

Comments on Gianni Amisano and Oreste
Tristani
**Euro-area inflation persistence
in an estimated nonlinear DSGE model**

Michael Dueker

Federal Reserve Bank of St. Louis

Culling the herd: A Markov switching analogy

With the Kalman filter or a typical Markov switching filter, one filters the unobserved state vector forward and is able to integrate out the previous period's values of the state vector from the current period's density function in an elegant fashion.

- In some Markov switching models, this is not possible because $P(S_t = 1 | I_{t-1})$ depends on all past values of the state variable S , e.g., Markov switching in a GARCH proces. Technically we would need to consider 2^T combinations of past values of the state variable to evaluate the likelihood for observation T .

Integrating out past values of state variables

- In these cases, the collapsing procedure for integrating out the past state variables as we proceed through the filtering causes the likelihood function to be evaluated subject to an approximation. The densities, f_{it} are, in effect,

$$f(h(S_t = i)) = f(P(S_{t-1} = 0 | I_{t-1})h(S_t = i, S_{t-1} = 0) + P(S_{t-1} = 1 | I_{t-1})h(S_t = i, S_{t-1} = 1))$$

and it is this passing of an integration operator through the nonlinear density function that brings the approximation error. The path-dependency problem is not solved cleanly as the collapsing procedure induces approximation error. But the approximate likelihood can be evaluated without being a function of all past values of the state variable, i.e., the collapsing procedure “culls the herd.”

Culling the herd

- A similar “culling of the herd” takes place in the particle filtering here. In general we can write a likelihood function for a variable that depends on unobserved state variables as

$$p(y_t | I_{t-1}) = \int \int \int \cdots \int p(y_t | S_t) p(S_t | S_{t-1}, I_{t-1}) \\ p(S_{t-2} | S_{t-3}, I_{t-1}) \cdots p(S_1 | S_0, I_{t-1}) \\ dS_t dS_{t-1} dS_{t-2} \cdots dS_1 \quad (1)$$

As in the GARCH with Markov switching example, the particle filter ‘freezes’ the inference of past states by substituting

$$p(S_{t-2} | S_{t-3}, I_{t-2}) \text{ for } p(S_{t-2} | S_{t-3}, I_{t-1}).$$

Culling the herd, cont'd

- That is, we do not go back and update our inferences of past values of the state variable when calculating the likelihood of today's data; we hold those inferences fixed.
- Some papers treat the particle filter likelihood as the true value and see how closely the linearized model matches this 'true' likelihood. This might overstate the extent to which the particle filtered likelihood is the 'truth.'
- No discussion appears of a trade-off between approximation error from linearizing the model and approximation error from integrating out past values of the state variables: mis-specification from linearizing vs. mis-filtering through the collapsing procedure.

M-H sampling algorithm

- The random-walk M-H sampler is very popular among DSGE models.
- This algorithm centers the proposal draw of the parameter vector at the current value.
- This proposal density tends to lower the acceptance rate and raises the no. of MCMC iterations needed to sample the full parameter space. When the particle filtering procedure is computationally intensive, it is worthwhile to reduce the number of MCMC iterations needed to produce satisfactory estimates.
- Belaygorod and Dueker (2006) follow a suggestion by Chib and Greenberg (1994) that a single ‘tailored’ proposal density (with a fixed mean) be used as an independence proposal generator for all draws. They obtain very low inefficiency factors across parameters in an estimated DSGE model.

Previous literature

- The paper does not cite Fernández-Villaverde and Rubio-Ramírez (J. Appl. Econometrics, 2005) when it claims to be the first paper to take a DSGE with nonlinear likelihood to real data.

Nonlinearity or heterogeneity?

- I am very sympathetic with the idea that different initial conditions lead to different impulse responses. But the differences in initial conditions coincide with differing levels of heterogeneity across Euroland monetary policies. Therefore, it is hard to tell which force is greater in driving the results.
- Could Model 1 have been written with $(1 - \rho_I)(\pi^* - \ln\beta)$ in the interest rate rule? This difference seems to be driving the result, even though the discussion of the two models centers on the random-walk nature of the inflation target. In fact, with the estimate of ρ_π so high, this property barely differs across the two models.
- It would be more helpful to characterize Model 1 as a fixed endpoint model in the parlance of Kozicki and Tinsley and Model 2 as a moving endpoint model.