Online Appendix for "Changes in Federal Reserve Preferences"

Aeimit Lakdawala *

Department of Economics

Michigan State University

May 2016

*aeimit@msu.edu

1 Time-varying inflation target

A alternative approach to model preferences of the Fed is to consider a framework where the Fed's inflation target rather than the weight parameter changes over time. This has been used by Ireland (2007) and Liu et al. (2011) among others. Here, I argue that the time-varying weight approach is more compelling and give a detailed explanation of its relation to a time-varying inflation target within the context of this model.

First, it is important to clarify what the inflation target means in this model's framework and its relation to the measure of inflation targets in other studies. From a loss function framework, an interpretation of the inflation target is the level of inflation that the central bank would want if the other variables in the loss function were equal to their targets, i.e. it is π^* in the loss function function in (**??**). This implies that if the output gap is zero then the central bank would want inflation equal to the inflation target.[1] I refer to π^* as the unconditional inflation target. This is distinct from the level of inflation that the central bank actually chooses (either directly, as in Sargent et al. (2006) and others, or indirectly by setting the interest rate, as in this paper). I will refer to this second concept as the conditional inflation target. Unfortunately these two concepts have been used interchangeably in the literature.[2] The important point is that the model in this paper allows for a time-varying conditional inflation target. The central bank may choose a level of inflation higher or lower than its unconditional target depending on the state of the economy.

Second, I argue that a framework which allows for large changes in the unconditional inflation target (π^*) may not be the best way to think about Federal Reserve behavior. For example, this would imply that the Federal Reserve would have wanted high inflation in the 1970s even if the output gap was zero. The following quote from Meltzer (2006) about the thoughts of then chairman Arthur Burns corroborates this view and the framework of time-varying preferences in general.

[1]The interest rate smoothing term is ignored here for ease of exposition.

[2]For example, in the abstract of Sargent et al. (2006), the authors refer to an "inflation target" which corresponds to the conditional inflation target as define here. While on page 1197 they refer to π^* as the "targeted level of inflation".

> "During the Great Inflation, the Federal Reserve also held the view that more than a modest increase in unemployment, even if temporary, was unacceptable as a way of reducing inflation. As Burns said, in principle, the Federal Reserve could have slowed money growth to end inflation at any time. In practice, it reduced its independence by acceding to the fashion that interpreted the Employment Act as giving greater weight to unemployment and lesser weight to inflation."

I think the issue is that the literature using Taylor-type rules does not model optimal central bank behavior and thus it is not clear how to (and often not important to) differentiate between a central bank primitive (unconditional inflation target) and something the central bank can affect (conditional inflation target). But the difference is crucial when analyzing the motivation behind Federal Reserve actions. In the framework of this paper, even with a low unconditional inflation target the central bank can choose a high rate of inflation depending upon, among other things, where the variables in the loss function are relative to their targets and the preference parameters. Thus high inflation in the 1970s can be compatible with a low Federal Reserve unconditional inflation target.

Finally, recall that in this framework a change in the unconditional inflation target would only change the intercept term in the optimal interest rate rule. The coefficients (response to inflation, output gap and lagged interest rates) are unaffected by the unconditional inflation target, while a change in the preference parameter affects both the intercept and the coefficients. A similar point is made by Nelson (2005).[3] Given the evidence supporting time variation in the coefficients of the Fed's reaction function it seems reasonable to have a model that allows changes in Fed preferences to affect these coefficients.

[3] This quote from page 9 suggests that a change in the inflation target should show up in the intercept of the monetary policy rule rather than the coefficients: "If the rise in inflation in the 1970s reflected a shift to a higher inflation target, it should imply an interest-rate rule with a sizable intercept term together with a greater than one-for-one response to deviations of inflation from target..."

2 Bayesian MCMC: Metropolis Hastings

This section explains the Bayesian estimation procedure that is used to estimate all the parameters of the model jointly. Bayesian estimation involves combining the prior and the likelihood to characterize the posterior. Given the non-linear nature of the state space system (**??**) and (**??**), a non-linear filtering method is required to evaluate the likelihood. I use the Extended Kalman filter and motivate its use in Appendix D. Starting with the non-linear state space form, (**??**) and (**??**), the following equations show how to evaluate the likelihood using the EKF.

$$\begin{aligned}
\alpha_{t|t-1} &= \alpha_{t-1|t-1} \\
P_{t|t-1} &= P_{t-1|t-1} + Q \\
\eta_{t|t-1} &= y_t - h(\alpha_{t|t-1}, X_t, \Gamma, 0) \\
f_{t|t-1} &= H_t P_{t|t-1} H_t' + M_t \Sigma_t M_t' \\
\alpha_{t|t} &= \alpha_{t|t-1} + K_t \eta_{t|t-1} \\
P_{t|t} &= P_{t|t-1} - K_t H_t P_{t|t-1}
\end{aligned}$$

where

$$\begin{aligned}
H_t &= \left.\frac{\partial h()}{\partial \alpha}\right|_{\alpha_{t|t-1}} \\
M_t &= \left.\frac{\partial h()}{\partial \varepsilon}\right|_{\varepsilon_{t|t-1}} \\
K_t &= P_{t|t-1} H_t' f_{t|t-1}^{-1}
\end{aligned}$$

With the filtered values in hand the log-likelihood is evaluated using the error decomposition method

$$llf = -\frac{Tn}{2} log(2\pi) - \frac{1}{2}\sum_{t=1}^{T} log(|f_{t|t-1}|) - \frac{1}{2}\sum_{t=1}^{T} \eta_{t|t-1}' f_{t|t-1}^{-1} \eta_{t|t-1}$$

The likelihood is then combined with the prior to get the posterior. Here, the estimation follows a two step procedure. In the first step we numerically maximize the log posterior distribution to get an estimate of the posterior mode. In the second step, using the posterior mode calculated in the first step as a starting value, we use the Metropolis-Hastings algorithm to completely characterize the posterior distribution. Let θ be the parameters to be estimated. The M-H algorithm involves generating a draw from a candidate generating density, $q(.)$. Let this candidate draw be called $\theta^{(g+1)}$. Then this new draw is accepted with the following probability.

$$\alpha(\theta^{(g+1)}, \theta^{(g)}) = min\left(\frac{p(\theta^{(g+1)}|Y).q(\theta^{(g)})}{p(\theta^{(g)}|Y).q(\theta^{(g+1)})}, 1\right)$$

As is standard in this literature I use the inverse of the Hessian at the posterior mode (that comes out of the numerical optimization procedure) as the candidate generating density which is centered around the current draw $\theta^{(g)}$.

$$\delta^{(g+1)} = \delta^{(g)} + c\tilde{H}^{-1}$$

where c is a scale factor and $\tilde{H}$ is the Hessian at the posterior mode. This is known as a random walk Metropolis-Hastings step. I then tune the parameter c to get an acceptance rate of between 25% and 35%. The full parameter vector θ is sampled in one block. I have also tried blocking by splitting the parameter vector θ into 2 or more blocks but found that the Metropolis-Hastings algorithm ran most efficiently with one block and had good convergence properties. The RWMH is run for 1.25 million draws with a burn-in of sample of 250,000 (this means that the first 250,000 draws are discarded). Then I use a thinning factor of 10 (thinning factor of i means that only the i^{th} draws are stored). This gives an effective number of draws equal to 100,000. I perform numerous convergence checks to ensure that there is satisfactory convergence.[4]

[4]Some of the convergence checks include checking that 20^{th} order autocorrelations are low enough, there are enough draws as recommended by Raftery and Lewis (1992) method and that the inefficiency factors (the inverse of the relative numerical efficiency of Geweke (1992)) are low enough.

3 Justification for Extended Kalman Filter

I start by reproducing the non-linear state space form of the model represented by equations

$$y_t = h(\alpha_t, X_t, \Gamma, \varepsilon_t) \quad \text{where } \varepsilon_t \sim N(0, \Sigma_t)$$
$$\alpha_{t+1} = \alpha_t + v_{t+1} \qquad \text{where } v_t \sim N(0, Q)$$

α_t is the time-varying preference parameter, X_t is lagged data, Γ contain the constant parameters and ε_t and v_t are the shocks. Here the errors are normal but the observation equation is not linear. Specifically since the unobservable variable α_t enters non-linearly we cannot use the standard Kalman Filter. The Extended Kalman Filter deals with the non-linearity by taking a first order Taylor expansion around the current filtered estimate and then uses the regular Kalman Filter recursions. As long as the non-linearity is not severe the EKF gives a good approximation to the optimal estimate. Thus if the function $h(.)$ is not too nonlinear in α_t then the EKF will perform well.

Remember in this model α_t only appears in the interest rate equation. Fixing the constant parameters Γ at their posterior means, Figure 1 plots the coefficients of the interest rate equation as functions of α_t, these are the $F_{i,t}$ from equation (**??**). The range of α_t on the x-axis includes the maximum and the minimum of the estimated values of α_t . It is apparent that the non-linearities are indeed not very severe and thus the EKF should be a good approximation.

To further confirm this result I have computed filtered and smoothed values using two other popular non-linear filters, the Unscented Kalman Filter (UKF) and the Particle Filter (PF). Again in this exercise I fix the values of the rest of the parameters at their posterior means. The filtered and smoothed values from these alternative filters are very similar to the Extended Kalman Filter. This is not surprising as these two filters tend to perform better in more extreme nonlinear and non-normal situations. Thus the EKF, UKF and PF give very similar results in this situation.

References

Geweke, J. (1992). Evaluating the Accuracy of Sampling-Based Approaches to the Calculation of Posterior Moments. In *Bayesian Statistics*.

Ireland, P. N. (2007). Changes in the Federal Reserve's Inflation Target: Causes and Consequences. *Journal of Money, Credit and Banking 39*(8), 1851–1882.

Liu, Z., D. F. Waggoner, and T. Zha (2011). Sources of Macroeconomic Fluctuations: A Regime-switching DSGE Approach. *Quantitative Economics 2*(2), 251–301.

Meltzer, A. H. (2006). From Inflation to More Inflation, Disinflation, and Low Inflation. *The American Economic Review 96*(2), pp. 185–188.

Nelson, E. (2005). The Great Inflation of the Seventies: What Really Happened? *Advances in Macroeconomics 5*(3).

Raftery, A. E. and S. Lewis (1992). How Many Iterations in the Gibbs Sampler? In *In Bayesian Statistics 4*, pp. 763–773. Oxford University Press.

Sargent, T., N. Williams, and T. Zha (2006). Shocks and Government Beliefs: The Rise and Fall of American Inflation. *American Economic Review 96*(4), 1193–1224.

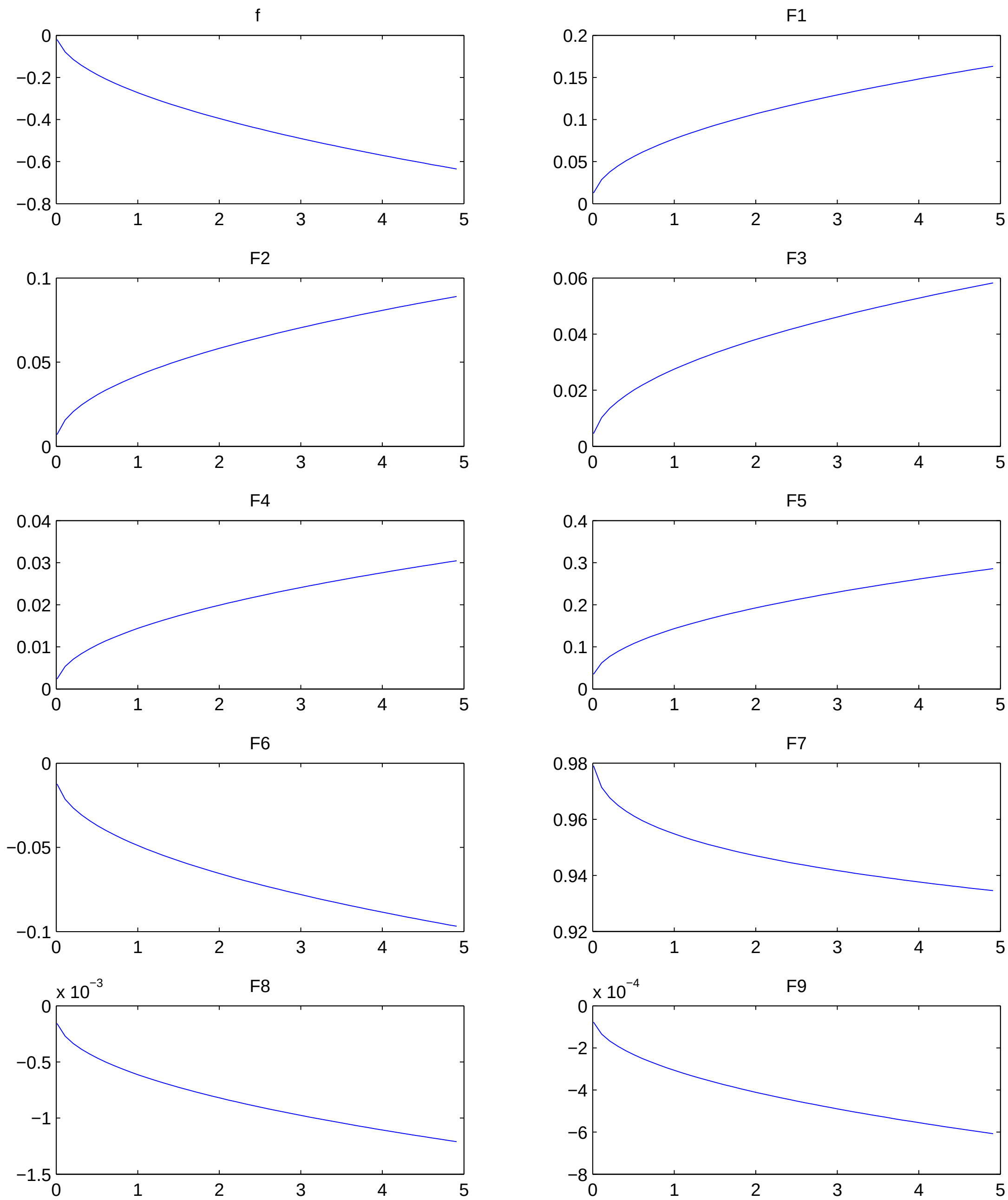

Figure 1: Coefficients of the optimal interest rate equation as functions of the time-varying preference parameter