987).

[17] This is easy to see by looking at (6), and assuming the AES process above for M_t suggested by Gourieroux et al. (1982). Note that the economy will eventually converge, because $\lim_{i->\infty} \xi_{t+i} = \lim_{i->\infty} E_{t+i}(y_{t+1+i}) = 0$.

where N_t is a martingale and $\Im_{\{.\}}$ is the indicator function. In order to satisfy the the transversality condition, the martingale process N_t multiplies the indicator function $\Im_{\{\|y^B_{t-1}\|<\bar{U}\}}$. The latter is equal to one if the norm of the backward looking solution at time $t-1$ is less than a certain scalar $\bar{U} < \infty$, and is equal to zero otherwise. Since $\theta > 1$, there exists a random date $\bar{T}$ in which $\left\|y^B_{\bar{T}-1}\right\|$ becomes greater than $\bar{U}$ and the indicator function will be equal to zero for all $t > \bar{T}$. After this date, the stochastic process $M_{1,t}$ will also be equal to to zero, and the dynamics will coincide with the unique stable solution (i.e., the forward-looking solution). The indicator function in (10) is a random variable which realization is known at time t, since it depends on the past value of the backward looking solution. Then, in general (i.e., except in period $\bar{T}$) $E_t M_{1,t+1} = M_{1,t}$, which implies the expected value of the forecast error is equal to zero. However, $\lim_{s\to\infty} E_t M_{1,t+s} = 0$, so that also the transversality condition holds. The presence of the indicator function is a simple expedient. In other words, the probability that RE requirement may be violated in any near future is zero. However, it will be violated in a finite future, whatever far, depending on how large is $\bar{U}$. In order to allow for temporarily unstable paths, we implicitly assume that this possibility in the very distant future is disregarded by the agents.[18] In the practical implementation of the econometric procedure, one can choose $\bar{U}$ so large that both conditions are satisfied by any estimate or simulated impulse response paths, such that the indicator function is equal to one for all the draws and the time periods considered in our sample. For example, in Section 5 we specify the following process for N_t:

$$N_t = \begin{cases} N_{t-1}/\gamma + \zeta_t & \text{with probability } \gamma \\ 0 & \text{with probability } 1-\gamma \end{cases} \tag{11}$$

where $\zeta_t \sim N(0, \sigma^2_\zeta)$. Without the shock ζ_t the process for N_t would have zero as absorbing state so that the stable solution would be forever active once selected. The presence of the shock makes it possible for an economy that is on the stable path, to jump on an explosive trajectory. This "walk on unstable paths" will only be temporary, either because with

[18]As Blanchard and Watson (1982, , p. 8) put it, the argument that rules out this possibility: *"may be pushing rationality too far. [...] the probability [...] may be so small, and the future time so far as to be considered nearly rationally irrelevant for market participants."*

probability $(1 - \gamma)$ the solution will become stable again or because of the indicator.

Implications and discussion. There are three main differences between our approach and the standard sunspot literature. First, in the standard additive sunspot approach M is constant and thus it can be different from zero only if $\theta \leq 1$. Hence, sunspots are allowed only if $\theta \leq 1$, i.e., under indeterminacy. Our approach, instead, allows temporarily unstable paths, even if $\theta > 1$, because M_t varies with time, and stability can be imposed asymptotically on the process for M_t. Our proposed solution makes it possible to consider infinitely many possible asymptotically stable solutions: the stability criterion is not anymore enough to select a unique possible equilibrium even if $\theta > 1$. Hence, the model is not "determinate" anymore even if $\theta > 1$: *indeterminacy, in the sense of an infinite number of admissible paths, is the natural case.*

Second, as already stressed, multiplicative sunspot implies stochastic volatility and drifting parameters within a RE framework. As clear by comparing the forecast errors (2) and (9), our approach implies that the structural dynamics of the solution changes over time with M_t, in terms of the response to both the current realization of the shock, leading to stochastic volatility, and past data, leading to drifting parameters.

Last but not least, we provide a way to take our framework to the data. If M is time-varying, theoretically it is harder to rule out equilibria that are only temporarily unstable. We simply acknowledge that in the empirical analysis, it should be appropriate to consider this possibility. We want to allow temporary "walks along unstable paths" by estimating the latent process for M_t and then ask to the data which kind of equilibria they prefer. At the very least, our approach could be seen as a test of RE, or of the transversality conditions, as normally applied. Hence, we are not taking a stand a priori on the possible equilibria in our estimation strategy, by allowing for all the possible cases: indeterminacy, determinacy and (temporary) instability. We then propose a methodology to let the data choose the preferred equilibria, and thus to test the empirical validity of these temporarily unstable paths. This is what we turn to next, explaining our proposed methodology in a more general context.

2.4 Implementation: The general solution

To implement our proposed solution in the simple case, given an exogenous process for M_t, recursively define the solution for the expectation error in (7) using the backward-looking solution $\xi_t^B = -\theta \sum_{i=0}^{t-1} \theta^i \varepsilon_{t-i}$ so that: $\eta_t = (1 + M_{t-1})\, \varepsilon_t - (M_t - M_{t-1})\xi_t^B/\theta$ where $\xi_t^B = -\theta\varepsilon_t + \theta\xi_{t-1}^B$. Once solved for the expectation error, the solution for ξ_t is simply given by (3). More compactly using (6), we can write our solution as:

$$\xi_t = M_t \xi_t^B \tag{12}$$

$$\xi_t^B = -\theta\varepsilon_t + \theta\xi_{t-1}^B. \tag{13}$$

plus a stochastic process for M_t.

The multivariate case is a relative straightforward extension of the simple case. Appendix A.1 describes it in details, following similar steps as above, involving: (i) parameterizing the system using M (now a matrix); (ii) introducing time variation in M; (iii) imposing stability. As in LS, we follow the approach of Sims (2002) and we write a general linear RE system as:

$$y_t = \Gamma_1^* y_{t-1} + \Psi^* \varepsilon_t + \Pi^* \eta_t. \tag{14}$$

where y_t is the vector of the n endogenous variables (including the expectations as in (3)), ε_t is the vector of the j exogenous fundamental shocks, and η_t is the vector of the $k \leq n$ RE forecast errors. As usual, we need to partition the system. Use Jordan decomposition to diagonalize $\Gamma_1^* = J\Lambda J^{-1}$ and define the vector of transformed variables $\tilde{y}_t = J^{-1}y_t$. We depart from Sims (2002) and LS by partitioning the system according to the number of forward-looking variables/expectation errors, rather than the number of explosive eigenvalues. Then:

$$\tilde{y}_t = \begin{bmatrix} \Lambda_1 & \mathbf{0} \\ ((n-k)\times(n-k)) & ((n-k)\times k) \\ \mathbf{0} & \Lambda_2 \\ (k\times(n-k)) & (k\times k) \end{bmatrix} \tilde{y}_{t-1} + \begin{bmatrix} J_{\mu 1} \\ ((n-k)\times n \\ J_{\mu 2} \\ (k \times n) \end{bmatrix} [\Psi^*\varepsilon_t + \Pi^*\eta_t]\,. \tag{15}$$

Let m be the number of explosive eigenvalues (i.e., such that $\lambda_i \geq 1$). As usual, we assume that the number of explosive eigenvalues is smaller or equal to the number of forecast errors. This case $m \leq k$, is the usual one in the literature where one can have either determinacy ($m = k$) or indeterminacy ($m < k$).[19] That means that in our partition, the first $(n - k)$ rows contain only stable eigenvalues, while the last k rows contain both $(k - m)$ stable and m unstable eigenvalues. Hence, we do not need to impose any stability condition on the first block of the system (15). However, we will do on the second block of equation in (15):

$$\tilde{y}_{k,t} = \Lambda_2 \tilde{y}_{k,t-1} + J_{\mu 2} \left[\Psi^* \varepsilon_t + \Pi^* \eta_t \right] \tag{16}$$

where $\tilde{y}_{k,t}$ denotes a vector of dimension k.

Define M_t as a (kxk) diagonal matrix whose elements on the principal diagonal are changing over time. Generalizing (12) and (13), Appendix A.1 shows that the solution to the system of disconnected difference equations (16) can be written recursively using the backward-looking variable $\tilde{y}^B_{k,t}$ as:

$$\tilde{y}^B_{k,t} = \Lambda_2 \tilde{y}^B_{k,t-1} + J_{\mu 2} \Psi^* \varepsilon_t \tag{17}$$

$$\tilde{y}_{k,t} = -M_t \tilde{y}^B_{k,t} = -M_t \left(\Lambda_2 \tilde{y}^B_{k,t-1} + J_{\mu 2} \Psi^* \varepsilon_t \right) \tag{18}$$

and the expectation error is equal to:

$$\eta_t = \left(J_{\mu 2} \Pi^* \right)^{-1} \left[-(I + M_{t-1}) J_{\mu 2} \Psi^* \varepsilon_t - \left(M_t - M_{t-1} \right) \tilde{y}^B_{k,t} \right], \tag{19}$$

assuming that the (kxk) matrix $J_{\mu 2} \Pi^*$ is invertible.

Note that if all the elements $M_{i,t}$ on the principal diagonal of M_t are equal to zero for all t, then $\tilde{y}_{k,t} = 0$, so that we recover the forward looking solution. As in the simple model, we impose stability by allowing only particular processes for $M_{i,t}$'s. In general, if we have k non predetermined variables, the cardinality of the set of solutions is infinite to the power of k. However, the stability requirement imposes a restriction on the $M_{i,t}$'s

[19] We rule out the case $m > k$, where the number of unstable eignevalues, m, is bigger than the number of forward-looking variables, k. In this case, a stable solution does not exist.

that correspond to eigenvalues of the system that are outside the unit circle. As in the simple example, we will restrict the processes for these $M_{i,t}$ to randomly converge to the stationary forward-looking solution in finite time, such that $M_{i,t} = 0, \forall t > \bar{T}$, where $\bar{T}$ is a random variable. The stability condition, instead, does not impose any restrictions on the stochastic processes governing the $(k-m)$ elements of M_t corresponding to stable eigenvalues. The stability condition reduces the degrees of freedom in the matrix M_t, and it downsizes the set of solutions because the degree of indeterminacy is only $k-m$. The limiting case is the determinacy case, when the Blanchard-Kahn condition is satisfied and $k = m$: the stability condition will force all the elements in the main diagonal of M_t to converge to 0, that is, to converge to the unique stable forward-looking solution.

Appendix A.1 shows that the solution for the original variables is:

$$\begin{bmatrix} y_t \\ y_t^B \end{bmatrix} = \begin{bmatrix} J & \mathbf{0} \\ \mathbf{0} & J \end{bmatrix} G_t^* \begin{bmatrix} J^{-1} & \mathbf{0} \\ \mathbf{0} & J^{-1} \end{bmatrix} \begin{bmatrix} y_{t-1} \\ y_{t-1}^B \end{bmatrix} + \begin{bmatrix} J & \mathbf{0} \\ \mathbf{0} & J \end{bmatrix} H_t^* \varepsilon_t, \tag{20}$$

where

$$G_t^* = \begin{bmatrix} \Lambda_1 & \mathbf{0} & \mathbf{0} & -B_{t,t-1} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & -M_t\Lambda_2 \\ \mathbf{0} & \mathbf{0} & \Lambda_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \Lambda_2 \end{bmatrix}, \qquad H_t^* = \begin{bmatrix} A_t \\ -M_t J_{\mu 2}\Psi^* \\ J_{\mu 1}\Psi^* \\ J_{\mu 2}\Psi^* \end{bmatrix}, \tag{21}$$

and A_t is a $(n-k) \times l$ matrix and $B_{t,t-1}$ is a $(n-k) \times k$ matrix equal to, respectively:

$$A_t = J_{\mu 1} \left[\Psi^* - \Pi^* \left(J_{\mu 2}\Pi^* \right)^{-1} (I + M_t) J_{\mu 2}\Psi^* \right], \tag{22}$$

$$B_{t,t-1} = J_{\mu 1}\Pi^* \left(J_{\mu 2}\Pi^* \right)^{-1} \left(M_t - M_{t-1} \right) \Lambda_2. \tag{23}$$

Note that $M_t = \mathbf{0}, \forall t$ implies $A_t = J_{\mu 1} \left[\Psi^* - \Pi^* \left(J_{\mu 2}\Pi^* \right)^{-1} J_{\mu 2}\Psi^* \right]$ and $B_{t,t-1} = J_{\mu 1}\Pi^* \left(J_{\mu 2}\Pi^* \right)^{-1}$ $\left(M_t - M_{t-1} \right) \Lambda_2 = \mathbf{0}$, so that the system does not depend on $y_{k,t}^B$ anymore, and the solution coincides with:

$$y_t = J \begin{bmatrix} \Lambda_1 & \mathbf{0} \\ ((n-k)\times(n-k)) & ((n-k)\times k) \\ \mathbf{0} & \mathbf{0} \\ (k\times(n-k)) & (k\times k) \end{bmatrix} J^{-1} y_{t-1} + J \begin{bmatrix} J_{\mu 1}\left[\Psi^* - \Pi^*\left(J_{\mu 2}\Pi^*\right)^{-1} J_{\mu 2}\Psi^*\right] \\ (n-k)\times l \\ \mathbf{0} \\ (k\times l) \end{bmatrix} \varepsilon_t, \tag{24}$$

which is the usual Blanchard-Khan solution in case of a determinate system or the minimum state variable solution for an indeterminate system.

3 Econometric Strategy

In this section we take a Bayesian approach to make inference regarding the parameters and the latent processes of a DSGE model when considering the class of solutions (20). The presence of stochastic volatility in the reduced form of the model, related to the time-varying characteristic of the latent state M_t, leads to a non Gaussian, analytically intractable likelihood function. In such situations, when estimating non linear or non Gaussian DSGE models, a well-known approach proposed by Fernández-Villaverde and Rubio-Ramírez (2007), is performed in two-steps. In the first step, the integrated likelihood of the parameters is approximated through the implementation of a particle filter. Then, in the second step, one uses the approximated likelihood within a Markov chain Monte Carlo (MCMC) scheme that samples from the posterior distribution of the parameters.[20] We depart from this tradition and suggest the use of an efficient particle filtering strategy directly to approximate the joint posterior distribution of both the parameters and the latent state variables such as M_t. In particular we follow Chen et al. (2010) who introduced a particle strategy for DSGE models that combines the sequential Monte Carlo (SMC) algorithms of Liu and West (2001) and Carvalho et al. (2010).

In what follows, we first show how to write the solution (20) in a convenient state space form, and then we illustrate and discuss the main aspects of our particle filtering

[20] Recent papers in which this approach is implemented are Carvalho et al. (2017) and Gust et al. (2017).

strategy, refering to Appendix A.2 for an in depth description. Finally, we motivate our choice by comparing it with possible alternatives.

3.1 The state space form

In the class of solutions (20), we need to keep track of the pure backward looking solution y_t^B. This vector contains both endogenous and exogenous variables, and since the latter do not depend on the agent's expectations, their evolution will be the same as the analogous variables in y_t. When estimating the model, it is convenient to rewrite the solution (20) so that the exogenous variables appear only once, that is, using a compact notation:

$$l_t = G_t l_{t-1} + H_t \varepsilon_t \tag{25}$$

with

$$l_t = \begin{bmatrix} y_t \\ y_t^{B,E} \end{bmatrix},$$

where $y_t^{B,E}$ is a vector with the endogenous variables in the pure backward looking solution. The matrices G_t and H_t are appropriate transformations of the matrices in (20).

At each time t we observe a vector of data, which will be simply denoted by D_t. Then, the solution of model (14) has the following state space representation:

$$\begin{cases} D_t = c + F l_t + v_t & v_t \sim N\left(\mathbf{0}, \Sigma_v\right) \\ l_t = G_t l_{t-1} + H_t \varepsilon_t & \varepsilon_t \sim N\left(\mathbf{0}, \Sigma_\varepsilon\right) \end{cases} \tag{26}$$

where c is a vector of constants, F is a matrix with appropriate dimensions and v_t is a vector of measurement errors.

3.2 The particle filter

The parameters in c, F, G_t, H_t, Σ_v, Σ_ε are collected in the vector θ. As already mentioned, let D_t be the vector of observed data at time t, and $D_{m:n}$ denote the set of all observations from $t = m$ to $t = n$ for $m \leq n$. We perform posterior Bayesian inference via Monte Carlo methods to approximate the joint posterior distribution of parameters and latent states of the model by a sufficiently large number of sample draws, or particles. More precisely, our econometric strategy is based on Bayesian sequential learning via particle filtering: at time $t-1$, we start with a particle set $\{(l_{t-1}, M_{t-1}, \theta)^{(i)}\}_{i=1}^{N}$ and associated particle weights $\{w_{t-1}^{(i)}\}_{i=1}^{N}$ that summarize, via Monte Carlo, the full joint posterior of states (l_{t-1}, M_{t-1}) and parameters θ, i.e. $p(l_{t-1}, M_{t-1}, \theta|D_{1:(t-1)})$. The goal is to arrive at the end of time t with a similar set of particles $\{(l_t, M_t, \theta)^{(i)}\}_{i=1}^{N}$ and weights $\{w_t^{(i)}\}_{i=1}^{N}$ representing the joint posterior distribution

$$p(l_t, M_t, \theta|D_{1:t}). \tag{27}$$

Loosely speaking, a particle filter is a sampling importance resampling (SIR) scheme implemented iteratively over time: since it is not possible to extract the particles directly from the posterior distribution, we draw from another distribution, say $q(l_t, M_t, \theta|D_{1:t})$, commonly refereed to as an importance distribution, and we approximate the target density (27) assigning appropriate weights to each particle. The re-weighting of a particle from the importance distribution gives that particle the "status" of an actual draw from the posterior distribution.[21] If the support of the target $p(\cdot)$ is included in the support of proposal $q(\cdot)$, then for each particle i the appropriate weight is given by

$$w_t^{(i)} = \frac{p(l_t^{(i)}, M_t^{(i)}, \theta^{(i)}|D_{1:t})}{q(l_t^{(i)}, M_t^{(i)}, \theta^{(i)}|D_{1:t})}. \tag{28}$$

The essence of a particle filter ultimately depends on the design of the importance distribution $q(l_t, M_t, \theta|D_{1:t})$. Our choice is tailored on the peculiar aspects related to the set of solutions we analyze through equation (20).

The most important peculiarity is that, conditionally on M_t, the state space form (26) is linear and Gaussian, which implies that, given a set of particles for M_t, both the predictive likelihood and the full conditional distribution of the other latent states are analytically available through the standard Kalman filter recursion. This practice increases the efficiency of our particle filter through analytical integration, as it follows

[21] See, for instance, Cappe et al. (2007) and Lopes and Tsay (2011) (and the references therein) for a review of particle methods for Bayesian inference.

from the Rao-Blackwell theorem (see Lopes et al., 2011, for further details).

The posterior distribution of the parameters can be updated sequentially combining two different methodologies for parameter learning. In particular, it is useful to divide the parameters in two sets: one with the variances and the covariances of the exogenous disturbances, and one with all the other structural parameters. For the variances and covariances we assume the prior distributions are Inverse Gamma or Inverse Wishart. Then, we are able to characterize the posterior distribution analytically (up to a normalizing constant), using sufficient statistics computed as functions of the data and the latent processes of the model. This is the idea of the Particle Learning approach introduced by Carvalho et al. (2010). The posterior distribution of the other parameters is, in general, not available analytically. It can be approximated using mixtures of Gaussian densities, as in the flexible and general particle filtering with parameter learning algorithm proposed by Liu and West (2001).[22]

The use of SMC methods to approximate the posterior distribution of the parameters of a DSGE model is not very common in the literature: Creal (2007); Chen et al. (2010) and Herbst and Schorfheide (2014) are exceptions to the usual practice based on MCMC. Nevertheless, the particle filtering approach is the most suited for our framework, given the peculiarity of the class of solutions we are considering. First of all, a time-varying M_t makes the model non Gaussian, and it is well known that in this context MCMC methods may have serious limitations due to the high time-dependence of the latent variables. In particular, the convergence of the Markov chain generated through MCMC to the posterior distribution can be very slow and difficult to achieve.

Moreover, within the class of solutions we analyze, the case corresponding to $M_t = 0$ implies a completely different reduced form compared to all the other cases: as shown in the previous section this is the minimum state variable solution, characterized by a simpler lag structure. The shape of the likelihood function, conditional on $M_t = 0$, may be substantially different from the one under $M_t \neq 0$, making MCMC-based inference more complicated: the Markov chain explores accurately the parameter space around the mode

[22]This methodology builds on the resample propagation scheme of the Auxiliary Particle Filter proposed by Pitt and Shephard (1999).

of the distribution, but in practice it is less able to approximate the posterior when the latter is not well shaped, or has multiple modes. This is not a mere technicality, because $M_t = 0$ is a very important case: it is the unique stable solution when the Blanchard Kahn conditions are satisfied, and it characterizes the dynamics implied by the model under the case that the literature labels determinacy. In order to deal with this issue, LS first run a test to determine which is the relevant case, and then use MCMC to estimate the model under the specific assumption of determinacy or indeterminacy, exploring only the corresponding subset of the parameter space. Our suggestion is to estimate the model considering all the relevant cases simultaneously using particle filters instead of MCMC. In general, Sequential Monte Carlo methods are more appropriate when the posterior distribution displays irregular patterns. We show the ability of our econometric strategy to deal with this specific problem in the empirical application described in the next section.

Another advantage of particle filters is computational: the use of multi-core processors makes it possible to increase the speed and the accuracy of the estimation through parallel computing. The gains one can achieve are substantial for SMC, while they are limited for MCMC even if parallelization is implemented in an efficient way, as in the prefetching approach described by Strid (2010).[23]

There are two approaches in the SMC literature to estimate the"static" parameters of a model. One uses all the data available in the sample in each iteration of the SMC to approximate a sequence of distributions, starting from a very simple case (i.e. the prior distribution), and ending with the posterior distribution of interest. We follow a second alternative: we construct particle approximations to the posterior distribution augmenting, at each iteration, the sample data we use. In this case each step of the SMC corresponds to an additional observation, as if new data become available sequentially. We prefer this second technique because it gives us the possibility to study how the inference on the unknowns evolves over time. We show, in the empirical application below, how this "learning" perspective unveils additional information on the role of sunspots and temporarily unstable paths in describing the data. Moreover this approach makes it

[23] For a discussion see also Herbst and Schorfheide (2014).

simpler for us to deal with the filtering problem related to the estimation of M_t.

4 Multiplicative Sunspots and Unstable Paths at Work: the Great Inflation and the New Keynesian Model

We apply our new methodology to inflation dynamics through the lens of the following prototypical New Keynesian model:

$$x_t = E_t(x_{t+1}) - \tau(R_t - E_t(\pi_{t+1})) + g_t, \tag{29}$$

$$\pi_t = \beta E_t(\pi_{t+1}) + \kappa(x_t - z_t), \tag{30}$$

$$R_t = \rho_R R_{t-1} + (1-\rho_R)(\psi_1 \pi_t + \psi_2(x_t - z_t)) + \varepsilon_{R,t}, \tag{31}$$

where x is output, π is inflation and R the nominal interest rate. π and R are expressed in deviation from the steady state, and x in deviation from the steady state trend path. The model admits 3 shocks: (i) a demand shock, g, that can be interpreted as a time-varying government spending shock or a preference shock; (ii) a shock to the marginal costs of production, z; (iii) a monetary policy shock, ε_R. The model and the notation are exactly the same as the one in the seminal paper by LS, that is the natural paper to compare the results of our methodology. The first equation is the New Keynesian IS curve (NKIS), that relates the dynamics of the output x_t to the real interest rate, given by the nominal interest rate, R_t, minus expected inflation, $E_t(\pi_{t+1})$. The dynamics of the inflation rate π_t are described by the second equation, the New Keynesian Phillips curve (NKPC). 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 with inertia closes the model. It describes how the central bank conducts the monetary policy, moving the nominal interest rate R_t, in response to the deviations of inflation and output gap from their targets.

As in LS, we also suppose that the shocks in the NKIS and in the NKPC are autocorrelated, that is:

$$g_t = \rho_g g_{t-1} + \varepsilon_{g,t}; \qquad z_t = \rho_z z_{t-1} + \varepsilon_{z,t} \tag{32}$$

and we allow for non-zero correlation, ρ_{gz}, between the two innovations $\varepsilon_{g,t}$ and $\varepsilon_{z,t}$. The standard deviations of the zero-mean innovations $\varepsilon_{g,t}, \varepsilon_{z,t}$ and $\varepsilon_{R,t}$ are denoted σ_g, σ_z and σ_R, respectively.

The parameters of the model are also standard: $\beta \in (0, 1)$ is the households' subjective discount factor, τ is the elasticity of intertemporal substitution in consumption, κ is the slope of the NKPC, that ultimately depends on the degree of nominal price stickiness and the labour supply elasticity, ρ_R is the inertial parameter in the Taylor rule while ψ_1 and ψ_2 measure the response of the nominal interest rate to the inflation and the output targets, respectively.

The model has five variables: three predetermined (R_t, g_t and z_t) and two non predetermined (x_t, π_t). Then, the matrix M_t has dimension two. We also know that among the five eigenvalues of the dynamic system, three of them are inside the unit circle (because ρ_g, ρ_z, and ρ_R are less than one in absolute value), and one is always outside the unit circle (for sensible values of the parameters, see Bullard and Mitra, 2002). The remaining eigenvalue can be inside or outside the unit circle, depending on the following condition (i.e. the Taylor principle):

$$\psi_1 > 1 - \frac{1-\beta}{\kappa}\psi_2. \tag{33}$$

The literature usually imposes the stability criterion to select valid equilibria and thus it distinguishes two possible cases. If (33) holds, the model has two eigenvalues greater than one in absolute value. This is the determinacy case: there is a unique *stable* RE equilibrium, i.e. the forward looking one, because the number of eigenvalues outside the unit circle is equal to the number of non predetermined variables. Otherwise, if (33) does not hold, there will be an infinite number of *stable* RE equilibria and this case is normally labelled indeterminacy.

Note, however, that in both cases, due to the presence of at least one unstable eigenvalue, there is an infinite number of *unstable* RE equilibria that the literature usually does not consider because of the way the stability criterion is imposed. Our framework, instead, imposes stability in the long run, but it admits temporary walks on these unstable paths by allowing for time variation in the way agents are setting their expectations. We can test the validity of our framework in a particular sample comparing the relative performance of the New Keynesian model, under different hypotheses on the set of admissible solutions: standard RE and our time-varying RE framework. Hence, we compare two assumptions: one in which the stability criterion is imposed as normally in the literature, so that the economy needs to be on a stable trajectory in any point in time, and one in which we also consider solutions normally excluded by the same criterion, allowing temporary instability. The aim is to let the data speak about their preferred assumption.

We distinguish two cases.

Model M_S: stable solutions. When the stability criterion is imposed and no time variation is allowed, we exclude unstable solutions. We label this case as model M_S, and the matrix M_t is:

$$
\begin{aligned}
M_t &= \begin{bmatrix} M_{1,t} & 0 \\ 0 & 0 \end{bmatrix} \\
M_{1,t} &= \begin{cases} 0 & \text{if } \psi_1 > 1 - \frac{1-\beta}{\kappa}\psi_2 \\ M_{1,t-1} + \zeta_t & \zeta_t \sim N(0, \sigma_\zeta^2) \quad \text{otherwise.} \end{cases}
\end{aligned}
$$

The south east element in M_t is imposed to be 0 because there is always one explosive eigenvalue. For the first element, $M_{1,t}$, instead, we distinguish the two cases described above. When the Taylor principle is satisfied, the corresponding eigenvalue is outside the unit circle and we need to select the forward looking solution for all t, because the economy needs to be on the unique stable path in each period t. Hence, $M_{1,t} = 0, \forall t$. When the Taylor principle is not satisfied, the eigenvalue is then inside the unit circle and thus there is an infinite number of stable solutions and the stability condition poses no restrictions on $M_{1,t}$. $M_{1,t}$ needs to be a martingale, so we will assume that $M_{1,t}$ follows a random walk driven by a sunspot shock.

Model M_U: temporarily unstable solutions. In this case, we allow for time variation in M_t even if the eigenvalues are outside the unit circle, and the stability criterion is imposed only in the long run. We define the matrix M_t as:

$$M_t = M_{1,t} I$$

where I is the identity matrix of size two. Since we assume that the elements in the main diagonal of M_t are the same, the two models, M_S and M_U, explore two different sets of solutions that intersect only in one point: $M_{1,t} = 0$. This is the unique stable solution under model M_U: if $M_{1,t} \neq 0$ the dynamics are unstable independently of the Taylor principle given that at least one eigenvalue lies always outside the unit circle.

The process $M_{1,t}$ is defined as in (10) and (11), where we assume that the martingale process N_t multiplies the indicator function $\Im_{\{\|y_{t-1}^B\| < \bar{U}\}}$ in order for the transversality condition to hold. Recall that the backward looking solution is always unstable because of the presence of one eigenvalue greater than one in absolute value. Then, there exists a random date $\bar{T}$ in which $\|y_t^B\|$ bacomes greater than $\bar{U}$ and the indicator function will be equal to zero for all $t > \bar{T}$. After this date, the stochastic process $M_{1,t}$ will also be equal to to zero, and the dynamics will coincide with the unique stable solution allowed under M_U. As discussed before, under this hypotesis RE may be violated in the future time period T. However, if $\bar{U}$ is very big, the probability that this can happen in the near future is approximately zero, and in order to allow for temporarily unstable paths we need to assume that it is disregarded by the agents. When we estimate the model under M_U we choose $\bar{U} = 10^{300}$, and this ensures that the indicator function is equal to one for all the draws and the times considered in our sample.

5 Empirical Results

5.1 Data and subsamples

To compare our results with the seminal work by LS, we estimate the New Keynesian model (29) - (31) on the same quarterly postwar data for inflation, output and nominal interest rate used by LS, as available from the AER website. Inflation and interest rates are annualized, and the HP filter is used to get a measure of the output gap.[24]

Figure 1 plots the inflation series. As it is clear, from the mid Sixties until the end of the '70s, the US experienced a period of price instability, also known as Great Inflation. Then, the Volcker disinflation took place and prices came back under control: inflation became low, as did the volatility of prices and of other macroeconomic variables. By contrast to the previous period, these times are known as the Great Moderation. One popular explanation of this change through the lens of the New Keynesian model (e.g., Clarida et al., 2000) ascribes it to the shift in the monetary policy: from a passive (i.e., (33) not satisfied) to an active (i.e., (33) satisfied) monetary policy. As we previously underlined, this interpretation excludes a priori unstable paths, even though inflation reached 15%. Here we want to answer the following question: would the data prefer an explanation of the Great Inflation based on a stable system with sunspot shocks, as in LS, or one based on unstable dynamics? Again we closely follow LS in considering two subsamples: the pre-Volcker period, from 1960:I to 1979:II, and a post-82 period from 1982:IV to 1997:IV.[25]

5.2 Priors

Table 1 collects the prior distributions for the parameters. We chose them in accordance with LS, in the same spirit that we chose the model specification and the data.

Differently from LS, we specify the prior for the variance covariance matrix of the shock $\varepsilon_{g,t}$ and $\varepsilon_{z,t}$ as an Inverse Wishart with scale matrix and degrees of freedom as in Table 1. The Inverse Wishart prior allows us to update the posterior of the parameters using sufficient statistics, as in the Particle Learning approach described above. This is a big advantage in terms of the efficiency of our particle filter. On the other hand, our choice is very similar to the one of LS in terms of mean and variances of the three

[24] As from footnote 9 at p. 202 in LS: (i) output is log real per capita GDP HP detrended over the period 1955:I to 1998:IV; (ii) inflation is annualized percentage change of CPI-U; (iii) Nominal interest rate is the average Federal Funds Rate in percent.

[25] As in LS, we exclude the Volcker disinflation period where monetary policy is characterized by nonborrowed-reserve targeting rather than by an interest rate rule.

parameters involved (σ_g, σ_z and ρ_{gz}).

The standard deviation of our sunspot shock is distributed as an Inverse Gamma with mean equal to 0.1 and standard deviation to 0.05. This value is lower than the one in LS because our sunspot shock enters in a multiplicative way.

Under model M_U we estimate the probability that the economy is on a temporarily unstable path, that is the parameter γ, for which the prior density is a Beta distribution with mean 0.8 and standard deviation 0.15. A mean of 0.8 implies that when an unstable trajectory is selected, this temporary situation is expected to last for five quarters. Since this is a new parameter we try different values both for the mean and for the standard deviation: results are robust, and some are discussed below.

Finally, the process $M_{1,t}$ at $t = 0$ is supposed to be Normally distributed, with mean 0, and standard deviation 0.1, in accordance to the prior of the standard deviation of the sunspot shock.

Table 1: Prior Distributions

Parameter	Density	Mean	Standard Deviation
ψ_1	Gamma	1.1	0.5
ψ_2	Gamma	0.25	0.15
ρ_R	Beta	0.5	0.2
π^*	Gamma	4	2
r^*	Gamma	2	1
κ	Gamma	0.5	0.2
τ^{-1}	Gamma	2	0.5
ρ_g	Beta	0.7	0.1
ρ_z	Beta	0.7	0.1
γ	Beta	0.8	0.15
σ_R	Inverse Gamma	0.31	0.16
σ_ζ	Inverse Gamma	0.1	0.05
Variance Covariance	**Density**	**Scale**	**Degrees of freedom**
Σ_{gz}	Inverse Wishart	$5 \begin{bmatrix} 0.38^2 & 0 \\ 0 & 1 \end{bmatrix}$	8

5.3 Estimation results

Table 2 reports the estimates of the parameters in the two subsamples. For each subsample, Table 2 shows the estimates for both the stable (M_S) and the unstable (M_U)

model and, for comparison, the correspondent estimates in the paper by LS (see Table 3, p. 206 therein).

5.3.1 Great Inflation subsample

The model under stability: M_S. Let us first analyze the results for the model under stability (model M_S) where we impose the stability criterion. Contrary to LS, our methodology allows us not to impose a determinate or an indeterminate equilibrium prior to the estimation, but lets the data choose which one to select during the estimation. Despite this, Table 2 shows that under stability (model M_S) our methodology recovers results very similar to LS. This is particularly true for the crucial policy rule parameters. Figure 2 displays our prior and posterior distributions and the 90% intervals in LS for these parameters. It shows that our estimation method yields posterior distributions, which are very close and statistically indistinguishable from the ones in LS.[26] This is reassuring as we interpret this finding as corroborating our estimation methodology.

Hence, in accordance with the literature, our method also points to indeterminacy as the most plausible explanation of the Great Inflation period once the stability criterion is imposed on the model. It suggests that the Fed did not respect the Taylor principle, and thus movements in inflation (and output) were due to shifts in expectations due to sunspot shocks. The estimated standard deviation of the sunspot shock for M_S is lower (one third) than the one estimated by LS. However, note that our sunspot is a multiplicative sunspot shock that interacts and amplifies the structural shocks, rather than an additive one as in LS's approach. Hence, these standard deviations are not really comparable due to the different assumption about how the sunspot affects the model.

Figure 3 displays the transmission mechanism of the structural shocks, by showing the generalized impulse response functions (GIRFs) and the 90% intervals of R, x and π to the structural shocks: to the monetary policy shock in the first row, to the demand shock in the second row and to the supply shock in the third row.[27] These GIRFs are

[26]The 90 percent intervals do not overlap only for the slope of the Phillips Curve, κ, and of the elasticity of intertemporal substitution, τ^{-1}.

[27]The GIRFs show the impulse responses to one standard deviation of each shock, and are computed conditioning on the distribution of $M_{1,t}$ at the end of the first subsample, that is the second quarter

Table 2: Posterior Estimates

	Pre- Volcker 1960:I - 1979:II			Post-82 1982:IV - 1997:IV			Sample: 1960:I - 1997:IV
Parameter	M_S	M_U	LS	M_S	M_U	LS	Stochastic Volatility
ψ_1	0.80 [0.66 0.92]	0.76 [0.61 0.91]	0.77 [0.64 0.91]	2.18 [1.53 3.07]	2.32 [1.44 3.58]	2.19 [1.38 2.99]	1.25 [1.12 1.39]
ψ_2	0.16 [0.11 0.20]	0.20 [0.16 0.34]	0.17 [0.04 0.30]	0.17 [0.06 0.38]	0.23 [0.07 0.66]	0.30 [0.07 0.51]	0.21 [0.11 0.41]
ρ_R	0.68 [0.65 0.71]	0.60 [0.53 0.68]	0.60 [0.42 0.78]	0.86 [0.81 0.90]	0.85 [0.80 0.9]	0.84 [0.79 0.89]	0.75 [0.70 0.80]
π^*	1.90 [1.62 2.25]	1.73 [1.31 2.47]	4.28 [2.21 6.21]	3.28 [2.73 3.82]	3.25 [2.82 3.73]	3.43 [2.84 3.99]	2.88 [2.41 3.41]
r^*	1.41 [1.29 1.58]	1.23 [0.93 1.74]	1.13 [0.63 1.62]	2.81 [2.17 3.59]	3.00 [2.40 3.69]	3.01 [2.21 3.80]	2.11 [1.69 2.59]
κ	0.14 [0.10 0.18]	0.10 [0.07 0.14]	0.77 [0.39 1.12]	0.3 [0.22 0.39]	0.48 [0.30 0.81]	0.58 [0.27 0.89]	0.36 [0.26 0.51]
τ^{-1}	3.41 [2.65 4.51]	3.02 [2.46 3.74]	1.45 [0.85 2.05]	2.56 [1.97 3.37]	1.69 [1.20 2.45]	1.86 [1.04 2.64]	2.07 [1.55 3.41]
ρ_g	0.64 [0.59 0.69]	0.68 [0.63 0.74]	0.68 [0.54 0.81]	0.76 [0.69 0.81]	0.78 [0.71 0.84]	0.83 [0.77 0.89]	0.80 [0.77 0.83]
ρ_z	0.76 [0.72 0.80]	0.75 [0.67 0.81]	0.82 [0.72 0.92]	0.72 [0.59 0.83]	0.73 [0.61 0.82]	0.85 [0.77 0.93]	0.81 [0.75 0.85]
ρ_{gz}	0.26 [0.19 0.37]	0.16 [0.06 0.25]	0.14 [−0.4 0.71]	0.03 [0.00 0.07]	0.04 [0.01 0.08]	0.36 [0.06 0.67]	0.14 [0.08 0.19]
γ	–	0.96 [0.85 0.99]	–	–	0.04 [0.01 0.12]	–	–
σ_R	0.22 [0.2 0.26]	0.19 [0.16 0.22]	0.23 [0.19 0.27]	0.16 [0.13 0.19]	0.16 [0.13 0.2]	0.18 [0.14 0.21]	$\delta_R =$ 0.11 [0.09 0.13]
σ_g	0.35 [0.3 0.4]	0.31 [0.24 0.37]	0.27 [0.17 0.36]	0.20 [0.16 0.25]	0.21 [0.17 0.26]	0.18 [0.14 0.23]	$\delta_g =$ 0.12 [0.010 0.014]
σ_z	1.11 [0.97 1.29]	1.00 [0.85 1.31]	1.13 [0.95 1.30]	0.67 [0.55 0.87]	0.63 [0.53 0.76]	0.64 [0.52 0.76]	$\delta_z =$ 0.02 [0.016 0.026]
σ_ς	0.08 [0.07 0.1]	0.06 [0.05 0.08]	0.20 [0.12 0.27]	–	–	–	–

Note: 90% credibility interval in brackets very similar in shape to the ones of a determinate equilibrium, and to the IRFs in LS under their prior 2. Note that the technology shock is the only one that moves output and inflation in opposite directions, as required to explain the stagflation episode during the last part of the Great Inflation period. This explains why the standard deviation of the technology shocks is much bigger than the other shocks for both M_S and LS.

Recall that the non-linear multiplicative sunspot shock affects the model only in the presence of a structural shock. Hence, to understand how the sunspot shock affects the transmission mechanism of our model, we plot in Figure 4 the GIRFs for two different values of M: the solid line corresponds to the pure forward looking solution ($M_{1,t} = 0$), and it shows the impulse response functions estimated under the stable model before the fourth quarter of 1974; the dashed line refers to $M_{1,t} = 0.49$, that is the value estimated in 1974Q4. The difference between the two lines shows how the impulse response functions change precisely when sunspot shocks start playing a role. The sunspot shock amplifies the effects of the structural shock, hence: (i) it does not qualitatively change the response of the variables; (ii) it acts as a stochastic volatility shifter. We interpret this shock as a shift in the way people form expectations after a structural shock hits the economy.

We think that one of the most interesting aspects of our methodology is the estimated path for $M_{1,t}$ that measures of how much expectations deviate from the standard forward looking RE solution. Recall that when $M_{1,t} = 0$ then expectations are selecting the forward looking solution, otherwise they are selecting a combination of the backward and forward looking ones. Figure 5 shows the estimated path for $M_{1,t}$ in the case of M_S, and the corresponding sequential estimate of the policy parameter ψ_1. Figure 5 clearly depicts the challenge faced by the New Keynesian model in this subsample: to simultaneously explain the stable output and inflation paths in the first part of the subsample and the stagflation in the second part of the subsample, where output and inflation move in opposite directions, and inflation accelerates. Up the first oil shock, the estimate of $M_{1,t}$ points toward expectations aligned on the "standard" forward looking solution and, correspondingly, ψ_1 is estimated to satisfy the Taylor principle. Until to that point the

of 1979. The uncertainty of the GIRFs, summarized by the 90% probability interval, reflects also the posterior distribution of the parameters.

data would favour a determinate stable model. However, such a model has hard times in explaining the data in the second part of the subsample. Then, the data switch to favour the only alternative model available under stability: a model with sunspot shocks. The extra degree of freedom provided by the sunspot makes the data choose the indeterminate model both in LS's and in our estimation. $M_{1,t}$ drifts away from 1, when inflation starts to grow in the data.

Of course, another plausible possibility explored in the literature to make a stable determinate model able to match such behaviour in the data would be to have a stochastic volatility model, where the standard deviation of technology shocks increases in the second part of the subsample (e.g., Justiniano and Primiceri, 2008). We will consider such a model in Section 6. Our multiplicative sunspot shock yields a similar effect, as explained above, but the sunspot shock occurs only if the model is indeterminate under M_S.

The model under instability: M_U. The model M_U makes the data consider also *temporarily unstable* paths. The point estimate in Table 2 are very similar between the two cases M_S and M_U. However, it should by now be clear to the reader that this does not imply indeterminacy as usually intended in the literature, that is, an infinite number of *stable* RE trajectories. It does imply another sort of indeterminacy, in the sense that we let the data choose among an infinite number of *unstable, but temporary* trajectories, irrespective of whether the Taylor principle is satisfied or not. (33) is a condition for one eigenvalue to be inside or outside the unit circle, but whatever the value of ψ_1, there is always an unstable eigenvalue. However, we do not force the model to the forward looking solution with respect to this unstable eigenvalue in the $M_{U,}$ case, as explained in Section 4. It follows that, despite the parameter estimates being very similar between the M_S and M_U cases, M_U gives a completely different interpretation about the instability of that period. Independently from the Fed policy, the dynamics of M_U are structurally unstable.

Figure 6 shows the GIRFs in this case. Again a supply shock generates stagflation. Most importantly, however, stagflation could also now be generated by a monetary policy shock. In particular, a contractionary monetary policy shock can be inflationary: inflation drops on impact but then starts rising and it is above steady state from the fourth quarter onward. Interestingly, a somewhat similar behaviour is highlighted in LS under their preferred prior 1: *"an increase in the nominal interest rate can have a slightly inflationary effect"* (p. 207, see Figure 3, p. 208 and the discussion at p. 207-208 therein). They conclude that *"before 1979 indeterminacy substantially altered the propagation of shocks"* (LS, abstract).[28] Similarly, instability in our framework substantially alters the transmission mechanism. However, in our case, output remains below steady state, so that a monetary policy shock could generate an opposite response of output and inflation. In LS case, instead, inflation and output move in the same direction after a monetary policy shock: after dropping on impact, they *both* become slightly positive. The same consideration applies also to the demand shock: both output and inflation increase on impact, but inflation turns negative in the fourth quarter. Our framework therefore seems to be able to provide a transmission mechanism more prone to accommodate stagflation under instability.

The transmission mechanism of the sunspot shock is also quite different in our case. In LS, the impulse response function to a sunspot shock under indeterminacy does imply (again) that output and inflation move in the same direction (see Figure 2, p. 207 in LS), not in an opposite one, as in a stagflation episode. Intuitively, if a sunspot shock leads to a self-fulfilling increase in inflation, then the real interest rate decreases, due to the passive monetary policy, and thus output increases, rather than decreases. Thus the structural dynamics implied by an indeterminate stable model do not seem to be well suited to explain stagflation episodes after an additive sunspot shock. In our setup, instead, the non-linear multiplicative sunspot shock amplifies the responses of the model to a structural shock. Similarly to Figure 4, Figure 7 shows the GIRFs for two different values of M in the M_U case: the solid line corresponds to the pure forward looking solution ($M_{1,t} = 0$), and it shows the impulse response functions estimated under the stable model before the fourth quarter of 1974; the dashed line refers to $M_{1,t} = 0.52$, that is the value estimated in 1974Q4. As in the M_S case, the sunspot shock amplifies the GIRfs, but the implied dynamics is very different in the M_U case. The amplification is similar on impact

[28] *"This finding suggests that the fit of the model can be improved by deviating from the baseline solution and altering the propagation of the structural shocks."* (LS, p. 205)

between the two cases, but then the unstable root induce an explosive dynamics such that initially the distance between the two lines increases over time. The economy is travelling on an explosive trajectory and it diverges away from the stable forward-looking solution. However, the walk on the unstable trajectory is temporary and the GIRFs exhibits a boom-bust type of behaviour: given the assumed process for $M_{1,t}$, at a certain stochastic date the economy converges back to the unique stable forward-looking solution.

As before, it is instructive to look at the estimated path for $M_{1,t}$ that measures how much expectations deviate from the standard forward looking RE solution. Recall that we let the data choose: the data could still choose a stable forward looking solution where $M_{1,t}$ is estimated to equal zero. Figure 8 shows the estimated path for the latent process $M_{1,t}$ under model M_U. Similarly as before, it initially fluctuates around zero and then it drifts from zero (the 90 percent interval exhibits a mass above zero), exactly when inflation starts increasing and drifting away from its steady state value from 3% to 15%. If we allow for temporarily unstable paths, the estimation then unambiguously selects those to explain the data in this period.

It is possible to compare the relative fit of the stable (M_S) and unstable (M_U) models by computing the Sequential Bayes factor as in West (1986). The Bayes factor is the model likelihood ratio:

$$H_t = \frac{p(y_t|y_{0:t-1}, M_S)}{p(y_t|y_{0:t-1}, M_U)} \tag{34}$$

and measures the relative success of M_S and M_U in predicting the data: values lower than one of H_t indicate worse predictive performance of M_S than the alternative M_U. West (1986) suggests to compute the Bayes factor sequentially as $W_t(k) = H_t H_{t-1} ... H_{t-k+1}$. $W_t(k)$ is called the *cumulative* Bayes factor and it assesses the relative fit of the two models by considering the most recent k observations. Figure 9 shows twice the natural logarithm of the cumulative Bayes factor $W_t(k)$ (as suggested by Kass and Raftery, 1995) together with the path of inflation in the sample for a 10 year window (i.e., k is equal to 40). Therefore, a value of zero of the logarithm of the cumulative Bayes factor means that the two models have the same performance in terms of predictive likelihood; while a positive value means that M_S is preferred (and vice versa for negative values). The advantage of the cumulative Bayes factor, with respect to the conventional measures in Bayesian Econometrics, is that we can compare two models over time, and verify the sub-periods in which a model performs better than another in terms of predictive likelihood. In our specific case, as expected, the unstable model is much preferred from the '70s onwards, when inflation starts drifting away reaching high values. According to the Kass and Raftery (1995, p.777) classification, there is "very strong" evidence in favour of M_U from the beginning of the '70s. In particular, the cumulative Bayes factor reaches a very low level from the first quarter of 1974 onwards.

To conclude, our methodology allows the data to choose between different possible alternatives: determinacy, indeterminacy and temporary instability. When the data are allowed this possibility, they unambiguously select the unstable model to explain the stagflation period in the '70s.

5.3.2 Post-82 Subsample

In the second subsample, our estimates under stability again reproduce the same results as in LS (see Table 2). There is no statistically significant difference between our parameter estimates and the ones in LS, again signalling the reliability of our estimation methodology (see Figure 10). The Taylor principle is satisfied and hence the data choose the unique determinate forward looking solution under M_S: there is no sunspot shock and the process for $M_{1,t}$ degenerates to the value of zero.

Also in the case of model M_U, the estimation yields results similar to LS. However, recall that in our framework, despite the Taylor principle being satisfied, the data could still choose a temporarily unstable path, as well as the unique standard stable manifold under determinacy (or MSV solution), when $M_{1,t}$ is equal to zero. The upper panel in Figure 11 shows that $M_{1,t}$ is indeed estimated to be equal zero for the whole period, meaning that the estimation chooses the standard MSV solution under determinacy. Coherently, the estimation returns a negligible probability of the economy travelling on a temporarily unstable path, as evident from the sequential estimate of the parameter γ in the lower panel in Figure 11. From (11), γ represents the probability of $M_{1,t}$ being different from zero, that is, the probability that the economy is on an unstable trajectory. In the Great Moderation sample, the final point estimate of γ is extremely low (0.05 in Table 2), despite the prior was set to 0.8, as in the Great Inflation sample (where the posterior point estimate is 0.96). The data, thus, are in this case extremely informative, and strongly points towards the stable MSV solution, which is the same as the one imposed by the standard RE methods.

Comparing the two models as in the previous case using the cumulative Bayes factor presents mild evidence in favour of the (less parameterized) stable model (see Figure 12). The evidence is not strong though, but "weak" till 1992 and then "positive", because the two models delivers very similar estimates.

5.3.3 Prior on γ

In terms of point estimates, our results are very robust to changing the priors of our parameters . A prior of 0.9 for γ would deliver very similar results, but the cumulative Bayes factor would speak in favour of our benchmark choice. A tighter prior on γ (i.e., standard error prior equals to 0.05 rather than 0.15) improves the fit of the model in the Great Inflation sample, because the sequential estimate of γ is very stable around 0.8. For the same reason, however, the estimation performs worse in the Great Moderation period. In particular, it is not able to recover the standard rational expectation MSV solution for that sub-sample, because the tighter prior does not allow the particles to sufficiently explore that region of the parameter space, and the sequential estimate of γ fluctuates quite tightly around 0.8. Hence, it is very important in our approach to allow for a sufficiently wide prior over the parameter γ to give the estimation a chance to explore adequately all the different regions of the parameter space corresponding to the there cases of determinacy, indeterminacy and temporarily explosive paths.

6 A Comparison with a Stochastic Volatility Model

A large empirical literature shows how stochastic volatility is an important feature of US macroeconomic variables in the sample we analyze. Cogley and Sargent (2005); Primiceri (2005); Justiniano and Primiceri (2008) find evidence in favor of high volatility in the Seventies and a subsequent decrease during the Great Moderation. The methodology we present rationalizes this evidence through the hypothesis of time variation in the agents expectations formation process. In particular, the estimates of $M_{1,t}$ are such that the effects of structural shocks during the last part of the first subsample are amplified (see Figure 4 and 7).

On the other hand our framework imposes a strong link between the "walks on unstable trajectories" and stochastic volatility: under the unstable model M_U, stochastic volatility always occurs in the presence of temporarily unstable paths, and it is absent only when the unique stable solution is selected. This restriction may be too tight, and it might be the case that our M_U model is favourite by the data because of the implied stochastic volatility rather than for the temporary intrinsic unstable dynamics. In other words, a model with stochastic volatility without unstable dynamics might be sufficient to adequately interpret the data.

To investigate this issue we compare the fit of the unstable model M_U with a stochastic volatility model under determinacy during the Great Inflation. Thus, we estimate a model with time-varying variances of each error under determinacy. The results show that while modelling the heteroskedasticity of shocks in a flexible way leads to some improvements, temporarily unstable paths remain a key feature to interpret the behavior of inflation, GDP and interest rate during the '70s. We therefore conclude that stochastic volatility alone, without explosive dynamics, is not able to fully capture the unstable behavior of the data during the Great Inflation period.

In specifying the assumptions on the model with determinacy and stochastic volatility we follow closely Justiniano and Primiceri (2008). In particular we suppose that the logarithm of the standard error of each shock is described by a random walk process:

$$\log \sigma_{i,t} = \log \sigma_{i,t-1} + \nu_{i,t} \tag{35}$$

where $\nu_{i,t} \sim N(0, \delta_i^2)$ and $i = g, z, R$. The model is estimated under determinacy, then $M_{1,t} = 0, \forall t$ and we only explore the region of the parameters that satisfy the Taylor principle. Moreover we estimate the model considering the entire sample from 1960:I to 1997:IV.

Inference on the parameters and on the time-varying volatilities is performed using the same econometric strategy we presented above. Note that conditional on the values of the volatilities, the model is linear and Gaussian. Then, we simply proceed in analogy with the estimation of models M_S and M_U, and we treat the time variation in the variances in the same way we did with the time variation in $M_{1,t}$ (see Appendix A.3.5 for details).

The prior distributions on the parameters are the same as in Table 1, with the exception that now we only allow for determinacy. In practice, this is simply done by setting the particle weight equal to zero whenever the parameters are such that the Taylor principle is not satisfied. For the variances of the shocks to the volatilities, we assume an Inverse Gamma distribution with mean equal to 0.02 and 3 degrees of freedom.[29] Finally, we assume that the standard deviations at time zero have the same prior distribution as in the time invariant case, reported in Table 1.

The last column of Table 2 displays the posterior distribution of the parameters, and Figure 13 shows the estimated pattern of the time-varying standard deviations of the different shocks. With respect to Justiniano and Primiceri (2008) we work with a smaller model and a shorter sample period. Nevertheless, we find very similar results: first, the model accounts for the reduction in the volatility of the US macroeconomic variables during the Great Moderation with a substantial decrease in the volatility of exogenous disturbances. Second, the degree of stochastic volatility is not the same for all the shocks,

[29] Justiniano and Primiceri (2008) set the prior mean equal to 0.01, a half of what we assume. In our model we find that this specification restricts too much the time variation in the standard deviations, penalizing the model with determinacy and stochastic volatility. Under our prior, instead, we find results that are very similar to Justiniano and Primiceri (2008), as described below.

and in particular the disturbance with the biggest variation in the standard deviation is the one to monetary policy. Moreover, the latter is the unique shock directly comparable, and the pattern of stochastic volatility is remarkably similar to the one in Justiniano and Primiceri (2008). Finally, for the other two shocks we find a decline of roughly one third in the last part of the sample, again in line with the results in Justiniano and Primiceri (2008).

In Table 3 we compare the overall fit of this model during the Great Inflation period, with both the stable model M_S and the unstable model M_U. The model with determinacy and stochastic volatility is favored by the Bayes factor when compared with the stable model M_S. In M_S the variations in the variances are all related to one common component, that is $M_{1,t}$, while the standard deviation of the monetary policy shock behaves differently with respect to the other two, when more flexibility is allowed. This finding does not necessarily imply that the restrictions imposed by our method are in general too tight. The size of the model we consider is small, allowing for only one element in the matrix M_t to be time-varying (i.e., only indeterminacy of order one), when stability is imposed.

Table 3: Model Comparison with Determinacy and Stochastic Volatility

Sample: 1960:I-1979:II	
Alternative Model	$2\log(\text{Bayes Factor})$
M_S	−7.2611
M_U	16.2346

A positive value means evidence in favor of the alternative model

In estimating the model under instability we chose to limit ourselves to the case of only one degree of freedom, setting the elements in the main diagonal of the matrix M_t to the same stochastic process. Then, also the unstable model M_U penalizes the variability of the variances in the same way as model M_S. Despite this limit, Table 3 shows that the Bayes factor clearly favours the unstable model: the evidence for model M_U is labelled as "very strong" in Kass and Raftery's (1995) classification.

This result suggests that temporarily unstable paths are a key feature to describe the unstable pattern of the US macroeconomic variables during the Great Inflation period. In our setting the heteroskedasticity of shocks emerges as a consequence of our assumption about the expectations formation process. However, the unstable nature of the dynamics remains crucial to interpret the data during the ‘70s.

7 Conclusions

We propose a novel framework to consider a broader class of solutions to stochastic linear RE models.

Theoretically, we provide two main generalizations: our framework generates time-varying parameter solutions and stochastic volatility, as well as it allows for the possibility of the economy walking on temporarily unstable paths. First, we show how all the possible RE solutions could be parameterized by one single parameter that has a natural interpretation as the way agents weight past data to form their RE. Then, we introduce multiplicative sunspots by assuming that this parameter follows a stochastic process, so that agents randomly select one of the possible RE fundamental solutions. Under indeterminacy, there is an infinite number of admissible stable solutions. Under determinacy, instead, only one value of this parameter is coherent with the economy converging to the equilibrium in the long run. Appropriate restrictions on the stochastic process driving this parameter allows temporary walks on unstable trajectories and stability on the long run.

Empirically, we propose an econometric methodology that allows the data to choose among the different RE alternatives: determinacy, indeterminacy and temporary instability, without imposing them a priori in the estimation. This methodology can be used to test the empirical relevance of temporarily unstable dynamics.

Finally, we apply this approach to the data to explain US inflation dynamics in the Great Inflation and Great Moderation period. The empirical evidence suggests that the Great Inflation in the US can be explained by temporarily unstable paths, while the usual practice of excluding a priori unstable solutions seems not to be supported by the data. When allowed, the data unambiguously select the unstable model to explain the stagflation period in the ‘70s. Our framework provides a different interpretation of the Great Inflation from a policy perspective. Despite our estimates point to a passive monetary policy behaviour in the ‘70s, our framework implies that this is not the cause in itself of unstable inflation dynamics, that was instead due to drifting expectations, independently from the stance of monetary policy.

Our analysis therefore suggests that unstable paths can be empirically relevant. This result may call for a rethinking of the stability criterion as the selection mechanism among all the possible RE paths, and for theoretically considering the possibility that RE could push the economy to walk along unstable paths, at least temporarily.

This line of research is still in its infancy and can be expanded in many directions. A first important direction would be to endogenize the process for the multiplicative sunspot. The process for the drifting expectations is taken as exogenous in this paper (as in the sunspot literature) and then estimated on the data. We would like to be able to say something about why agents RE start to drift, by endogenizing this expectation formation process and then estimating it on the data, in a spirit similar to the escape dynamics literature. Moreover, the estimation indicates a link between unstable paths and the monetary policy parameter, that is, between the wandering of M away from the value of the stable solution and the feature of policy. This is reminiscent of the debate about monetary policy and the anchoring of inflation expectations, because M determines the way agents combine past data to form their expectations.

Extending the framework to non-linear models and non-linear solution methods is a second direction for future research. The linear approximation of a model could become unreliable if the system drifts too far away from the steady state by following a temporarily unstable path. The extension should be feasible because we have available methods to solve non-linear models and the econometric strategy does not depend on the model being linear . An important application, then, would be to use a model with the zero lower bound (see e.g., Gust et al., 2017), to investigate how the zero lower bound affects the process of expectation formation and hence the stability of the economy, in this case in the direction of deflation, rather than inflation as in the ‘70s.

Third, one could modify the framework to allow a subset of variables to explode. Following the insights in Cochrane (2011), for example, nominal variables do not need to satisfy a transversality condition. More generally, the framework could be generalized characterizing the relationship between different possible stochastic processes for M and the implied deviation from RE.

Finally, there are many potential application of our framework, notably, but not exclusively, finance, where boom and bust episodes of asset prices (stock, houses, etc..) is a pervasive phenomenon.

References

Benati, Luca and Paolo Surico, "VAR Analysis and the Great Moderation," *American Economic Review*, 2009, *99(4)*, 1636–1652.

Benhabib, Jess and Roger E. A. Farmer, "Indeterminacy and sunspots in macroeconomics," in John B. Taylor and Michael Woodford, eds., *Handbook of Macroeconomics*, Vol. 1, Elsevier Science, North Holland, 1999, chapter 6, pp. 387–448.

Blanchard, Olivier J., "Backward and Forward Solutions for Economies with Rational Expectations," *American Economic Review*, 1979, *69*, 114–118.

_ **and Mark W. Watson**, "Bubbles, Rational Expectations and Financial Markets," NBER Working Papers 0945, National Bureau of Economic Research, Inc July 1982.

Blanchard, Olivier Jean and Charles M. Kahn, "The Solution of Linear Difference Models under Rational Expectations," *Econometrica*, 1980, *48*, 1305–1311.

Boivin, J. and M. Giannoni, "Has Monetary Policy Become More Effective?," *Review of Economics and Statistics*, 2006, *88(3)*, 445–462.

Brock, William A., "Money and Growth: The Case of Long Run Perfect Foresight," *International Economic Review*, 1974, *15*, 750–777.

Bullard, James and Kaushik Mitra, "Learning about monetary policy rules," *Journal of Monetary Economics*, 2002, *49(6)*, 1105–1129.

Burmeister, Edwin, Robert P. Flood, and Peter M. Garber, "On the Equivalence of Solutions in Rational Expectations Models," *Journal of Economic Dynamics and Control*, 1983, *5*, 311–21.

Cappe, Olivier, Simon J. Godsill, and Eric Moulines, "An Overview of Existing Methods and Recent Advances in Sequential Monte Carlo," *Proceedings of the IEEE*, May 2007, *95* (5), 899–924.

Carboni, Giacomo and Martin Ellison, "The Great Inflation and the Greenbook," *Journal of Monetary Economics*, 2009, *56(6)*, 831–841.

Carvalho, Carlos M., Michael S. Johannes, Hedibert F. Lopes, and Nicholas G. Polson, "Particle Learning and Smoothing," *Statistical Science*, 02 2010, *25* (1), 88–106.

Carvalho, Carlos, Stefano Eusepi, Emanuel Moench, and Bruce Preston, "Anchored Inflation Expectations," 2017, *Available at SSRN: https://ssrn.com/abstract=3018198 or http://dx.doi.org/10.2139/ssrn.3018198.*

Castelnuovo, Efrem and Luca Fanelli, "Monetary Policy Indeterminacy and Identification Failures in the U.S.: Results from a Robust Test," *Journal of Applied Econometrics*, 2015, *30*, 924–47.

_ , Luciano Greco, and Davide Raggi, "Policy Rules, Regime Switches, and Trend Inflation: An Empirical Investigation for the U.S.," *Macroeconomic Dynamics*, 2014, *18*, 920–42.

Chen, Hao, Francesca Petralia, and Hedibert F. Lopes, "Sequential Monte Carlo Estimation of DSGE Models," Working Papers June 2010.

Cho, In-Koo, Noah Williams, and Thomas J. Sargent, "Escaping Nash Inflation," *The Review of Economic Studies*, 2002, *69*, 1–40.

Clarida, Richard, Jordi Galí, and Mark Gertler, "Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory," *Quarterly Journal of Economics*, 2000, *115*, 147–180.

Cochrane, John H., "Determinacy and identification with Taylor rules," *Journal of Political Economy*, 2011, *119*, 565–615.

Cogley, Timothy and Thomas J. Sargent, "Drifts and Volatilities: Monetary Policies and Outcomes in the Post War U.S.," *Review of Economic Dynamics*, 2005, *8*, 262–302.

Creal, Drew, "Sequential Monte Carlo samplers for Bayesian DSGE models," Working Papers, Vrije Universitiet Amsterdam August 2007.

Fernández-Villaverde, Jesús and Juan Rubio-Ramírez, "Estimating Macroeconomic Models: A Likelihood Approach," *Review of Economic Studies*, 2007, *74*, 1059–1087.

_ , P.A. Guerrón-Quintana, and Juan F. Rubio-Ramírez, "Fortune or Virtue: Time-Invariant Volatilities Versus Parameter Drifting in U.S. Data," 2010, *NBER Working Paper No. 15928.*

Flood, Robert P. and Peter M. Garber, "Market Fundamentals versus Price Level Bubbles: The First Tests," *Journal of Political Economy*, 1980, *88*, 745–70.

Gourieroux, Christian, Jean-Jacques Laffont, and Alain Monfort, "Rational Expectations in Dynamic Linear Models: Analysis of the Solutions," *Econometrica*, 1982, *50*, 409–425.

Gust, Christopher, Edward Herbst, David López-Salido, and Matthew E. Smith, "The Empirical Implications of the Interest-Rate Lower Bound," *American Economic Review*, July 2017, *107* (7), 1971–2006.

Herbst, Edward and Frank Schorfheide, "Sequential Monte Carlo Sampling For Dsge Models," *Journal of Applied Econometrics*, November 2014, *29* (7), 1073–1098.

Justiniano, Alejandro and Giorgio E. Primiceri, "The Time-Varying Volatility of Macroeconomic Fluctuations," *American Economic Review*, 2008, *98*, 604–641.

Kass, Robert E. and Adrian E. Raftery, "Bayes Factors," *Journal of the American Statistical Association*, 1995, *90*, 773–795.

Liu, Jane and Mike West, "Combined parameters and state estimation in simulation-based filtering," in Springer, ed., *Sequential Monte Carlo Methods in Practice*, New York: Arnaud Doucet and Nando de Freitas and Neil Gordon, 2001, chapter 10, pp. 197–223.

Lopes, Hedibert F. and Ruey S. Tsay, "Particle filters and Bayesian inference in financial econometrics," *Journal of Forecasting*, 2011, *30* (1), 168–209.

_ **, Carlos M. Carvalho, Michael S. Johannes, and Nicholas G. Polson**, "Particle learning for sequential Bayesian computation (with discussion)," *Bayesian Statistics*, 2011, *9*, 317–360.

Lubik, Thomas A. and Christian Matthes, "Indeterminacy and learning: An analysis of monetary policy in the Great Inflation," *Journal of Monetary Economics*, 2016, *82* (C), 85–106.

_ **and Frank Schorfheide**, "Testing for Indeterminacy: An Application to U.S. Monetary Policy," *American Economic Review*, 2004, *94(1)*, 190–217.

Mavroeidis, Sophocles, "Monetary Policy Rules and Macroeconomic Stability: Some New Evidence," *American Economic Review*, 2010, *100(1)*, 491–503.

Muth, John F., "Optimal Properties of Exponentially Weighted Forecasts," *Journal of the American Statistical Association*, 1960, *55*, 299–306.

_ , "Rational Expectations and the Theory of Price Movements," *Econometrica*, 1961, *29*, 315–335.

Pesaran, M. Hashem, *The Limits to Rational Expectations*, Oxford, UK: Basil Blackwell, 1987.

Phelps, Edmund S. and John B. Taylor, "Stabilizing Powers of Monetary Policy under Rational Expectations," *Journal of Political Economy*, 1977, *85*, 163–190.

Pitt, Michael K. and Neil Shephard, "Filtering via Simulation: Auxiliary Particle Filters," *Journal of the American Statistical Association*, 1999, *94* (446), 590–599.

Primiceri, Giorgio E., "Time-Varying Structural Vector Autoregressions and Monetary Policy," *Review of Economic Studies*, 2005, *72*, 821–852.

Sargent, Thomas J., *The Conquest of American Inflation*, Princeton, NJ: Princeton University Press, 1999.

_ **and Neil Wallace**, "The Stability of Money and Growth with Perfect Foresight," *Econometrica*, 1973, *41*, 1043–48.

_ **, Noah Williams, and Tao Zha**, "Shocks and Government Beliefs: The Rise and Fall of American Inflation," *American Economic Review*, 2006, *96(4)*, 1193–1224.

Sims, Christopher A., "Solving Linear Rational Expectations Models," *Computational Economics*, 2002, *20(1-2)*, 1–20.

Strid, Ingvar, "Efficient parallelisation of Metropolis-Hastings algorithms using a prefetching approach," *Computational Statistics & Data Analysis*, November 2010, *54* (11), 2814–2835.

Taylor, John B., "On the Conditions for Unique Solutions in Stochastic Macroeconomic Models with Rational Expectations," *Econometrica*, 1977, *45*, 1377–85.

West, Kenneth D., "A Specification Test for Speculative Bubbles*," *The Quarterly Journal of Economics*, 1987, *102* (3), 553–580.

West, Mike, "Bayesian Model Monitoring," *Journal of the Royal Statistical Society. Series B (Methodological)*, 1986, *48*, 70–78.

Woodford, Michael, *Interest and Prices: Foundations of a Theory of Monetary Policy*, Princeton University Press. Princeton, New Jersey, 2003.

A Appendix

A.1 Implementation: The general solution

As in LS, we follow the approach of Sims (2002) and we write a general linear RE system as:

$$\Gamma_0 y_t = \Gamma_1 y_{t-1} + \Psi \varepsilon_t + \Pi \eta_t, \tag{A1}$$

where y_t is the vector of the n endogenous variables (including the expectations as in (3)), ε_t is the vector of the h exogenous fundamental shocks, and η_t is the vector of the $k \leq n$ RE forecast errors. For simplicity, we assume that Γ_0 is invertible,[30] so to write as in (14) :

$$y_t = \Gamma_1^* y_{t-1} + \Psi^* \varepsilon_t + \Pi^* \eta_t. \tag{A2}$$

The multivariate case is a relative straightforward extension of the simple case, so the description follows similar steps as above, involving: (i) parameterizing the system using M (now a matrix); (ii) introducing time variation in M, (iii) imposing stability. As usual, however, first we need to decouple the system through a variable transformation.

Partitioning. As in the main text, use Jordan decomposition to partition the system, and define the vector of transformed variables $\tilde{y}_t = J^{-1} y_t$. Let the ith element of $\tilde{y}_t$ be $\tilde{y}_{it}$, the ith element on the principal diagonal of Λ be λ_i and denote the ith row of $J^{-1}\Pi^*$ and $J^{-1}\Psi^*$ by $[J^{-1}\Pi^*]_i$ and $[J^{-1}\Psi^*]_i$, respectively. The model can then be written as a collection of AR(1) processes as in the univariate case: $\tilde{y}_{it} = \lambda_i \tilde{y}_{it-1} + [J^{-1}\Psi^*]_i \varepsilon_t + [J^{-1}\Pi^*]_i \eta_t$. Order the eigenvalues (and the corresponding eigenvectors) in descending order, and partition the system in two blocks, of dimensions $(n-k)$ and k, respectively. As explained in the main text, we depart from Sims (2002) and LS because we partition the system as in (15), according to the number of forward-looking variables/expectation errors, rather than the number of explosive eigenvalues. Let m be the number of explosive eigenvalues (i.e., such that $\lambda_i \geq 1$). As usual, we assume that the number of explosive eigenvalues is smaller or equal to the number of forecast errors, to rule out instability. Hence, the first $(n-k)$ rows contain only stable eigenvalues, while the last k rows contain both $(k-m)$ stable and m unstable eigenvalues. Hence, we do not need to impose any stability condition on the first block of the system (15), but we do need on the second block of equation, i.e., (16).

Parameterization. Note that the system is decoupled, so it is just a collection of independent AR(1) processes. Each row in (16) corresponds to our simple example above (3). As for the case of the simple model, it is possible to parameterize the fundamental solutions, i.e., where the expectation error is just a function of the structural shock, by modifying the stability condition under determinacy. In matrix notation, the usual stability condition under determinacy would be $J_{\mu 2}[\Psi^* \varepsilon_t + \Pi^* \eta_t] = 0$, and, as in the simple case, we modify it to $(I+M)J_{\mu 2}\Psi^* \varepsilon_t = -J_{\mu 2}\Pi^* \eta_t$, when we restrict the matrix M to be diagonal, with M_i being the ith element on the principal diagonal of M. Hence:

$$\tilde{y}_{k,t} = \Lambda_2 \tilde{y}_{k,t-1} + J_{\mu 2}\Psi^* \varepsilon_t - (I+M)J_{\mu 2}\Psi^* \varepsilon_t = \Lambda_2 \tilde{y}_{k,t-1} - M J_{\mu 2}\Psi^* \varepsilon_t. \tag{A3}$$

[30] This is the case in the LS's model, that we will use in our empirical analysis below. If Γ_0 is singular, it is trivial to generalize the method to use the Schur decomposition (QZ).

Iterate (A3) backward to find:

$$\tilde{y}_{k,t} = -M \sum_{i=0}^{t-1} \Lambda_2^i \left(J_{\mu 2} \Psi^*\right) \varepsilon_{t-i}. \tag{A4}$$

This expression corresponds to (4), and as (4), it exists assuming that we start from steady state (it exists a time 0, such that $\tilde{y}_{-k,i} = \varepsilon_{-i} = \eta_{-i} = 0, \forall i \geq 0$). Moreover, some of the solutions for $\tilde{y}_{i,k,t}$ in (A4) will be stable and some will be unstable, depending on the values of the M_i's and on the stability properties of the system, i.e., depending on the values of the $\lambda_{2,i}$'s, where $\lambda_{2,i}$ is the ith element on the principal diagonal of Λ_2.

Time variation. Assume now that the M_i elements on the principal diagonal of the matrix M are changing over time following indipendently distibuted and uncorrelated stochastic processes. Our proposed solution then is:

$$\tilde{y}_{k,t} = -M_t \sum_{i=0}^{t-1} \Lambda_2^i \left(J_{\mu 2} \Psi^*\right) \varepsilon_{t-i}, \tag{A5}$$

which corresponds to (6). Note that in each period t, the solution just depends on the current realization of M_t. A solution pins down the expectations errors, actually $J_{\mu 2}\Pi^*\eta_t$. As in Sims (2002), a solution pins down the expectations errors, actually $J_{\mu 2}\Pi^*\eta_t$. Plugging (A5) in (16) yields:

$$J_{\mu 2}\Pi^*\eta_t = -(I + M_t) J_{\mu 2}\Psi^*\varepsilon_t - (M_t - M_{t-1}) \sum_{i=1}^{t-1} \Lambda_2^i \left(J_{\mu 2} \Psi^*\right) \varepsilon_{t-i}. \tag{A6}$$

The RE condition implies $E_{t-1}\left(J_{\mu 2}\Pi^*\eta_t\right) = 0$, so that each $M_{i,t}$ must be: 1) a martingale; 2) uncorrelated with ε_t. Again, it is easy to recognise two particular solutions: 1) the forward-looking solution, given by $M_t = 0 => \tilde{y}^F_{k,t} = 0 => \eta_t = -\left(J_{\mu 2}\Pi^*\right)^{-1} J_{\mu 2}\Psi^*\varepsilon_t, \forall t$; 2) the backward-looking solution, given by $M_t = -I => \tilde{y}^B_{k,t} = \sum_{i=0}^{t-1} \Lambda_2^i \left(J_{\mu 2}\Psi^*\right) \varepsilon_{t-i}$ and $\eta_t = 0, \forall t$. The forward-looking solution always exists and it is always (under our assumption) a stable solution: it is the only stable one under determinacy ($m = k$), while it is one out of many possible stable ones under indeterminacy ($m < k$). However, in this latter case, the forward-looking solution is a special one given how we partition the system: it coincides with the minimum state variable solution, because it delivers a solution which is just a linear function of the state variables.

Then the solution to the system of disconnected difference equations (A3) can be written recursively almost as in Blanchard (1979), but actually using only the backward-looking variable $\tilde{y}^B_{k,t}$ as:

$$\tilde{y}_{k,t} = -M_t \sum_{i=0}^{t-1} \Lambda_2^i \left(J_{\mu 2} \Psi^*\right) \varepsilon_{t-i} = -M_t \tilde{y}^B_{k,t} \tag{A7}$$

so that:

$$\tilde{y}^B_{k,t} = \Lambda_2 \tilde{y}^B_{k,t-1} + J_{\mu 2}\Psi^*\varepsilon_t \tag{A8}$$

$$\tilde{y}_{k,t} = -M_t \tilde{y}^B_{k,t} = -M_t \left(\Lambda_2 \tilde{y}^B_{k,t-1} + J_{\mu 2}\Psi^*\varepsilon_t\right) \tag{A9}$$

which are (17) and (18) in the main text.

Note that since: $\tilde{y}_{k,t}^{B}=\sum_{i=0}^{t-1}\Lambda_2^i\left(J_{\mu2}\Psi^*\right)\varepsilon_{t-i}=J^{-1}\Psi^*\varepsilon_t+\sum_{i=1}^{t-1}\Lambda_2^i\left(J_{\mu2}\Psi^*\right)\varepsilon_{t-i}$, the expectation error could be written as:

$$\begin{aligned}
J_{\mu2}\Pi^*\eta_t &= -(I+M_t)J_{\mu2}\Psi^*\varepsilon_t-(M_t-M_{t-1})\sum_{i=1}^{t-1}\Lambda_2^i\left(J_{\mu2}\Psi^*\right)\varepsilon_{t-i}\\
&= -(I+M_t)J_{\mu2}\Psi^*\varepsilon_t-(M_t-M_{t-1})\left(\tilde{y}_{k,t}^{B}-J_{\mu2}\Psi^*\varepsilon_t\right)\\
&= -(I+M_{t-1})J_{\mu2}\Psi^*\varepsilon_t-(M_t-M_{t-1})\,\tilde{y}_{k,t}^{B}
\end{aligned}$$

which yields (19), assuming that the (kxk) matrix $J_{\mu2}\Pi^*$ is invertible.

We discuss stability in the main text. Again as in the simple model, and we impose stability by allowing only particular processes for $M_{i,t}$'s.

Recompose the system and solve for original variables. Having solved for the forward-looking variables, we now need to recompose the system from the original partition. First, we need to substitute for $J_{\mu1}\left[\Psi^*\varepsilon_t+\Pi^*\eta_t\right]$ into (15), given the η_t implied by our proposed solution from (19). Substitute (A9) in the system (15) adding the auxiliary variable $\tilde{y}_{k,t}^{B}$:

$$\begin{bmatrix}\tilde{y}_{(n-k),t}\\ ((n-k)\times 1)\\ \tilde{y}_{k,t}\\ (k\times 1)\\ \tilde{y}_{k,t}^{B}\\ (k\times 1)\end{bmatrix}=\begin{bmatrix}\Lambda_1 & \mathbf{0} & 0\\ ((n-k)\times(n-k)) & ((n-k)\times k) & ((n-k)\times k)\\ \mathbf{0} & 0 & -M_t\Lambda_2\\ (k\times(n-k)) & (k\times k) & (k\times k)\\ 0 & 0 & \Lambda_2\\ (k\times(n-k)) & (k\times k) & (k\times k)\end{bmatrix}\begin{bmatrix}\tilde{y}_{(n-k),t-1}\\ ((n-k)\times 1)\\ \tilde{y}_{k,t-1}\\ (k\times 1)\\ \tilde{y}_{k,t-1}^{B}\\ (k\times 1)\end{bmatrix}+$$
$$+\begin{bmatrix}J_{\mu1}\left[\Psi^*\varepsilon_t+\Pi^*\eta_t\right]\\ ((n-k)\times 1)\\ -M_tJ_{\mu2}\Psi^*\varepsilon_t\\ (k\times 1)\\ J_{\mu2}\Psi^*\varepsilon_t\\ (k\times 1)\end{bmatrix}$$

Then the problem is pin down $J_{\mu1}\Pi^*\eta_t$, but we know η_t, given our proposed solution from (19), so:

$$\begin{aligned}
& J_{\mu1}\left[\Psi^*\varepsilon_t+\Pi^*\eta_t\right]\\
= \;& J_{\mu1}\left[\Psi^*\varepsilon_t+\Pi^*\left(J_{\mu2}\Pi^*\right)^{-1}\left[-(I+M_{t-1})J_{\mu2}\Psi^*\varepsilon_t-(M_t-M_{t-1})\,\tilde{y}_{k,t}^{B}\right]\right]\\
= \;& J_{\mu1}\left[\Psi^*-\Pi^*\left(J_{\mu2}\Pi^*\right)^{-1}(I+M_{t-1})J_{\mu2}\Psi^*\right]\varepsilon_t-J_{\mu1}\Pi^*\left(J_{\mu2}\Pi^*\right)^{-1}(M_t-M_{t-1})\,\tilde{y}_{k,t}^{B}
\end{aligned}$$

Then given (A8), we can write:

$$\begin{aligned}
& J_{\mu1}\left[\Psi^*\varepsilon_t+\Pi^*\eta_t\right]\\
= \;& J_{\mu1}\left[\Psi^*-\Pi^*\left(J_{\mu2}\Pi^*\right)^{-1}(I+M_{t-1})J_{\mu2}\Psi^*\right]\varepsilon_t+\\
- \;& J_{\mu1}\Pi^*\left(J_{\mu2}\Pi^*\right)^{-1}(M_t-M_{t-1})\left(\Lambda_2\tilde{y}_{k,t-1}^{B}+J_{\mu2}\Psi^*\varepsilon_t\right)\\
= \;& J_{\mu1}\left[\Psi^*-\Pi^*\left(J_{\mu2}\Pi^*\right)^{-1}(I-M_t)J_{\mu2}\Psi^*\right]\varepsilon_t-J_{\mu1}\Pi^*\left(J_{\mu2}\Pi^*\right)^{-1}(M_t-M_{t-1})\,\Lambda_2\tilde{y}_{k,t-1}^{B}
\end{aligned}$$

So we can write:

$$J_{\mu1}\left[\Psi^*\varepsilon_t+\Pi^*\eta_t\right]=A_t\varepsilon_t-B_{t,t-1}\tilde{y}_{k,t-1}^{B},\tag{A10}$$

where A_t is the $(n-k) \times l$ matrix and $B_{t,t-1}$ is a $(n-k) \times k$ matrix, respectively given by (22) and (23) in the main text, that is:

$$\begin{aligned} A_t &= J_{\mu 1}\left[\Psi^* - \Pi^* \left(J_{\mu 2}\Pi^*\right)^{-1}(I+M_t)J_{\mu 2}\Psi^*\right]; && \text{(A11)} \\ B_{t,t-1} &= J_{\mu 1}\Pi^* \left(J_{\mu 2}\Pi^*\right)^{-1}\left(M_t - M_{t-1}\right)\Lambda_2 && \text{(A12)} \end{aligned}$$

The final system therefore is:

$$\begin{aligned} \tilde{y}_{(n-k),t} &= \Lambda_1 \tilde{y}_{(n-k),t-1} - B_{t,t-1}\tilde{y}^B_{k,t-1} + A_t\varepsilon_t \\ \tilde{y}_{k,t} &= -M_t\Lambda_2\tilde{y}^B_{k,t-1} - M_t J_{\mu 2}\Psi^*\varepsilon_t \\ \tilde{y}^B_{(n-k),t} &= \Lambda_1 \tilde{y}^B_{(n-k),t-1} + J_{\mu 1}\Psi^*\varepsilon_t \\ \tilde{y}^B_{k,t} &= \Lambda_2 \tilde{y}^B_{k,t-1} + J_{\mu 2}\Psi^*\varepsilon_t, \end{aligned}$$

which in matrix notation is:

$$\begin{bmatrix} \tilde{y}_{(n-k),t} \\ \tilde{y}_{k,t} \\ \tilde{y}^B_{(n-k),t} \\ \tilde{y}^B_{k,t} \end{bmatrix} = \underbrace{\begin{bmatrix} \Lambda_1 & 0 & 0 & -B_{t,t-1} \\ 0 & 0 & 0 & -M_t\Lambda_2 \\ 0 & 0 & \Lambda_1 & 0 \\ 0 & 0 & 0 & \Lambda_2 \end{bmatrix}}_{G^*} \begin{bmatrix} \tilde{y}_{(n-k),t-1} \\ \tilde{y}_{k,t-1} \\ \tilde{y}^B_{(n-k),t-1} \\ \tilde{y}^B_{k,t-1} \end{bmatrix} + \underbrace{\begin{bmatrix} A_t \\ -M_t J_{\mu 2}\Psi^* \\ J_{\mu 1}\Psi^* \\ J_{\mu 2}\Psi^* \end{bmatrix}}_{H^*} \varepsilon_t. \quad \text{(A13)}$$

Finally, to recover the original variables use $= \tilde{y}_t = J^{-1}y_t$ to obtain (20) in the main text.

A.2 The econometric strategy

Inference regarding the structural parameters of the model, collected in the vector θ, as well as the latent states is fully Bayesian. The time-varying characteristic of the latent state M_t leads to a non linear and analytically intractable non Gaussian likelihood function for the unknowns. This motivates the use of the sequential Monte Carlo strategy described below.

A.2.1 Preliminaries

The class of solution we propose in equation (20), parametrized by the matrix M_t, has state space representation (26) that we repeat below for convenience:

$$\begin{cases} D_t = c + Fl_t + v_t & v_t \sim N\left(\mathbf{0}, \Sigma_v\right) \\ l_t = G_t l_{t-1} + H_t\varepsilon_t & \varepsilon_t \sim N\left(\mathbf{0}, \Sigma_\varepsilon\right) \end{cases} \quad \text{(A14)}$$

D_t is the vector with data at time t, and $D_{m:n}$ the set of observations from m to n for $m \leq n$. The parameters of the model are collected in the vector $\theta = (\theta_1, \theta_2)$, where we group in θ_1 all the parameters other than the variances and the covariances of the shocks, wich are in turn collected in the vector θ_2. Finally, we assume that the dynamics of M_t are described by a transition law:

$$M_t = f\left(M_{t-1}, \zeta_t\right) \quad \text{(A15)}$$

where ζ_t is a multiplicative sunspot shock. The properties of the stochastic process for M_t are discussed in the paper.

Our econometric strategy is based on sequential learning: suppose the posterior distribution of the unknowns is approximated at time $t-1$ by a set of particles $\{(l_{t-1}, M_{t-1}, \theta_1, \theta_2)^{(i)}\}_{i=1}^N$ and associated weights $\{w_{t-1}^{(i)}\}_{i=1}^N$. Given the new observed data D_t, we want to generate an updated set of particles $\{(l_t, M_t, \theta_1, \theta_2)^{(i)}\}_{i=1}^N$ and weights $\{w_t^{(i)}\}_{i=1}^N$ that approximate the posterior distribution:

$$p\left(l_t, M_t, \theta_1, \theta_2 | D_{1:t}\right). \tag{A16}$$

The way we group the latent processes (distinguishing M_t from all the other states l_t) and the parameters (dividing them in θ_1 and θ_2) has a specific reason: as a general principle of our econometric strategy, we implement analytical computation whenever is possible. To this aim, note that given a value for M_t, the state space (A14) is linear and Gaussian: we can compute the posterior distribution of the latent processes in l_t analytically, using the Kalman filter. Moreover, an analytical expression for the posterior distribution can also be derived for some of the parameters, that we collect in θ_2. For DSGE models this is typically the case of the variances and covariances of the shocks, when the prior distributions are Inverse Gamma or Inverse Wishart. Then, following Carvalho et al. (2010), we keep track of a set of sufficient statistics collected in s_t that we will use to update the posterior distribution of θ_2.

To approximate the posterior distribution of the parameters in θ_1 we use the Liu and West (2001) filter. Since the method uses mixtures of Normal distributions we make sure that all the parameters have the right support, that is from $-\infty$ to $+\infty$. Then, we define a new vector ϕ where each element of θ_1 is appropriately transformed when needed. In the description of the algorithm we will add a time t subscript to this parameter, writing ϕ_t. This notation is introduced simply to reinforce the notion that sequential inference regarding ϕ is performed at time t, and does not mean that the parameters are time-varying.

A.2.2 The particle filter

The algorithm we use is based on two main steps: an *updating* step, in which an appropriate number of particles N is drawn from an importance distribution $q\left(\vartheta_t, M_t, \theta_1, \theta_2 | D_{1:t}\right)$, and a *re-weighting* step in which the weights are computed as:

$$w_t^{(i)} = \frac{p\left(l_t^{(i)}, M_t^{(i)}, \theta_1^{(i)}, \theta_2^{(i)} | D_{1:t}\right)}{q\left(l_t^{(i)}, M_t^{(i)}, \theta_1^{(i)}, \theta_2^{(i)} | D_{1:t}\right)}. \tag{A17}$$

Step 1: Drawing from the importance distribution

Drawing from the importance distribution involves two sub-steps, following the schema in Pitt and Shephard (1999): a resampling step in which we select "the most fit particles", and the actual propagation step in which these particles are updated.

Resampling. Once new data are arrived we start selecting the particles with higher predictive ability. We perform a resampling step using weights $\tilde{w}_t^{(i)}$ proportional to:

$$\tilde{w}_t^{(i)} \propto w_{t-1}^{(i)} p\left(D_t | l_{t-1}^{(i)}, g_M(M_{t-1}^{(i)}), m_{t-1}^{(i)}, \theta_2^{(i)}\right) \tag{A18}$$

Following Pitt and Shephard (1999) and Liu and West (2001), the predictive likelihood in equation (A18) is conditional on $g_M(M_{t-1}^{(i)})$, that is a *best guess* of M_t^i at time $t-1$ like

$E\left(M_t|M_{t-1}^{(i)}\right)$, and on m_{t-1}^{i} defined as:[31]

$$m_{t-1}^{(i)} = a\phi_{t-1}^{(i)} + (1-a)\bar{\phi}_{t-1} \tag{A19}$$

where $\bar{\phi}_{t-1}$ is the weighted sample mean of $\phi_{t-1}^{(i)}$. Define also V_{t-1} as the sample weighted covariance matrix of $\phi_{t-1}^{(i)}$, that we will use later. Then, given the state space (A14), the predictive likelihood is a Normal distribution with mean $\hat{f}_t^{(i)}$ and variance $\hat{Q}_t^{(i)}$ where:

$$\hat{f}_t^{(i)} = \hat{c}^{(i)} + \hat{F}\hat{G}_{t-1}^{(i)} l_{t-1}^{(i)} \tag{A20}$$

$$\hat{Q}_t^{(i)} = \hat{F}\left(\hat{G}_{t-1}^{(i)} C_{t-1}^{(i)} \hat{G}_{t-1}^{(i)\prime} + \hat{H}_{t-1}^{(i)} \Sigma_{\varepsilon}^{(i)} \hat{H}_{t-1}^{(i)\prime}\right)\hat{F}' \tag{A21}$$

and $C_{t-1}^{(i)}$ is the variance of the latent process $l_{t-1}^{(i)}$. Note that the matrices $\hat{F}$, $\hat{G}_{t-1}^{(i)}$ and $\hat{H}_{t-1}^{(i)}$ and the vector $\hat{c}^{(i)}$ are function of $g_M(M_{t-1}^{(i)})$ and of the parameters in $m_{t-1}^{(i)}$.

At this point we have a set of resampled particles that, for convenience, we accentuate with a tilde: $\{(\tilde{l}_{t-1}, \tilde{M}_{t-1}, \tilde{m}_{t-1}, \tilde{\theta}_2, \tilde{s}_{t-1}, \tilde{C}_{t-1}, \tilde{f}_t, \tilde{Q}_t)^{(i)}\}_{i=1}^{N}$.

Propagation. The resampled particles are then propagated starting from the set of parameters $\phi_t^{(i)}$. Following the schema of Liu and West (2001) we update this vector drawing its new values from the normal distribution:

$$\phi_t^{(i)} \sim N\left(\tilde{m}_{t-1}^{(i)}, (1-a^2)V_{t-1}\right). \tag{A22}$$

Then, we proceed with the propagation of $M_{1,t}^{(i)}$ from the distribution implied by it's low of motion (A15):

$$M_t^{(i)} \sim p\left(M_t|\tilde{M}_{t-1}^{(i)}, \phi_t^{(i)}, \tilde{\theta}_2^{(i)}\right) \tag{A23}$$

Given $M_t^{(i)}$ the state space (A14) becomes linear and Gaussian. We can draw $l_t^{(i)}$ from its posterior distribution:

$$l_t^{(i)} \sim p\left(l_t|\tilde{l}_{t-1}^{(i)}, M_t^{(i)}, \phi_t^{(i)}, \tilde{\theta}_2^{(i)}, D_t\right) \tag{A24}$$

that is a Normal distribution with mean $\mu_t^{(i)}$ and variante $C_t^{(i)}$ computed through the Kalman filter recursion:

$$f_t^{(i)} = c^{(i)} + FG_t^{(i)}\tilde{l}_{t-1}^{(i)} \tag{A25}$$

$$Q_t^{(i)} = F\left(G_t^{(i)}\tilde{C}_{t-1}^{(i)}G_t^{(i)\prime} + H_t^{(i)}\tilde{\Sigma}_{\varepsilon}^{(i)}H_t^{(i)\prime}\right)F' \tag{A26}$$

$$\mu_t^{(i)} = G_t^{(i)}\tilde{l}_{t-1}^{(i)} + \left(G_t^{(i)}\tilde{C}_{t-1}^{(i)}G_t^{(i)\prime} + H_t^{(i)}\tilde{\Sigma}_{\varepsilon}^{(i)}H_t^{(i)\prime}\right)F'\left(Q_t^{(i)}\right)^{-1}\left(D_t - f_t^{(i)}\right) \tag{A27}$$

$$\begin{aligned} C_t^{(i)} = &\left(G_t^{(i)}\tilde{C}_{t-1}^{(i)}G_t^{(i)\prime} + H_t^{(i)}\tilde{\Sigma}_{\varepsilon}^{(i)}H_t^{(i)\prime}\right) + \\ &- \left(G_t^{(i)}\tilde{C}_{t-1}^{(i)}G_t^{(i)\prime} + H_t^{(i)}\tilde{\Sigma}_{\varepsilon}^{(i)}H_t^{(i)\prime}\right)F'\left(Q_t^{(i)}\right)^{-1}F\left(G_t^{(i)}\tilde{C}_{t-1}^{(i)}G_t^{(i)\prime} + H_t^{(i)}\tilde{\Sigma}_{\varepsilon}^{(i)}H_t^{(i)\prime}\right) \end{aligned} \tag{A28}$$

[31] The parameter a in equation (A19), that accounts for the amount of shrinkage, is suggested to be set between 0.974 and 0.995 (see Liu and West, 2001, for details)

Note that the matrices F, $G_t^{(i)}$ and $H_t^{(i)}$, and the vector $c^{(i)}$ are function of $M_t^{(i)}$ and of the updated parameters $\phi_t^{(i)}$. Then, the mean and the covariance matrix of the predictive distribution, respectively $f_t^{(i)}$ and $Q_t^{(i)}$, are different from those defined in (A20) and (A21).

Finally, we propagate the vector $\theta_2^{(i)}$ following the Particle Learning approach of Carvalho et al. (2010). The latent processes $l_t^{(i)}$ and $M_t^{(i)}$ and the parameters $\phi_t^{(i)}$ are used to update the set of sufficient statistics $s_t^{(i)}$.[32] Hence, we can draw $\theta_2^{(i)}$ from its posterior distribution:

$$\theta_2^{(i)} \sim p\left(\theta_2 | s_t^{(i)}\right) \tag{A29}$$

We have drawn a new set of particles from the importance distribution obtained combining equations (A18), (A22), (A23), (A24) and (A29).

Step 2: Re-weighting the particles

In order to approximate the target density we need to compute the appropriate weight for each particle, according to equation (A17).

Start from the joint posterior distribution (A16) which is proportional to:

$$p\left(l_t, M_t, \theta | D_{1:t}\right) \propto p\left(D_t | l_t, M_t, \theta\right) p\left(l_t, M_t, \theta | D_{1:(t-1)}\right), \tag{A30}$$

where the second term on the right hand side is written as

$$\begin{aligned} & p\left(l_t, M_t, \theta | D_{1:(t-1)}\right) = \\ = & \int p\left(l_t, M_t, \theta | l_{1:(t-1)}, M_{1:(t-1)}\right) p\left(l_{1:(t-1)}, M_{1:(t-1)} | D_{1:(t-1)}\right) dl_{1:(t-1)} dM_{1:(t-1)} \\ \approx & \sum_{i=1}^{N} w_{t-1}^{(i)} p\left(l_t, M_t, \theta | l_{1:(t-1)}^{(i)}, M_{1:(t-1)}^{(i)}\right). \end{aligned} \tag{A31}$$

Consequently, the posterior is approximated by

$$p\left(l_t, M_t, \theta | D_{1:t}\right) \propto \sum_{i=1}^{N} w_{t-1}^{(i)} p\left(D_t | l_t, M_t, \theta\right) p\left(l_t, M_t, \theta | l_{1:(t-1)}^{(i)}, M_{1:(t-1)}^{(i)}\right). \tag{A32}$$

Assuming that the latent processes are Markov chains, we can write the numerator in equation (A17) as:

$$p\left(l_t^{(i)}, M_t^{(i)}, \theta^{(i)} | D_{1:t}\right) = w_{t-1}^{(i)} p\left(D_t | l_t^{(i)}, M_t^{(i)}, \theta^{(i)}\right) p\left(l_t^{(i)}, M_t^{(i)}, \theta^{(i)} | l_{t-1}^{(i)}, M_{t-1}^{(i)}\right). \tag{A33}$$

Following Carvalho et al. (2010) we compute the weights before propagating the parameters in θ_2. Taking this into account and combining equations (A18), (A22), (A23), (A24), (A29) and (A33) in equation (A17) we get:[33]

$$w_t^{(i)} \propto \frac{p\left(D_t | l_t^{(i)}, M_t^{(i)}, \theta_1^{(i)}, \tilde{\theta}_2^{(i)}\right) p\left(l_t^{(i)} | \tilde{l}_{t-1}^{(i)}, M_t^{(i)}, \theta_1^{(i)}, \tilde{\theta}_2^{(i)}\right)}{p\left(D_t | \tilde{l}_{t-1}^{(i)}, g_M(\tilde{M}_{t-1}^{(i)}), \tilde{m}_{t-1}^{(i)}, \tilde{\theta}_2^{(i)}\right) p\left(l_t^{(i)} | \tilde{l}_{t-1}^{(i)}, M_t^{(i)}, \theta_1^{(i)}, \tilde{\theta}_2^{(i)}, D_t\right)} \tag{A34}$$

[32] For example, if the variance of a shock is a priori distributed as an Inverse Gamma, to compute the conjugate posterior we need the sum of the squared errors.

[33] The weights are expressed as "proportional to" instead of "equal to" because they need to be normalized such that their sum is equal to one.

Note that the density $p\left(l_t^{(i)}|\tilde{l}_{t-1}^{(i)}, M_t^{(i)}, \theta_1^{(i)}, \tilde{\theta}_2^{(i)}, D_t\right)$ in the denominator can be rewritten as

$$p\left(l_t^{(i)}|\tilde{l}_{t-1}^{(i)}, M_t^{(i)}, \theta_1^{(i)}, \tilde{\theta}_2^{(i)}, D_t\right) = \frac{p\left(D_t|l_t^{(i)}, M_t^{(i)}, \theta_1^{(i)}, \tilde{\theta}_2^{(i)}\right) p\left(l_t^{(i)}|\tilde{l}_{t-1}^{(i)}, M_t^{(i)}, \theta_1^{(i)}, \tilde{\theta}_2^{(i)}\right)}{p\left(D_t|\tilde{l}_{t-1}^{(i)}, M_t^{(i)}, \theta_1^{(i)}, \tilde{\theta}_2^{(i)}\right)} \tag{A35}$$

Substituting this equation in (A34) we find that the weights to approximate the joint posterior distribution at time t are:

$$w_t^{(i)} \propto \frac{p\left(D_t|\tilde{l}_{t-1}^{(i)}, M_t^{(i)}, \theta_1^{(i)}, \tilde{\theta}_2^{(i)}\right)}{p\left(D_t|\tilde{l}_{t-1}^{(i)}, g_M(\tilde{M}_{t-1}^{(i)}), \tilde{m}_{t-1}^{(i)}, \tilde{\theta}_2^{(i)}\right)}. \tag{A36}$$

At the numerator we have the Normal distribution with mean $f_t^{(i)}$ and covariance matrix $Q_t^{(i)}$ defined in equations (A25) and (A26). The distribution at the denominator is the Normal with mean $\tilde{f}_t^{(i)}$ and covariance matrix $\tilde{Q}_t^{(i)}$ defined in (A20) and (A21), and resampled according to weights $\tilde{w}_t^{(i)}$ computed in (A18). Both densities are evaluated in D_t.

Equation (A36) is very intuitive: the weight of each particle is computed comparing two predictive likelihoods. The particle i has higher weight if, after propagation of $M_t^{(i)}$ and $\theta_1^{(i)}$, leads to higher improvement in predicting D_t.

Step 3 (optional): Resampling

The approximation of the posterior distribution obtained in the two steps described above is good if the the particle weights in (A36) are distributed Uniformly. It is well known in the literature that the variance of the distribution of the weights tends to increase over time since a subset of particles will have higher predictive power. Then, an additional resampling step using the weights computed in (A36) can be added to mitigate this problem. After a resampling step is performed, all the weights are set equal to $1/N$.

Usually the final resampling step is implemented when a certain criterion suggests that the distribution of weights became too uneven. A common practice is to check the *effective sample size* defined as:

$$N_t^e = \left(\sum_{i=1}^{N} \left(w_t^{(i)}\right)^2\right)^{-1}. \tag{A37}$$

N_t^e takes values from 1 (very uneven distribution) to N (Uniform distribution), so the resampling step is performed when N_t^e is less then a certain threshold $\bar{N}$.

The procedure to implement our particle filter is summarized in the algorithm below.

THE ALGORITHM

Initialization: t=0

Draw a set of particles $\{(l_0, M_0, \theta_1, \theta_2, s_0, C_0)^{(i)}\}_{i=1}^N$ from a prior

Recursion: for $t = 1, 2, ...T$ repeat steps 1 to 6

1. Approximate $p(\phi|y_{1:(t-1)})$

   1a) Consider a transformation of the vector θ_1 and call it ϕ_t

   1b) Compute the weighted sample mean $\bar{\phi}_{t-1}$ and covariance matrix V_{t-1}

   1c) Compute $m_{t-1}^{(i)} = a\phi_{t-1}^{(i)} + (1-a)\bar{\phi}_{t-1}$

2. Resample

   2a) Compute $\tilde{w}_t^{(i)} \propto w_{t-1}^{(i)} p\left(D_t | l_{t-1}^{(i)}, g_M(M_{t-1}^{(i)}), m_{t-1}^{(i)}, \theta_2^{(i)}\right)$

   2b) Resample $\{(l_{t-1}, M_{t-1}, m_{t-1}, \theta_2, s_t, C_t)^{(i)}\}_{i=1}^N$ with weights $\tilde{w}_t^{(i)}$

   Let the new particles be $\{(\tilde{l}_{t-1}, \tilde{M}_{t-1}, \tilde{m}_{t-1}, \tilde{\theta}_2, \tilde{s}_{t-1}, \tilde{C}_{t-1})^{(i)}\}_{i=1}^N$.

3. Propagate

   3a) Sample $\phi_t^{(i)}$ from $N\left(\tilde{m}_{t-1}^{(i)}, (1-a^2)V_{t-1}\right)$

   3b) Sample $M_t^{(i)}$ from $p\left(M_t | \tilde{M}_{t-1}^{(i)}, \phi_t^{(i)}, \tilde{\theta}_2^{(i)}\right)$

   3c) Sample $l_t^{(i)}$ from $N\left(\mu_t^{(i)}, C_t^{(i)}\right)$

   where $\mu_t^{(i)}$ and $C_t^{(i)}$ are defined in (A27) and (A28).

4. Compute new weights

$$w_t^{(i)} \propto \frac{p(D_t | \tilde{l}_{t-1}^{(i)}, M_t^{(i)}, \theta_1^{(i)}, \tilde{\theta}_2^{(i)})}{p(D_t | \tilde{l}_{t-1}^{(i)}, g_M(\tilde{M}_{t-1}^{(i)}), \tilde{m}_{t-1}^{(i)}, \tilde{\theta}_2^{(i)})}.$$

5. Update sufficient statistics and propagate θ_2

   5a) Compute $s_t^{(i)} = \mathcal{S}\left(l_t^{(i)}, \tilde{l}_{t-1}^{(i)}, M_t^{(i)}, \tilde{M}_t^{(i)}, \phi_t^{(i)}, D_t\right)$

   5b) Sample $\theta_2^{(i)}$ from $p\left(\theta_2 | s_t^{(i)}\right)$

6. Decide to resample or not

   if $\bar{N} < \left(\sum_{i=1}^N \left(w_t^{(i)}\right)^2\right)^{-1}$

   6a) Resample with weights $w_t^{(i)}$

   6b) Re-set weights $w_t^{(i)} = \frac{1}{N}$

A.3 Estimating the New Keynesian model

We show how to apply our estimation strategy to estimate the model of LS described in Section 4.

A.3.1 The model and its state space representation

The model consists of equations (29), (30), (31) and (32).

In order to write the model in the Sims (2002) canonical form (A1) define $\eta_t^x = x_t - E_{t-1}(x_t)$, $\eta_t^\pi = \pi_t - E_{t-1}(\pi_t)$, $\xi_t^x = E_t(x_{t+1})$ and $\xi_t^\pi = E_t(\pi_{t+1})$. Then the NK model can be expressed as:

$$\eta_t^x + \xi_{t-1}^x = \xi_t^x - \tau(R_t - \xi_t^\pi) + g_t \tag{A38}$$

$$\eta_t^\pi + \xi_{t-1}^\pi = \beta\xi_t^\pi + \kappa(\eta_t^x + \xi_{t-1}^x - z_t) \tag{A39}$$

$$R_t = \rho_R R_{t-1} + (1-\rho_R)(\psi_1\left(\eta_t^\pi + \xi_{t-1}^\pi\right) + \psi_2(\eta_t^x + \xi_{t-1}^x - z_t)) + \varepsilon_{R,t} \tag{A40}$$

Defining the vector $y_t = [\ x_t \quad \pi_t \quad R_t \quad \xi_t^x \quad \xi_t^\pi \quad g_t \quad z_t\ ]'$, the system in matrix form is:

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & (1-\rho_R)\psi_2 \\ 0 & 0 & -\tau & 1 & \tau & 1 & 0 \\ 0 & 0 & 0 & 0 & \beta & 0 & -\kappa \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_t \\ \pi_t \\ R_t \\ \xi_t^x \\ \xi_t^\pi \\ g_t \\ z_t \end{bmatrix} =$$

$$= \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & \rho_R & (1-\rho_R)\psi_2 & (1-\rho_R)\psi_1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\kappa & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \rho_g & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \rho_z \end{bmatrix} \begin{bmatrix} x_{t-1} \\ \pi_{t-1} \\ R_{t-1} \\ \xi_{t-1}^x \\ \xi_{t-1}^\pi \\ g_{t-1} \\ z_{t-1} \end{bmatrix} +$$

$$+ \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_{R,t} \\ \varepsilon_{g,t} \\ \varepsilon_{z,t} \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ (1-\rho_R)\psi_2 & (1-\rho_R)\psi_1 \\ 1 & 0 \\ -\kappa & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \eta_t^x \\ \eta_t^\pi \end{bmatrix}$$

The class of solution we propose, parametrized by the matrix M_t, is written in equation (20) and it is expressed in terms of the vector $\tilde{l}_t = \begin{bmatrix} y_t \\ y_t^B \end{bmatrix}$, where y_t^B describes the evolution of the variables in the backward looking solution. Note that in the vector $\tilde{l}_t$ the exogenous state variables g_t and z_t appear twice, since their dynamics are independent of M_t. For practical purposes it is convenient to rewrite the solution in terms of a vector l_t where each exogenous shock is reported only once. First, define the following vectors:

$$y_{1,t} = [\ x_t \quad \pi_t \quad R_t \quad \xi_t^x \quad \xi_t^\pi\ ]'; \qquad y_{2,t} = [g_t \quad z_t\ ]'.$$

The solution can be partitioned as

$$\begin{bmatrix} y_{1,t} \\ y_{2,t} \\ y_{1,t}^B \\ y_{2,t} \end{bmatrix} = \begin{bmatrix} \tilde{G}_{1,t} & \tilde{G}_{2,t} & \tilde{G}_{3,t} & \tilde{G}_{4,t} \\ \mathbf{0} & \tilde{G}_{5,t} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \tilde{G}_{6,t} & \tilde{G}_{7,t} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \tilde{G}_{5,t} \end{bmatrix} \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \\ y_{1,t-1}^B \\ y_{2,t-1} \end{bmatrix} + \begin{bmatrix} \tilde{H}_{1,t} \\ \tilde{H}_{2,t} \\ \tilde{H}_{3,t} \\ \tilde{H}_{2,t} \end{bmatrix} \varepsilon_t$$

where $\varepsilon_t = [\ \varepsilon_{R,t} \quad \varepsilon\prime_{g,t} \quad \varepsilon_{z,t}\ ]$. The endogenous variables in $y_{1,t}$ depend on the entire vector $\tilde{l}_{t-1}$, while the same variables in the backward looking solution depend only on the backward looking components of $\tilde{l}_{t-1}$. The exogenous variables, instead, are described by their own dynamics. It is straightforward, then, to rewrite the solution as:

$$\begin{bmatrix} y_{1,t} \\ y_{2,t} \\ y_{1,t}^B \end{bmatrix} = \begin{bmatrix} \tilde{G}_{1,t} & \left(\tilde{G}_{2,t} + \tilde{G}_{4,t}\right) & \tilde{G}_{3,t} \\ \mathbf{0} & \tilde{G}_{5,t} & \mathbf{0} \\ \mathbf{0} & \tilde{G}_{7,t} & \tilde{G}_{6,t} \end{bmatrix} \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \\ y_{1,t-1}^B \end{bmatrix} + \begin{bmatrix} \tilde{H}_{1,t} \\ \tilde{H}_{2,t} \\ \tilde{H}_{3,t} \end{bmatrix} \varepsilon_t$$

that is, using a compact notation:

$$l_t = G_t l_{t-1} + H_t \varepsilon_t \tag{A41}$$

that is the state equation of system (A14), where the latent vector is:

$$l_t = [\ x_t \quad \pi_t \quad R_t \quad \xi_t^x \quad \xi_t^\pi \quad g_t \quad z_t \quad x_t^B \quad \pi_t^B \quad R_t^B \quad \xi_t^{xB} \quad \xi_t^{\pi B}\ ]'.$$

The observation equation is:

$$D_t = c + F l_t \tag{A42}$$

where D_t is a column vector with output gap, inflation and interest rate,

$$c = \begin{bmatrix} 0 \\ \pi^* \\ \pi^* + r^* \end{bmatrix}, \quad \text{and} \quad F = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

A.3.2 The parameters updated through the Liu and West filter

The set of parameters has two components: $\theta = (\theta_1, \theta_2)$, where θ_1 contains all the parameters of the model except the variances:[34]

$$\theta_1 = [\ \rho_g \quad \rho_z \quad \rho_R \quad \kappa \quad \psi_1 \quad \psi_2 \quad \tau^{-1} \quad \pi^* \quad r^* \quad \gamma\ ]'.$$

Define the vector ϕ as a transformation of the vector θ_1 such that every element has support from $-\infty$ to $+\infty$. In particular we use the logit function for the parameters that

[34] The parameter γ is estimated only under the unstable model, and it is not included in the vector θ_1 under the stable model M_S.

can take values in $[-1\ 1]$, and the logarithm for the parameters with positive support:

$$\phi^{(i)} = \begin{bmatrix} h(\rho_g^{(i)}) \\ h(\rho_z^{(i)}) \\ h(\rho_R^{(i)}) \\ \log(\kappa^{(i)}) \\ \log(\psi_1^{(i)}) \\ \log(\psi_2^{(i)}) \\ \log(\tau^{-1(i)}) \\ \log(\pi^{*(i)}) \\ \log(r^{*(i)}) \\ h(\gamma^{(i)}) \end{bmatrix}$$

where h is the logit function.

Finally, the parameter a in equation (A19) is set equal to 0.99.

A.3.3 The multiplicative sunspots

The latent process $M_{1,t}$ is updated using it's low of motion. Under the stable model M_S we distinguish two cases: if condition (33) is not satisfied $M_{1,t}^{(i)}$ can vary over time and we sample its values from the Normal distribution:

$$N\left(M_{t-1}^{(i)}, \sigma_\zeta^{2(i)}\right).$$

By contrary, if the Taylor principle is respected we set it equal to zero, that is the value corresponding to the unique stable solution.

Under the unstable model M_U, we first verify that the indicator function in (10) is equal to one. Then, with probability $\gamma^{(i)}$ we draw $M_{1,t}^{(i)}$ from the Normal distribution:

$$N\left(\frac{M_{t-1}^{(i)}}{\gamma^{(i)}}, \sigma_\zeta^{2(i)}\right)$$

while we set it equal to zero with probability $\left(1-\gamma^{(i)}\right)$.

A.3.4 The parameters updated through Particle Learning

The vector θ_2 collects all the error variances and covariances:

$$\theta_2 = [\ \sigma_R^2 \quad \sigma_\zeta^2 \quad \sigma_g^2 \quad \sigma_z^2 \quad \rho_{gz}\ ]'. \tag{A43}$$

We follow the Particle Learning approach by Carvalho et al. (2010). The latent processes and the parameters in $\theta_1^{(i)}$ are used to update a set of sufficient statistics $s_t^{(i)}$ that contains $T_R^{(i)}$, $T_\zeta^{(i)}$, $T_{gz}^{(i)}$, $n_{\zeta,t}^{(i)}$ and, where:

$$T_R^{(i)} = \sum_{j=1}^{t}\left(\varepsilon_{R,j}^{(i)}\right)^2; \qquad T_\zeta^{(i)} = \sum_{j=1}^{n_{\zeta,t}^{(i)}}\left(\varepsilon_{\zeta,j}^{(i)}\right)^2; \qquad T_{gz}^{(i)} = \sum_{j=1}^{t}\left(\begin{bmatrix} \varepsilon_{g,j}^{(i)} \\ \varepsilon_{z,j}^{(i)} \end{bmatrix}\begin{bmatrix} \varepsilon_{g,j}^{(i)} & \varepsilon_{z,j}^{(i)} \end{bmatrix}\right)$$

and $n_{\zeta,t}^{(i)}$ is he number of times $M_t^{(i)}$ has been drawn from a Normal distribution rather than being set equal to zero. The sufficient statistics are then used to update the posterior distributions of the parameters in θ_2, which are known analytically (up to a normalizing constant), given our assumptions on the prior distributions. In particular we assume that the priors for σ_R^2 and σ_ζ^2 have an Inverse Gamma distribution defined, respectively, by shape parameters a_R and a_ζ, and rate parameters b_R and b_ζ.[35] Their posterior distributions are also Inverse Gamma:

$$\left(\sigma_R^{2(i)}|D_t\right) \sim IG\left(a_R + \frac{t}{2}, b_R + \frac{T_R^{(i)}}{2}\right)$$
$$\left(\sigma_\zeta^{2(i)}|D_t\right) \sim IG\left(a_\zeta + \frac{n_{\zeta,t}^{(i)}}{2}, b_\zeta + \frac{T_\zeta^{(i)}}{2}\right).$$

Since the shocks to supply and demand are correlated, we assume that the prior for σ_g^2, σ_z^2 and the covariance ρ_{gz} is an Inverse Wishart with 8 degrees of freedom and scale matrix Σ_0. Given new data at time t, we can draw these parameters from their posterior distribution:

$$(\Sigma_{gz}|D_t) \sim IW\left(\Sigma_0 + T_{gz}, 8 + t\right).$$

A.3.5 The model under determinacy and stochastic volatility

In section 6 we compare the models M_S and M_U with a case in which we impose determinacy, but at the same time we allow the standard deviations of the structural shocks to vary over time. In this case we set $M_{1,t} = 1$ for every t, and we explore only the parameter space such that condition (33) is satisfied.

To estimate this model we use the same algorithm described above with some modifications. First the parameter vector θ is partitioned as:

$$\theta_1 = [\ \rho_g \quad \rho_z \quad \rho_R \quad \kappa \quad \psi_1 \quad \psi_2 \quad \tau^{-1} \quad \pi^* \quad r^* \quad \gamma \quad \rho_{gz}\ ]'.$$

and

$$\theta_2 = [\ \delta_R^2 \quad \delta_g^2 \quad \delta_z^2\ ]'. \tag{A44}$$

The latent processes are l_t, with dynamics described by equation (A41), and

$$\bar{\sigma}_t = [\ \log\sigma_{R,t} \quad \log\sigma_{g,t} \quad \log\sigma_{z,t}\ ]' \tag{A45}$$

with dynamics described by equation (13).

We take advantage of analytical integration, in analogy with the estimation of model M_S and M_U: conditional on $\bar{\sigma}_t$ the state space model for l_t is linear and Gaussian. Then, we modify the weights for the first resampling defined in equation (A18) (point 2a in the algorithm):

$$\tilde{w}_t^{(i)} \propto w_{t-1}^{(i)} p\left(D_t|l_{t-1}^{(i)}, g_{\bar{\sigma}}(\bar{\sigma}_{t-1}^{(i)}), m_{t-1}^{(i)}, \theta_2^{(i)}\right) \tag{A46}$$

where

$$g_{\bar{\sigma}}(\bar{\sigma}_{t-1}^{(i)}) = E\left(\bar{\sigma}_t^{(i)}|\bar{\sigma}_{t-1}^{(i)}\right) = \bar{\sigma}_{t-1}^{(i)} \tag{A47}$$

[35] These hyperparameters are such that the prior means and variances for σ_R^2 and σ_ζ^2 are the ones reported in Table 1.

Moreover, in the *propagation* step, we keep $M_t = 1$ and we propagate $\bar{\sigma}_t^{(i)}$ from the distribution implied by its law of motion (13) (point 3b in the algorithm). The distribution of the latent process $l_t^{(i)}$ is again Normal, with mean and covariance matrix computed through the Kalman recursion (A25) to (A28), appropriately modified.

Finally, the set of sufficient statistics $s_t^{(i)}$ contains the following variables:

$$T_R^{(i)} = \sum_{j=1}^{t} \left(\nu_{R,j}^{(i)}\right)^2; \qquad T_g^{(i)} = \sum_{j=1}^{t} \left(\nu_{g,j}^{(i)}\right)^2; \qquad T_z^{(i)} = \sum_{j=1}^{t} \left(\nu_{z,j}^{(i)}\right)^2;$$

These allow us to draw δ_R, δ_g and δ_z from their posterior distributions.

A.3.6 Computational details

We work with 500.000 particles: this number is big enough to guarantee that the filter explores well the parameter space and the support of the latent processes at any time t. However, as clear from Figure 5, when the inference on ψ_1 switches to the indeterminacy region we observe a reduction in the variance of the posterior distribution. In order to make sure that this change in the distribution reflects the likelihood implied by new data, and not a technical problem related to the filter, we increase the number of particles to 2.000.000 from 1972:IV to 1979:II.

The particles are distributed to 44 cores who run in parallel. We use a computer with two processors Intel Xeon E5-2699 v4. To estimate the first subsample the algorithm takes approximately 90 minutes.

B Figures

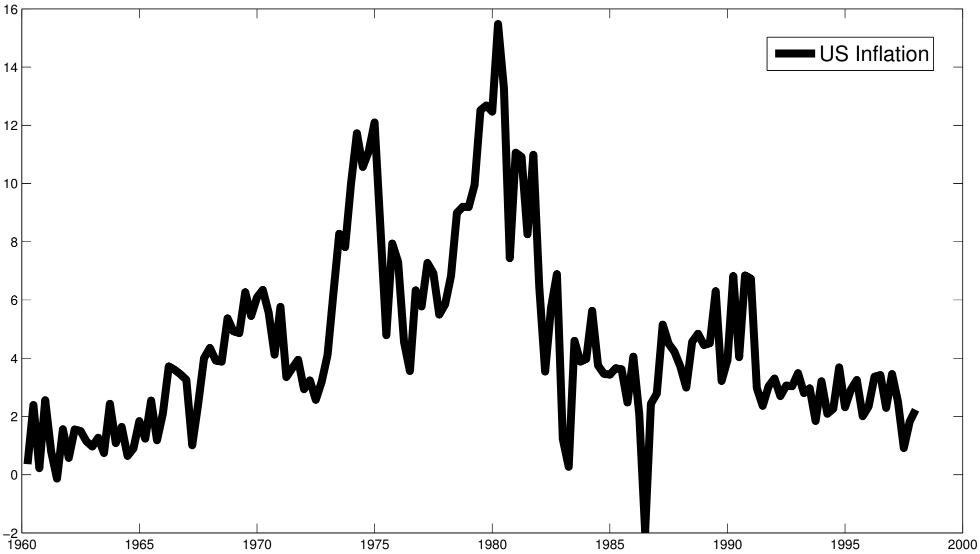

Figure 1: CPI inflation, quarterly data. Sample: 1955Q1 - 2006Q4.

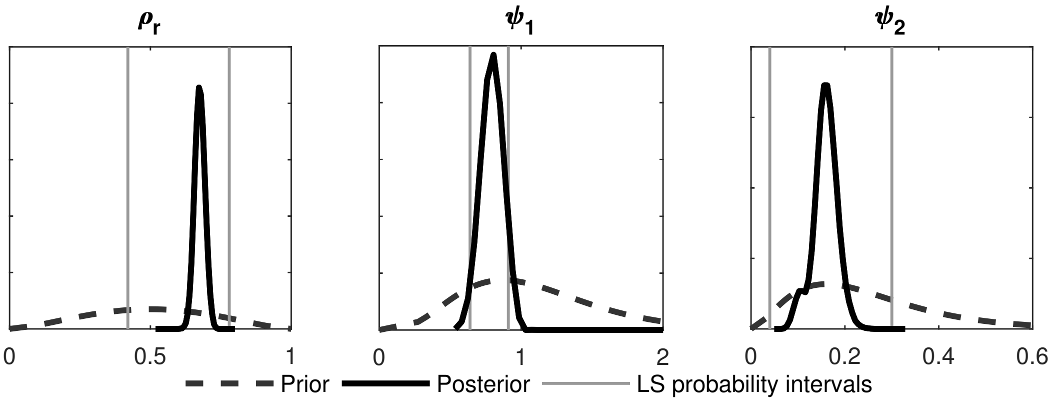

Figure 2: M_S : Comparison between the posterior distributions of the policy parameters and the probability intervals of LS.

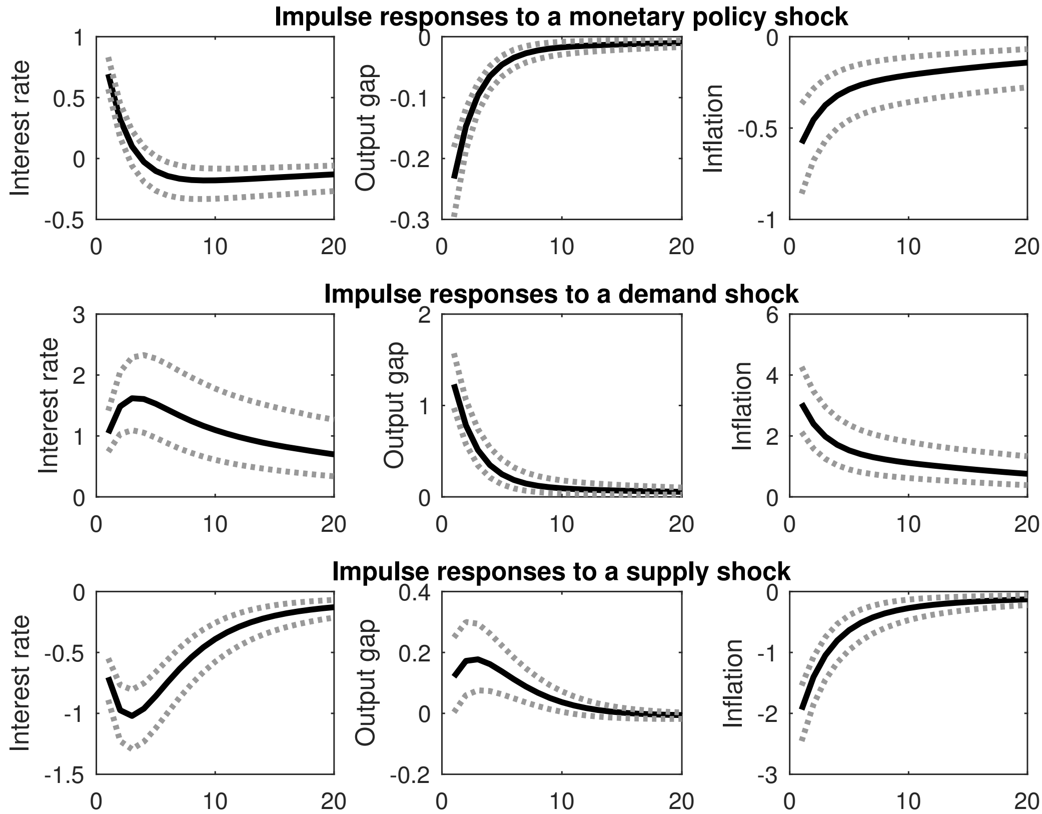

Figure 3: Generalized Impulse Response Function in the M_S model computed under the posterior distribution of $M_{1,t}$ in 1979:II.

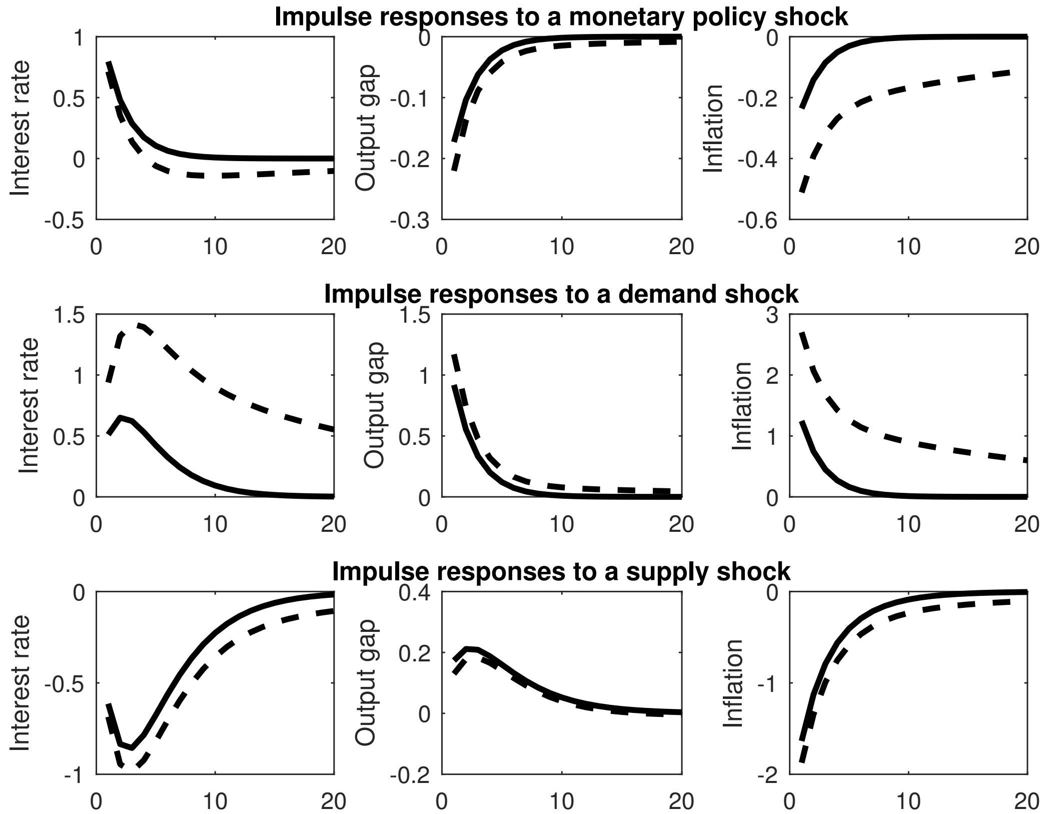

Figure 4: Generalized Impulse Response Function in the M_S model: solid line: $M_1 = 0$, dashed line: $M_1 = 0.49$

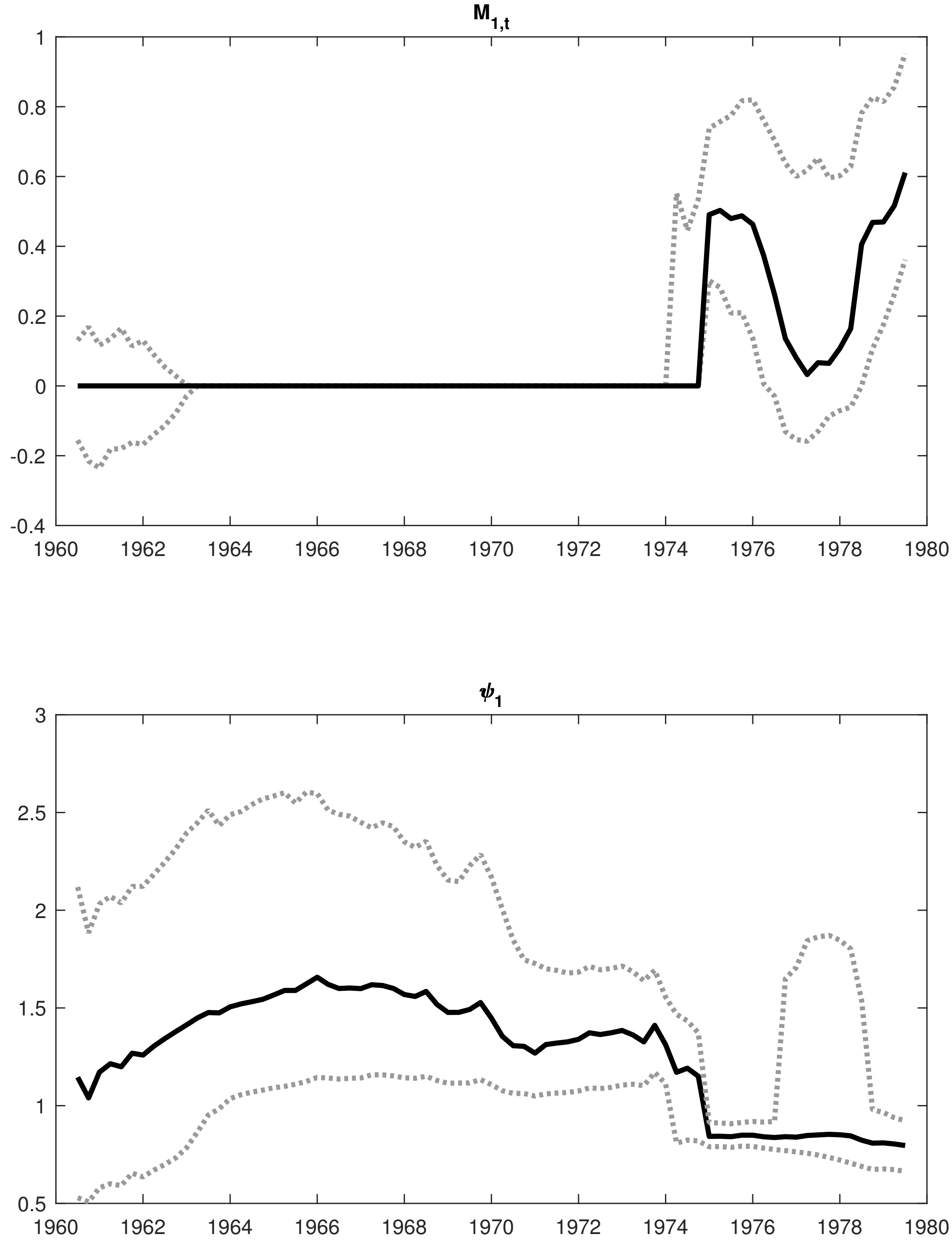

Figure 5: Estimated path of $M_{1,t}$ for the stable model M_S in the Great Inflation subsample (upper panel); sequential inference on the parameter ψ_1 (lower panel).

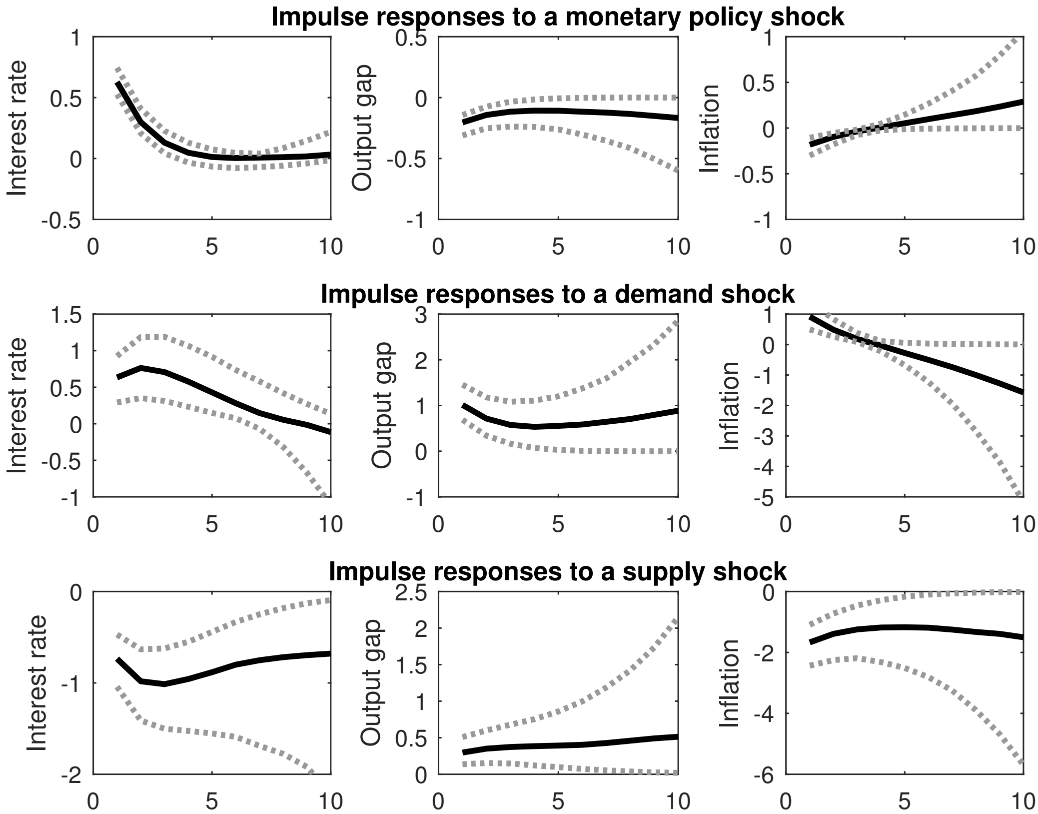

Figure 6: Generalized Impulse Response Function in the M_U model computed under the posterior distribution of $M_{1,t}$ in 1979:II.

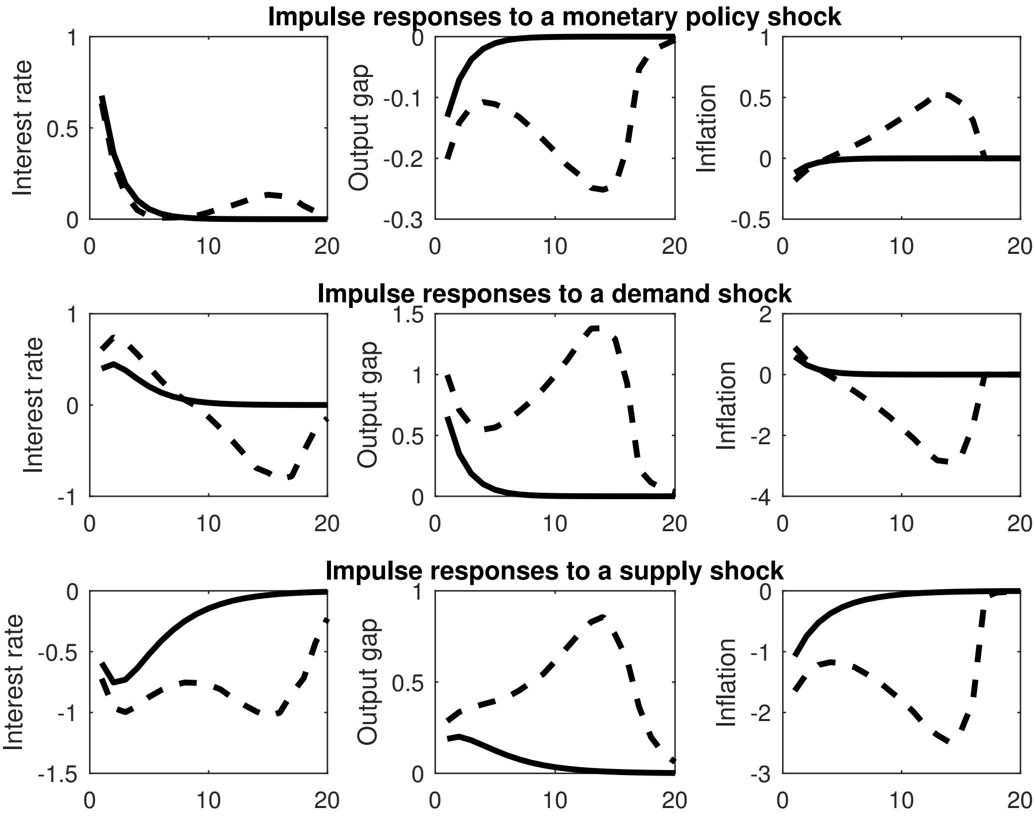

Figure 7: Generalized Impulse Response Function in the M_U model: solid line: $M_1 = 0$, dashed line: $M_1 = 0.52$

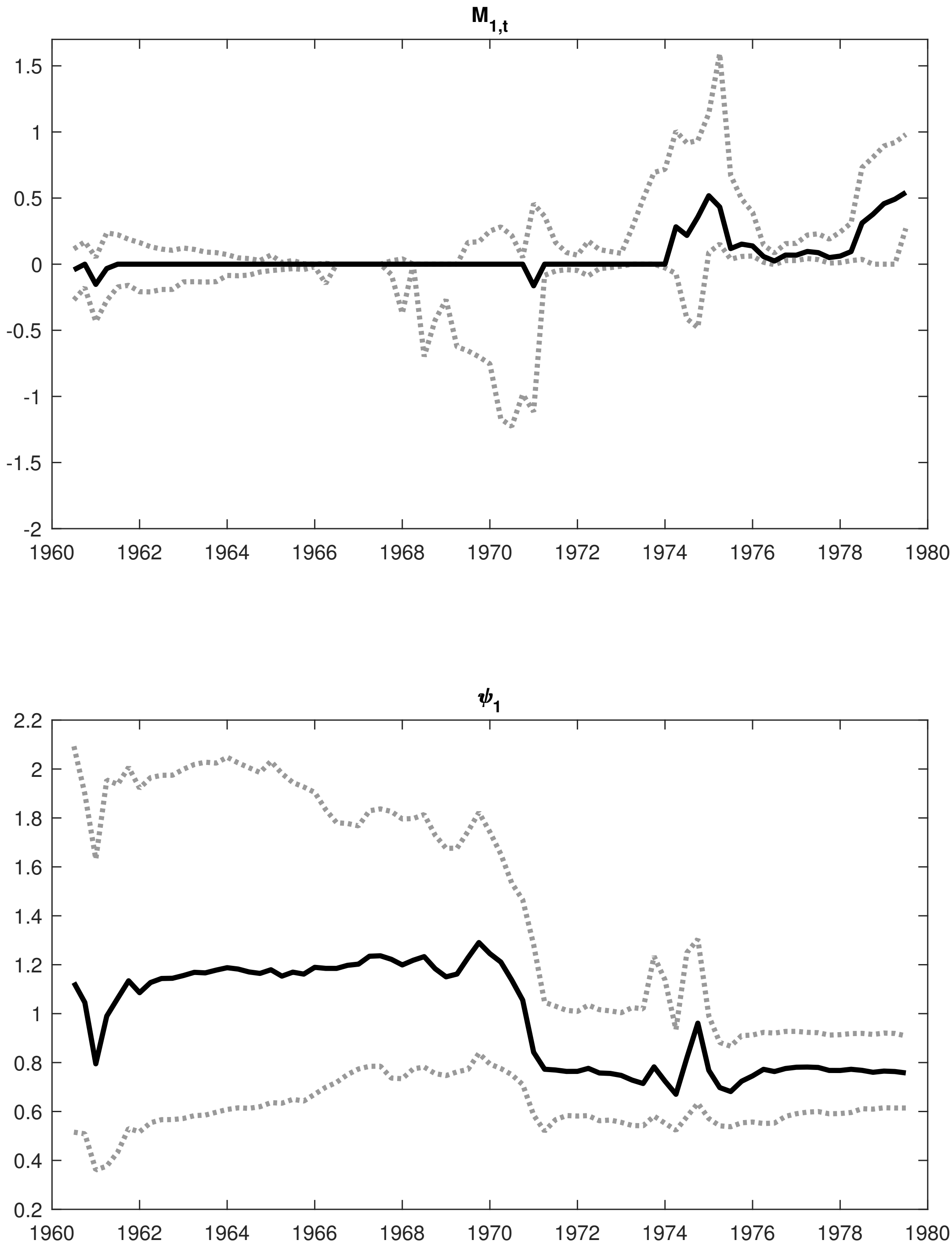

Figure 8: Estimated path of $M_{1,t}$ for the stable model M_S in the Great Inflation subsample (upper panel); sequential inference on the parameter ψ_1 (lower panel).

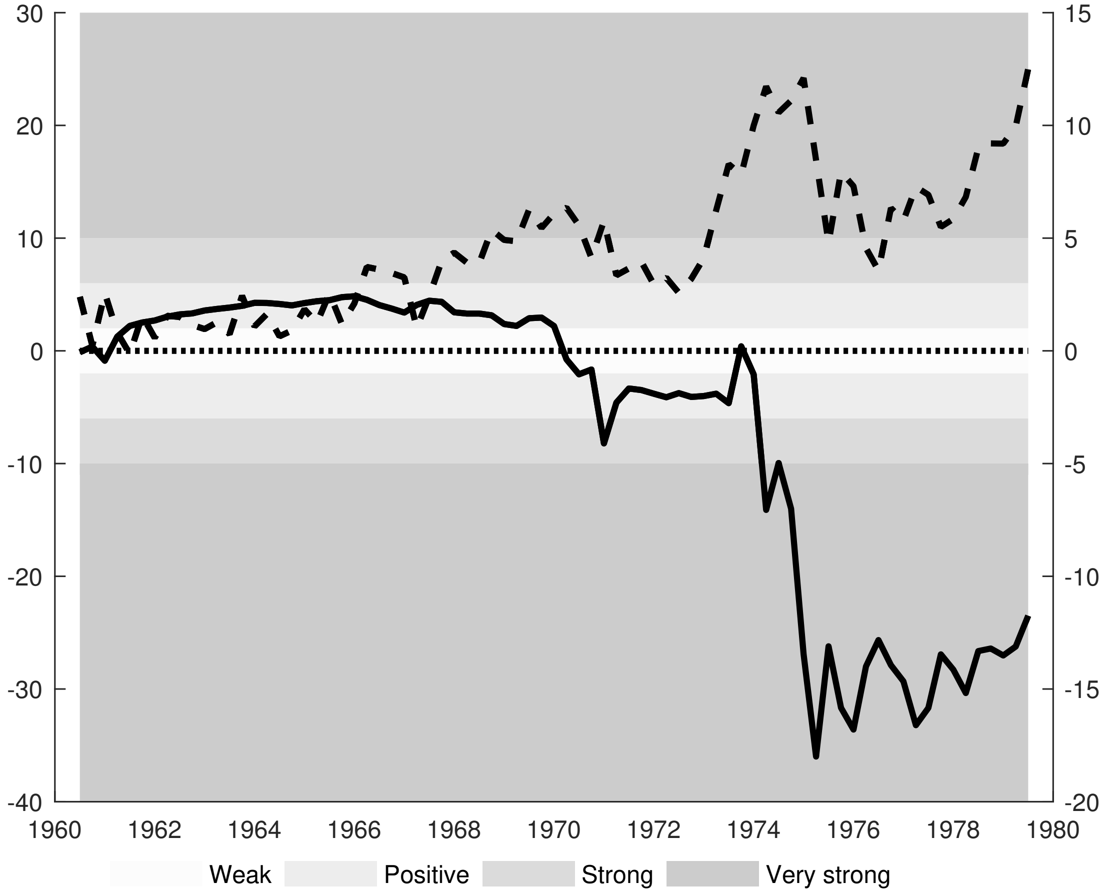

Figure 9: Comparing M_S - M_U, Great Inflation period. The panels show $2\ln(W_t)$ (solid line, scale on the left axis) and the inflation rate (dashed line, scale on the right axis)

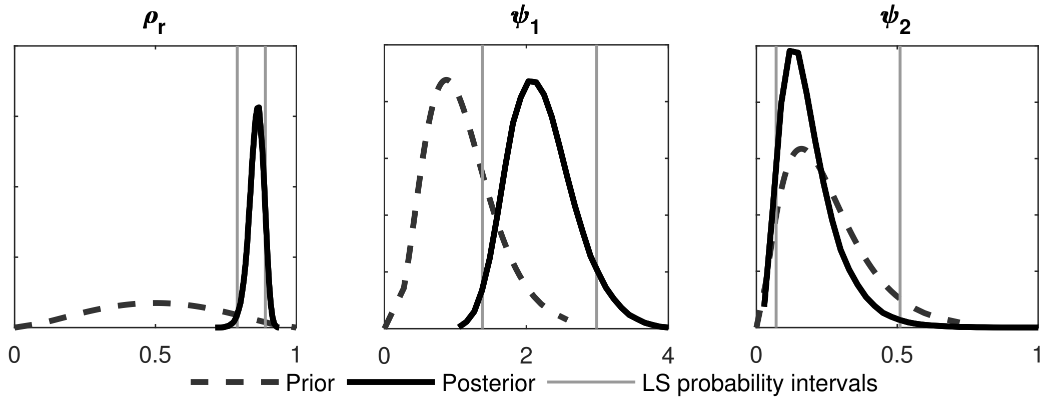

Figure 10: M_S : Comparison between the posterior distributions of the policy parameters and the probability intervals of LS.

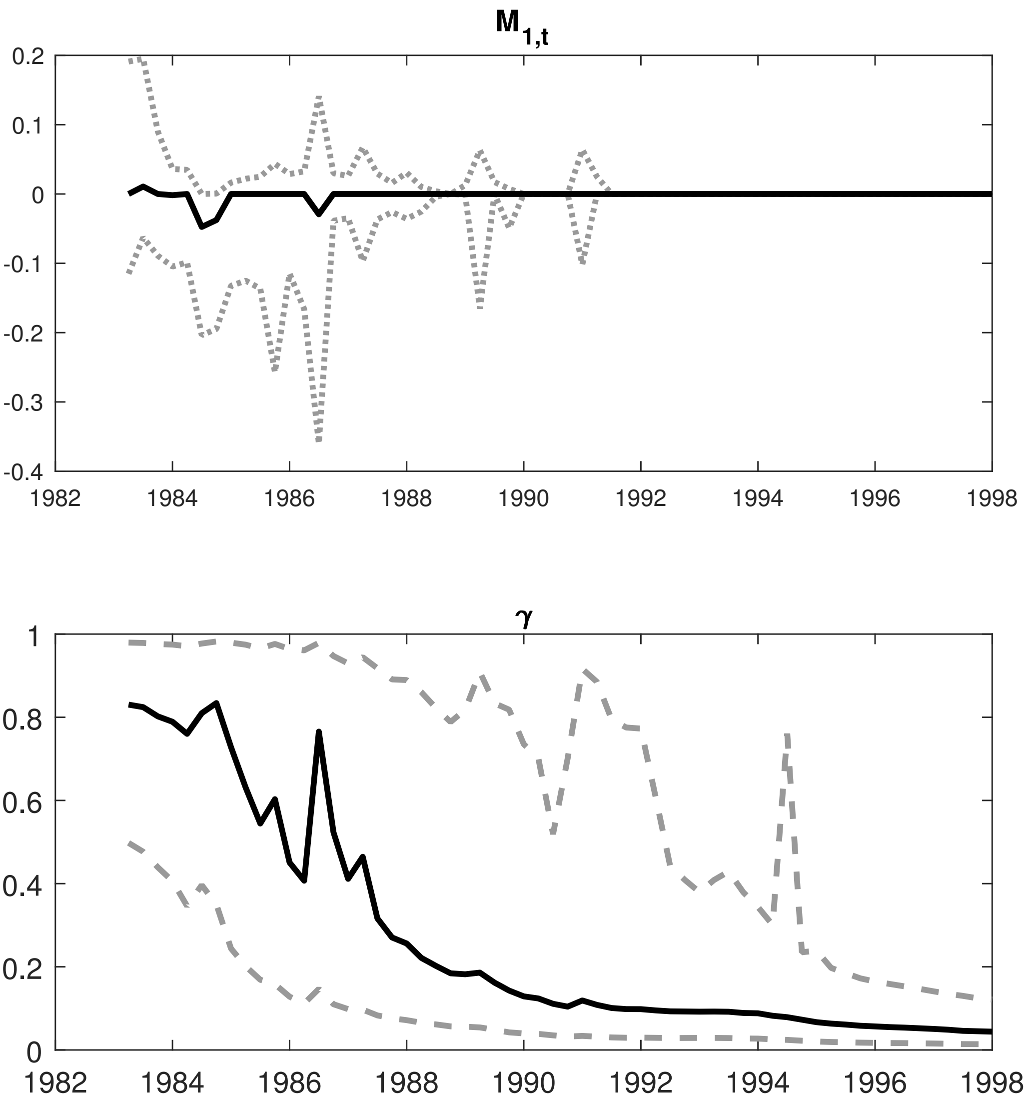

Figure 11: Estimated path of $M_{1,t}$ for the unstable model M_U in the Great Moderation subsample (upper panel); sequential inference on the parameter γ (lower panel).

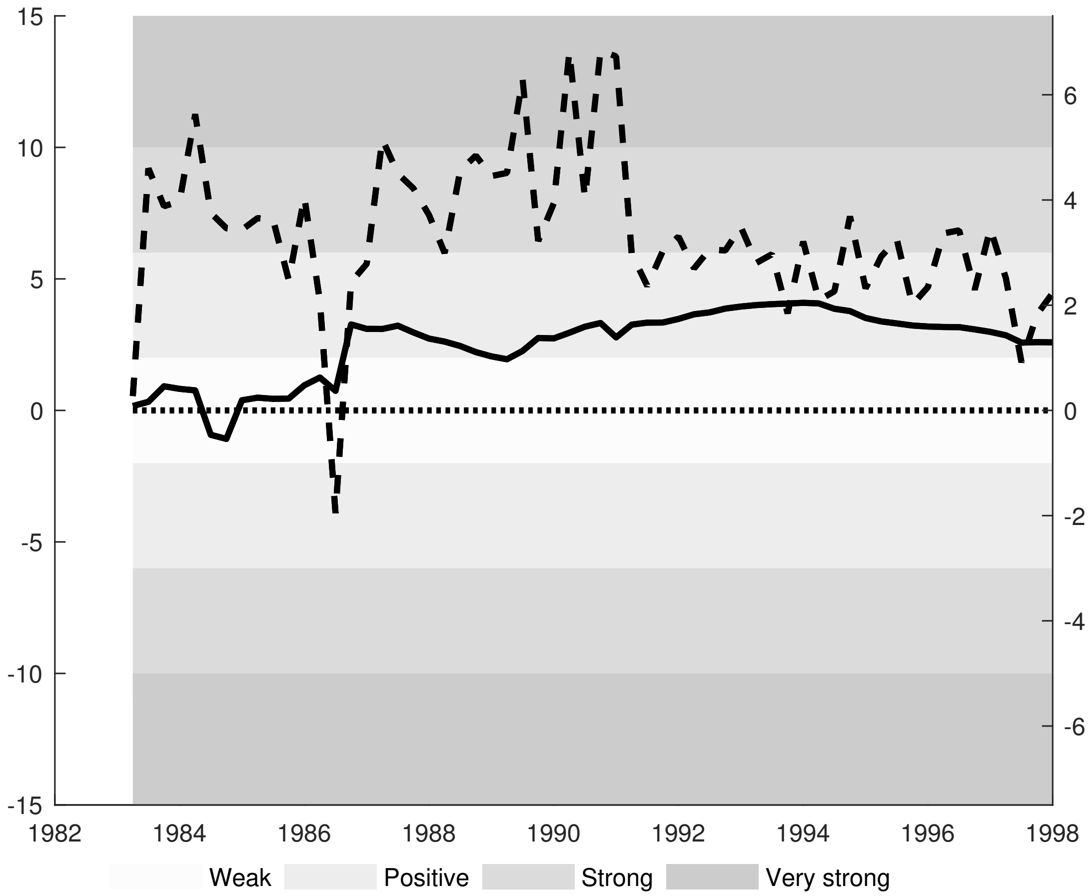

Figure 12: Comparing M_S - M_U, Great Moderation period. The panels show $2\ln(W_t)$ (solid line, scale on the left axis) and the inflation rate (dashed line, scale on the right axis)

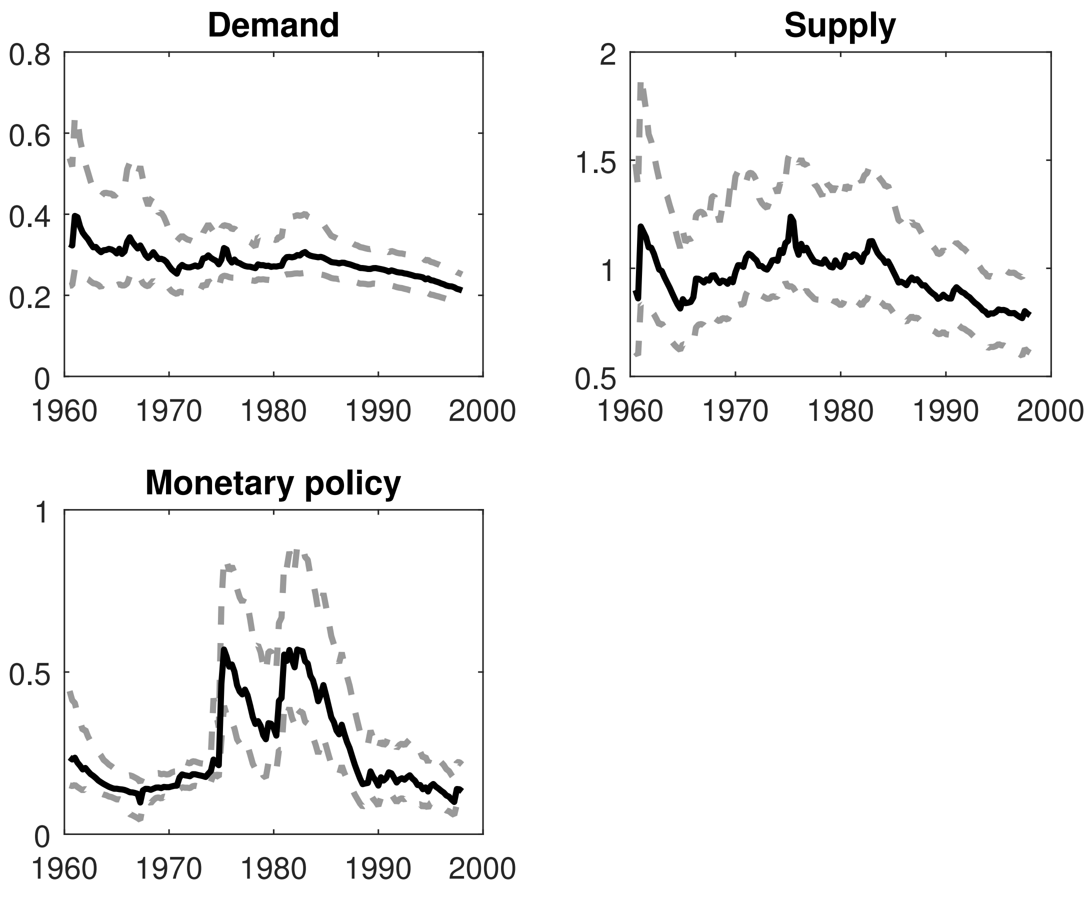

Figure 13: time-varying standard deviation of each shock - Model with determinacy and stochastic volatility

Previous DNB Working Papers in 2018

No. 583 **Dorinth van Dijk, David Geltner and Alex van de Minne**, Revisiting supply and demand indexes in real estate
No. 584 **Jasper de Jong**, The effect of fiscal announcements on interest spreads: Evidence from the Netherlands
No. 585 **Nicole Jonker**, What drives bitcoin adoption by retailers?
No. 586 **Martijn Boermans and Robert Vermeulen,** Quantitative easing and preferred habitat investors in the euro area bond market
No. 587 **Dennis Bonam, Jakob de Haan and Duncan van Limbergen,** Time-varying wage Phillips curves in the euro area with a new measure for labor market slack
No. 588 **Sebastiaan Pool**, Mortgage debt and shadow banks
No. 589 **David-Jan Jansen**, The international spillovers of the 2010 U.S. flash crash
No. 590 **Martijn Boermans and Viacheslav Keshkov**, The impact of the ECB asset purchases on the European bond market structure: Granular evidence on ownership concentration
No. 591 **Katalin Bodnár, Ludmila Fadejeva, Marco Hoeberichts, Mario Izquierdo Peinado, Christophe Jadeau and Eliana Viviano**, Credit shocks and the European labour market
No. 592 **Anouk Levels, René de Sousa van Stralen, Sînziana Kroon Petrescu and Iman van Lelyveld**, CDS market structure and risk flows: the Dutch case
No. 593 **Laurence Deborgies Sanches and Marno Verbeek**, Basel methodological heterogeneity and banking system stability: The case of the Netherlands
No. 594 **Andrea Colciago, Anna Samarina and Jakob de Haan**, Central bank policies and income and wealth inequality: A survey
No. 595 **Ilja Boelaars and Roel Mehlkopf**, Optimal risk-sharing in pension funds when stock and labor markets are co-integrated
No. 596 **Julia Körding and Beatrice Scheubel**, Liquidity regulation, the central bank and the money market

DeNederlandscheBank

EUROSYSTEEM

De Nederlandsche Bank N.V.
Postbus 98, 1000 AB Amsterdam
020 524 91 11
dnb.nl