Online Appendix (Not intended for Publication): Federal Reserve Credibility and the Term Structure of Interest Rates

Aeimit Lakdawala
Michigan State University

Shu Wu
University of Kansas

August 2017

1 Derivation of Bond Yields

1.1 No regime switching

The rational expectation solution to the DSGE model can be described by the following equation:

$$\xi_t = F\xi_{t-1} + Gv_t \tag{1}$$

$$v_t \sim N(0, Q) \tag{2}$$

Without loss of generality we can write the log of the stochastic discount factor m_{t+1} as,

$$m_{t+1} = -\lambda_0 - \lambda_1' \xi_{t+1} - \lambda_2' \xi_t \tag{3}$$

Let $P_t^{(n)}$ denote the price of a n-period zero-coupon bond at time t. From the no-arbitrage asset pricing equations

$$P_t^{(n)} = E_t \left[M_{t+1} P_{t+1}^{(n-1)} \right] \tag{4}$$

With the joint lognormal distributions we can write this as (lowercase letters denote logs)

$$P_t^{(n)} = exp \left\{ E_t m_{t+1} + E_t p_{t+1}^{(n-1)} + \frac{1}{2} Var_t m_{t+1} + \frac{1}{2} Var_t p_{t+1}^{(n-1)} + Cov_t(m_{t+1}, p_{t+1}^{(n-1)}) \right\} \tag{5}$$

using the property that if a random variable $X \sim logN(\mu, \sigma^2)$ then $logX \sim (\mu, \sigma^2)$. Moreover, $E[X] = exp\{\mu + (1/2)\sigma^2\} = exp\{E(logX) + (1/2)Var(logX)\}$ Assuming an affine structure for log bond prices gives

$$p_t^{(n)} = -A_n - B_n' \xi_t \tag{6}$$

Using the above equations we can derive

$$\begin{aligned}
E_t m_{t+1} &= -\lambda_0 - \lambda_1' F \xi_t - \lambda_2' \xi_t \\
E_t p_{t+1}^{(n-1)} &= -A_{n-1} - B_{n-1}' F \xi_t \\
Var_t m_{t+1} &= \lambda_1' G Q G' \lambda_1 \\
Var_t p_{t+1}^{(n-1)} &= B_{n-1}' G Q G' B_{n-1} \\
Cov_t(m_{t+1}, p_{t+1}^{(n-1)}) &= \lambda_1' G Q G' B_{n-1}
\end{aligned}$$

Taking log of equation (5) and plugging in the above gives

$$-A_n - B_n'\xi_t = -\lambda_0 - \lambda_1' F\xi_t - \lambda_2'\xi_t - A_{n-1} - B_{n-1}' F\xi_t$$
$$+\tfrac{1}{2}\left(\lambda_1' GQG'\lambda_1\right) + \tfrac{1}{2}\left(B_{n-1}' GQG' B_{n-1}\right) + \lambda_1' GQG' B_{n-1}$$

Matching coefficients on the constant and noting that $\lambda_1' GQG' B_{n-1} = B_{n-1}' GQG'\lambda_1 = \frac{1}{2}\left(\lambda_1' GQG' B_{n-1} + B_{n-1}' GQG'\lambda_1\right)$

$$\begin{aligned} A_n &= \lambda_0 + A_{n-1} - \frac{1}{2}\left(\lambda_1' GQG'\lambda_1\right) - \frac{1}{2}\left(B_{n-1}' GQG' B_{n-1}\right) - \lambda_1' GQG' B_{n-1} && (7)\\ &= \lambda_0 + A_{n-1} - \frac{1}{2}\left(\lambda_1' GQG'\lambda_1\right) - \frac{1}{2}\left(B_{n-1}' GQG' B_{n-1}\right) - \frac{1}{2}\left(\lambda_1' GQG' B_{n-1}\right) - \frac{1}{2}\left(B_{n-1}' GQG'\lambda_1\right) && (8)\\ &= \lambda_0 + A_{n-1} - \frac{1}{2}(\lambda_1 + B_{n-1})' GQG'(\lambda_1 + B_{n-1}) && (9) \end{aligned}$$

Matching coefficients on ξ_t we get

$$B_n = F' B_{n-1} + F'\lambda_1 + \lambda_2 \tag{10}$$

1.2 Regime switching: Loose commitment and regime switching shock variances

In this case the dynamics of state variables can be described by the following equation:

$$\begin{aligned} \xi_t &= F_{s_t}\xi_{t-1} + Gv_t && (11)\\ v_t &\sim N(0, Q_{s_t^v}) && (12) \end{aligned}$$

where s_t (with transition matrix P) governs the loose commitment switches and s_t^v (with transition matrix P^v) governs the switches in the shock variance, each being a two-state Markov chain. Let $\widetilde{s}_t$ be a composite regime indicator with 4 regimes and a transition matrix $\widetilde{P} = P \otimes P^v$

Now consider the same setup for the stochastic discount factor

$$m_{t+1} = -\lambda_0 - \lambda_1'\xi_{t+1} - \lambda_2'\xi_t \tag{13}$$

Let $P_{n,t}$ denote the price of a n-period zero-coupon bond at time t. It then follows that, for

$n \geq 0$,

$$P_{n,t} = e^{-A_{n,\widetilde{s}_t} - B'_{n,\widetilde{s}_t}\xi_t} \tag{14}$$

The coefficients $A_{n,\widetilde{s}_t}$ and $B_{n,\widetilde{s}_t}$ are allowed to depend on the regime indicator $\widetilde{s}_t$. Now consider p_{t+1}^{n-1}, conditional on $\widetilde{s}_{t+1}$, the mean and variance of the log bond price can be written as

$$\begin{aligned}
E\left[m_{t+1}|I_{t+1}\right] &= -\lambda_0 - \lambda_1' F_{\widetilde{s}_{t+1}}\xi_t - \lambda_2'\xi_t && (15)\\
Var\left[m_{t+1}|I_{t+1}\right] &= \lambda_1' G Q_{\widetilde{s}_{t+1}} G'\lambda_1 && (16)\\
E\left[p_{t+1}^{(n-1)}|I_{t+1}\right] &= -A_{n-1,\widetilde{s}_{t+1}} - B_{n-1,\widetilde{s}_{t+1}} F_{\widetilde{s}_{t+1}}\xi_t && (17)\\
Var\left[p_{t+1}^{(n-1)}|I_{t+1}\right] &= B'_{n-1,\widetilde{s}_{t+1}} G Q_{\widetilde{s}_{t+1}} G' B_{n-1,\widetilde{s}_{t+1}} && (18)\\
Cov\left[(m_{t+1}, p_{t+1}^{(n-1)})|I_{t+1}\right] &= \lambda_1' G Q_{\widetilde{s}_{t+1}} G' B_{n-1} && (19)
\end{aligned}$$

where Ω_t is all available information upto time t and $I_{t+1} = \Omega_t \cup \widetilde{s}_{t+1}$.

$$\begin{aligned}
P_t^{(n)} &= E\left[M_{t+1}P_{t+1}^{(n-1)}|\Omega_t\right] && (20)\\
&= \sum_{\widetilde{s}_{t+1}} \pi_{i,\widetilde{s}_{t+1}} E\left[M_{t+1}P_{t+1}^{(n-1)}|I_{t+1}\right] && (21)\\
&= \sum_{\widetilde{s}_{t+1}} \pi_{i,\widetilde{s}_{t+1}} exp\{E\left[m_{t+1}|I_{t+1}\right] + E\left[p_{t+1}^{(n-1)}|I_{t+1}\right] + \frac{1}{2}Var\left[m_{t+1}|I_{t+1}\right] + Cov\left[(m_{t+1}, p_{t+1}^{(n-1)})|I_{t+1}\right]\} && (22)
\end{aligned}$$

Following Bansal & Zhou (2002), use the log-linear approximation $exp(x) \approx 1 + x$. The RHS of the above equation becomes

$$\sum_{\widetilde{s}_{t+1}} \pi_{i,s_{t+1}} \left(E\left[m_{t+1}|I_{t+1}\right] + E\left[p_{t+1}^{(n-1)}|I_{t+1}\right] + \frac{1}{2}Var\left[m_{t+1}|I_{t+1}\right] + Cov\left[(m_{t+1}, p_{t+1}^{(n-1)})|I_{t+1}\right] + 1\right) \tag{23}$$

Moving the 1 on the other side, the LHS becomes

$$\begin{aligned}
&= P_t^{(n)} - 1 && (24)\\
&= e^{-A_{n,\widetilde{s}_t} - B'_{n,\widetilde{s}_t}\xi_t} - 1 && (25)\\
&\approx -A_{n,\widetilde{s}_t} - B'_{n,\widetilde{s}_t}\xi_t && (26)
\end{aligned}$$

Plugging in equations (15-19) into the RHS we get

$$\sum_{\widetilde{s}_{t+1}} \pi_{i,\widetilde{s}_{t+1}}(-\lambda_0 - \lambda_1' F_{\widetilde{s}_{t+1}}\xi_t - \lambda_2'\xi_t - A_{n-1,\widetilde{s}_{t+1}} - B_{n-1,\widetilde{s}_{t+1}} F_{\widetilde{s}_{t+1}}\xi_t$$
$$+\frac{1}{2}\lambda_1' GQ_{\widetilde{s}_{t+1}} G'\lambda_1 + \frac{1}{2}B'_{n-1,\widetilde{s}_{t+1}} GQ_{s_{t+1}} G' B_{n-1,\widetilde{s}_{t+1}} + \lambda_1' GQ_{\widetilde{s}_{t+1}} G' B_{n-1})$$

Matching coefficients we get

$$\begin{pmatrix} A_{n,\widetilde{s}_t=1} \\ A_{n,\widetilde{s}_t=2} \\ A_{n,\widetilde{s}_t=3} \\ A_{n,\widetilde{s}_t=4} \end{pmatrix} = \begin{pmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \\ p_{41} & p_{42} & p_{43} & p_{44} \end{pmatrix} \begin{pmatrix} \lambda_0 + A_{n-1,\widetilde{s}_{t+1}=1} - \frac{1}{2}(\lambda_1 + B_{n-1,s_{t+1}=1})' GQ_{\widetilde{s}_{t+1}=1} G'(\lambda_1 + B_{n-1,s_{t+1}=1} \\ \lambda_0 + A_{n-1,\widetilde{s}_{t+1}=2} - \frac{1}{2}(\lambda_1 + B_{n-1,s_{t+1}=2})' GQ_{\widetilde{s}_{t+1}=2} G'(\lambda_1 + B_{n-1,s_{t+1}=2} \\ \lambda_0 + A_{n-1,\widetilde{s}_{t+1}=3} - \frac{1}{2}(\lambda_1 + B_{n-1,s_{t+1}=3})' GQ_{\widetilde{s}_{t+1}=3} G'(\lambda_1 + B_{n-1,s_{t+1}=3} \\ \lambda_0 + A_{n-1,\widetilde{s}_{t+1}=4} - \frac{1}{2}(\lambda_1 + B_{n-1,s_{t+1}=4})' GQ_{\widetilde{s}_{t+1}=4} G'(\lambda_1 + B_{n-1,s_{t+1}=4} \end{pmatrix} \tag{27}$$

$$\begin{pmatrix} B_{n,\widetilde{s}_t=1} \\ B_{n,\widetilde{s}_t=2} \\ B_{n,\widetilde{s}_t=3} \\ B_{n,\widetilde{s}_t=4} \end{pmatrix} = \begin{pmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \\ p_{41} & p_{42} & p_{43} & p_{44} \end{pmatrix} \begin{pmatrix} F'_{\widetilde{s}_{t+1}=1} B_{n-1,\widetilde{s}_{t+1}=1} + F'_{\widetilde{s}_{t+1}=1}\lambda_1 + \lambda_2 \\ F'_{\widetilde{s}_{t+1}=2} B_{n-1,\widetilde{s}_{t+1}=2} + F'_{\widetilde{s}_{t+1}=2}\lambda_1 + \lambda_2 \\ F'_{\widetilde{s}_{t+1}=3} B_{n-1,\widetilde{s}_{t+1}=3} + F'_{\widetilde{s}_{t+1}=3}\lambda_1 + \lambda_2 \\ F'_{\widetilde{s}_{t+1}=4} B_{n-1,\widetilde{s}_{t+1}=4} + F'_{\widetilde{s}_{t+1}=4}\lambda_1 + \lambda_2 \end{pmatrix} \tag{28}$$

Note for the specific setup here the transition matrices are given as follows.

$$P = \begin{bmatrix} \gamma & 1-\gamma \\ \gamma & 1-\gamma \end{bmatrix}$$
$$P^v = \begin{bmatrix} p_1 & (1-p_1) \\ (1-p_2) & p_2 \end{bmatrix}$$

The composite regime switching variable is $\widetilde{s}_t$, noting that s_t and s_t^v are independent.

$$\widetilde{s}_t = 1 \text{ if } s_t = 1 \text{ and } s_t^v = 1$$
$$\widetilde{s}_t = 2 \text{ if } s_t = 1 \text{ and } s_t^v = 2$$
$$\widetilde{s}_t = 3 \text{ if } s_t = 0 \text{ and } s_t^v = 1$$
$$\widetilde{s}_t = 4 \text{ if } s_t = 0 \text{ and } s_t^v = 2$$

1.3 Regime switching: Loose commitment

In this case the dynamics of state variables can be described by the following equation:

$$\xi_t = F_{s_t}\xi_{t-1} + Gv_t \tag{29}$$
$$v_t \sim N(0, Q) \tag{30}$$

where s_t governs the loose commitment switches with transition matrix P depending on the probability of commitment γ

$$P^m = \begin{bmatrix} \gamma & 1-\gamma \\ \gamma & 1-\gamma \end{bmatrix}$$

Thus probability that next period $s_t = 1$ (continuing plans) is γ regardless of the current state. Similarly probability that next period $s_t = 0$ (re-optimization) is $1-\gamma$. Following the same procedure above we can derive

$$\begin{pmatrix} A_{n,s_t=1} \\ A_{n,s_t=0} \end{pmatrix} = \begin{pmatrix} \gamma & 1-\gamma \\ \gamma & 1-\gamma \end{pmatrix} \begin{pmatrix} \lambda_0 + A_{n-1,s_{t+1}=1} - \frac{1}{2}(\lambda_1 + B_{n-1,s_{t+1}=1})' GQ_{s_{t+1}=1}G'(\lambda_1 + B_{n-1,s_{t+1}=1}) \\ \lambda_0 + A_{n-1,s_{t+1}=0} - \frac{1}{2}(\lambda_1 + B_{n-1,s_{t+1}=0})' GQ_{s_{t+1}=0}G'(\lambda_1 + B_{n-1,s_{t+1}=0}) \end{pmatrix} \tag{31}$$

$$\begin{pmatrix} B_{n,s_t=1} \\ B_{n,s_t=0} \end{pmatrix} = \begin{pmatrix} \gamma & 1-\gamma \\ \gamma & 1-\gamma \end{pmatrix} \begin{pmatrix} F'_{s^m_{t+1}=1} B_{n-1,s_{t+1}=1} + F'_{s_{t+1}=1}\lambda_1 + \lambda_2 \\ F'_{s^m_{t+1}=0} B_{n-1,s_{t+1}=0} + F'_{s_{t+1}=0}\lambda_1 + \lambda_2 \end{pmatrix} \tag{32}$$

Since the rows of the transition matrix are the same, we conclude that $A_{n,s_t=1} = A_{n,s_t=0}$ and $B_{n,s_t=1} = B_{n,s_t=0}$

2 Understanding the estimated probability of commitment

The estimated probability of commitment is 0.6, which is closer to discretion relative to the estimate of 0.8 obtained in the Debortoli and Lakdawala (2016) paper that does not use yield data. This difference can be understood by analyzing in a little more detail the exact source of identification in this setting. We should mention that a priori, we had no reason to expect whether this estimate would be lower or higher relative to the model without yield data. In other words, this is purely an empirical issue and the theory does not point us to a specific effect of adding yields to the structural model. In the loose commitment framework, the probability of commitment shows up in two different places. First, the probability of commitment governs the transition matrix for the re-optimization shock process, as is seen in equation 32. But agents take this into account when forming expectations and the probability of commitment also affects how agents form expectations, as is seen in equation 9. Thus intuitively, the estimation is going to choose the probability of commitment to best fit both these aspects.

To evaluate the second effect one option is to study the one-step ahead forecast errors for the macro variables from the model with yields and compare it to the model without yields. The columns titled "1Q" in Table 1 below (also now added to the online appendix) present the root mean squared forecast error for both these models. Note that the model without yields is estimated with the fed funds rate, while the model with yields replaces the fed funds rate with the 3 month Treasury bill rate. We can see that overall the one-step ahead forecast errors are of similar magnitude for both the models with forecast errors being slightly smaller for the model without yields. Going back to the two effects above, the similar forecast errors for both models is suggesting that the difference in the estimated probability of commitment is likely due to the yield data providing information about the re-optimization episodes.

To understand why the yield curve can provide information about re-optimization episodes, it is helpful to revisit the simple model from section 2.2. Figure 2 shows the effect of a re-optimization shock on inflation and output (which determine the nominal stochastic discount factor). A policy re-optimization, therefore, is essentially a persistence shock to the stochastic discount factor, which in turn determines long term yields. The model's prediction of the effects of re-optimization are then tested against the long term yield data by the estimation. Note that this additional information is missing from the model without yields. There the model's prediction of future short rates is never tested against the data. We evaluate the implications of this effect by studying the difference in the model implied behavior between the following two cases: i) when a re-optimization shock occurs and ii) when one does not occur. Intuitively, if the data on average is closer to the models prediction from case i) then this should cause the estimation to find a lower probability of commitment (i.e. closer to discretion). Adding the yield data provides more information for the estimation algorithm, which can be seen from a figure already presented in the draft, i.e. figure 14. This figure shows the hypothetical effects on yields of a re-optimization shock occurring in every period. Notice that there are several occasions where this difference is sizeable. For example in the early 1990s and early 2000s this difference is as high as 25 basis points (which is greater than one half of the standard deviation of the estimated measurement error shocks to the yields). Notably, this includes the 2003-2004 period of "Greenspan's conundrum". Thus our estimates suggest that when we add yield data, the model identifies more re-optimization episodes. This feature also explains why the re-optimization episodes do not exactly align in our baseline model relative to the model without yields.

3 Fit of macro variables

Table 2 compares the standard deviation and cross correlations in the data with the baseline model and a model without yield data. Overall, the model without yield data does a better job of matching the moments in the data. Both models imply a roughly similar standard deviation for output but the model without yields generates a worse fit for inflation, consumption and investment. These results underlie a tension that is commonly found in DSGE models using yield data. Parameter values that create a good fit for the yield curve tend to come at the cost of degrading the fit for macro data. As mentioned in the paper, the inverse of the elasticity of intertemporal substitution and the labor elasticity are estimated to be higher in our model relative to models that do not use the yield data. This is to be expected given that these utility function parameters are closely tied to the stochastic discount factor and bond pricing.

However, the yield curve has been shown to capture information about future economic performance. To evaluate this, we have calculated the root mean squared error for different forecast horizons. These results are reported in table 1 below. For the one-quarter ahead forecast we notice that the model without yields performs slightly better. However as we increase the forecast horizon to 1 or 2 years the model with yields tends to perform better for most of the variables. At the longer 4 year horizon, both models tend to have a similar performance for forecasting macro variables. Finally, as is to be expected, the model with yields provides a far superior forecast of the short rate relative to the model without yields. Overall, while adding yield data to the model slightly hurts the in-sample fit of macro variables it actually leads to small improvements in the forecasting performance of the model.

	Root Mean Squared Forecast Error							
	Model with yields				Model without yields			
	1Q	4Q	8Q	16Q	1Q	4Q	8Q	16Q
Output	0.52	0.61	0.62	0.63	0.55	0.64	0.64	0.64
Consumption	0.50	0.61	0.63	0.60	0.51	0.63	0.64	0.62
Investment	1.52	2.07	2.02	2.19	1.51	2.04	2.18	2.18
Price Inflation	0.80	0.91	0.75	0.72	0.64	0.69	0.68	0.71
Wage Inflation	0.19	0.22	0.25	0.26	0.20	0.24	0.26	0.26
Short Rate	0.18	0.39	0.56	0.58	0.17	0.40	0.57	0.71

Table 1: Root Mean Squared Forecast Error from baseline model and model without yields

	Standard Deviation		
	Data	Model with yields	Model without yields
Output	0.59	0.78	0.75
Consumption	0.54	0.79	0.66
Investment	1.78	2.36	1.88
Price Inflation	0.24	0.65	0.25
Wage Inflation	0.67	1.64	0.74
	Cross-correlation with output		
	Data	Model with yields	Model without yields
Consumption	0.63	0.77	0.73
Investment	0.66	0.52	0.56
Price Inflation	-0.16	-0.09	-0.18
Wage Inflation	-0.08	0.49	0.29

Table 2: Moments for baseline model and model without yields