Term Paper: On the Policy Preferences of the US Federal Reserve

John Hejkal
6E:222 Spring 2006

July 31, 2006

1 Introduction

I replicated portions of "The Policy Preferences of the US Federal Reserve" by Richard Dennis (2006) and used it as a basis for two related extensions.[1] In Dennis (2006), hereafter Dennis, a loss function is ascribed to the Fed, which depends on deviations from a target inflation rate, a target interest rate, and potential GDP. The Fed minimizes loss subject to the constraints of a linear economy.

I perform two extensions, both dealing with the output gap component of the Fed's objective function. First, I use a measure of the "unemployment rate gap" instead of the output gap in estimating what is otherwise the same model as Dennis uses. Second, I replace the Congressional Budget Office's estimate of the output gap with the cyclical component of real GDP, using the Hodrick-Prescott filter.

[1] The paper can be found on the internet at http://www3.interscience.wiley.com/cgi-bin/fulltext/110525961/PDFSTART)

I use maximum likelihood to estimate the Federal Reserves policy preference parameters. I learned a great deal about how to perform ML estimation in practice. I also learned about the importance of closely examining the results of a paper, and choosing data appropriately. I learned how to set up and solve maximization problems posed in a linear quadratic form.

2 Motivation

The original motivation for Dennis is that estimated policy rules (without any objective function for the Fed) match the data well, but do not provide information about policy formulation as such (see Clarida, *et al.* 2000). Such models are also unable to provide information about an implied inflation target. On the other hand, studies of optimal policy have usually been parameterized independently of data (according to Dennis), and hence perform poorly when trying to match data. Dennis parameterizes the Fed's loss function subject to the constraints of a linear economy and subject to the data. In this way, the Fed's inflation target can be estimated.

However, he finds, as has other research (e.g. Castelnuovo and Surico (2004)), that the Fed's preference parameter for keeping the output gap small cannot be distinguished from zero. According to Söderlind *et al.* (2002), the high volatility in output, compared to the volatility in inflation, requires a low or zero weight to be placed on the reduction of the output gap. Dennis uses potential real GDP as calculated by the Congressional Budget Office as the Fed's implied target level of real GDP. I will see if his result is robust to the choice of some alternative targets.

One of the main objectives of Dennis is to find the implied inflation target of the Fed. Dennis reports the point estimate, but he does not report the standard error of this estimate. I find a high standard error in my replication, so it is of interest to see whether alternative data will bring about a better estimate for the inflation target.

3 Data

The data for Dennis consist of the annualized quarterly inflation rate, π_t, the output gap, y_t, and the annualized quarterly federal funds rate, i_t. That is, $y_t = 100 \times \log(Y_t/Y_t^p)$, where Y_t is real GDP and Y_t^p is potential GDP as measured by the Congressional Budget Office, and $\pi_t = 400 \times \log(P_t/P_{t-1})$, where P_t is the GDP chain-weighted price index. I also define an average of past inflation $\pi_t^a = \frac{1}{4}\sum_{j=0}^{4} \pi_{t-j}$ and of the federal funds rate $i_t^a = \frac{1}{4}\sum_{j=0}^{4} i_{t-j}$.

For my extension I introduce u_t, which is defined as $u_t = U_t^{nr} - U_t$, where U_t^{nr} is the CBO's natural rate of unemployment and U_t is the actual unemployment rate quarter t as measured by the Bureau of Labor Statistics. I also introduce y_t^c as the cyclical component of real GDP that has been HP-filtered with a "smoothing parameter" of 1600. Following the construction Dennis uses for the output gap, I define $y_t^c = 100\log(Y_t/Y_t^T)$, where Y_t^T is the trend component of the filtered real GDP data, and Y_t is the observed value of real GDP.

4 Reproduction of the Original Study

The loss function Dennis attributes to the Federal Reserve is

$$\text{Loss} = E_t \sum_{j=0}^{\infty} \beta^j [(\pi_{t+j}^a - \pi^*)^2 + \lambda y_{t+j}^2 + \nu(i_{t+j} - i_{t+j-1})^2] \qquad (1)$$

Where π^* is the target inflation rate, λ is the weight placed on the output gap, and ν is the weight on interest rate smoothing. β is a discount factor, set to 0.99 in Dennis and this paper. All other variables are defined in Section 3.

The Fed faces the constraints of a linear economy proposed in Rudebusch and Svensson (1999).

$$y_t = a_0 + a_1 y_{t-1} + a_2 y_{t-2} + a_3[i^a_{t-1} - \pi^a_{t-1}] + g_t \tag{2}$$

$$\pi_t = b_0 + b_1\pi_{t-1} + b_2\pi_{t-2} + b_3\pi_{t-3} + (1 - b_1 - b_2 - b_3)\pi_{t-4} + b_4 y_{t-1} + v_t \tag{3}$$

where g_t and v_t are disturbance terms, which Dennis interprets as demand shocks and supply shocks, respectively.

Before proceeding with an optimization-based estimation, Dennis simply estimates equations (2) and (3) taking the federal funds rate as given. An SUR is used, taking into account possible correlation between g_t and v_t. The results of this estimation are shown shown in Table I in Dennis. I am able to replicate the results exactly, to the number of digits reported in Dennis. I reproduce Table I in Dennis below, in my own Table 1. The data used are from 1966Q1-2000Q2.

Table 1

Parameter	Point Est.	Std. Err.		Parameter	Point Est.	Std. Err.
a_0	0.157	0.110		b_0	0.051	0.088
a_1	1.208	0.080		b_1	0.638	0.084
a_2	-0.292	0.079		b_2	0.023	0.100
a_3	-0.067	0.031		b_3	0.186	0.100
				b_4	0.146	0.035
σ^2_g	0.639			σ^2_π	1.054	

4.1 State-Space Form

To solve for optimal policy, the constraints must be in state-space form. Following the notation in Dennis, we have

$$\mathbf{z}_{t+1} = \mathbf{C} + \mathbf{A}\mathbf{z}_t + \mathbf{B}\mathbf{x}_t + \mathbf{u}_{t+1} \tag{4}$$

where $\mathbf{z}_t = [\pi_t \quad \pi_{t-1} \quad \pi_{t-2} \quad \pi_{t-3} \quad y_t \quad y_{t-1} \quad i_{t-1} \quad i_{t-2} \quad i_{t-3}]'$ is the state vector, $\mathbf{x}_t = i_t$ is the control variable, and $\mathbf{u}_{t+1} = [v_{t+1}\ 0\ 0\ 0\ g_{t+1}\ 0\ 0\ 0\ 0]'$ is the shock vector, with variance-covariance matrix $\mathbf{\Sigma}$. Referring to equations (2) and (3), we have

$$\mathbf{C} = [b_0 \quad 0 \quad 0 \quad 0 \quad a_0 \quad 0 \quad 0 \quad 0 \quad 0]'$$
$$\mathbf{B} = [0 \quad 0 \quad 0 \quad 0 \quad a_3/4 \quad 0 \quad 1 \quad 0 \quad 0]'$$

and

$$\mathbf{A} = \begin{bmatrix} b_1 & b_2 & b_3 & 1-b_1-b_2-b_3 & b_4 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{-a_3}{4} & \frac{-a_3}{4} & \frac{-a_3}{4} & \frac{-a_3}{4} & a_1 & a_2 & \frac{a_3}{4} & \frac{a_3}{4} & \frac{-a_3}{4} \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

Now I write the loss function in terms of $\mathbf{x}$ and $\mathbf{y}$.

$$E_t \sum_{j=0}^{\infty} \beta^j [(\mathbf{z}_{t+j} - \bar{\mathbf{z}})'\mathbf{W}(\mathbf{z}_{t+j} - \bar{\mathbf{z}}) + (\mathbf{x}_{t+j} - \bar{\mathbf{z}})'\mathbf{Q}(\mathbf{z}_{t+j} - \bar{\mathbf{z}})$$
$$+2(\mathbf{z}_{t+j} - \bar{\mathbf{z}})'\mathbf{H}(\mathbf{x}_{t+j} - \bar{\mathbf{x}}) + 2(\mathbf{x}_{t+j} - \bar{\mathbf{x}})'\mathbf{G}(\mathbf{z}_{t+j} - \bar{\mathbf{z}})]$$

which represents the loss function if I appropriately define $\mathbf{W}$, $\mathbf{Q}$, $\mathbf{H}$, and $\mathbf{G}$ to contain the policy parameters. $\bar{\mathbf{x}}$ and $\bar{\mathbf{z}}$ are the target vectors for the policy and state, respectively.

Set $\mathbf{Q} = \nu$, and $\mathbf{H}' = \mathbf{G} = [\mathbf{0}_{1\times 6} \quad -\nu/2 \quad \mathbf{0}_{1\times 2}]$. $\mathbf{W}$ contains penalty terms for state variables, so define matrices $\mathbf{P}$ and $\mathbf{R}$ as

$$\mathbf{P} = \begin{pmatrix} 1/4 & 1/4 & 1/4 & 1/4 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}$$

$$\mathbf{R} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \nu \end{pmatrix}.$$

Then $\mathbf{W} = \mathbf{P}'\mathbf{R}\mathbf{P}$. The optimal policy rule is

$$\mathbf{x}_t = \mathbf{f} + \mathbf{F}\mathbf{z}_t \tag{5}$$

where

$$\mathbf{f} = \bar{\mathbf{x}} - \mathbf{F}\bar{\mathbf{z}} \tag{6}$$

$$\mathbf{F} = -(\mathbf{Q} + \beta\mathbf{B}'\mathbf{M}\mathbf{B})^{-1}(\mathbf{H}' + \mathbf{G} + \beta\mathbf{B}'\mathbf{M}\mathbf{A}) \tag{7}$$

$$\mathbf{M} = \mathbf{W} + \mathbf{F}'\mathbf{Q}\mathbf{F} + 2\mathbf{H}\mathbf{F} + 2\mathbf{F}'\mathbf{G} + \beta(\mathbf{A} + \mathbf{B}\mathbf{F})'\mathbf{M}(\mathbf{A} + \mathbf{B}\mathbf{F}) \tag{8}$$

One method of solving the system is to use Gauss-Seidel on equations (7) and (8) to find $\mathbf{F}$ and $\mathbf{M}$. Noting that $\bar{\mathbf{z}} = [\pi^* \; \pi^* \; \pi^* \; \pi^* \; 0 \; 0 \; \bar{\mathbf{x}} \; \bar{\mathbf{x}} \; \bar{\mathbf{x}}]$, it is possible to solve for $\mathbf{f}$. Thus it has been shown that the system under optimal policy can be written as

$$\mathbf{z}_{t+1} = \mathbf{C} + \mathbf{A}\mathbf{z}_t + \mathbf{B}\mathbf{x}_t + \mathbf{u}_{t+1} \tag{9}$$

$$\mathbf{x}_t = \mathbf{f} + \mathbf{F}\mathbf{z}_t \tag{10}$$

4.2 Estimation

Once the problem above has been solved, the solution can be expressed as equations (2) and (3) and

$$i_t = f + F_1\pi_t + F_2\pi_{t-1} + F_3\pi_{t-2} + F_4\pi_{t-3} + F_5 y_t + F_6 y_{t-1} + F_7 i_{t-1} + F_8 i_{t-2} + F_9 i_{t-3} \tag{11}$$

Now define

$$\mathbf{d}_t = [\pi_t \quad y_t \quad i_t]', \quad \varepsilon_t = [v_t \quad g_t \quad \omega_t]'$$

where ω_t, while not part of the optimal policy rule, is a required addition (appended to equation (11)) to make the covariance matrix of error terms nonsingular in the ML estimation below.

Define also

$$\mathbf{A}_0 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -F_1 & -F_5 & 1 \end{pmatrix}, \mathbf{A}_2 = \begin{pmatrix} b_1 & b_4 & 0 \\ -a_3/4 & a_1 & a_3/4 \\ F_2 & F_6 & F_7 \end{pmatrix}, \mathbf{A}_3 = \begin{pmatrix} b_2 & 0 & 0 \\ -a_3/4 & a_2 & a_3/4 \\ F_3 & 0 & F_8 \end{pmatrix}$$

$$\mathbf{A}_4 = \begin{pmatrix} b_3 & 0 & 0 \\ -a_3/4 & 0 & a_3/4 \\ F_4 & 0 & F_9 \end{pmatrix}, \mathbf{A}_5 = \begin{pmatrix} 1 - b_1 - b_2 - b_3 & 0 & 0 \\ -a_3/4 & 0 & a_3/4 \\ 0 & 0 & 0 \end{pmatrix}, \mathbf{A}_1 = \begin{pmatrix} b_0 \\ a_0 \\ f \end{pmatrix}$$

Then write the model as

$$\mathbf{A}_0\mathbf{d}_t = \mathbf{A}_1 + \mathbf{A}_2\mathbf{d}_{t-1} + \mathbf{A}_3\mathbf{d}_{t-2} + \mathbf{A}_4\mathbf{d}_{t-3} + \mathbf{A}_5\mathbf{d}_{t-4} + \varepsilon_t \tag{12}$$

The data have joint probability density function

$$P(\{\mathbf{d}_t\}_1^T; \theta, \mathbf{\Psi}) = P(\{\mathbf{d}_t\}_5^T | \{\mathbf{d}_t\}_1^4; \theta, \mathbf{\Psi}) P(\{\mathbf{d}_t\}_1^4; \theta, \mathbf{\Psi})$$

where T is sample size, $\mathbf{\Psi}$ is the variance-covariance matrix of ε, and $\theta = [a_0, a_1, a_2, a_3, a_4, b_0, b_1, b_2, b_3, b_4, \lambda, \nu, \pi^*]$. Assume $\varepsilon_t | \{\mathbf{d}_j\}_1^t \sim N(0, \mathbf{\Psi})$ for all t, and that $\{\mathbf{d}_t\}_1^4$ are fixed initial conditions. The latter assumption means the last term in the above equation can be ignored when maximizing the likelihood function. Skipping the statement of the joint PDF itself, the log-likelihood function is

$$\ln L(\theta, \mathbf{\Psi}; \{\mathbf{d}_t\}_1^T) \propto -\frac{n(T-4)}{2} \ln(2\pi) + (T-4) \ln |\mathbf{A}_0| - \frac{T-4}{2} \ln |\mathbf{\Psi}| - \frac{1}{2} \sum_{t=5}^{T} (\varepsilon_t' \mathbf{\Psi}^{-1} \varepsilon_t)$$

The MLE of $\mathbf{\Psi}$ is $\hat{\mathbf{\Psi}}(\theta) = \sum_{t=5}^{T} \frac{\hat{\varepsilon}_t \hat{\varepsilon}_t'}{T-4}$. Using this, and dropping the first (constant) term, the concentrated log-likelihood function is

$$\ln L_c(\theta, \{\mathbf{d}\}_1^T) \propto (T-4)\ln|\mathbf{A}_0| - \tfrac{T-4}{2}\ln|\hat{\mathbf{\Psi}}(\theta)| \tag{13}$$

which is the expression to be maximized over θ. To obtain estimates of the standard errors of the elements of $\hat{\theta}$, I follow Dennis in constructing a covariance matrix for $\hat{\theta}$ by

$$\hat{\text{var}}(\hat{\theta}) = \mathbf{H}(\theta|_{\hat{\theta}})\mathbf{G}(\theta|_{\hat{\theta}})\mathbf{H}(\theta|_{\hat{\theta}}) \tag{14}$$

where $\mathbf{H}(\theta)$ is the information matrix, i.e.

$$\mathbf{H}(\theta) = -\Big[\frac{\partial^2 \ln L_c(\theta; \{\mathbf{d}_t\}_1^T)}{\partial\theta\partial\theta'}\Big]^{-1}$$

and

$$\mathbf{G}(\theta) = \frac{1}{T-4}\sum_{t=5}^{T}\Big(\frac{\partial \ln L_c^t(\theta; \{\mathbf{d}_j\}_{t-4}^t)}{\partial\theta}\frac{\partial \ln L_c^t(\theta; \{\mathbf{d}_j\}_{t-4}^t)}{\partial\theta'}\Big)$$

is the outer-product variance estimator.[2]

4.3 Main Replication Results

The procedure described above was used to replicate Table II in Dennis, which provides the point estimates and standard errors for $\hat{\theta}$. The data run from 1982Q1-2000Q2.

[2]See also Greene (2003) pp.180-181.

Table 2

Parameter	Point Est., Dennis	SE, Dennis		Point Est., Repl.	SE, Repl.
a_0	0.035	0.098		-0.0249	0.0302
a_1	1.596	0.073		1.6120	0.0059
a_2	-0.683	0.052		-0.6927	0.0121
a_3	-0.021	0.017		-0.0132	0.0111
b_0	0.025	0.092		0.0496	0.0180
b_1	0.401	0.104		0.2367	0.1230
b_2	0.080	0.111		0.1419	0.0595
b_3	0.407	0.115		0.3763	0.0195
b_4	0.144	0.042		0.1296	0.0217
λ	2.941	5.685		7.652	3.319
ν	4.517	1.749		4.415	1.267
$\hat{\pi}^*$	1.38	–		3.557	6.5963
$\hat{r}^*$	1.66	–		1.67	–

The results of my replication attempt are not exactly the same as in Dennis. However, the point estimates for all parameters with relatively low standard errors, as reported by Dennis, are very close. The value of equation (13) is $\log L_c = -160.33$.

Dennis does not report a standard error for his estimate of $\hat{\pi}^*$. In light of my results, I expect the standard error was quite high. $\hat{r}^*$, which corresponds to $\bar{\mathbf{x}}$ above, appears robust to my replication attempt, but this seems to be a spurious outcome.

The most surprising result of my replication attempt is that λ has a relatively low standard error, allowing it to pass conventional tests for significance.[3]

[3] A word about the standard errors themselves: closed-form expressions for the gradient and Hessian are not attainable, so numerical differentiation is used. Numerical derivatives

The slightly different estimates can probably be accounted for simply from differences in programming techniques, computer software, and the like. Since this is a fairly complex maximization problem for a computer to handle, it is not surprising that estimates are not exactly the same.

5 Extension

My extension involved the replacement of the output gap, as calculated using the CBO's estimate of potential GDP, with (A) a measure of the unemployment rate gap, using the CBO's estimate of the natural rate of unemployment, and (B) the cyclical component of GDP found by applying the Hodrick-Prescott filter. Extension (A) was motivated by seeing that the insignificance of λ was attributed to the high volatility of the output gap. Since the unemployment rate is less volatile that real GDP, the unemployment gap (the deviation of the real unemployment rate from the "natural" rate) seemed a reasonable choice. Extension (B) was motivated by the possibility that the Fed is concerned primarily with fluctuations of GDP at business cycle frequencies. Obviously, GDP fluctuates for reasons other than the business cycle. The focus on business cycle frequencies is the standard reason for using the HP filter.

6 Results

Table 3 gives the results from extensions (A) and (B). For extension (A), the series $\{y_t\}_1^T$ is replaced with $\{u_t\}_1^T$. For extension (B), the series $\{y_t\}_1^T$ is

are frequently unreliable, and may not be reliable here. Consider, I was able to maximize the concentrated log-likelihood function easily with Matlab's `fminsearch`, which uses direct search methods on a simplex. Attempts to maximize the same function using `fminunc`, which uses gradient-based methods, utterly failed.

replaced with $\{y_t^c\}_1^T$. Please refer to Section 3 for details on these series.

Table 3

Parameter	Point Est., u_t	SE, u_t		Point Est., y_t^c.	SE, y_t^c
a_0	0.0122	0.0013		-0.0051	0.0135
a_1	1.8217	0.1475		1.5020	0.0264
a_2	-0.8643	0.1486		-0.5981	0.0020
a_3	-0.0068	0.0002		-0.0135	0.0018
b_0	0.0637	0.1112		0.0462	0.0988
b_1	0.3381	0.0490		0.3034	0.0852
b_2	0.3757	0.4205		0.0907	0.0396
b_3	0.1931	0.1591		0.3452	0.0090
b_4	0.2771	0.2133		0.2535	0.0879
λ	3.1720	12.0492		11.2811	16.5438
ν	6.6218	3.8404		12.2302	5.2610
$\hat{\pi}^*$	-3.8338	0.6613		0.4705	1.7869
$\hat{r}^*$	-2.06	–		0.0948	–

The value of the concentrated log-likelihood function in each case is $\log L_{c,u} = -93.6428$ and $\log L_{c,y^c} = -168.29$. It is clear that the unemployment gap model does not produce economically meaningful results. The oddest result is that π^* is both unbelievable (being negative) and possesses a low standard error.

On the other hand, the results of the HP cyclical-component model look broadly similar to the results found in Dennis. My results lend further support to the claim that the Fed does not attempt to correct for deviations of output from trend.

7 Conclusion

In writing this paper, I learned how to solve a maximum-likelihood problem with what amounts to a complicated system of constraints. I also learned about the theory and practice of linear-quadratic forms, and why they are so useful.

In doing further work, it would perhaps be useful to calculate more reliable numerical derivatives, if possible. It would also be worthwhile to seek out an alternative description of the economy, i.e. use a different set of constraints from equations (2) and (3).

Bibliography

Castelnuovo E, Surico P. 2004. Model Uncertainty, Optimal Monetary Policy and the Preferences of the Fed. *Scottish Journal of Political Economy* **51**:105-126

Clarida R, Galí J, Gertler M. 2000. Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory. *Quarterly Journal of Economics* **115**: 147-180

Congressional Budget Office. 2001. CBO'S Method For Estimating Potential Output: An Update. `http://www.cbo.gov/showdoc.cfm?index=3020&sequence=0#N_`

Dennis, R. 2006. The Policy Preferences of the US Federal Reserve. *Journal of Applied Econometrics* **21**: 55-77.

Greene, W. 2003. Econometric Analysis, 5th ed. Pearson Education, Inc.

Rudebusch G, Svensson L. 1999 Policy Rules for Inflation Targeting. In *Monetary Policy Rules*, Taylor J. (ed.). University of Chicago Press: Chicago.

Söderlind P., Söderström U., Vredin A. 2002. Can Calibrated New-Keynesian Models of Monetary Policy Fit the Facts? Sveriges Riksbank Working Paper No. 140.