# Technical Appendix to "Why Inflation Rose and Fell: Policymakers' Beliefs and US Postwar Stabilization Policy" 

Giorgio E. Primiceri*

February 2005

## A Self-confirming equilibrium

Appendix A formally defines a self-confirming equilibrium and derives the self-confirming equilibrium of the model of section 2 , in the case of constant gain learning. Let $y_{t}^{\pi}, y_{t}^{u}, x_{t}^{\pi}, x_{t}^{u}, \hat{\beta}^{\pi}$ and $\hat{\beta}^{u}$ be the same objects defined in section 2.3. I follow Sargent (1999) in defining a self-confirming equilibrium in this framework:

Definition 1 A self-confirming equilibrium is a set of policymakers' beliefs about the models' parameters $\hat{\beta} \equiv\left[\hat{\beta}^{\pi}, \hat{\beta}^{u}, \hat{u}^{N}\right]$, a fixed optimal policy rule $g(\hat{\beta})$ and an associated stationary stochastic process for the vector $\left[\pi_{t}, u_{t}, V_{t}, u_{t}^{N}\right]$ such that: (a) $\hat{u}^{N}, \hat{\beta}^{\pi}$ and $\hat{\beta}^{u}$ satisfy

$$
\begin{align*}
E\left[u_{t}-\hat{u}^{N}\right] & =0  \tag{I}\\
E\left[x_{t}^{i}\left(y_{t}^{i}-x_{t}^{i} \hat{\beta}^{i}\right)\right] & =0, \quad i=\{\pi, u\} \tag{II}
\end{align*}
$$

where the expectations are taken with respect to the probability distribution generated by (5), (2), (3) and (9); (b) the vector $\left[\pi_{t}, u_{t}, V_{t}, u_{t}^{N}\right]$ is generated by the stationary stochastic process implied by (5), (2), (3) and (9).

It is straightforward to verify that the set of beliefs $\hat{u}^{N}=u^{*}, \hat{\beta}^{\pi}=\left[0 ; \alpha_{1} ; \alpha_{2} ; \theta_{1} ; \theta_{2}\right]$ and $\hat{\beta}^{u}=\left[0 ; \rho_{1} ; \rho_{2}\right]$ satisfy (I) and (II) and therefore represents a self-confirming equilibrium, in the case of $\sigma_{\tau}^{2}=0$. When $\sigma_{\tau}^{2}>0$, finding the self-confirming equilibrium is more involved and requires a numerical solution of the system of equations given by (I) and (II). The procedure works as follows: any given fixed value of $\hat{u}^{N}, \hat{\beta}^{\pi}$ and $\hat{\beta}^{u}$ implies a linear stochastic process for $\left[\pi_{t}, u_{t}, V_{t}, u_{t}^{N}\right]$ via equations (5), (2), (3) and (9). The linear process can be rewritten as a first order system of the form $z_{t}=C+A z_{t-1}+B \nu_{t}$. Thus $E\left(z_{t}\right)=(I-A)^{-1} C$ and $\operatorname{Var}\left(z_{t}\right)$ can

[^0]be found by solving the Lyapunov equation $\operatorname{Var}\left(z_{t}\right)=A \operatorname{Var}\left(z_{t}\right) A^{\prime}+B \operatorname{Var}\left(\nu_{t}\right) B^{\prime}$. The elements of $E\left(z_{t}\right)$ and $\operatorname{Var}\left(z_{t}\right)$ can be used to compute $E\left[u_{t}-\hat{u}^{N}\right]$ and $E\left[x_{t}^{i}\left(y_{t}^{i}-x_{t}^{i \ell} \hat{\beta}^{i}\right)\right]$, for $i=\{\pi, u\}$, which, in general, will not be equal to zero. A simple equation solver can be used to solve for the set of beliefs $\hat{u}^{N}, \hat{\beta}^{\pi}$ and $\hat{\beta}^{u}$, which satisfy (I) and (II). Of course the solution will depend on the value of the true parameters of the model. As an illustrative example I consider the case in which the true parameters of the model are the point estimates of the baseline specification, presented in the first column of table 1 . The self-confirming equilibrium in this case corresponds to $\hat{u}^{N}=6, \hat{\beta}^{\pi}=[0.0394 ; 0.7203 ; 0.2623 ;-0.8409 ; 0.7637]$ and $\hat{\beta}^{u}=[0 ; 1.5703 ;-0.6269]$. Furthermore, the eigenvalues of the Jacobian of the expressions contained in (I) and (II), evaluated at the selfconfirming equilibrium, have negative real parts. This guarantees the stability of the equilibrium.

As $\operatorname{Var}\left(u_{t}^{N}\right)$ increases the mistakes associated with the estimation of the current level of the natural rate of unemployment will bias toward zero the estimate the slope of the Phillips curve and of the persistence of unemployment deviations from the natural rate in the aggregate demand equation. This leads to self-confirming equilibria whose distance from the true values of the parameters is increasing in $\operatorname{Var}\left(u_{t}^{N}\right)$. Finally, when $\operatorname{Var}\left(u_{t}^{N}\right)$ is large with respect to $\sigma_{\varepsilon}^{2}$ and $\sigma_{\eta}^{2}$, the model does not admit a self confirming equilibrium anymore. Figure 1 plots the Euclidean distance between the true parameters ( $\left[0 ; \alpha_{1} ; \alpha_{2} ; \theta_{1} ; \theta_{2} ; 0 ; \rho_{1} ; \rho_{2} ; u^{*}\right]$ ) and the set of beliefs about these parameters corresponding to the self-confirming equilibria. These self-confirming equilibria are computed for different values of $\operatorname{Var}\left(u_{t}^{N}\right) \geq 0$. This graph confirms the intuition that the distance between equilibrium beliefs and true parameters increases with $\operatorname{Var}\left(u_{t}^{N}\right)$. The line is truncated at the value 4.63, because for $\operatorname{Var}\left(u_{t}^{N}\right)$ larger than this value a self confirming equilibrium cannot be found. All the found self-confirming equilibria are stable.

## B State space form for model's estimation

Appendix B gives the details of the state space form representation of the model for the estimation with the Kalman filter.

The canonical state space form is given by:

$$
\begin{align*}
y_{t} & =A Z_{t}+B X_{t}+R e_{t},  \tag{III}\\
X_{t} & =C+G X_{t-1}+Q s_{t},  \tag{IV}\\
\binom{e_{t}}{s_{t}} & \sim \text { i.i.d. } N(0, I) . \tag{V}
\end{align*}
$$

In our case, $y_{t}=\left[\pi_{t} ; u_{t}\right]^{\prime} ; Z_{t}=\left[\pi_{t-1} ; \pi_{t-2} ; u_{t-1} ; u_{t-2} ; V_{t-1}\right]^{\prime} ; X_{t}=\left[u_{t}^{N} ; u_{t-1}^{N} ; u_{t-2}^{N}\right]^{\prime} ;$

$$
A=\left[\begin{array}{ccccc}
\alpha_{1} & \alpha_{2} & -\theta_{1} & -\theta_{2} & 0 \\
0 & 0 & \rho_{1} & \rho_{2} & 1
\end{array}\right] ; B=\left[\begin{array}{ccc}
0 & \theta_{1} & \theta_{2} \\
1 & -\rho_{1} & \rho_{2}
\end{array}\right] ; C=\left[\begin{array}{c}
(1-\gamma) u^{*} \\
0 \\
0
\end{array}\right] ; G=\left[\begin{array}{ccc}
\gamma & 0 & 0 \\
1 & 0 & 0 \\
0 & 1 & 0
\end{array}\right]
$$

$R=\left[\begin{array}{cc}\sigma_{\varepsilon} & 0 \\ 0 & \sigma_{\eta}\end{array}\right] ; Q=\left[\begin{array}{ccc}\sigma_{\tau} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0\end{array}\right]$.
The standard Kalman filter recursion formulas can be found in Hamilton (1994). To start the recursion it is necessary to specify $E\left(X_{0} \mid \Omega_{0}\right)$ and $\operatorname{Var}\left(X_{0} \mid \Omega_{0}\right)$, where $\Omega_{0}$ represents the information set available at time 0 . Following a common practice, I set $E\left(X_{0} \mid \Omega_{0}\right)$ and $\operatorname{Var}\left(X_{0} \mid \Omega_{0}\right)$ to the unconditional values implied by the transition equation. In particular, this results in $E\left(X_{0} \mid \Omega_{0}\right)=$ $[6 ; 6 ; 6]^{\prime}$, which corresponds approximately to the estimate of the natural rate of unemployment of Staiger, Stock, and Watson (2001) in 1960 (which is the initial date of our sample).

## C The solution method for the forward looking model

Appendix C illustrates in more detail the method used to solve the forward looking model with fully rational agents. As mentioned in section 5.2, the model is hard to solve because it is a nonlinear system of rational expectation equations. The source of nonlinearity is the learning behavior of policymakers. I will rely on numerical methods. The adopted solution method is based on Fackler and Miranda (2001). A similar method is in Fernandez-Villaverde and Rubio (2002).

Consider the system of rational expectation equations given by (19) and (20). To simplify the analysis and only for the purposes of this section I will assume that $u_{t}^{N}$ is a deterministic function, known by the private sector, but, as usual, unknown by policymakers. I will set $u_{t}^{N}$ to be equal to the smoothed estimate of $u_{t}^{N}$ obtained in the estimation of the forward looking model with partially rational agents (section 5.1). Thus, let $\tilde{u}_{t} \equiv u_{t}-u_{t}^{N}$. Equations (19) and (20), the only ones involving expectations, can be rearranged and rewritten in the following compact form:

$$
\begin{equation*}
y_{t}=A E_{t-1} y_{t+1}+B X_{t-1}+v_{t} \tag{VI}
\end{equation*}
$$

where $y_{t} \equiv\left[\pi_{t}, \tilde{u}_{t}\right]^{\prime}$ is the vector of observed endogenous variables; $X_{t-1} \equiv\left[\pi_{t-1}, \tilde{u}_{t-1}, V_{t-1}\right]$ is the vector of observed predetermined variables; $v_{t} \equiv\left[\varepsilon_{t}, \eta_{t}\right]^{\prime}$ is the vector of unobservable shocks; $A$ and $B$ are matrices of coefficients, omitted for brevity. (VI) is linear, but the complete system, given by (VI), (9), (10), (11), (12) and (13) is nonlinear. The solution of the model is the unknown response function $E_{t-1} y_{t+1}=\Psi\left(\Omega_{t-1}\right)$, where $\Omega_{t}$ represents the information available at time $t$ and $\Psi(\cdot)$ satisfies

$$
\Psi\left(\Omega_{t-1}\right)=A E_{t-1} \Psi\left(\Omega_{t}\right)+B E_{t-1} X_{t}
$$

When the model is nonlinear in general there is not a closed form expression for $\Psi(\cdot)$ and it must be approximated numerically by projection methods. The basic idea of Fackler and Miranda (2001) is approximating $\Psi\left(\Omega_{t}\right)$ with a linear combination of basis functions of the state variables. This is given by $\Phi s_{t}$, where $s_{t}$ is an $m \times 1$ vector of basis functions and $\Phi$ is a $2 \times m$ matrix of coefficients. In the numerical procedure, also the expectation operator must be approximated using quadrature methods. Therefore, the expectation of a generic function $f(\cdot)$ of the model's source of randomness,
$v_{t}$, is approximated by a discrete version of the integral, given by

$$
E f\left(v_{t}\right) \approx \sum_{j=1}^{k} \omega_{j} f\left(v_{t}^{j}\right)
$$

For a given value of the innovations $v_{t}^{j}$,

$$
y_{t}^{j}=A \Phi s_{t-1}+B X_{t-1}+v_{t}^{j}
$$

and

$$
y_{t+1}^{j}=A \Phi s_{t}^{j}+B X_{t}+v_{t+1},
$$

where the superscript $j$ for $s$ indicates that the value of $s$ at time $t$ depends on the realization of the shocks at time $t$. Now we can compute $E_{t-1} y_{t+1}=E_{t-1} E_{t} y_{t+1} \approx \sum_{j=1}^{k} \omega_{j}\left(A \Phi s_{t}^{j}+B X_{t}\right) \equiv z_{t-1}$. Let $S \equiv\left[s_{1, t-1}, \ldots, s_{n, t-1}\right]$ be a collection of $n \geq m$ values of $s_{t-1}$ and $Z \equiv\left[z_{1, t-1}, \ldots, z_{n, t-1}\right]$ the collection of the corresponding $n$ values of $z_{t-1}$. The solution consists in the $\Phi^{*}$ which solves $\Phi^{*} S=Z$, or $\Phi^{*} S S^{\prime}=Z S^{\prime}$ in the case in which $n>m$. It can be done using standard equation solvers.

In this application, to approximate the integrals and expectation operators, I use a GaussHermite quadrature with 3 nodes (see Judd, 1998). As mentioned above, the dimension of the state vector is high. Thus, the use of tensor product bases or complete polynomial bases is unfeasible for any polynomial degree bigger or equal to 2 . For this reason I chose the following ad hoc collection of 21 basis functions of the states, which turned out to work well:

$$
\begin{aligned}
s_{t}= & {\left[1 ; \pi_{t} ; \tilde{u}_{t} ; \pi_{t-1} ; \tilde{u}_{t-1} ; V_{t} ; \hat{c}_{\pi, t} ; \hat{\alpha}_{1, t} ; \hat{\alpha}_{2, t} ; \hat{\theta}_{1, t} ; \hat{\theta}_{2, t} ; \hat{c}_{u, t} ; \hat{\rho}_{1, t} ; \hat{\rho}_{2, t} ; \hat{u}_{t}^{N} ;\right.} \\
& \left.\pi_{t}^{2} ; \tilde{u}_{t}^{2} ; \hat{\alpha}_{1, t} \pi_{t} ; \hat{\alpha}_{2, t} \pi_{t-1} ; \hat{\theta}_{1, t}\left(u_{t}-\hat{u}_{t}^{N}\right) ; \hat{\theta}_{1, t}\left(u_{t-1}-\hat{u}_{t-1}^{N}\right)\right] .
\end{aligned}
$$

Notice that the choice of $s_{t}$ includes all linear terms in the state variables of the problem and some potentially relevant second order terms. The dimension of $s_{t}$ is so large that the choice of $S$ based on standard grid methods is unfeasible, even specifying only two values for any state variable. To solve this problem I chose a collection of $n=86 s_{t}$ 's, corresponding to the actual values of $s_{t}$ observed in the data, every 2 quarters, from 1959:IV to 2002:IV. The results are only marginally affected by a different choice of values for $S$, like for example the observed data, every 2 quarters, 1960:I to 2002:III.

## D The MCMC algorithm for the stochastic volatility model

Appendix D illustrates the details of the MCMC algorithm used in section 6 for the estimation of the model with stochastic volatility. The parameters of interest are the coefficients $\Psi_{1} \equiv$ $\left[\alpha_{1} ; \theta_{1} ; \theta_{2} ; \rho_{1} ; \rho_{2}\right], \Psi_{2} \equiv[k ; \phi], \Psi_{3} \equiv\left[\sigma_{\nu_{\varepsilon}}^{2} ; \sigma_{\nu_{\eta}}^{2}\right]$ and the unobservable states $u^{N} \equiv\left\{u_{t}^{N}\right\}_{t=1}^{T}$, $\sigma_{\varepsilon} \equiv\left\{\sigma_{\varepsilon, t}\right\}_{t=1}^{T}$ and $\sigma_{\eta} \equiv\left\{\sigma_{\eta, t}\right\}_{t=1}^{T}$. The estimation consists of the simulation of the posterior
of the parameters of interest, conditional on the observed data. MCMC allows to simulate lower dimensional conditional posteriors instead of the high dimensional unconditional one.

Notice that the model can be rewritten like in (III), (IV) and (V), with the difference that now the elements of $R$ are time varying and follow the processes (21) and (22). The algorithm works in 5 steps.

## D. 1 Step 1: drawing $u^{N}$

Conditional on $\sigma_{\varepsilon}, \sigma_{\eta}, \Psi_{1}$ and $\Psi_{2}$, the observation equation (III) is linear and has Gaussian innovations with known variance. Therefore, the vector $u^{N}$ can be drawn using standard simulation smoothers, like, for instance, Carter and Kohn (1994) or Durbin and Koopman (2002). Details of this procedure can be also found in Kim and Nelson (1999).

## D. 2 Step 2: drawing $\sigma_{\varepsilon}$ and $\sigma_{\eta}$

Consider now the system of equations

$$
\begin{equation*}
y_{t}-A Z_{t}-B X_{t}=y_{t}^{*}=R_{t} e_{t} \tag{VII}
\end{equation*}
$$

where, taking $u^{N}, \Psi_{1}$ and $\Psi_{2}$ as given, $y_{t}^{*}$ is observable. This is a system of nonlinear measurement equations, but can be easily converted in a linear one, by squaring and taking logarithms of every element of (VII), which leads to the following approximating state space form:

$$
\begin{align*}
y_{t}^{* *} & =2 h_{t}+e_{t}^{* *}  \tag{VIII}\\
h_{t} & =h_{t-1}+\omega_{t} . \tag{IX}
\end{align*}
$$

$y_{i t}^{* *}=\log \left[\left(y_{i t}^{*}\right)^{2}+\bar{c}\right] ; \bar{c}$ is an offset constant (set to 0.001$) ; e_{i t}^{* *}=\log \left(e_{i t}^{2}\right) ; h_{t}=\log \left(\operatorname{diag}\left(R_{t}\right)\right)$. Observe that the $e^{* *}$ 's and the $\omega$ 's are not correlated. The system in this form has a linear, but non-Gaussian state space form, because the innovations in the measurement equations are distributed as a $\log \chi^{2}(1)$. In order to further transform the system in a Gaussian one, a mixture of normals approximation of the $\log \chi^{2}$ distribution is used, as described in Kim, Shephard, and Chib (1998). Observe that the variance covariance matrix of the $e$ 's is the identity matrix. This implies that the variance covariance matrix of the $e^{* *}$ 's is also diagonal, allowing to use the same (independent) mixture of normals approximation for any element of $e^{* *}$. Kim, Shephard and Chib (1998) select a mixture of 7 normal densities with component probabilities $q_{i}$, means $m_{i}$, and variances $v_{i}^{2}, i=1, \ldots, 7$. The constants $\left\{q_{i}, m_{i}, v_{i}^{2}\right\}$ are chosen to match a number of moments of the $\log \chi^{2}(1)$ distribution. The constants $\left\{q_{i}, m_{i}, v_{i}^{2}\right\}$ can be found in Kim, Shephard, and Chib (1998).

For the innovation to the variable $y_{j t}$, define as $s_{j}^{T}=\left[s_{j 1}, \ldots, s_{j T}\right]^{\prime}$ the vector of indicator variables selecting at every point in time which member of the mixture of normal approximation has to be used. Conditional on $u^{N}, \Psi$ and $s^{T}$ (which denotes the collection of $s_{j}^{T}$ ), the system has
an approximate linear and Gaussian state space form. Again, exactly like in the previous step of the sampler, this procedure allows to draw every $h_{t}$ using a simulation smoother.

Conditional on the data and the new series of $h_{t}$ 's, it is possible to sample the new $s_{j}^{T}$ vectors, to be used in the next iteration. This is easily done (separately for every $j$ ) by sampling from the discrete densities defined by

$$
\operatorname{Pr}\left(s_{j t}=i \mid y_{j t}^{* *}, h_{j t}\right) \propto q_{i} f_{N}\left(y_{j t}^{* *} \mid 2 h_{j t}+m_{i}, v_{i}^{2}\right), \quad i=1, \ldots, 7 .
$$

Further details can be found in Kim, Shephard, and Chib (1998) or Primiceri (2005).

## D. 3 Step 3: drawing $\Psi_{1}$

Conditional on $\Psi_{2}, \sigma_{\varepsilon}, \sigma_{\eta}$ and $u^{N}$, the objects $Z_{t}, X_{t}$ and $R_{t}$ are observable. Therefore, the elements of $\Psi_{1}$ (which correspond to the elements of $A$ and $B$ ) can be easily drawn from the posterior of the coefficients of a regression with known variance. This posterior is normally distributed with mean equal to the OLS coefficients and variance equal to the variance of the OLS coefficients.

## D. 4 Step 4: drawing $\Psi_{2}$

$\Psi_{2}$ enters the model non-linearly. Therefore, in order to draw from the conditional posterior of $\Psi_{2}$, I use a Metropolis step, nested in the Gibbs sampler. The procedure works as follows: I draw a candidate value $\Psi_{2}^{*}$ from a proposal distribution $\varphi\left(\Psi_{2}^{*} \mid \Psi_{2}^{i-1}\right)$, where $\Psi_{2}^{i-1}$ is the previous draw of the chain. At this point I compute the value of the posterior associated to the draw, $p\left(\Psi_{2}^{*} \mid \Psi_{1}, \sigma_{\varepsilon}, \sigma_{\eta}, u^{N},\left\{y_{t}\right\}_{t=1}^{T}\right)$, which, under flat prior, is proportional to the value of the likelihood. The new draw is accepted with probability

$$
a=\min \left\{\frac{p\left(\Psi_{2}^{*}\right) / \varphi\left(\Psi_{2}^{*}\right)}{p\left(\Psi_{2}^{i-1}\right) / \varphi\left(\Psi_{2}^{i-1}\right)}, 1\right\} .
$$

If the proposal value is rejected, the next element of the chain is set to be $\Psi_{2}^{i-1}$. In order to satisfy the constraints $\phi \geq 0$ and $\mathrm{m}, 0 \leq k \leq 1$, I chose the proposal distribution to be normal in $f\left(\Psi_{2}^{*}\right)$, where $f(a, b)=\left[\log (a), \log \left(\frac{b}{1-b}\right)\right]$. The mean is chosen to be $f\left(\Psi_{2}^{i-1}\right)$, while I fix the variance to a diagonal matrix with elements 0.001 and 0.005 on the main diagonal.

## D. 5 Step 5: drawing $\Psi_{3}$

Conditional on $\sigma_{\varepsilon}, \sigma_{\eta}$, each element of $\Psi_{3}$ has an inverse-Gamma posterior distribution, independent of the other element. Conditional on $\sigma_{\varepsilon}, \sigma_{\eta}$, it is easy to draw from these inverse-Gamma posteriors because the innovations are observable. ${ }^{1}$

[^1]
## References

Carter, C. K., and R. Kohn (1994): "On Gibbs Sampling for State Space Models," Biometrika, 81(3), 541-553.

Durbin, J., and S. J. Koopman (2002): "A Simple and Efficient Simulation Smoother for State Space Time Series Analysis," Biometrika, 89(3), 603-616.

Fackler, P. L., and M. J. Miranda (2001): "Solving Non-Linear Rational Expectations Models," mimeo, North Carolina State University.

Fernandez-Villaverde, J., and J. Rubio (2002): "Estimating Nonlinear Dynamic Equilibrium Economies: A Likelihood Approach," mimeo, University of Pennsylvania.

Gelman, A., J. B. Carlin, H. S. Stern, and D. B. Rubin (1995): Bayesian Data Analysis. Chapman and Hall, London.

Hamilton, J. D. (1994): Time Series Analysis. Princeton University Press, Princeton, New Jersey.
Judd, K. L. (1998): Numerical Methods in Economics. The MIT Press, Cambridge, MA.
Kim, C.-J., and C. R. Nelson (1999): State-Space Models with Regime Switching: Classical and Gibbs Sampling Approaches with Applications. M.I.T. Press, Cambridge, Massachussetts.

Kim, S., N. Shephard, and S. Chib (1998): "Sthochastic Volatility: Likelihood Inference and Comparison with ARCH Models," The Review of Economic Studies, 65(3), 361-393.

Primiceri, G. E. (2005): "Time Varying Structural Vector Autoregressions and Monetary Policy," The Review of Economic Studies, forthcoming.

Staiger, D., J. H. Stock, and M. W. Watson (2001): "Prices, Wages and the U.S. NAIRU in the 1990s," mimeo, Princeton University.


Figure 1: Euclidean norm of the distance between the true value of the coefficients and the value of policymakers beliefs in a self-confirming equilibrium (as a function of the variance of the noninflationary rate of unemployment).


[^0]:    *Northwestern University and NBER

[^1]:    ${ }^{1}$ See Gelman, Carlin, Stern, and Rubin (1995) for a description of the sampling procedure from an inverse-Gamma or inverse-Wishart distributions.

