Lecture Notes for Monetary Policy (PhD course at UNISG)

Paul Söderlind[1]

October 2003

[1]University of St. Gallen and CEPR. *Address:* s/bf-HSG, Rosenbergstrasse 52, CH-9000 St. Gallen, Switzerland. *E-mail:* Paul.Soderlind@unisg.ch. Document name: MonAll.TeX.

Contents

1 Traditional Models of Monetary Policy

Main references: Romer (1996) (Romer), Blanchard and Fischer (1989) (BF), Obstfeldt and Rogoff (1996) (OR), and Walsh (1998).

1.1 The IS-LM Model

Reference: Romer 5, BF 10.4, and King (1993).

The IS curve (in logs) is

$$y_t = -\gamma i_t + \varepsilon_{yt} \Rightarrow i_t = \frac{-y_t + \varepsilon_{yt}}{\gamma}, \tag{1.1}$$

where ε_{yt} is a real (demand) shock. The LM curve (in logs) is

$$m_t - p_t = \psi y_t - \omega i_t + \varepsilon_{mt} \Rightarrow i_t = \frac{\psi y_t + \varepsilon_{mt} - m_t + p_t}{\omega}, \tag{1.2}$$

where ε_{mt} is a money demand shock. Consider *fixed prices*, which amounts to assuming a perfect elastic aggregate supply schedule: income is demand driven, which is the opposite to RBC models where income is essentially supply driven. Increasing m_t lowers the interest rate, which increases output. An outward shift in the IS curve because of an increase in ε_{yt}, increases both output and the nominal interest rate.

The most important problem with this model is that there are no supply-side effects, that is, prices are fixed. As a logical consequence, the IS curve is written in terms of the nominal interest rate, which differs from the real interest rate by a constant only. At a minimum, this model need to be amended with a model for prices (and thus price expectations), and also a term $\gamma \mathrm{E}_t \Delta p_{t+1}$ in the IS curve to let demand depend on the ex ante real interest rate.

The IS-LM framework has, in spite of these problems, been used extensively to discuss many important monetary policy issues. The following examples summarize two of them.

Example 1 *(Monetary Policy: Interest Rate Targeting or Money Targeting? BF. 11.2, Poole (1970), Mishkin (1997) 23) Suppose the goal of monetary policy is to stabilize output. The central bank must set its instrument (either m_t or i_t) before the shocks have been observed. Which instrument should it choose? If i_t is kept fixed, then*

$$\frac{dy_t}{d\varepsilon_{mt}} = 0 \text{ and } \frac{dy_t}{d\varepsilon_{yt}} = 1, (i_t \text{ fixed})$$

since the money demand shocks are not allowed to spill over to output, and the interest rate is not allowed to cushion real shocks. If m_t is kept fixed, then

$$\frac{dy_t}{d\varepsilon_{mt}} = -\frac{1}{\omega/\gamma + \psi} < 0 \text{ and } \frac{dy_t}{d\varepsilon_{yt}} = \frac{1}{1 + \gamma\psi/\omega} < 1, (m_t \text{ fixed})$$

since money demand shocks now increase the nominal interest rate and thereby decreases output, but the real shocks are cushioned by the increase in interest rates. Poole's conclusion was that interest rate targeting is preferred if most shocks are money demand shocks, while money stock targeting is better if most shocks are real. This is illustrated in Figure 1.1.

Example 2 *(The Mundell-Flemming Model and choice of exchange rate regime, Reference: OR 9.4, Romer 5.3, and BF 10.4) Add a real exchange rate term to the IS curve (1.1)*

$$y_t = -\gamma i_t + \phi\left(s_t + p_t^* - p_t\right) + \varepsilon_{yt} \Rightarrow$$
$$i_t = \frac{-y_t + \varepsilon_{yt} + \phi\left(s_t + p_t^* - p_t\right)}{\gamma},$$

and let asset market equilibrium be given by the UIP condition

$$i_t = i_t^* + E\Delta s_{t+1}.$$

Assumptions: fixed prices, foreign and domestic goods are imperfect substitutes, foreign and domestic bonds are perfect substitutes. Assume also that $E\Delta s_{t+1} = 0$ so $i_t = i_t^*$ (this does, of course, allow s_t to change—and makes a lot of sense if all shocks are permanent). If m_t is fixed, so the exchange rate is floating (set $m_t = 0$, for simplicity), then the LM equation gives $i_t = (\psi y_t + \varepsilon_{mt})/\omega$ or $y_t = \left(\omega i_t^* - \varepsilon_{mt}\right)/\psi$ (since $i_t = i_t^*$) so

$$\frac{dy_t}{d\varepsilon_{mt}} = -\frac{1}{\psi} < 0 \text{ and } \frac{dy_t}{d\varepsilon_{yt}} = 0 \; (m_t \text{ fixed, } s_t \text{ floating}).$$

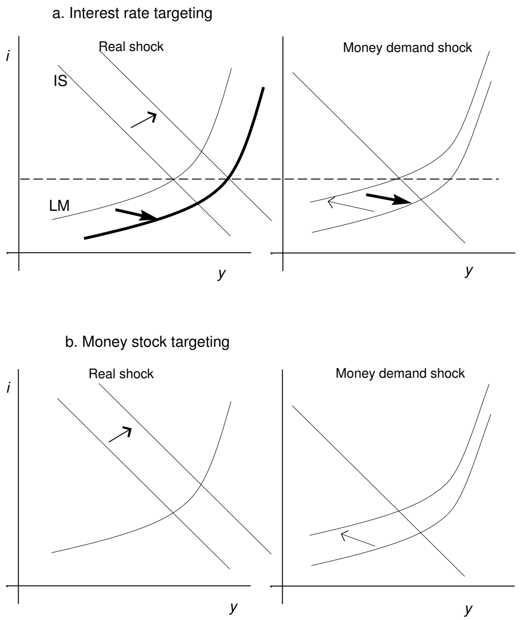

Figure 1.1: Poole's analysis of different monetary policy instruments in an IS-LM model. The real shock is a positive aggregate demand shock, and the money demand shock is a positive shock to money demand.

A money demand shock has a negative effect on output (similar to a closed economy model), while a real shock has not (different from a closed economy model). y_t *cannot increase unless* m_t, i_t^* *or* ε_{mt} *does. If they do not, then any real shock must simply spill over into an exchange rate appreciation. If the exchange rate is fixed, say* $s_t = 0$, *then the IS equation gives* $i_t = \left(-y_t + \varepsilon_{yt}\right)/\gamma$ *or* $y_t = -\gamma i_t^* + \varepsilon_{yt}$ *(since* $i_t = i_t^*$*) so*

$$\frac{dy_t}{d\varepsilon_{mt}} = 0 \text{ and } \frac{dy_t}{d\varepsilon_{yt}} = 1 \text{ (}s_t\text{ fixed).}$$

All shocks to the LM curve must be accommodated by corresponding changes in m_t *to keep* s_t *fixed. Any real shocks feed right through, since the money stock is expanded*

to accommodate the extra money demand to keep the exchange rate fixed (that is, the output shock is not allowed to increase the nominal interest rate). A fixed exchange rate (or a currency union) means that the country abandons the possibility to use monetary policy to buffer country specific real shocks (a common real shock among the participating countries can be buffered), but all money demand shocks are buffered. The extent of country-specific shocks is a main determinant behind optimum currency areas (the other is the degree of factor mobility). The conclusion from this analysis is that a floating exchange rate is better at stabilizing output if real shocks dominate, while a fixed exchange rate is better if money demand shocks dominate.

1.2 The Barro-Gordon Model

1.2.1 The Basic Model

References: Walsh 8, OR 9.5, BF 11.2 and 11.4, and Romer 9.4 and 9.5.

Use the LM curve (1.2) in the IS curve (1.1) to derive the *aggregate demand curve*

$$y_t^d = \frac{\gamma}{\omega + \gamma\psi}\left(m_t - p_t - \varepsilon_{mt}\right) + \frac{\omega}{\omega + \gamma\psi}\varepsilon_{yt}. \tag{1.3}$$

For simplicity, merge $-\varepsilon_{mt} + \omega/\left(\omega + \gamma\psi\right)\varepsilon_{yt}$ into a composite demand shock, ε_t^d,

$$y_t^d = \frac{\gamma}{\omega + \gamma\psi}\left(m_t - p_t\right) + \varepsilon_t^d. \tag{1.4}$$

This is a very common formulation of aggregate demand; it shows up in Lucas' model of the Phillips curve, and also in several monetary models with monopolistic competition (see, for instance, BF 8.1). Note, however, that if the IS curve depended on the ex ante real interest rate instead of the nominal interest rate, then a term $\mathrm{E}_t \Delta p_{t+1}\omega\gamma/(\omega + \gamma\psi)$ is added to (1.4).

We now also introduce an aggregate supply side inspired by Lucas' version of the Phillips curve or by a model with predetermined prices (or long nominal contracts)

$$y_t^s = b\left(p_t - p_{t|t-1}^e\right) + \varepsilon_t^s, \tag{1.5}$$

$$= b\left(\pi_t - \pi_{t|t-1}^e\right) + \varepsilon_t^s \text{ /* } \pm p_{t-1} \text{ */} \tag{1.6}$$

where $p_{t|t-1}^e$ is the log price level in t which private agents expect based on the information in $t-1$, and $\pi^e_{t|t-1}$ is the corresponding expected inflation rate, $\pi^e_{t|t-1} = p^e_{t|t-1} - p_{t-1}$.

Let expectations be rational, so $\pi^e_{t|t-1}$ in (1.6) is the mathematical expectation

$$\pi^e_{t|t-1} = \mathrm{E}_{t-1}\pi_t. \tag{1.7}$$

To simplify the algebra we note that the central bank can always generate any inflation it wants by manipulating the money supply, m_t. We therefore treat inflation π_t as the policy instrument (the required m_t can be backed out from the equilibrium).

The loss function of the central bank is

$$L_t = \pi_t^2 + \lambda\left(y_t - \bar{y}\right)^2, \tag{1.8}$$

so the central bank want to stabilize inflation around its natural level (normalized to zero), but output around $\bar{y}$, which may be different from the natural level (once again normalized to zero). The target level for output, $\bar{y}$, is typically positive—perhaps the natural level of output (zero) is not compatible with full employment (due to labour market imperfections) or because the natural level of output is affected by product market imperfections. Using monetary policy to solve such imperfections is probably not the best idea; in this model, it will not even work.

The central bank sets the monetary policy instrument after observing the shock, ε^s_t. (This is different from the two examples given at the beginning of this note, where policy had to be set before the shocks were realized.) In practice, monetary policy can react quickly, although perhaps not completely without a lag. However, the main point in this analysis is that the monetary policy can react more quickly than the private sector (price and wage setters). This is probably a realistic assumption. This opens a channel for monetary policy to have effect.

1.2.2 Monetary Policy with Commitment

In the commitment case, the central bank chooses a policy rule in $t-1$ and precommits to it. It will therefore choose a rule which minimizes $\mathrm{E}_{t-1}L_t$. Since the model is linear-quadratic, we can assume that the policy rule is linear. Since only innovations can affect output we can safely restrict attention to policy rules in terms of a constant (there is no dynamics in the model) and the shocks. We therefore assume (correctly, it can be shown)

that the policy rule is on the form

$$\pi_t = \alpha + \beta\varepsilon^s_t + \delta\varepsilon^d_t, \tag{1.9}$$

where we have to find the values of α, β, and δ. The public's expectations must be

$$\begin{aligned}\pi^e_{t|t-1} &= \mathrm{E}_{t-1}\pi_t \\ &= \alpha,\end{aligned} \tag{1.10}$$

provided the shocks are unpredictable. Note that α is not determined yet. The idea is that whatever value of α that the central bank would happen to choose, the public knows it and will adjust their expectations accordingly. This means that the central bank can influence the public's expectations and that it makes use of this in the optimization problem.

Using the supply function (1.6) and (1.9)-(1.10) in the loss function (1.8), and taking expectations as of $t-1$ gives the optimization problem

$$\mathrm{E}_{t-1}L_t = \mathrm{E}_{t-1}\left(\alpha + \beta\varepsilon^s_t + \delta\varepsilon^d_t\right)^2 + \lambda\mathrm{E}_{t-1}\left[b\left(\alpha + \beta\varepsilon^s_t + \delta\varepsilon^d_t - \alpha\right) + \varepsilon^s_t - \bar{y}\right]^2. \tag{1.11}$$

The first order condition with respect to α gives

$$\alpha = 0. \tag{1.12}$$

The first order condition with respect to δ is

$$2\delta\sigma_{dd} + 2\lambda b^2\delta\sigma_{dd} = 0 \text{ or } \delta = 0, \tag{1.13}$$

provided the shocks are unpredictable and also uncorrelated, $\mathrm{E}_{t-1}\varepsilon^d_t\varepsilon^s_t = 0$. Finally, the first order condition with respect to β is then

$$\begin{aligned}2\beta\sigma_{ss} + 2\lambda b\left(b\beta + 1\right)\sigma_{ss} = 0 \text{ or} \\ \beta = -\frac{\lambda b}{1 + \lambda b^2}.\end{aligned} \tag{1.14}$$

The policy rule (1.9) is therefore

$$\pi_t = \beta\varepsilon^s_t, \tag{1.15}$$

with β given by (1.14). Output is then

$$y_t = (b\beta + 1)\,\varepsilon_t^s. \tag{1.16}$$

If the central bank targets inflation only, $\lambda = 0$, then $\beta = 0$, which by (1.15) and (1.16) means that inflation is completely stable and that output shocks are not cushioned. Conversely, if the central bank targets output only, $\lambda \to \infty$, then $\beta = -1/b$ (apply l'Hôpital's rule) so output is now completely stable, but inflation varies.

More generally, note that

$$\frac{1}{\mathrm{Var}(\varepsilon_t^s)}\frac{\partial}{\partial\lambda}\mathrm{Var}(\pi_t) = \frac{\partial}{\partial\lambda}\beta^2 = \frac{\partial}{\partial\lambda}\left(-\frac{\lambda b}{1+\lambda b^2}\right)^2 = \frac{2\lambda b^2}{\left(1+\lambda b^2\right)^3} > 0 \text{ and} \tag{1.17}$$

$$\frac{1}{\mathrm{Var}(\varepsilon_t^s)}\frac{\partial}{\partial\lambda}\mathrm{Var}(y_t) = \frac{\partial}{\partial\lambda}(b\beta+1)^2 = \frac{\partial}{\partial\lambda}\left(-\frac{\lambda b^2}{1+\lambda b^2}+1\right)^2 = -2\frac{b^2}{\left(1+\lambda b^2\right)^3} < 0. \tag{1.18}$$

As expected, the variance of π is therefore increasing in λ. Conversely, the variance of output decreasing in λ.

Example 3 *When $b = 1$, then $\pi_t = -\lambda/(1+\lambda)\varepsilon_t^s$ and $y_t = 1/(1+\lambda)\varepsilon_t^s$ so $Var(\pi_t)/Var(y_t) = \lambda^2$, which is clearly increasing in λ.*

The policy rule implies that average inflation is zero, $\alpha = 0$. There is no point in creating a non-zero average inflation, since anticipated inflation does not affect output.

The policy rule also implies that *demand shocks should always be completely offset*: they do not enter either inflation (1.15) or output (1.16). The reason is that demand shocks push prices and output in the same direction, so there is no trade-off between price and output stability. Only supply shocks, which push inflation and output in different directions, gives a trade-off.

To see this, let us simplify by setting price expectations in (1.5), $p^e_{t|t-1}$, to zero and also revert to considering m_t as the policy instrument (there is a one-to-one relation to the inflation rate). We can then solve the system (1.4) and (1.5) for output and price as

$$[b(\omega+\gamma\psi)+\gamma]\begin{bmatrix} y_t \\ p_t \end{bmatrix} = \begin{bmatrix} \gamma b \\ \gamma \end{bmatrix} m_t + \begin{bmatrix} b(\omega+\gamma\psi) & \gamma \\ \omega+\gamma\psi & -(\omega+\gamma\psi) \end{bmatrix}\begin{bmatrix} \varepsilon_t^d \\ \varepsilon_t^s \end{bmatrix}$$

All parameters are positive. A positive shock to ε_t^d increases both output and price proportionally, so a decrease in m_t can stabilize the effects completely. This can also be seen directly from (1.4). In contrast, a positive shock to ε_t^s increases output but decreases the price. Since the effect of m_t on output and price has the same sign, the central bank cannot use monetary supply to stabilize both when the economy is hit by a supply shock. If it opts for increasing m_t, then this may stabilize the price but destabilizes output further, and vice versa.

1.2.3 Monetary Policy without Commitment (Discretionary)

One problem with the commitment equilibrium is that the policy rule announced in $t-1$ may no longer be the optimal rule in t. At that time, inflation expectations can be treated as given (for instance, inflation expectations might enter the model because they represent nominal contracts written in $t-1$). The central bank could have an incentive to exploit this: the policy rule is then not "time consistent." If the central bank cannot commit to a policy rule, then the time inconsistent rule is not credible, and the commitment equilibrium falls apart.

We now assume that the central bank cannot commit to a rule. Instead, we look for a policy that is optimal in t (after the shocks have been observed), when $\pi_{t|t-1}$ is taken as given. If this happens to be the same decision rule as above, then there is no time inconsistency problem—otherwise there is. With discretionary monetary policy, the choice of inflation minimizes

$$\pi_t^2 + \lambda\left(b\pi_t - b\pi^e_{t|t-1} + \varepsilon_t^s - \bar{y}\right)^2. \tag{1.19}$$

There is no expectations operator, since the central bank makes its decision after the shocks are realized, and it does not precommit (before the shock) to follow any particular decision rule.

The first order condition with respect to π_t is

$$\pi_t = -\pi_t\lambda b^2 + \lambda b^2\pi^e_{t|t-1} - \lambda b\varepsilon_t^s + \lambda b\bar{y}, \tag{1.20}$$

with (two times the) marginal cost of inflation on the left hand side and (two times the) marginal benefits on the right hand side. The public knows that (1.20) will determine how the central bank acts. They therefore form their expectations in $t-1$ by rationally using all available information. Taking mathematical expectations of (1.20) based on the information available in $t-1$ and rearranging gives that expectations formed in $t-1$ must be

$$\pi^e_{t|t-1} = \lambda b \bar{y}. \tag{1.21}$$

Combine this with (1.20) to get

$$\begin{aligned} \pi_t &= \lambda b \bar{y} - \frac{\lambda b}{1+\lambda b^2}\varepsilon^s_t \\ &= \lambda b \bar{y} + \beta \varepsilon^s_t \end{aligned} \tag{1.22}$$

This rule has the same response to the output shock as the commitment rule, but a higher average inflation (if both λ and $\bar{y}$ are positive). The first of these results means that the variances are the same as in the commitment equilibrium. The reason is that there is no persistence in this model. In a model with more dynamics this will no longer be true—in that case we can intuitively think of the natural output level, here normalized to zero, as time varying. This makes the difference between commitment and discretionary equilibrium more complicated.

The second of the results, the higher average inflation, is due to the incentive to deviate from the commitment rule—and that the public incorporates that when forming inflation expectations. To understand the incentives to inflate consider (1.20) when $\pi^e_{t|t-1} = \varepsilon^s_t = 0$. If the central bank then sets $\pi_t = 0$ (so there is no policy surprise), then the marginal cost of inflation (left hand side) is zero, but the marginal benefit (right hand side) is $\lambda b \bar{y}$. If both λ and $\bar{y}$ are positive, then there is an incentive to inflate. Private agents will realize this and form their expectations accordingly. The equilibrium is where $\mathrm{E}_t \pi_t = \pi^e_{t|t-1}$ and marginal cost and benefits are equal.

It is often argued that making the central bank more independent of the government is quite similar to a lower λ, that is, to a lower relative weight on output. From (1.22) we see that this should lower the average inflation rate. At the same time, it should lower the variability of inflation, but increase the variability of output, see (1.17)-(1.18).

It is still unclear if the inflation bias is important. There are many other cases where the logic of the discretionary equilibrium seems unappealing, for instance, in capital income taxation (why is not all capital confiscated every year?). It might be the case that society has managed to set up institutions and informal rules which create some kind of commitment technology.

The high inflation between mid 1960s and early 1980s could possibly be due to the lack of commitment technology combined with more ambitious employment goals. An alternative explanation is that the policy makers believed in a *long run* trade-off between unemployment and inflation.

1.2.4 Empirical Illustration

Walsh Fig 8.5 (relation between central bank independence and average inflation).

Bibliography

Blanchard, O. J., and S. Fischer, 1989, *Lectures on Macroeconomics*, MIT Press.

King, R. G., 1993, "Will the New Keynesian Macroeconomics Resurrect the IS-LM Model?," *Journal of Economic Perspectives*, 7, 67–82.

Mishkin, F. S., 1997, *The Economics of Money, Banking, and Financial Markets*, Addison-Wesley, Reading, Massachusetts, 5th edn.

Obstfeldt, M., and K. Rogoff, 1996, *Foundations of International Macroeconomics*, MIT Press.

Romer, D., 1996, *Advanced Macroeconomics*, McGraw-Hill.

Walsh, C. E., 1998, *Monetary Theory and Policy*, MIT Press, Cambridge, Massachusetts.

2 Microfoundations of Monetary Policy Models

Main references: Romer (1996) (Romer), Blanchard and Fischer (1989) (BF), Obstfeldt and Rogoff (1996) (OR), and Walsh (1998).

2.1 Money Demand

Roles of money: medium of exchange, unit of account, and storage of value (often dominated by other assets).

Money is macro model is typically identified with currency which gives no interest. The liquidity service of money (medium of exchange) is emphasized, rather than store of value or unit of account.

2.1.1 Traditional money demand equations

References: Romer 5.2, BF 4.5, OR 8.3, Burda and Wyplosz (1997) 8.

The standard money demand equation

$$\ln \frac{M_t}{P_t} = \text{constant } + \psi \ln Y_t - \omega i_t \tag{2.1}$$

are used in many different models, for instance as the LM curve is IS-LM models. M_t in (2.1) is often a money aggregate like M1 or M3. In most of the models on this course, we will assume that the central bank have control over this aggregate.

2.1.2 Money Demand and Monetary Policy

There are many different models for why money is used. The common feature of these models is that they all generate something pretty close to (2.1). But why is this broader money aggregate related to the monetary base, which the central bank may control? Short answer: the central bank creates a demand for narrow money by forcing banks to hold it (reserve requirements) and by prohibiting private substitutes to narrow money (banks are not allowed to print bills).

The idea behind central bank interventions is to affect the money supply. However, most central banks use short interest rates as their operating target. In effect, the central bank has monopoly over supply over narrow money which allows it to set the short interest rate, since short debt is a very close substitute to cash. In terms of (2.1), the central bank may set i_t, which for a given output and price level determines the money supply as a residual.

2.1.3 Different Ways to Introduce Money in Macro Models

Reference: OR 8.3 and Walsh (1998) 2.3 and 3.3.

The *money in the utility function* (MIU) model just postulates that real money balances enter the utility function, so the consumer's optimization problem is

$$\max_{\{C_t, M_t\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t u\left(C_t, \frac{M_t}{P_t}\right). \tag{2.2}$$

One motivation for having the real balances in the utility function is that having cash may save time in transactions. The correct utility function would then be $u\left(C_t, \bar{L} - L_t^{shopping}\right)$, where $L_t^{shopping}$ is a decreasing function of M_t/P_t.

Cash-in-advance constraint (CIA) means that cash is needed to buy (some) goods, for instance, consumption goods

$$P_t C_t \leq M_{t-1}, \tag{2.3}$$

where M_{t-1} was brought over from the end of period $t-1$. Without uncertainty, this restriction must hold with equality since cash pays no interest: no one would accumulate more cash than strictly needed for consumption purposes since there are better investment opportunities. In stochastic economies, this may no longer be true.

The simple CIA constraint implies that "money demand equation" does not include the nominal interest rate. If the utility function depends on consumption only, then all rates of inflation gives the same steady state utility. This stands in sharp contrast to the MIU model, where the optimal rate of inflation is minus one times the real interest rate (to get zero nominal interest rate). However, this is not longer true if the cash-in-advance constraint applies only to a subset of the arguments in the utility function. For instance, if we introduce leisure or credit goods.

Shopping-time models typically have a utility function is terms of consumption and leisure

$$\sum_{s=0}^{\infty} \beta^s U \left(C_t, 1 - l_t - n_t\right), \tag{2.4}$$

where l_t is hours worked, and n_t hours spent on shopping (supposed to give disutility). The latter is typically modelled as some function which is increasing in consumption and decreasing in cash holdings

2.1.4 An Example of Money in the Utility Function

Reference: BF 4.5; OR 8.3; Walsh (1998) 2.3; and Lucas (2000)

The consumer's optimization problem is

$$\max_{\{C_t, M_t\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t u \left(C_t, \frac{M_t}{P_t}\right) \tag{2.5}$$

subject to the real budget constraint

$$K_{t+1} + \frac{M_t}{P_t} = (1 + r_t) K_t + \frac{M_{t-1}}{P_t} + w_t - C_t - T_t, \tag{2.6}$$

where r_t is the (net) real interest rate (from investing in $t-1$ and receiving the return in t), and w_t the real wage rate. Labor supply is normalized to one. The consumer rents his capital stock to competitive firms in each period. T_t denotes lump sum taxes.

Use (2.6) in (2.5) to get the unconstrained problem for the consumer

$$\max_{\{K_{t+1}, M_t\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t u \left[(1 + r_t) K_t + \frac{M_{t-1}}{P_t} + w_t - T_t - K_{t+1} - \frac{M_t}{P_t}, \frac{M_t}{P_t}\right]. \tag{2.7}$$

The first order condition for K_{t+1} is

$$u_C \left(C_t, \frac{M_t}{P_t}\right) = (1 + r_{t+1}) \beta u_C \left(C_{t+1}, \frac{M_{t+1}}{P_{t+1}}\right), \tag{2.8}$$

which is the traditional Euler equation for real bonds (with uncertainty we need to take the expected value of the right hand side, conditional on the information in t). It would also hold for any other financial asset.

The first order condition for M_t is

$$u_C \left(C_t, \frac{M_t}{P_t}\right) = u_{M/P} \left(C_t, \frac{M_t}{P_t}\right) + \beta u_C \left(C_{t+1}, \frac{M_{t+1}}{P_{t+1}}\right) \frac{P_t}{P_{t+1}}. \tag{2.9}$$

If money would not enter the utility function, then this is a special case of (2.8) since the real gross return on money is P_t/P_{t+1}. It is not obvious, however, that we get an interior solution to money holdings unless money gives direct utility.

The left hand side of (2.9) is the marginal utility lost because some resources are taken from time t consumption, and the right hand side is the marginal utility gained by having more cash today and the extra consumption this allows tomorrow (cash provides utility and is also a form of saving, whose purchasing power depends on the inflation).

Substitute for $\beta u_C \left(C_{t+1}, M_{t+1}/P_{t+1}\right)$ from (2.8) in (2.9) and rearrange to get

$$u_C \left(C_t, \frac{M_t}{P_t}\right) \left(1 - \frac{1}{1 + r_{t+1}} \frac{P_t}{P_{t+1}}\right) = u_{M/P} \left(C_t, \frac{M_t}{P_t}\right). \tag{2.10}$$

The Fisher equation is

$$1 + i_t = \mathrm{E}_t \left(1 + r_{t+1}\right) \frac{P_{t+1}}{P_t}, \tag{2.11}$$

where the convention is that the nominal interest rate is dated t since it is known as of t.

Under perfect foresight, (2.10) can then be written

$$\frac{i_t}{1 + i_t} = u_{M/P} \left(C_t, \frac{M_t}{P_t}\right) / u_C \left(C_t, \frac{M_t}{P_t}\right), \tag{2.12}$$

which highlights that the nominal interest rate is the relative price of the "money services" we get by holding money one period instead of consuming it. Note that (2.12) is a relation between real money balances, the nominal interest rate, and an activity level (here consumption), which is very similar to the LM equation.

Example 4 *(Explicit money demand equation from Cobb-Douglas/CRRA.) Let the utility function be*

$$u \left(C_t, \frac{M_t}{P_t}\right) = \frac{1}{1 - \gamma} \left[C_t^{\alpha} \left(\frac{M_t}{P_t}\right)^{1-\alpha}\right]^{1-\gamma},$$

in which case (2.12) can be written

$$\frac{M_t}{P_t} = C_t \frac{1-\alpha}{\alpha} \frac{1+i_t}{i_t},$$

which is decreasing in i_t and increasing in C_t. This is quite similar to the standard money demand equation (2.1). Take logs and make a first-order Taylor expansion of $\ln [(1+i_t)/i_t]$ *around i_{ss}*

$$\ln \frac{M_t}{P_t} = constant \ + \ln C_t - \frac{1}{i_{ss}(1+i_{ss})} i_t.$$

Compared with the money demand equation (2.1), $\psi \ln Y_t$ is replaced by $\ln C_t$ and $\omega = 1/[i_{ss}(1+i_{ss})]$. If $i_{ss} = 5\%$, then $\omega \approx 20$, which appears to be very high compared to empirical estimates.

2.2 The Effect of Money vs the Effect of Price Stickiness

Reference: Cooley and Hansen (1995)

2.2.1 Inflation Tax Model

This is a fairly standard real business cycle model, with some additional features. A stochastic money supply interacts with a cash-in-advance transaction technology to create some real effects of money supply shocks. The key equations are listed below. (Lower case letters denote values for a representative household, whereas upper case letters denote aggregates.)

$$\begin{aligned}
\text{Utility function}: &\ \mathrm{E}_0 \sum_{t=0}^{\infty} \beta^t \left[a \ln c_{1t} + (1-a) \ln c_{2t} - \gamma h_t\right] \\
\text{Real budget constraint}: &\ c_{1t} + c_{2t} + x_t + \frac{m_{t+1}}{P_t} = \frac{w_t}{P_t} h_t + r_t k_t + \frac{m_t}{P_t} + \frac{T_t}{P_t}. \\
\text{Cash-in-advance constraint}: &\ P_t c_{1t} = m_t + T_t \\
\text{Production function}: &\ Y_t = e^{z_t} K_t^{\theta} H_t^{1-\theta}. \\
\text{Capital accumulation}: &\ k_{t+1} = (1-\delta) k_t + x_t. \\
\text{Government budget constraint}: &\ T_t = \Delta M_{t+1}. \\
\text{Money supply}: &\ \Delta \ln M_{t+1} = 0.49 \Delta \ln M_t + \xi_{t+1}, \ \ln \xi_{t+1} \sim N, \text{known at } t. \\
\text{Log productivity}: &\ z_{t+1} = 0.95 z_t + \epsilon_{t+1}, \ \epsilon_{t+1} \sim N\left(0, 4.9 \times 10^{-5}\right)
\end{aligned}$$

(Note: it should be T_t/P_t in the real budget constraint; there is a typo in the book.) The notation is: capital stock (K), money stock (M), price level (P), wage rate (W), hours worked (H), output (Y), investment (X), and productivity (z). Note the notation: the money stock held at the end of period t is denoted M_{t+1} (M_t in Benassy).

Private consumption consists of a "cash good," c_{1t}, and a "credit good," c_{2t}. One interpretation of the trading sequence within a time period t is the following.

1. In the beginning of the period, the household carries over m_t from $t-1$, and gets T_t is cash transfers from the government. Households also own all physical capital (k_t). Firms hold no cash or physical capital. The government finances the transfers by printing new money.

2. Firms rent capital and labor (the rent and wages are paid somewhat later in the period), and produce goods.

3. The household buys the cash good with the available cash, where the cash-in-advance restriction $P_t c_{1t} \leq m_t + T_t$ must hold. (The log-normal distribution of the money supply shock ξ_t means that the money stock can never decrease, which is enough to ensure that the CIA constraint always binds: positive nominal interest rate with probability one.) Firms now hold $m_t + T_t$ in cash.

4. The household receives nominal factor payments $w_t h_t + P_t r_t k_t$ from the firms (ex-

hausts all profits), and buys credit goods ($P_t c_{2t}$) and investment goods ($P_t x_t$). The firms now hold no cash; households own the physical capital $k_{t+1} = (1-\delta) k_t + x_t$, and the cash $m_{t+1} = w_t h_t + P_t r_t k_t - P_t c_{2t} - P_t x_t$.

5. In equilibrium, the money stock held by the households (m_{t+1}) must equal money supply by the central bank ($m_t + T_t = M_{t+1}$).

Calibration

The parameters in the production function, depreciation, Solow residual, and time preference are chosen as in standard RBC models. The money supply process (for M1) is estimated with least squares. The a parameter is estimated from the first-order condition

$$\frac{C_{1t} + C_{2t}}{C_{1t}} = \frac{P_t (C_{1t} + C_{2t})}{P_t C_{1t}} = \frac{P_t C_t}{m_t} = \frac{1}{\alpha} + \frac{1-\alpha}{\alpha} \text{*interest rate}, \tag{2.13}$$

where the paper uses the portion of $M1$ held by households as a proxy for m_t (this differs from how they estimate the AR(1) for money supply, where they use all of M1). Identifying a from the intercept, they get $a = 0.85$. (If they had identified α from the slope instead, then they would have got $\alpha = 0.9$.)

To sum up, they use $\theta = 0.4$, $\delta = 0.019$, $\beta = 0.989$, $\gamma = 2.53$, and $a = 0.84$.

Solving the Model

The inflation tax means that the competitive solution will not coincide with the social planners's solution. The solution algorithm is therefore a based on the concept of recursive competitive equilibrium. Solving a quadratic approximation (in logs) of the model results in a set of linear decision rules in terms of the state of the economy. Productivity is stationary ($|\rho| < 1$), but the money supply is not, so prices will also be non-stationary. It is therefore very convenient to "detrend" all nominal variables by dividing by M_t before the solution algorithm is applied.

2.2.2 A Model with Nominal Wage Stickiness

The wage contract is based on the one-period ahead expectation of the marginal product of labor. The first order condition for profit maximization is

$$w_t = (1-\theta) P_t e^{z_t} \left(\frac{K_t}{H_t}\right)^\theta \Rightarrow \tag{2.14}$$

$$\ln w_t = \ln(1-\theta) + \ln P_t + z_t + \theta (\ln K_t - \ln H_t). \tag{2.15}$$

It is assumed (ad hoc) that $\ln w_t$ is set equal to the expectation of the right hand side of (2.15), conditional on the information in $t-1$. Note that K_t is in the information set at $t-1$, while P_t, H_t, and z_t are not. The deterministic steady state of the economy with the this type of wage contracts is the same as in the economy without wage contracts (simplifies a lot).

The nominal wage is fixed in $t-1$, and the price level is observed in t. Money supply shocks may therefore affect the real wage by affecting the price level. Workers are assumed to supply inelastically at the going real wage (firms are on their labor demand schedules). A positive money supply shock will decrease the real wage and therefore increase labor demand and output. As usual, this effect lasts as long as some prices remain fixed: here it is one period since we have one-period labor contracts. Consumers (which own both the firms and the labor resources and therefore get all output) choose to consume only a fraction of the temporary income increase, so most of output increase spills over to investment (saving).

The most important difference between these two models is that only the model with nominal stickiness shows a quantitatively interesting response of real variables to money supply shocks. See Cooley and Hansen (1995) Figures 7.6–7.

2.3 Dynamic Models of Sticky Prices

References: BF. 8.2, Romer 6.7, Rotemberg (1987).

This section deals with the effect of price rigidities in dynamic models. Prices are set in advance and firms are assumed to supply whatever demand happens to be (which is reasonable only as long as demand shocks do not force marginal costs above the price). This clearly assumes that firms can expand production, for instance, by hiring more labour, so there must be a fairly elastic factor supply. If factor supply is not particularly elastic, then marginal costs will increase rapidly so the assumption that marginal cost is always below the price becomes implausible.

Aggregate demand shocks (or money supply) will usually have real effects when prices adjust slowly. This is certainly the case when prices are changed with prespecified intervals (time-dependent rules), and the main issue is instead how long the effects last. It is typically also the case when prices are changed when the old prices are too far from the frictionless optimum (state dependent rules).

In general, we would like to find a reasonable model which can explain both why average prices seem to adjust gradually to monetary expansions and why price changes of individual firms appear to be "lumpy." This is hard.

2.3.1 Quadratic Costs of Price-Adjustment

Reference: Rotemberg (1982a), Rotemberg (1982b), and Walsh 5.5.

Firm i is a monopolist on its market and sets the log price, p_{it}, to maximize the value of the firm: the expected discounted sum of profits. If there were no costs of adjusting this price, then the price would be equal to some value, p^*_{it}, which we call the flex price optimum.

With costs of adjusting the price we formulate the maximization problem in two steps. First, find the flex price optimum, p^*_{it}. Second, minimize the loss from not being at p^*_{it} and from incurring adjustment costs. For the moment, we will take the time series process of p^*_{it} as given and focus on the second part of the maximization problem. To make any progress, we also approximate the objective function in the second step by a quadratic function

$$\min_{\{p_{it+s}\}_{s=0}^{\infty}} \mathrm{E}_t \sum_{s=0}^{\infty} \beta^s \left[\left(p_{it+s} - p^*_{it+s} \right)^2 + c \left(p_{it+s} - p_{it+s-1} \right)^2 \right] \text{ or} \tag{2.16}$$

$$\min_{\{p_{it+s}\}_{s=0}^{\infty}} \left\{ \left(p_{it} - p^*_{it} \right)^2 + c \left(p_{it} - p_{it-1} \right)^2 + \beta \mathrm{E}_t \left(p_{it+1} - p^*_{it+1} \right)^2 + \beta c \mathrm{E}_t \left(p_{it+1} - p_{it} \right)^2 + ... \right\}.$$

The first order condition with respect to p_{it} is

$$p_{it} - p^*_{it} + c \left(p_{it} - p_{it-1} \right) - \beta c \mathrm{E}_t \left(p_{it+1} - p_{it} \right) = 0 \text{ or} \tag{2.17}$$

$$\beta \mathrm{E}_t \Delta p_{it+1} + \frac{1}{c} \left(p^*_{it} - p_{it} \right) = \Delta p_{it}. \tag{2.18}$$

There is no lumpiness in individual price changes. Since both deviations from the p^*_{it} and prices changes are much more costly when they are large (the loss function is quadratic), the optimal policy will be to converge to p^*_{it} by taking many small steps rather than a few large. In a symmetric equilibrium $p_{it} = p_t$ and $p^*_{it} = p^*_t$. It can also be noted that situations with a high surprise inflation will lead to a higher $p^*_{it} - p_{it}$, so the price adjustment is then faster.

The smooth individual price changes carry over to the average prices, since all firms

are similar. Let $p_{it} = p_t$ and $p^*_{it} = p^*_t$ be the common prices and write (2.18) as

$$\Delta p_t = \beta \mathrm{E}_t \Delta p_{t+1} + \frac{1}{c} \left(p^*_t - p_t \right). \tag{2.19}$$

Special Case: No Adjustment Cost ($c = 0$)

If $c = 0$, then (2.17) shows that $p_{it} = p^*_{it}$, so the firm will always set its actual price equal to the unrestricted optimal price (quite obvious since the price is then unrestricted).

2.3.2 The Flex Price Optimum with Monopolistic Competition

What is the unrestricted optimal price, p^*_{it}, which plays such an important role in the previous model? A typical formulation is that it represents a monopolist's price in a flex-price equilibrium. That price is typically an increasing function of aggregate demand and a decreasing function of the productivity level. In logs, we write

$$p^*_t = p_t + \phi y_t + \varepsilon_t, \tag{2.20}$$

where ε_t is interpreted as the negative of a productivity shock (negative "supply shock"). Note that $\phi > 0$. It is typically increasing in slope of the marginal cost curve (the degree of decreasing returns to scale) and decreasing in the elasticity of substitution between goods in consumer preferences. In most models, we need an upward sloping marginal cost curve to get $\phi > 0$, which could be motivated by some fixed factors of production. If these fixed factors are not completely fixed, but can be accumulated over time, then the problem becomes more complicated (dynamic) and (2.20) can only be interpreted as an approximation that might be valid for short to medium run horizons (a business cycle, say).

Using (2.20) in (2.19) gives

$$\Delta p_t = \beta \mathrm{E}_t \Delta p_{t+1} + \delta \left(\phi y_t + \varepsilon_t \right), \text{ where } \delta = 1/c, \tag{2.21}$$

which can be thought of as an expectations-augmented Phillips curve. It is in an sense similar to the Keynesian AS curve, which has positive relation between output and the price level.

Recursion forward gives

$$\Delta p_t = \delta \sum_{s=0}^{\infty} \beta^s \mathrm{E}_t \left(\phi y_{t+s} + \varepsilon_{t+s} \right), \tag{2.22}$$

provided $\lim_{s\to\infty} \beta^{s+1} \mathrm{E}_t \Delta p_{t+s} = 0$. Note that $\mathrm{E}_t y_{t+s}$ has a large effect on inflation is ϕ is high (strong decreasing returns to scale and/or strong market power), and $\mathrm{E}_t \left(\phi y_{t+s} + \varepsilon_{t+s}\right)$ has a large effect if δ is high (small c in (2.21)).

As in any Phillips curve, it appears as if inflation is a real phenomenon! This is quite the opposite to the Cagan model, where it is assumed that both output and the real interest rate are constant. This suggests that this model of price setting is certainly not suitable for understanding a permanent change in the money supply trend. It is not plausible that the model parameters, for instance q and c, would remain unchanged in such a case.

2.3.3 Example: Calvo Model in a Very Simple Macro Model

For simplicity, assume that the quantity equation holds. In logs we have

$$m_t = p_t + y_t. \tag{2.23}$$

This can be taken to represent aggregate demand. Aggregate supply is represented by the price setting rule, and it is assumed that firms supply whatever the market demands at the going price: output is demand determined. In traditional monetarist models, the quantity equation *is* aggregate demand, without much discussion of where it comes from. In a Keynesian model, the quantity equation would be an approximation to the Keynesian AD curve (the combination of the IS and LM curves which traces out the relation between output and prices). Both these interpretations assume a negative relation between the price level and output. In some modern dynamic general equilibrium models, the quantity equation can be shown to be the money demand equation (see, for instance, Bénassy (1995)).

We now use this very simple model of "demand" to illustrate some properties of the sticky price model. Substitute for y_t in (2.21) by using (2.23)

$$\Delta p_t = \beta \mathrm{E}_t \Delta p_{t+1} + \delta\phi \left(m_t - p_t \right) + \delta \varepsilon_t$$

$$-p_{t-1} + p_t \left(1 + \beta + \delta\phi\right) - \beta \mathrm{E}_t p_{t+1} = \delta \left(\phi m_t + \varepsilon_t\right). \tag{2.24}$$

This is a second-order expectational difference equation, which can be solved with a variety of methods. The perhaps most straightforward one is to specify a time-series process for the exogenous driving process, and transform the system to a vector first-order system and then use a decomposition of the resulting matrix to decouple the variables in those that are predetermined in t (typically the exogenous variables and values determined in previous periods like the capital stock and lagged variables) and those that can jump in t in response to changes in expectations about future values (typically asset prices and anything else that depend on expected future values).

A trivial step is to note that (2.24) can be rewritten

$$\mathrm{E}_t p_{t+1} = -\frac{1}{\beta} p_{t-1} + p_t \frac{1+\beta+\delta\phi}{\beta} - \frac{\delta}{\beta} \left(\phi m_t + \varepsilon_t\right). \tag{2.25}$$

Suppose $\varepsilon_t = 0$ and that m_t is an AR(1)

$$m_t = \rho m_{t-1} + \varepsilon_{mt}. \tag{2.26}$$

We can then write the model on state space form as

$$\begin{bmatrix} m_{t+1} \\ p_t \\ \mathrm{E}_t p_{t+1} \end{bmatrix} = \begin{bmatrix} \rho & 0 & 0 \\ 0 & 0 & 1 \\ -\frac{\delta}{\beta}\phi & -\frac{1}{\beta} & \frac{1+\beta+\delta\phi}{\beta} \end{bmatrix} \begin{bmatrix} m_t \\ p_{t-1} \\ p_t \end{bmatrix} + \begin{bmatrix} \varepsilon_{mt+1} \\ 0 \\ 0 \end{bmatrix}. \tag{2.27}$$

Some impulse response functions (dynamic simulations obtained from setting $\varepsilon_{mt} = 1$ in $t = 0$ but zero in all other periods) are shown in Figure 2.1. In Figure 2.1.a, price adjustment is fairly slow (many prices are fixed in spite of an increase in nominal demand), so a monetary shock leads to a relatively large effect on output: money is far from neutral. In Figure 2.1.b, price adjustment is much faster (the rate at which an occasion to change the price arrives is much higher), so the monetary shock has almost no effect on output: money is almost neutral. In Figure 2.1.c also has fat price adjustment, but now because ϕ is high (quickly decreasing returns to scale or strong monopoly power), which makes it too costly for firms to keep their old prices.

2.3.4 The Calvo Model and the "Natural Rate Hypothesis"

Reference: Walsh 5.5.

The "natural rate hypothesis" states that the mean of output cannot be affected by

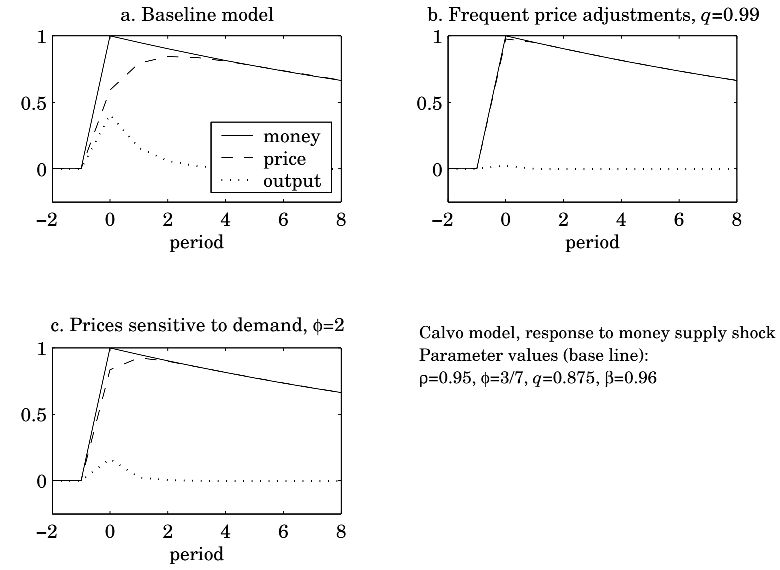

Figure 2.1: Impulse responses in the Calvo model

any monetary policy. Suppose the central bank can change the inflation rate by changing its policy instrument. Take the unconditional expectation of the Rotemberg/Calvo model (2.21) and use iterated expectations and $\mathrm{E}\varepsilon_t = 0$ to get

$$\mathrm{E}y_t = \frac{\mathrm{E}\Delta p_t - \beta \mathrm{E}\Delta p_{t+1}}{\delta\phi}. \tag{2.28}$$

If $\beta = 1$ ($\beta < 1$), and inflation is a stationary series so $\mathrm{E}\Delta p_t = \mathrm{E}\Delta p_{t+1}$, then this means that inflation cannot (can) affect average output. Irrespective of whether $\beta = 1$ or not, a drifting inflation rate ($\mathrm{E}\Delta p_t \neq \mathrm{E}\Delta p_{t+1}$) can certainly affect average output.

This should probably be regarded as an artifact of the Calvo model. It puts restrictions on which type of policy experiments which are meaningful to analyze with the help of this model: we should probably only use this model for policy changes which keeps the average inflation rate unchanged. In many applications, the Phillips equation is assumed

to refer to detrended output (as a measure of the business cycle). The main reason is that the Phillips effect is typically only relevant for as long as the production function has decreasing returns to scale, see the discussion of (2.20). Since detrended output per definition has a zero mean the kind of experiments that changes $\mathrm{E}y_t$ must be ruled out.

2.4 Aggregate Demand

The period utility function is

$$U(C_t) = \frac{A_t}{1-\gamma} C_t^{1-\gamma}, \tag{2.29}$$

where A_t is a taste shift parameter. The Euler equation for optimal consumption is

$$\frac{\partial U(C_t)}{\partial C_t} = \beta \mathrm{E}_t \left[\frac{\partial U(C_{t+1})}{\partial C_{t+1}} Q_{t+1} \right], \tag{2.30}$$

where Q_{t+1} is the gross real return.

The marginal utility of C_t is

$$\frac{\partial U(C_t)}{\partial C_t} = A_t C_t^{-\gamma}, \tag{2.31}$$

so the optimality condition can be written

$$\begin{aligned} 1 &= \beta \mathrm{E}_t Q_{t+1} \frac{A_{t+1}}{A_t} \left(\frac{C_{t+1}}{C_t} \right)^{-\gamma} \\ &= \beta \mathrm{E}_t \exp\left(\ln Q_{t+1} + \Delta \ln A_{t+1} - \gamma \ln C_{t+1} + \gamma \ln C_t \right). \end{aligned} \tag{2.32}$$

Assume that $\ln Q_{t+1}$, $\ln A_{t+1}$, and $\ln C_{t+1}$ are jointly normally distributed. (Recall $\mathrm{E}\exp(x) = \exp(\mathrm{E}x + \mathrm{Var}(x)/2)$ is x is normally distributed.) Take logs of (2.32) and rewrite it as

$$\begin{aligned} 0 &= \ln\beta + \mathrm{E}_t \ln Q_{t+1} + \mathrm{E}_t \Delta \ln A_{t+1} - \gamma \mathrm{E}_t \ln C_{t+1} + \gamma \ln C_t \\ &\quad + \mathrm{Var}_t \left(\ln Q_{t+1} + \ln A_{t+1} - \gamma \ln C_{t+1} \right)/2, \text{ or} \\ \mathrm{E}_t \ln C_{t+1} &= \ln C_t + \frac{1}{\gamma} \mathrm{E}_t \ln Q_{t+1} + \frac{1}{\gamma} \mathrm{E}_t z_{t+1}, \end{aligned} \tag{2.33}$$

where $\mathrm{E}_t z_{t+1} = \ln\beta + \mathrm{E}_t \Delta \ln A_{t+1} + \mathrm{Var}_t(.)$. The most important part of $\mathrm{E}_t z_{t+1}$ is $\mathrm{E}_t \Delta \ln A_{t+1}$. If $\ln A_{t+1} = \rho \ln A_t + u_{t+1}$, then $\mathrm{E}_t \Delta \ln A_{t+1} = (\rho - 1) \ln A_t$, so the AR(1) formulation carries over to the expected change, but the sign is reversed if $\rho > 0$.

2.5 Recent Models for Studying Monetary Policy

This section gives an introduction to more recent models of monetary policy. Such models typically combine a forward looking Phillips curve, for instance, from a Calvo model, with an aggregate demand equation derived from an optimizing consumer's intertemporal consumption/savings decision, and some kind of policy rule or objective function for the central bank.

2.5.1 A Simple Model

Price are set as in the *Calvo model*. In this model, a fraction q of the firms are allowed to set a new price in a period, and the fraction $1 - q$ must keep their old price. When allowed to change the price, the firms chooses a price to minimize a discounted sum of the squared deviations of the actual price and the flex price. We also assume that the flex price is determined as in model of monopolistic competition, $p_{it}^* = p_t + \phi y_t + \varepsilon_{\pi t}$, where ϕ measures how much price setters wants to increase the relative price when demand increases (ϕ is high when the substitution elasticities between goods is low and when the marginal cost curve is steep). The supply side of the economy can then be summarized by the "Phillips curve"

$$\pi_t = \beta \mathrm{E}_t \pi_{t+1} + \delta \left(\phi y_t + \varepsilon_{\pi t} \right), \tag{2.34}$$

where δ is increasing in the fraction q.

The "aggregate demand" curve is derived from an Euler condition for optimal consumption choice with taste shocks, combined with the assumption that consumption equals output. It is

$$\mathrm{E}_t y_{t+1} = y_t + \frac{1}{\gamma} \left(i_t - \mathrm{E}_t \pi_{t+1} \right) + \varepsilon_{yt}, \tag{2.35}$$

where ε_{yt} is a *negative* shock to current (time t) demand.

The central bank sets short interest rate, i_t. This can have effect on output since prices are sticky, so the nominal interest rate affects the real interest rate. This, in turn, affects demand, and thus inflation through the "Phillips effect." Suppose the *reaction function*, also called *simple policy rule*, of the central bank is a "Taylor rule"

$$i_t = \chi \pi_t + \upsilon y_t. \tag{2.36}$$

This is a sub-optimal commitment policy. It is a commitment rule since the policy setter will stick to this rule, even if it would be optimal to deviate from it in certain states. The optimal commitment rule, however, would not restrict the decision rule to be a function of y_t and π_t only.

Note that there is no money demand function in this model. The reason is that monetary policy is specified in terms of the interest rate, so the money stock becomes demand determined (the money supply curve is flat at the chosen nominal interest rate). Of course, in order for the central bank to control anything of importance, there must be a demand for money. The money demand function could be added to the model, but its only role is to determine the money stock.

Suppose the shocks in (2.34) and (2.35) follow

$$\begin{aligned} \varepsilon_{\pi t+1} &= \tau_\pi \varepsilon_{\pi t} + \zeta_{\pi t+1} \\ \varepsilon_{yt+1} &= \tau_y \varepsilon_{yt} + \zeta_{yt+1}. \end{aligned} \tag{2.37}$$

We can write (2.34)–(2.37) as

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \beta & 0 \\ 0 & 0 & \frac{1}{\gamma} & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_{\pi t+1} \\ \varepsilon_{yt+1} \\ \mathrm{E}_t \pi_{t+1} \\ \mathrm{E}_t y_{t+1} \end{bmatrix} = \begin{bmatrix} \tau_\pi & 0 & 0 & 0 \\ 0 & \tau_y & 0 & 0 \\ -\delta & 0 & 1 & -\delta\phi \\ 0 & 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_{\pi t} \\ \varepsilon_{yt} \\ \pi_t \\ y_t \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ \frac{1}{\gamma} \end{bmatrix} i_t + \begin{bmatrix} \zeta_{\pi t+1} \\ \zeta_{yt+1} \\ 0 \\ 0 \end{bmatrix}, \tag{2.38}$$

with

$$i_t = \begin{bmatrix} 0 & 0 & \chi & \upsilon \end{bmatrix} \begin{bmatrix} \varepsilon_{\pi t} \\ \varepsilon_{yt} \\ \pi_t \\ y_t \end{bmatrix}. \tag{2.39}$$

This system is in *state space form* and could be summarized as

$$\tilde{A}_0 \begin{bmatrix} x_{1t+1} \\ \mathrm{E}_t x_{2t+1} \end{bmatrix} = \tilde{A} \begin{bmatrix} x_{1t} \\ x_{2t} \end{bmatrix} + \tilde{B} i_t + \tilde{\xi}_{t+1}, \text{ and} \tag{2.40}$$

$$i_t = -F \begin{bmatrix} x_{1t} \\ x_{2t} \end{bmatrix}. \tag{2.41}$$

where x_{1t} is a vector of predetermined variables (here $\varepsilon_{\pi t}$ and ε_{yt}, which happens to be exogenous, but also endogenous variables can be predetermined) and x_{2t} a vector of forward looking variables (here π_t and y_t). Premultiply (2.40) with $\tilde{A}_0^{-1}$ to get

$$\begin{bmatrix} x_{1t+1} \\ \mathrm{E}_t x_{2t+1} \end{bmatrix} = A \begin{bmatrix} x_{1t} \\ x_{2t} \end{bmatrix} + B i_t + \xi_{t+1}, \text{ where} \tag{2.42}$$

$$A = \tilde{A}_0^{-1} \tilde{A},\ B = \tilde{A}_0^{-1} \tilde{B}, \text{ and } \mathrm{Cov}(\xi_t) = \tilde{A}_0^{-1} \mathrm{Cov}\left(\tilde{\xi}_t\right) \tilde{A}_0^{-1\prime}. \tag{2.43}$$

By using the policy rule (2.41) in (2.42) we get

$$\begin{bmatrix} x_{1t+1} \\ \mathrm{E}_t x_{2t+1} \end{bmatrix} = (A - BF) \begin{bmatrix} x_{1t} \\ x_{2t} \end{bmatrix} + \xi_{t+1}. \tag{2.44}$$

This system of expectational difference equations (with stable and unstable roots) can be solved in several different ways. For instance, a decomposition of $A - BF$ in terms of eigenvalues and eigenvectors will work if the latter are linearly independent. Otherwise, other techniques must be used (see, for instance, Söderlind (1999)). A necessary condition for a unique saddle path equilibrium is that $A - BF$ has as many stable roots (inside the unit circle) as there are predetermined variables (that is, elements in x_{1t}).

To solve the model numerically, parameter values are needed. The following values have been used in most of *Figures 2.2-2.4* (exceptions are indicated)

β	δ	ϕ	γ	τ_π	τ_y	υ	χ	λ_y	λ_i
0.99	2.25	2/7	2	0.5	0.5	0.5	1.5	0.5	0

The choice of δ implies relatively little price stickiness. The choice of ϕ means that a 1% increase in aggregate demand leads to a desired increase of the relative price of 2/7%. The choice of the relative risk aversion γ implies an elasticity of intertemporal substitution of 1/2. The υ and χ are those advocated by Taylor. The loss function parameters (see next

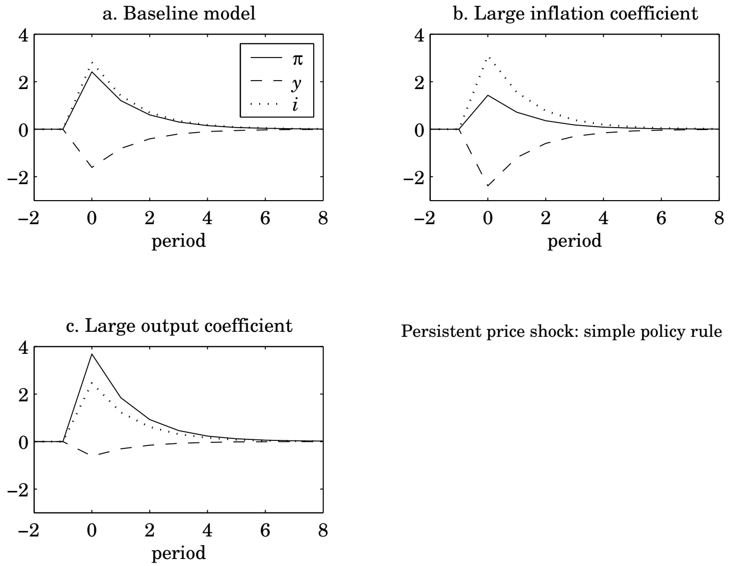

Figure 2.2: Impulse responses to price shock; simple policy rule

section) means that inflation is twice as important as output, and that the policy maker does not care about fluctuations in the nominal interest rate.

The first subfigure in Figure 2.2 illustrates how the model with the policy rule (2.36) works. An inflation shock in period $t = 0$ increases inflation. The policy maker reacts by raising the nominal interest even more in order to increase the real interest rate. This, in turn, has a negative effect on output and therefore on inflation via the "Phillips curve." The central bank creates a recession to bring down inflation. The other subfigures illustrates what happens if the coefficients in the reaction function (2.36) are changed.

2.5.2 Optimal Monetary Policy

Suppose the central bank's loss function is

$$\mathrm{E}_t \sum_{s=0}^{\infty} \beta^s L_{t+s}, \text{ where} \tag{2.45}$$

$$L_{t+s} = \left(\pi_{t+s} - \pi^*\right)^2 + \lambda_y \left(y_{t+s} - y^*\right)^2 + \lambda_i \left(i_{t+s} - i^*\right)^2 . \tag{2.46}$$

A particularly straightforward way to proceed is to optimize (2.45), by restricting the policy rule to be of the simple form discussed above, (2.36). Optimization then proceeds as follows: guess the coefficients υ and χ, solve the model, use the time series representation of the model to calculate the loss function value. Then try other coefficients υ and χ, and see if they give a lower loss function value. Continue until the best coefficients have been found.

The unrestricted optimal commitment policy and the optimal discretionary policy rule are a bit harder to find. Methods for doing that are discussed in, among other places, Söderlind (1999).

Figure 2.3 compares the equilibria under the simple policy rule, unrestricted optimal commitment rule, and optimal discretionary rule, when it is assumed that $\pi^* = y^* = 0$. It is clear that the optimal commitment rule achieves a much more stable inflation and output, in spite of a less vigorous increase in the nominal interest rate. This is achieved by *credibly* promising to keep interest rates high in the future (and even raise further), which gives expectations of lower future output and therefore future inflation. This, in turn, gives lower inflation and output today. The discretionary equilibrium is fairly similar to the simple rule in this model. Note that there is no constant "inflation bias" when target levels are at their natural levels (zero) as they are in these figures. The discretionary rule is still different from the commitment rule (they are, after all, outcomes of different games). The intuition is that there is a time-varying "bias" since the conditional expectations of output and inflation in the next periods (their "conditional natural rates") typically differ from the target rates (here zero).

Figure 2.4 makes the same type of comparison, but for a positive demand shock, $-\varepsilon_{yt}$. In this case, both optimal rules "kill" the demand shock, which is seen almost directly from (2.35): any shock ε_{yt} could be met by increasing i_t by $\gamma\varepsilon_{yt}$. In this way output is unaffected by the shock, and there will then be no effect on inflation either, since the only

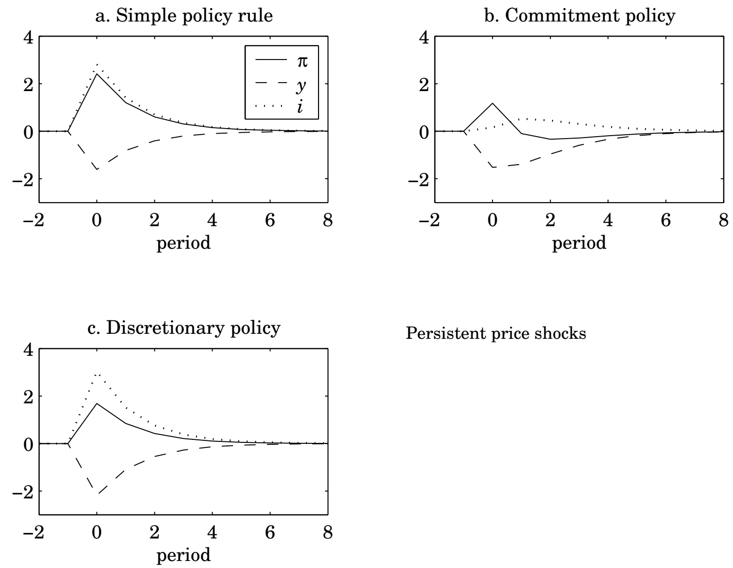

Figure 2.3: Impulse responses to price shock: simple rule, optimal commitment policy, and discretionary policy

way the demand shock can affect inflation is via output (see (2.34)). This is very similar to the static model discussed above: the demand shock drives both prices and output in the same direction and should, if possible, neutralized. Of course, the result hinges on the assumption that the policy maker is not averse to movements in the nominal interest rate, that is, $\lambda_i = 0$ in (2.46). (It can be shown that this case can be approximated in the simple policy rule (2.36) by setting the coefficients very high.) Many studies indicate that central banks are unwilling to let the nominal interest rate vary much. This is sometimes interpreted as a concern for the banking sector, and sometimes as due to uncertainty about the state of the economy and/or the effect of policy changes on output/inflation. In any case, $\lambda_i > 0$ is often necessary in order to make this type of model fit the observed variability in nominal interest rates.

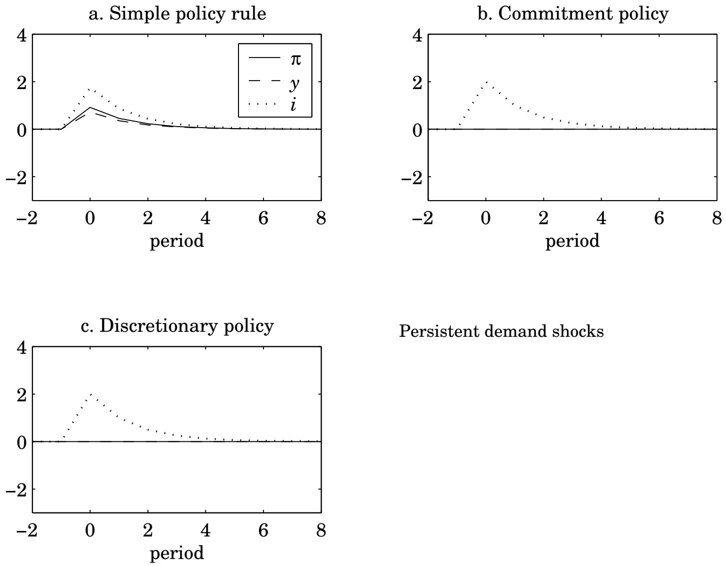

Figure 2.4: Impulse responses to positive demand shock: simple rule, optimal commitment policy, and discretionary policy

Bibliography

Bénassy, J.-P., 1995, "Money and Wage Contracts in an Optimizing Model of the Business Cycle," *Journal of Monetary Economics*, 35, 303–315.

Blanchard, O. J., and S. Fischer, 1989, *Lectures on Macroeconomics*, MIT Press.

Burda, M., and C. Wyplosz, 1997, *Macroeconomics - A European Text*, Oxford University Press, 2nd edn.

Cooley, T. F., and G. D. Hansen, 1995, "Money and the Business Cycle," in Thomas F. Cooley (ed.), *Frontiers of Business Cycle Research*, Princeton University Press, Princeton, New Jersey.

Lucas, R. E., 2000, "Inflation and Welfare," *Econometrica*, 68, 247–274.

Obstfeldt, M., and K. Rogoff, 1996, *Foundations of International Macroeconomics*, MIT Press.

Romer, D., 1996, *Advanced Macroeconomics*, McGraw-Hill.

Rotemberg, J. J., 1982a, "Monopolistic Price Adjustment and Aggregate Output," *Review of Economic Studies*, 49, 517–531.

Rotemberg, J. J., 1982b, "Sticky Prices in the United States," *Journal of Political Economy*, 60, 1187–1211.

Rotemberg, J. J., 1987, "New Keynesian Microfoundations," in Stanley Fischer (ed.), *NBER Macroeconomics Annual* . pp. 69–104, NBER.

Söderlind, P., 1999, "Solution and Estimation of RE Macromodels with Optimal Policy," *European Economic Review*, 43, 813–823.

Walsh, C. E., 1998, *Monetary Theory and Policy*, MIT Press, Cambridge, Massachusetts.

3 Looking into Some Recent Models of Monetary Policy

3.1 A Baseline Model

Reference: Paul Söderlind's lecture notes MacPol.TeX; Clarida, Galí, and Gertler (1999)

Prices are set as in the Calvo model (See Rotemberg (1987) and MacPri.TeX for derivations)

$$\pi_t = \beta \mathrm{E}_t \pi_{t+1} + \delta \left(\phi y_t + \varepsilon_{\pi t} \right). \tag{3.1}$$

The parameter ϕ captures the degree to which monopolistic competitor j wants to increase its relative price as demand increases, so the log desired price is $p^{j*} = p + \phi y$, where p is the average price level and y is log aggregate demand. Increasing marginal costs and a low demand elasticity make ϕ large. The parameter δ is determined as

$$\delta = \frac{q}{1-q} \left[1 - \beta \left(1 - q \right) \right], \tag{3.2}$$

where β is the discount rate, and q the fraction of firms that can change the price in each period. δ is increasing in q.

The "aggregate demand" curve is derived in Section 3.4.1 from an Euler condition for optimal consumption choice with taste shocks, combined with the assumption that consumption equals output

$$\mathrm{E}_t y_{t+1} = y_t + \frac{1}{\gamma} \left(i_t - \mathrm{E}_t \pi_{t+1} \right) + \varepsilon_{yt}, \tag{3.3}$$

where ε_{yt} is a *negative* shock to current (time t) demand.

The central bank sets short interest rate, i_t. This can affect output since prices are sticky, so changes in the nominal interest rate change the real interest rate. This will then influence price setting via the "Phillips effect" in (3.1).

Suppose the shocks in (3.1) and (3.3) follows

$$\begin{aligned} \varepsilon_{\pi t+1} &= \tau_\pi \varepsilon_{\pi t} + \zeta_{\pi t+1} \\ \varepsilon_{yt+1} &= \tau_y \varepsilon_{yt} + \zeta_{yt+1}. \end{aligned} \tag{3.4}$$

We can then write (3.1)-(3.4) as

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \beta & 0 \\ 0 & 0 & \frac{1}{\gamma} & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_{\pi t+1} \\ \varepsilon_{yt+1} \\ \mathrm{E}_t \pi_{t+1} \\ \mathrm{E}_t y_{t+1} \end{bmatrix} = \begin{bmatrix} \tau_\pi & 0 & 0 & 0 \\ 0 & \tau_y & 0 & 0 \\ -\delta & 0 & 1 & -\delta\phi \\ 0 & 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_{\pi t} \\ \varepsilon_{yt} \\ \pi_t \\ y_t \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ \frac{1}{\gamma} \end{bmatrix} i_t + \begin{bmatrix} \zeta_{\pi t+1} \\ \zeta_{yt+1} \\ 0 \\ 0 \end{bmatrix}. \tag{3.5}$$

This system is in *state space form* and could be summarized as

$$\tilde{A}_0 \begin{bmatrix} x_{1t+1} \\ \mathrm{E}_t x_{2t+1} \end{bmatrix} = \tilde{A} \begin{bmatrix} x_{1t} \\ x_{2t} \end{bmatrix} + \tilde{B} i_t + \tilde{\xi}_{t+1}, \text{ and} \tag{3.6}$$

where x_{1t} is a vector of predetermined variables (here $\varepsilon_{\pi t}$ and ε_{yt}, which are both exogenous, but also endogenous state variables could be predetermined), and x_{2t} a vector of forward looking variables (here π_t and y_t). Premultiply (3.6) with $\tilde{A}_0^{-1}$ to get

$$\begin{bmatrix} x_{1t+1} \\ \mathrm{E}_t x_{2t+1} \end{bmatrix} = A \begin{bmatrix} x_{1t} \\ x_{2t} \end{bmatrix} + B i_t + \xi_{t+1}, \text{ where} \tag{3.7}$$

$$A = \tilde{A}_0^{-1} A_1,\ B = \tilde{A}_0^{-1} \tilde{B}, \text{ and } \mathrm{Cov}\left(\xi_t\right) = \tilde{A}_0^{-1} \mathrm{Cov}\left(\tilde{\xi}_t\right) \tilde{A}_0^{-1\prime}. \tag{3.8}$$

3.1.1 Optimal Monetary Policy

Suppose the loss function is

$$\mathrm{E}_t \sum_{s=0}^{\infty} \beta^s L_{t+s}, \text{ where} \tag{3.9}$$

$$L_{t+s} = \left(\pi_{t+s} - \pi^* \right)^2 + \lambda_y \left(y_{t+s} - y^* \right)^2 + \lambda_i \left(i_{t+s} - i^* \right)^2. \tag{3.10}$$

A particularly straightforward way to proceed is to optimize (3.9), by restricting the policy rule to be a *simple (commitment) rule* like the Taylor rule

$$i_t = \chi \pi_t + \omega y_t. \tag{3.11}$$

Solution methods for this case, as well as for the unrestricted optimal commitment rule and the optimal discretionary rule is discussed in, among other places, Söderlind (1999).

3.1.2 Impulse Response Functions

See MacPol.TeX for a detailed discussion, but note also the following.

If the policy maker does not care about the volatility of the nominal interest rate, $\lambda_i = 0$, then it is always optimal to counter balance any demand (output) shock entirely. This is seen directly from the aggregate demand curve (3.3): any shock ε_{yt} could be counter balanced by changing i_t by $\gamma\varepsilon_{yt}$. In this way output is unaffected by the shock, and there will then be no effect on inflation either, since the only way the demand shock can affect inflation is via output (see (3.1)). The model extensions discussed below share this feature, and so will most models where policy have a contemporaneous effect on output.

The parameters used in these and subsequent figures are

β	δ	ϕ	γ	τ_π	τ_y	ω	χ	λ_y	λ_i	η
0.99	2.25	3/7	2	0 or 0.5	0 or 0.5	0.5	1.5	0.5	0	0.5

The choice of δ implies that q in (3.2) is around 0.75, which implies relatively little price stickiness. The choice of ϕ means that a 1% increase in aggregate demand leads to a desired increase of the relative price of $3/7$%. The choice of the relative risk aversion γ implies an elasticity of intertemporal substitution of 1/2. The ω and χ are those advocated by Taylor. The loss function parameters means that inflation is twice as important as output, and that the policy maker does not care about fluctuations in the nominal interest rate.

Figure 3.1.a-f show some impulse response functions for the price shock ε_{π_t}. In subfigures a-c, the shock is not autocorrelated, but in subfigures d-f it has an autocorrelation coefficient of 0.5. The "endogenous" dynamics of the model is quite weak: We need autocorrelated shocks to replicate data, or, as an alternative, some partial adjustment structure. It can be shown that this is true even if the degree of price inertia is increased (δ lowered). It will also be true in the model extensions discussed below.

There is, however, somewhat more dynamics in the commitment equilibrium. (This is also the theme in a recent paper by Woodford.) The intuition is that the policy maker in this case can make credible promises about future policy and thereby affect the ex-

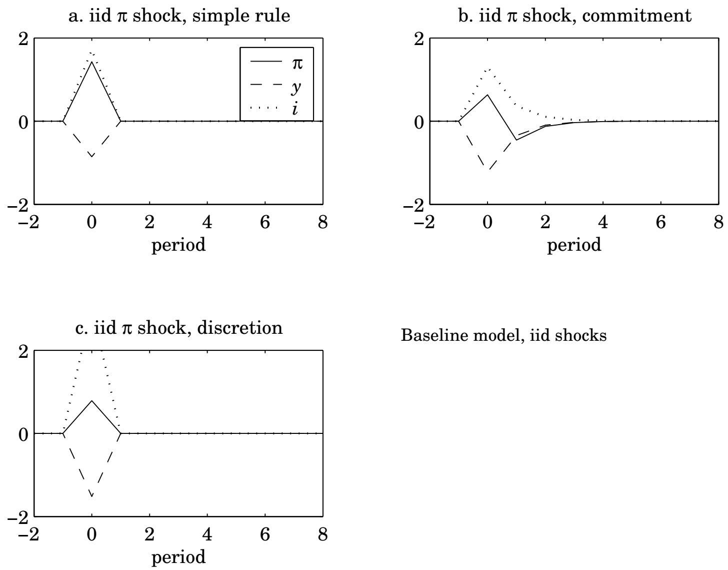

Figure 3.1: Baseline model, iid shocks

pectations of private agents, which in turn affects behavior today. In this way, the policy maker is able to stabilize the economy more effectively. Some clues to this can be gained by comparing the algebraic expressions of the policy rules in, for instance, (Söderlind (1999)). The discretionary case can be summarized by

$$x_{1t+1} = M^d x_{1t} + \varepsilon_{t+1}, \tag{3.12}$$

$$u_t = -F^d x_{1t}, \tag{3.13}$$

so the policy instrument, u_t, depends on x_{1t} only, which is a VAR(1). In contrast, the

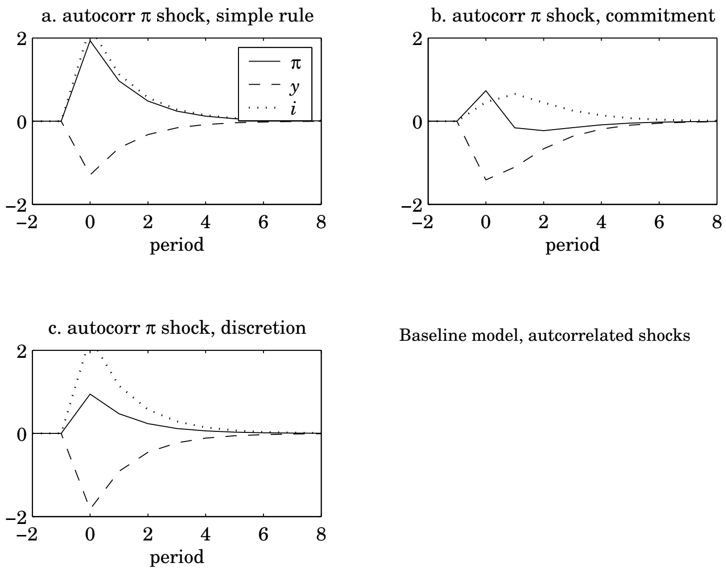

Figure 3.2: Baseline model, autocorrelated shocks

commitment case can be summarized by

$$\begin{bmatrix} x_{1t+1} \\ \rho_{2t+1} \end{bmatrix} = M^c \begin{bmatrix} x_{1t} \\ \rho_{2t} \end{bmatrix} + \begin{bmatrix} \varepsilon_{t+1} \\ \mathbf{0}_{n_2 \times 1} \end{bmatrix}, \text{ and} \tag{3.14}$$

$$u_t = -F^c \begin{bmatrix} x_{1t} \\ \rho_{2t} \end{bmatrix}. \tag{3.15}$$

The initial shadow prices of the forward looking variables are zero, $\rho_{20} = \mathbf{0}_{n_2 \times 1}$ and there are no shocks to ρ_{2t}. It is therefore possible to rewrite (3.14) in terms of $\{x_{1s}\}_{s=0}^{t}$ only, but this representation involves many more lags than the VAR(1) in the discretionary solution, (3.12).

As a final remark, the AR(1) representation of the discretionary equilibrium (3.12)-(3.13) looks deceptively similar to "simple policy rule" case, which also gives an AR(1)

of x_{1t} and a decision rule which is linear in x_{1t}. Note, however, that the two equilibria are quite different. For instance, plugging in F^d from (3.13) into a simple policy rule and solving for the equilibrium will not give the same AR(1) matrix as in (3.12).

3.1.3 Handling of Shocks to Forward Looking Equations*

Note that both forward looking equations (3.1) and (3.3) have shocks ($\varepsilon_{\pi t}$ and ε_{yt}). This is most easily handled by making both these shocks part of the state vector, so the state space formulation expresses the expected values of next periods forward looking variables ($\mathrm{E}_t\, \pi_{t+1}$ and $\mathrm{E}_t\, y_{t+1}$) in terms of today's state variables ($\varepsilon_{\pi t}$ and ε_{yt}) and forward looking variables (π_t and y_t). Note that this continues to be true even if there is no autocorrelation in $\varepsilon_{\pi t}$ and ε_{yt} (which in itself makes it natural to put them in the state vector).

3.1.4 Handling Identities*

Suppose we want to have the price level in state space form (3.5), for instance, because the loss function includes the price level. In other models, it may be the case that some equations are more easily expressed in inflation rates, while other equations include the price level. There are two ways to include the price levels. First, the model can be rewritten in terms of the price levels only. Second, add an identity.

As a very simple illustration, consider the simplified model of exogenous output ($y_{t+1} = \rho y_t + \varepsilon_{yt+1}$) and forward looking Phillips curve with no price shock ($\pi_t = \beta \mathrm{E}_t \pi_{t+1} + \psi y_t$). The state space formulation is

$$\begin{bmatrix} 1 & 0 \\ 0 & \beta \end{bmatrix} \begin{bmatrix} y_{t+1} \\ \mathrm{E}_t\, \pi_{t+1} \end{bmatrix} = \begin{bmatrix} \rho & 0 \\ -\psi & 1 \end{bmatrix} \begin{bmatrix} y_t \\ \pi_t \end{bmatrix} + \begin{bmatrix} \varepsilon_{yt} \\ 0 \end{bmatrix}. \tag{3.16}$$

To substitute the price levels for the inflation rate we rewrite is as

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -\beta & 0 & \beta \end{bmatrix} \begin{bmatrix} p_t \\ y_{t+1} \\ \mathrm{E}_t\, p_{t+1} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & \rho & 0 \\ -1 & -\psi & 1 \end{bmatrix} \begin{bmatrix} p_{t-1} \\ y_t \\ p_t \end{bmatrix} + \begin{bmatrix} 0 \\ \varepsilon_{yt} \\ 0 \end{bmatrix}. \tag{3.17}$$

The first equation is just a dynamic identity ($p_7 = p_7$ even if we are in $t = 8$). In this formulation, the lagged price level, p_{t-1}, is predetermined and the current price level, p_t, is forward looking. Note that (3.17) is still on the form (3.6). It is clear that an output shock, ε_{yt}, will have a temporary effect on output and inflation, but a permanent effect on the price level: the price level is non-stationary. This type of non-stationarity can sometimes be a problem in the solution algorithms and should perhaps be avoided if possible.

The second possibility is to add the price level (current and lagged) to the state space form, and to link it to the inflation rate by an identity ($\pi_t = p_t - p_{t-1}$)

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \beta & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} p_t \\ y_{t+1} \\ \mathrm{E}_t\, \pi_{t+1} \\ \mathrm{E}_t\, p_{t+1} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & \rho & 0 & 0 \\ 0 & -\psi & 1 & 0 \\ -1 & 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} p_{t-1} \\ y_t \\ \pi_t \\ p_t \end{bmatrix} + \begin{bmatrix} 0 \\ \varepsilon_{yt} \\ 0 \\ 0 \end{bmatrix}. \tag{3.18}$$

This is just an extension of the original system. In fact, the second and third equations are exactly the same as (3.16). The first equation is the same a dynamic identity as before, and the fourth equation is a static identity ($\pi_t = p_t - p_{t-1}$). The problem with this formulation is that the matrix on the left hand side is singular, so we cannot write the model on the form (3.6) with an invertible $\tilde{A}_0$ matrix. However, the singularity is confined to the forward looking equations so we can write (3.18) on the form

$$\tilde{A}_0 \begin{bmatrix} I_{n_1} & \mathbf{0}_{n_1\times n_2} \\ \mathbf{0}_{n_2\times n_1} & H \end{bmatrix} \begin{bmatrix} x_{1t+1} \\ \mathrm{E}_t x_{2t+1} \end{bmatrix} = \tilde{A} \begin{bmatrix} x_{1t} \\ x_{2t} \end{bmatrix} + \tilde{B} i_t + \tilde{\xi}_{t+1}, \tag{3.19}$$

where H is singular. This model formulation can often be handled with a slightly modified solution algorithm.

3.1.5 Adding Monetary Policy Shocks*

VAR models of monetary policy typically emphasize that the impulse response to a monetary policy shocks can tell us a lot about how the economy works. So far, there is no monetary policy shock in this model. One way of getting such a shock is to add a stochastic element to the loss function. Another (crude and simple) way is to simply postulate that the interest rate that affect the private sector is $i_{2t} = i_t + \varepsilon_{it}$, where ε_{it} is an exogenous disturbance. Since i_t represents the systematic policy, i_{2t} is systematic policy plus the shock. With this interpretation, i_{2t} is the interest rate and i_t is what the interest rate would have been in absence of the policy shock. For instance, with the Taylor rule (3.11)

we get $i_{2t} = \chi\pi_t + \omega y_t + \varepsilon_{it}$.

It is straightforward to modify (3.5) to incorporate such a shock: add ε_{it} to the vector of predetermined variables and make sure that it affects all other variables in the same way as i_t does. If we let ε_{it} be an AR(1), then we get

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & \beta & 0 \\ 0 & 0 & 0 & \frac{1}{\gamma} & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_{it+1} \\ \varepsilon_{\pi t+1} \\ \varepsilon_{yt+1} \\ \mathrm{E}_t \pi_{t+1} \\ \mathrm{E}_t y_{t+1} \end{bmatrix} = \begin{bmatrix} \tau_i & 0 & 0 & 0 & 0 \\ 0 & \tau_\pi & 0 & 0 & 0 \\ 0 & 0 & \tau_y & 0 & 0 \\ 0 & -\delta & 0 & 1 & -\delta\phi \\ \frac{1}{\gamma} & 0 & 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_{it} \\ \varepsilon_{\pi t} \\ \varepsilon_{yt} \\ \pi_t \\ y_t \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \frac{1}{\gamma} \end{bmatrix} i_t + \begin{bmatrix} \zeta_{it+1} \\ \zeta_{\pi t+1} \\ \zeta_{yt+1} \\ 0 \\ 0 \end{bmatrix}. \tag{3.20}$$

We can now solve the model and trace out the impulse response with respect to ζ_{it+1} (possibly with $\tau_i = 0$), which can be compared with the results from a VAR.

3.2 Model Extension 1: Predetermined Prices

We keep the demand curve in (3.3), but assume that prices are set as in the Calvo model, but that they have to be set one period in advance. It is straightforward to see that this changes (3.1) to (after forwarding one period)

$$\pi_{t+1} = \beta \mathrm{E}_t \pi_{t+2} + \delta \mathrm{E}_t \left(\phi y_{t+1} + \varepsilon_{\pi t+1} \right). \tag{3.21}$$

Since π_{t+1} is know already in t, we can replace $\mathrm{E}_t \pi_{t+1}$ by π_{t+1} in (3.3). If we use the fact that $E_t \varepsilon_{\pi t+1} = \tau_\pi \varepsilon_{\pi t}$ in (3.21), then state space for can be written

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & \beta & \delta\phi \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_{\pi t+1} \\ \varepsilon_{yt+1} \\ \pi_{t+1} \\ E_t \pi_{t+2} \\ E_t y_{t+1} \end{bmatrix} = \begin{bmatrix} \tau_\pi & 0 & 0 & 0 & 0 \\ 0 & \tau_y & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ -\delta\tau_\pi & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & -\frac{1}{\gamma} & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_{\pi t} \\ \varepsilon_{yt} \\ \pi_t \\ \pi_{t+1} \\ y_t \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \frac{1}{\gamma} \end{bmatrix} i_t + \begin{bmatrix} \zeta_{\pi t+1} \\ \zeta_{yt+1} \\ 0 \\ 0 \\ 0 \end{bmatrix}. \quad (3.22)$$

In this case, $\varepsilon_{\pi t}$, ε_{yt}, and π_t are predetermined, and π_{t+1} and y_t are forward looking.

The Taylor rule (3.11) can be written

$$i_t = \begin{bmatrix} 0 & 0 & \chi & 0 & \omega \end{bmatrix} \begin{bmatrix} \varepsilon_{\pi t} \\ \varepsilon_{yt} \\ \pi_t \\ \pi_{t+1} \\ y_t \end{bmatrix}. \quad (3.23)$$

3.2.1 Impulse Response Functions

Figures 3.3 show impulse responses to a persistent price shock of the model with predetermined prices.

Prices cannot change in $t = 0$ (the time of the shock), but output and the nominal interest rate can. In both optimal rules, the nominal interest rate is indeed changed in $t = 0$. In the simple rule, it is too, but only by little as long as the reaction function coefficient of output is low - since output moves only very little in $t = 0$.

The commitment response to a price shock is interesting. In $t = 0$, the nominal interest rate is lowered in order to drive down the real interest rate (magnified by the price shock which hits π_1, so $E_0\pi_1$ is high). This tilts the consumption schedule in favor of consuming in $t = 0$ instead of $t = 1$. The low (expected) output in $t = 1$ drives down π_1 by the Phillips effect in (3.21). It can be shown that we get this type of result also for

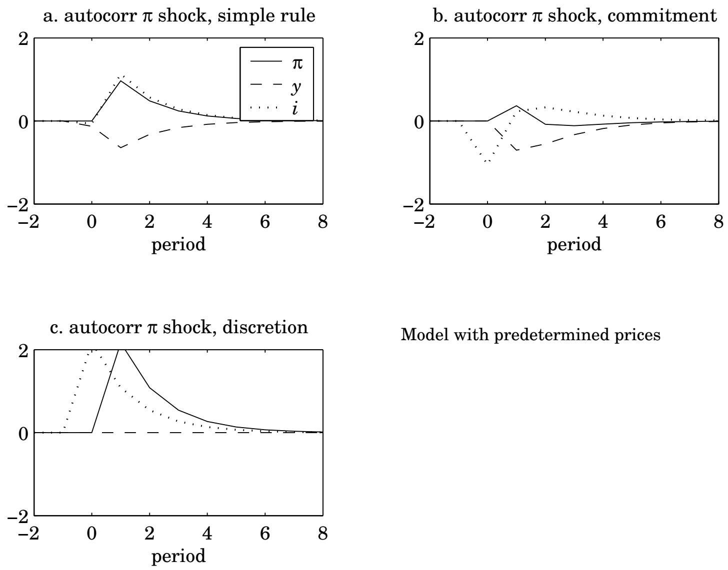

Figure 3.3: Predetermined prices

other parameter values, as long as the policy maker does not care too much about output.

The discretionary response to a price shock is also interesting. If the policy maker does not care about the fluctuations in the nominal interest rate, then output will be completely stabilized, and inflation is, effectively, left on its own. It can be shown that we get this type of result also for other parameter values, as long as the policy makers cares about output, $\lambda_y > 0$. To understand this result, consider the final period in this game, T. Then, the policy maker can affect y_T, but not π_T (set in $T-1$), by setting the nominal interest rate. Hence, it will set i_T in order to stabilize y_T, which it does by setting the nominal interest rate equal to expected inflation, that is, the ex ante real interest rate to zero (measured as a deviation from steady state). Private agents realize this in $T-1$, and form expectations accordingly. Of course, the situation is essentially the same in $T-1$, and so forth.

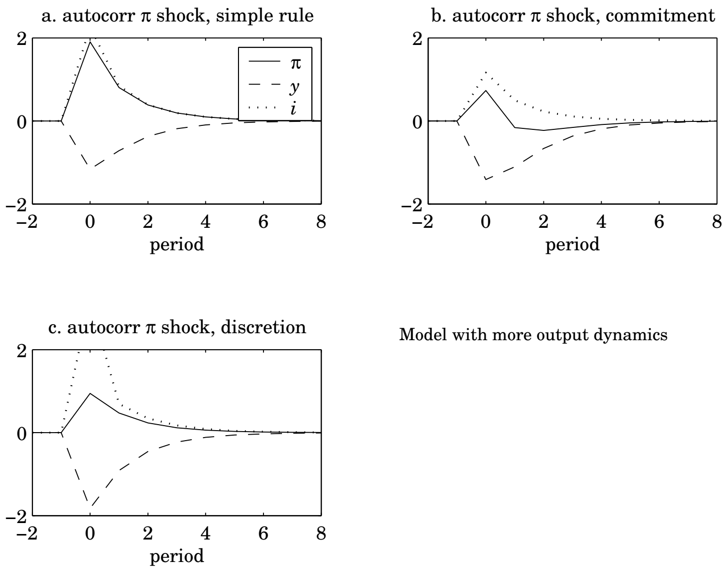

Figure 3.4: More output dynamics

3.3 Model Extension 2: More Output Dynamics

We now add some extra aggregate demand dynamics by assuming that period t utility depends negatively on aggregate consumption in $t-1$ (using the "Catching up with the Jonses" model in (Abel (1990))). If we once again assume that consumption equals output, then we get the following equation instead of (3.3)

$$\mathrm{E}_t y_{t+1} = \frac{\gamma - \eta(1-\gamma)}{\gamma} y_t + \eta \frac{1-\gamma}{\gamma} y_{t-1} + \frac{1}{\gamma}(i_t - \mathrm{E}_t \pi_{t+1}) + \varepsilon_{yt}, \eta > 0, \tag{3.24}$$

See Section 3.4.1 for a derivation. This aggregate demand curve is combined with standard Calvo model of price setting, (3.1).

The state space form can then be written

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & \beta & 0 \\ 0 & 0 & 0 & \frac{1}{\gamma} & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_{\pi t+1} \\ \varepsilon_{yt+1} \\ y_t \\ \mathrm{E}_t \pi_{t+1} \\ \mathrm{E}_t y_{t+1} \end{bmatrix} = \begin{bmatrix} \tau_\pi & 0 & 0 & 0 & 0 \\ 0 & \tau_y & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ -\delta & 0 & 0 & 1 & -\delta\phi \\ 0 & 1 & \eta\frac{1-\gamma}{\gamma} & 0 & \frac{\gamma-\eta(1-\gamma)}{\gamma} \end{bmatrix} \begin{bmatrix} \varepsilon_{\pi t} \\ \varepsilon_{yt} \\ y_{t-1} \\ \pi_t \\ y_t \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \frac{1}{\gamma} \end{bmatrix} i_t + \begin{bmatrix} \zeta_{\pi t+1} \\ \zeta_{yt+1} \\ 0 \\ 0 \\ 0 \end{bmatrix}. \tag{3.25}$$

In this case, $\varepsilon_{\pi t}$, ε_{yt}, and y_{t-1} are predetermined, and π_t and y_t are forward looking.

The Taylor rule (3.11) can be written

$$i_t = \begin{bmatrix} 0 & 0 & 0 & \chi & \omega \end{bmatrix} \begin{bmatrix} \varepsilon_{\pi t} \\ \varepsilon_{yt} \\ y_{t-1} \\ \pi_t \\ y_t \end{bmatrix}. \tag{3.26}$$

3.3.1 Impulse Response Functions

Figures 3.4 show impulse responses to a persistent price shock in the model with "Catching up with the Jonses." The results are very similar to the basic model, except that, in the optimal rules, the nominal interest rate is raised more at the time of the shock, but lowered back somewhat more quickly.

The intuition for why the nominal interest rate has to be increased more in $t=0$ is the following. Already in the basic model, the ex ante real interest rate is increased in $t=0$ in order to bring down consumption/output and thereby affect inflation via the Phillips curve. This means that $C_0 < \mathrm{E}_0 C_1$. In the "Catching up with the Jonses" model, the low average consumption in $t=0$ affects the (expected) marginal utility in $t=1$ negatively - a larger increase in the real interest rate is therefore required in order to make agents postpone consumption.

3.4 Appendix: Derivation of the Aggregate Demand Equation

3.4.1 Derivation of the Output Equation

The period utility function is

$$U(C_t) = \frac{A_t}{1-\gamma} C_t^{1-\gamma}, \tag{3.27}$$

where A_t is a taste shift parameter. The Euler equation for optimal consumption is

$$\frac{\partial U(C_t)}{\partial C_t} = \beta \mathrm{E}_t \left[\frac{\partial U(C_{t+1})}{\partial C_{t+1}} Q_{t+1} \right], \tag{3.28}$$

where Q_{t+1} is the gross real return.

The marginal utility of C_t is

$$\frac{\partial U(C_t)}{\partial C_t} = A_t C_t^{-\gamma}, \tag{3.29}$$

so the optimality condition can be written

$$\begin{aligned} 1 &= \beta \mathrm{E}_t Q_{t+1} \frac{A_{t+1}}{A_t} \left(\frac{C_{t+1}}{C_t} \right)^{-\gamma} \\ &= \beta \mathrm{E}_t \exp\left(\ln Q_{t+1} + \Delta \ln A_{t+1} - \gamma \ln C_{t+1} + \gamma \ln C_t \right). \end{aligned} \tag{3.30}$$

Assume that $\ln Q_{t+1}$, $\ln A_{t+1}$, and $\ln C_{t+1}$ are jointly normally distributed. (Recall $\mathrm{E}\exp(x) = \exp(\mathrm{E}x + \mathrm{Var}(x)/2)$ is x is normally distributed.) Take logs of (3.30) and rewrite it as

$$\begin{aligned} 0 &= \ln \beta + \mathrm{E}_t \ln Q_{t+1} + \mathrm{E}_t \Delta \ln A_{t+1} - \gamma \mathrm{E}_t \ln C_{t+1} + \gamma \ln C_t + \\ &\quad \mathrm{Var}_t \left(\ln A_{t+1} - \gamma \ln C_{t+1} \right)/2, \text{ or} \\ \mathrm{E}_t \ln C_{t+1} &= \ln C_t + \frac{1}{\gamma} \mathrm{E}_t \ln Q_{t+1} + \frac{1}{\gamma} \mathrm{E}_t z_{t+1}. \end{aligned}$$

The most important part of $\mathrm{E}_t z_{t+1}$ is $\mathrm{E}_t \Delta \ln A_{t+1}$. If $\ln A_{t+1} = \rho \ln A_t + u_{t+1}$, then $\mathrm{E}_t \Delta \ln A_{t+1} = (\rho - 1) \ln A_t$, so the AR(1) formulation carries over to the expected change, but the sign is reversed (assuming $|\rho| < 1$).

The case with "*Catching up with the Jonses*" is when the utility function is

$$U\left(C_t, \bar{C}_{t-1}\right) = \frac{A_t}{1-\gamma} \left(\frac{C_t}{\bar{C}_{t-1}^{\eta}} \right)^{1-\gamma}, \eta > 0, \tag{3.31}$$

where $\bar{C}_{t-1}$ is average consumption of other consumers in the previous period. The marginal utility of C_t is

$$\frac{\partial U(C_t)}{\partial C_t} = A_t C_t^{-\gamma} \bar{C}_{t-1}^{-\eta(1-\gamma)},$$

so the optimality condition can be written

$$\begin{aligned} 1 &= \beta \mathrm{E}_t \left[Q_{t+1} \frac{A_{t+1}}{A_t} \left(\frac{C_{t+1}}{C_t} \right)^{-\gamma} \left(\frac{C_t}{C_{t-1}} \right)^{-\eta(1-\gamma)} \right] \\ &= \beta \mathrm{E}_t \exp\left[\ln Q_{t+1} + \Delta \ln A_{t+1} - \gamma \Delta \ln C_{t+1} - \eta (1-\gamma) \Delta \ln C_t \right], \end{aligned} \tag{3.32}$$

which changes (3.31) to

$$\begin{aligned} 0 &= \ln \beta + \mathrm{E}_t \ln Q_{t+1} + \mathrm{E}_t \Delta \ln A_{t+1} - \gamma \Delta \mathrm{E}_t \ln C_{t+1} - \eta (1-\gamma) \Delta \ln C_t + \\ &\quad \mathrm{Var}_t \left(\ln Q_{t+1} + \ln A_{t+1} - \gamma \ln C_{t+1} \right)/2, \text{ or} \\ \mathrm{E}_t \ln C_{t+1} &= \frac{\gamma - \eta(1-\gamma)}{\gamma} \ln C_t + \eta \frac{1-\gamma}{\gamma} \ln C_{t-1} + \frac{1}{\gamma} \mathrm{E}_t \ln Q_{t+1} + \frac{1}{\gamma} \mathrm{E}_t z_{t+1}. \end{aligned} \tag{3.33}$$

Bibliography

Abel, A., 1990, "Asset Prices under Habit Formation and Catching Up with the Jonses," *American Economic Review*, 80, 38–42, Papers and Proceedings.

Clarida, R., J. Galí, and M. Gertler, 1999, "The Science of Monetary Policy: A New Keynesian Perspective," *Journal of Economic Literature*, 37, 1661–1707.

Rotemberg, J. J., 1987, "New Keynesian Microfoundations," in Stanley Fischer (ed.), *NBER Macroeconomics Annual* . pp. 69–104, NBER.

Söderlind, P., 1999, "Solution and Estimation of RE Macromodels with Optimal Policy," *European Economic Review*, 43, 813–823.

4 Solving Linear Expectational Difference Equations

References: Blanchard and Kahn (1980), King and Watson (1995), and Klein (2000).

4.1 The Model

The model is

$$\begin{bmatrix} x_{1t+1} \\ \mathrm{E}_t x_{2t+1} \end{bmatrix} = A \begin{bmatrix} x_{1t} \\ x_{2t} \end{bmatrix} + \begin{bmatrix} \varepsilon_{t+1} \\ \mathbf{0}_{n_2 \times 1} \end{bmatrix}, \tag{4.1}$$

where x_{1t} is an $n_1 \times 1$ vector of predetermined variables with x_{10} given, x_{2t} is an n_2 vector of "forward looking" variables, and ε_t is a white noise process with covariance matrix Σ. All dynamics of the exogenous processes have been placed in x_{1t}.

Example 5 *(Cagan model.) Consider the "Cagan model" (See, for instance, Blanchard and Kahn (1980) 4) where the price level, P_t, behaves like an asset price, and the money supply, M_t, is an exogenous AR(1)*

$$\ln P_t = (1-a)\ln M_t + aE_t \ln P_{t+1}, \text{ with } 0 < a < 1, \text{ and}$$
$$\ln M_{t+1} = \rho \ln M_t + \epsilon_{t+1}.$$

This can be rewritten on the same form as (4.1)

$$\begin{bmatrix} \ln M_{t+1} \\ E_t \ln P_{t+1} \end{bmatrix} = \begin{bmatrix} \rho & 0 \\ \frac{a-1}{a} & \frac{1}{a} \end{bmatrix} \begin{bmatrix} \ln M_t \\ \ln P_t \end{bmatrix} + \begin{bmatrix} \epsilon_{t+1} \\ 0 \end{bmatrix}.$$

Take expectations of (4.1), based in information in t, of both sides

$$\mathrm{E}_t \begin{bmatrix} x_{1t+1} \\ x_{2t+1} \end{bmatrix} = A \begin{bmatrix} x_{1t} \\ x_{2t} \end{bmatrix}. \tag{4.2}$$

We will first try to find the solution to (4.2), then reintroduce the shocks ε_{1t}.

4.2 Matrix Decompositions

Remark 6 *(Complex matrices.) Let A^H denote the transpose of the complex conjugate of A, so that if*

$$A = \begin{bmatrix} 1 & 2+3i \end{bmatrix} \text{ then } A^H = \begin{bmatrix} 1 \\ 2-3i \end{bmatrix}.$$

A square matrix A is unitary *(similar to orthogonal) if $A^H = A^{-1}$, for instance,*

$$A = \begin{bmatrix} \frac{1+i}{2} & \frac{1+i}{2} \\ \frac{1-i}{2} & \frac{-1+i}{2} \end{bmatrix} \text{ gives } A^H = A^{-1} = \begin{bmatrix} \frac{1-i}{2} & \frac{1+i}{2} \\ \frac{1-i}{2} & \frac{-1-i}{2} \end{bmatrix}.$$

Remark 7 *(Schur decomposition.) The decomposition of the $n \times n$ matrix A gives the matrices T and Z such that*

$$A = ZTZ^H \tag{4.3}$$

where Z is a unitary $n \times n$ matrix and T is an $n \times n$ upper triangular Schur form with the eigenvalues along the diagonal. Note that premultiplying (4.3) with $Z^{-1} = Z^H$ and postmultiplying with Z gives

$$T = Z^H AZ, \tag{4.4}$$

which is upper triangular. The ordering of the eigenvalues in T can be reshuffled, although this requires that Z is reshuffled conformably to keep (4.3) to hold - this involves a bit of tricky "book keeping."

Remark 8 *(Upper triangular matrices.) If T is upper triangular, then TT is as well.*

Example 9 *(Cagan again.) If $a = 0.5$ and $\rho = 0.9$ in the Cagan model so*

$$A = \begin{bmatrix} 0.9 & 0 \\ -1 & 2 \end{bmatrix}, \text{ and } Z \approx \begin{bmatrix} -0.74 & 0.673 \\ -0.673 & -0.74 \end{bmatrix}, T = \begin{bmatrix} 0.9 & 1 \\ 0 & 2 \end{bmatrix}, \text{ and } Z^H \approx \begin{bmatrix} -0.74 & -0.673 \\ 0.673 & -0.74 \end{bmatrix}.$$

Note that T is upper triangular, with the eigenvalues along the diagonal, and Z is unitary. Note that $A = ZTZ^H$ holds. In this example, T and Z are real since all eigenvalues are real (unique).

4.2.1 Why not a Spectral Decomposition?

Remark 10 *(Spectral decomposition.) The n eigenvalues (λ_i) and associated eigenvectors (z_i) of the $n \times n$ matrix A satisfies*

$$(A - \lambda_i I_n) z_i = \mathbf{0}_{n\times 1}. \tag{4.5}$$

If the eigenvectors are linearly independent, then we can decompose A as

$$A = Z\Lambda Z^{-1}, \text{ where } \Lambda = diag(\lambda_1, ..., \lambda_1) \text{ and } Z = \begin{bmatrix} z_1 & z_2 & \cdots & z_n \end{bmatrix},$$

so Λ is a matrix with the eigenvalues along the diagonal and zeros elsewhere. To see why the spectral decomposition works, note that by (4.5) $AZ = Z\Lambda$, which can be premultiplied by Z^{-1}. (Note that this decomposition can be quite convenient since the fact that Λ is diagonal implies $A^2 = AA = Z\Lambda Z^{-1} Z\Lambda Z^{-1} = Z\Lambda\Lambda Z^{-1} = Z\Lambda^2 Z^{-1}$.)

Why should we not decompose A with the help of eigenvalues and eigenvectors instead? We could if the eigenvectors were linearly independent (distinct eigenvalues is a sufficient, not necessary, condition for this). In this case, the approach in Section 4.3 still applies, but where we let $T = \Lambda$.

Often the eigenvectors are linearly dependent. This would create a fundamental problem when we try to "decouple" the system of difference equations (see below). We then have to use some other decomposition. The Jordan decomposition used by Blanchard and Kahn (1980) is perhaps the neatest, but also very difficult to calculate accurately (see Golub and van Loan (1989)). The calculation of the Schur decomposition is fairly robust, and is therefore widely implemented in software libraries.

Example 11 *Consider the process $x_t - x_{t-1} = x_{t-1} - x_{t-2} + \varepsilon_t$. It can be written as a VAR(1) as*

$$\begin{bmatrix} x_t \\ x_{t-1} \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_{t-1} \\ x_{t-2} \end{bmatrix} + \begin{bmatrix} \varepsilon_t \\ 0 \end{bmatrix}.$$

The VAR matrix has a repeated eigenvalue (1) and eigenvectors ($\begin{bmatrix} 1 & 1 \end{bmatrix}$).

Example 12 *(Cagan again.) The A matrix in Example 5 has the following spectral decomposition*

$$\begin{bmatrix} \rho & 0 \\ \frac{\alpha-1}{\alpha} & \frac{1}{\alpha} \end{bmatrix} = \begin{bmatrix} \frac{\alpha\rho-1}{\alpha-1} & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \rho & 0 \\ 0 & \frac{1}{\alpha} \end{bmatrix} \begin{bmatrix} \frac{\alpha\rho-1}{\alpha-1} & 0 \\ 1 & 1 \end{bmatrix}^{-1}.$$

4.3 Solving

4.3.1 "Decoupling"

Calculate the Schur decomposition (4.3) of A and reorder (both T and Z, a bit tricky) so the n_θ eigenvalues with modulus smaller than one comes first. (Note that T and Z may include complex elements.) Partition T accordingly

$$T = \begin{bmatrix} T_{\theta\theta} & T_{\theta\delta} \\ \mathbf{0} & T_{\delta\delta} \end{bmatrix}. \tag{4.6}$$

If there are n_θ stable and n_δ unstable eigenvalues, then $T_{\theta\theta}$ is $n_\theta \times n_\theta$, $T_{\theta\delta}$ is $n_\theta \times n_\delta$, and $T_{\delta\delta}$ is $n_\delta \times n_\delta$.

Introduce the auxiliary variables

$$\begin{bmatrix} \theta_t \\ \delta_t \end{bmatrix} = Z^H \begin{bmatrix} x_{1t} \\ x_{2t} \end{bmatrix}. \tag{4.7}$$

Use the Schur decomposition (4.3), $A = ZTZ^H$, in (4.2). Then, premultiply with the non-singular matrix Z^H ("no information is lost," that is, we get an equivalent system), use (4.7) and (4.3)

$$\begin{aligned} Z^H \mathrm{E}_t x_{t+1} &= Z^H ZTZ^H \begin{bmatrix} x_{1t} \\ x_{2t} \end{bmatrix}, \\ \mathrm{E}_t \begin{bmatrix} \theta_{t+1} \\ \delta_{t+1} \end{bmatrix} &= Z^H ZT \begin{bmatrix} \theta_t \\ \delta_t \end{bmatrix} \text{ /*from (4.7)*/} \\ &= \begin{bmatrix} T_{\theta\theta} & T_{\theta\delta} \\ \mathbf{0} & T_{\delta\delta} \end{bmatrix} \begin{bmatrix} \theta_t \\ \delta_t \end{bmatrix}. \end{aligned} \tag{4.8}$$

4.3.2 Solving the System of $\mathrm{E}_t\,\theta_{t+1}$ and $\mathrm{E}_t\,\delta_{t+1}$

Since $T_{\delta\delta}$ contains the roots outside the unit circle, δ_t will diverge as t increases unless $\delta_0 = \mathbf{0}$. Any stable solution will therefore require that $\delta_t = \mathbf{0}$ for all t. The system (4.8) can therefore be written as

$$\begin{aligned} &\delta_t = \mathbf{0}, \text{ and} \\ &\mathrm{E}_t\theta_{t+1} = T_{\theta\theta}\theta_t. \end{aligned} \tag{4.9}$$

4.3.3 Initial Values of θ_0

Invert (4.7) and partition as

$$\begin{aligned} \begin{bmatrix} x_{1t} \\ x_{2t} \end{bmatrix} &= \begin{bmatrix} Z_{k\theta} & Z_{k\delta} \\ Z_{\lambda\theta} & Z_{\lambda\delta} \end{bmatrix} \begin{bmatrix} \theta_t \\ \delta_t \end{bmatrix} \\ &= \begin{bmatrix} Z_{k\theta} \\ Z_{\lambda\theta} \end{bmatrix} \theta_t, \end{aligned} \tag{4.10}$$

since $\delta_t = \mathbf{0}$.

The initial conditions are that x_{10} is given. From (4.10) we have

$$x_{10} = Z_{k\theta}\theta_0, \tag{4.11}$$

which can be solved for θ_0 if $Z_{k\theta}$ is invertible. It has n_1 rows and n_θ columns (as many as stable roots), so a necessary condition is that the number of stable roots equal the number of backward looking variables (Blanchard and Kahn, proposition 1). If that is the case, then

$$\theta_0 = Z_{k\theta}^{-1} x_{10}. \tag{4.12}$$

If the number of stable roots is less than the number of predetermined variables, n_1, then there is no stable solution. In contrast, if the number of stable roots is larger than the number of predetermined variables, , n_1, then there is an infinite number of stable solutions. See Blanchard and Kahn (1980) for details.

Example 13 *(Cagan model with too many stable roots.) Consider the Cagan model in Example 5 again, but change the price equation to*

$$\ln P_t = \ln M_t + aE_t \ln P_{t+1}, \text{ with } a > 0. \tag{4.13}$$

$$\begin{bmatrix} \ln M_{t+1} \\ E_t \ln P_{t+1} \end{bmatrix} = \begin{bmatrix} \rho & 0 \\ \frac{-1}{a} & \frac{1}{a} \end{bmatrix} \begin{bmatrix} \ln M_t \\ \ln P_t \end{bmatrix} + \begin{bmatrix} \epsilon_{t+1} \\ 0 \end{bmatrix}, \tag{4.14}$$

with the eigenvalues ρ (still assuming $|\rho| < 1$) and $1/a$. If $a > 1$, then there are two stable eigenvalues, so we have an infinite number of solutions. To illustrate this, suppose $|a\rho| < 1$. Then, iterating on the price equation gives the stable fundamental solution

$$\begin{aligned} \ln P_t^* &= \sum_{s=0}^{\infty} a^s E_t \ln M_{t+s} \\ &= \frac{1}{1 - a\rho} \ln M_t. \end{aligned}$$

However, the full set of solutions is $\ln P_t = \ln P_t^* + b_t$, where b_t is a "bubble." Try this in (4.13) to get

$$\begin{aligned} \frac{1}{1 - a\rho} \ln M_t + b_t &= \ln M_t + aE_t \left(\frac{1}{1 - a\rho} \ln M_{t+1} + b_{t+1} \right) \\ &= \ln M_t + \frac{a\rho}{1 - a\rho} \ln M_t + aE_t b_{t+1}, \end{aligned}$$

which requires $b_t = aE_t b_{t+1}$. When $|a| < 1$ this means that the bubble is unstable, and we choose $b_t = 0$ to get an economically meaningful (stable) solution. However, with $|a| > 1$, there is an infinity of stable bubbles and we have no good reason to choose one over another.

Example 14 *(Cagan model with too many stable roots, continued.) The matrix in (4.14) can be decomposed in terms of the eigenvectors and eigenvalues*

$$\begin{bmatrix} \rho & 0 \\ \frac{-1}{a} & \frac{1}{a} \end{bmatrix} = \begin{bmatrix} 1 - a\rho & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \rho & 0 \\ 0 & \frac{1}{a} \end{bmatrix} \begin{bmatrix} 1 - a\rho & 0 \\ 1 & 1 \end{bmatrix}^{-1}.$$

When $|a| > 1$, then both eigenvalues are stable so (4.10) can be written

$$\begin{bmatrix} \ln M_t \\ \ln P_t \end{bmatrix} = \begin{bmatrix} 1 - a\rho & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \theta_{1t} \\ \theta_{2t} \end{bmatrix},$$

where θ_{1t} and θ_{2t} are the elements of the vector θ_t. We can identify $\theta_{1t} = \ln M_t/(1 - a\rho)$ from the first equation. The second equation then says that $\ln P_t = \ln M_t/(1 - a\rho) + \theta_{2t}$, where by (4.9) $E_t\theta_{2t+1} = \theta_{2t}/a$, so θ_{2t} is indeed a stable variable. However, beyond that *we cannot say much about what it represents—this is the "bubble" discussed above.*

Example 15 *(Square, but singular, $Z_{k\theta}$.) It is possible that we have the right number of stable roots, but that $Z_{k\theta}$ is singular. This is a fairly odd case, but we cannot rule it out. For instance, consider a slight variation on the example by Stock and Watson (1995)*

$$\begin{bmatrix} x_{1t+1} \\ x_{2t+1} \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ \alpha & \frac{1}{2} \end{bmatrix} \begin{bmatrix} x_{1t} \\ x_{2t} \end{bmatrix}, \text{ with } x_{10} \text{ given.}$$

The model therefore has one stable root and one initial condition (it will turn out to be in the wrong place, however). The spectral decomposition is

$$\begin{bmatrix} 2 & 0 \\ \alpha & \frac{1}{2} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & \frac{2}{3}\alpha \end{bmatrix} \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & \frac{2}{3}\alpha \end{bmatrix}^{-1}.$$

(4.9) and (4.10) become

$$\delta_t = 0 \text{ and } \theta_{t+1} = \frac{1}{2}\theta_t,$$
$$\begin{bmatrix} x_{1t} \\ x_{2t} \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \theta_t.$$

The stable auxiliary variable, θ_t, is not related to the variable with an initial condition, x_{1t}, so $Z_{k\theta}$ is indeed singular. It is clear this model cannot have a stable solution since the solution for the first variable must be $x_{1t} = 2^t x_{10}$.

Example 16 *(Fixing Example 15.) Change the model to*

$$\begin{bmatrix} x_{1t+1} \\ x_{2t+1} \end{bmatrix} = \begin{bmatrix} 2 & \beta \\ 0 & \frac{1}{2} \end{bmatrix} \begin{bmatrix} x_{1t} \\ x_{2t} \end{bmatrix}, \text{ with } x_{10} \text{ given.}$$

The spectral decomposition is

$$\begin{bmatrix} 2 & \beta \\ 0 & \frac{1}{2} \end{bmatrix} = \begin{bmatrix} -\frac{2}{3}\beta & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} -\frac{2}{3}\beta & 1 \\ 1 & 0 \end{bmatrix}^{-1}.$$

In this case (4.9) and (4.10) become

$$\delta_t = 0 \text{ and } \theta_{t+1} = \frac{1}{2}\theta_t,$$
$$\begin{bmatrix} x_{1t} \\ x_{2t} \end{bmatrix} = \begin{bmatrix} -\frac{2}{3}\beta \\ 1 \end{bmatrix} \theta_t.$$

If $\beta \neq 0$, then $Z_{k\theta}$ is non-singular and we have a unique stable solution. It is (since $x_{2t} = \theta_t$)

$$x_{1t+1} = \frac{1}{2}x_{1t} \text{ with } x_{10} \text{ given, and } x_{2t} = -\frac{3}{2\beta}x_{1t}.$$

This model has a unique stable solution when $\beta \neq 0$ since x_{20} adjusts so that x_{1t} will not explode. When $\beta = 0$, then x_{1t} does not depend on x_{20}, so there is no possibility to put the system on a stable path (recall that x_{1t} has an inherent tendency of being unstable).

The evolution of the deterministic system is (4.9) with (4.12) as starting values. (4.10) shows to transform to expected values of x_{1t} and x_{2t}.

4.3.4 Putting the Innovations Back

From (4.1) we know that $x_{1t+1} - \mathrm{E}_t x_{1t+1} = \varepsilon_{t+1}$. Using (4.10) to rewrite this gives

$$Z_{k\theta}\left(\theta_{t+1} - \mathrm{E}_t\theta_{t+1}\right) = \varepsilon_{t+1}, \tag{4.15}$$

which under the same conditions as above can be inverted

$$\theta_{t+1} = \mathrm{E}_t\theta_{t+1} + Z_{k\theta}^{-1}\varepsilon_{t+1}. \tag{4.16}$$

Combined with (4.9) we have

$$\theta_{t+1} = T_{\theta\theta}\theta_t + Z_{k\theta}^{-1}\varepsilon_{t+1}, \tag{4.17}$$

which with (4.12) and (4.10) is a complete solution of the stochastic model.

4.3.5 Dynamics in Terms of x_{1t} and x_{2t}

Using $\theta_t = Z_{k\theta}^{-1}x_{1t}$ from (4.10) in (4.17) gives

$$x_{1t+1} = Z_{k\theta}T_{\theta\theta}Z_{k\theta}^{-1}x_{1t} + \varepsilon_{t+1}. \tag{4.18}$$

Similarly, combining $x_{2t} = Z_{\lambda\theta}\theta_t$ with $\theta_t = Z_{k\theta}^{-1} x_{1t}$ (both from (4.10)) gives

$$x_{2t} = Z_{\lambda\theta} Z_{k\theta}^{-1} x_{1t}. \tag{4.19}$$

Example 17 *(Solving the Cagan model.) From the Cagan model above, and (4.19) we have*

$$\begin{aligned}\ln P_t &= Z_{\lambda\theta} Z_{k\theta}^{-1} \ln M_t \\ &\approx -0.673\,(-0.74)^{-1} \ln M_t \\ &\approx 0.909 \ln M_t, \text{ (exact answer is } 10/11 \ln M_t\text{)},\end{aligned}$$

and (4.18) recovers the AR(1) of money supply

$$\begin{aligned}\ln M_{t+1} &= Z_{k\theta} T_{\theta\theta} Z_{k\theta}^{-1} \ln M_t + \epsilon_{t+1} \\ &= -0.74 * 0.9 * (0.74)^{-1} \ln M_t + \epsilon_{t+1} \\ &= 0.9 \ln M_t + \epsilon_{t+1}.\end{aligned}$$

4.4 Singular Dynamic Equations*

Instead of (4.1), let the dynamic equations be

$$\begin{bmatrix} x_{1t+1} \\ H \mathrm{E}_t x_{2t+1} \end{bmatrix} = A \begin{bmatrix} x_{1t} \\ x_{2t} \end{bmatrix} + \begin{bmatrix} \varepsilon_{t+1} \\ \mathbf{0}_{n_2 \times 1} \end{bmatrix}, \tag{4.20}$$

where H can be singular (if not, premultiply by H^{-1} to get the system on standard form). This type of model can often be solved by using the Generalized Schur decomposition. This is discussed in Chaper 7.

5 A "Simple" Policy Rule

Reference: Currie and Levine (1993), and Söderlind (1999).

5.1 Model and Solution

We now change (4.1) by adding an effect of a vector of policy instruments, u_t,

$$\begin{bmatrix} x_{1t+1} \\ \mathrm{E}_t x_{2t+1} \end{bmatrix} = A \begin{bmatrix} x_{1t} \\ x_{2t} \end{bmatrix} + B u_t + \begin{bmatrix} \varepsilon_{t+1} \\ \mathbf{0}_{n_2 \times 1} \end{bmatrix}, \tag{5.1}$$

with x_{10} given.

The policy instrument is assumed to be a linear function of x_{1t} and x_{2t} (this might force you to change the definition of x_{1t} - you can always *add* variables)

$$u_t = -F \begin{bmatrix} x_{1t} \\ x_{2t} \end{bmatrix}. \tag{5.2}$$

Substituting for u_t in (5.1) gives

$$\begin{bmatrix} x_{1t+1} \\ \mathrm{E}_t x_{2t+1} \end{bmatrix} = (A - BF) \begin{bmatrix} x_{1t} \\ x_{2t} \end{bmatrix} + \begin{bmatrix} \varepsilon_{t+1} \\ \mathbf{0}_{n_2 \times 1} \end{bmatrix}, \tag{5.3}$$

which is on the same form as (4.1), but where $A - BF$ replaces A. We can therefore apply the solution algorithm above.

5.2 Time Series Representation

Rewrite the VAR(1) in (4.18) as

$$x_{1t+1} = M x_{1t} + \varepsilon_{t+1}, \tag{5.4}$$

where ε_{t+1} has the covariance matrix Σ. We know that the impulse response function is

$$x_{1t} = \varepsilon_t + M\varepsilon_{t-1} + M^2 \varepsilon_{t-2} + ... + M^t x_{10}, \tag{5.5}$$

which immediately gives unconditional covariance matrix of x_{1t} as (since $\mathrm{E}\varepsilon_t \varepsilon_{t-s} = 0$ if $s \neq 0$)

$$\mathrm{Cov}\,(x_{1t}) = \Sigma + M\Sigma M' + M^2 \Sigma \left(M^2\right)' + ... \tag{5.6}$$

This is easily calculated by iterating until convergence on $\mathrm{Cov}(x_{1t})$ in

$$\mathrm{Cov}\,(x_{1t}) = \Sigma + M \mathrm{Cov}\,(x_{1t})\, M'. \tag{5.7}$$

The iteration could start by setting $\text{Cov}(x_{1t})$ on the right hand side to a matrix of zeros. (An exact formula exist, but it usually gives longer computation time and less accuracy.)

Example 18 *(Why (5.7) works). Consider iterating on $A_{s+1} = B + CA_sC'$ by starting from $A_0 = \mathbf{0}$. The first iteration gives $A_1 = B$, the second $A_2 = B+CA_1C' = B+CBC'$, the third $A_3 = B + CA_2C' = B + C\left(B + CBC'\right)C' = B + CBC' + CCBC'C'$, and so forth. Continue until $A_{s+1} \approx A_s$.*

Other variables of interest can be expressed as linear functions of x_{1t}, so it is straightforward to calculate the impulse response and the covariance. For instance, (4.19) can be written as

$$x_{2t} = Cx_{1t}, \tag{5.8}$$

so the policy rule (5.2) is

$$u_t = -F\begin{bmatrix} I \\ C \end{bmatrix} x_{1t}. \tag{5.9}$$

5.3 Value of Loss Function

Given the time series representation of the model, we now find the value of the loss function by guess and verify. The loss function is

$$J_0 = \text{E}_0 \sum_{t=0}^{\infty} \beta^t \begin{bmatrix} x_t \\ u_t \end{bmatrix}' \begin{bmatrix} Q & U \\ U' & R \end{bmatrix} \begin{bmatrix} x_t \\ u_t \end{bmatrix}. \tag{5.10}$$

From (5.8) and (5.9), note that x_t and u_t can be expressed in terms of x_{1t} as

$$x_t = \begin{bmatrix} I \\ C \end{bmatrix} x_{1t} \text{ and } u_t = -F\begin{bmatrix} I \\ C \end{bmatrix} x_{1t}. \tag{5.11}$$

We can therefore write

$$\begin{aligned} \begin{bmatrix} x_t \\ u_t \end{bmatrix} &= \begin{bmatrix} I \\ C \\ -F\begin{bmatrix} I \\ C \end{bmatrix} \end{bmatrix} x_{1t}. \\ &= Px_{1t}, \end{aligned} \tag{5.12}$$

so the loss function can be written

$$\begin{aligned} J_0 &= \text{E}_0 \sum_{t=0}^{\infty} \beta^t x_{1t}' P' \begin{bmatrix} Q & U \\ U' & R \end{bmatrix} P x_{1t} \\ &= \text{E}_0 \sum_{t=0}^{\infty} \beta^t x_{1t}' W x_{1t}, \text{ where } W = P' \begin{bmatrix} Q & U \\ U' & R \end{bmatrix} P. \end{aligned} \tag{5.13}$$

Remark 19 *(Target variables.) In some cases the loss function is formulated in terms of target variables,*

$$\begin{aligned} Y_t &= K \begin{bmatrix} x_{1t} \\ x_{2t} \\ u_t \end{bmatrix} \\ &= KPx_{1t}. \end{aligned} \tag{5.14}$$

The loss function is then

$$\begin{aligned} J_0 &= E_0 \sum_{t=0}^{\infty} \beta^t Y_t' D Y_t \\ &= E_0 \sum_{t=0}^{\infty} \beta^t x_{1t}' P' K' D K P x_{1t} \\ &= E_0 \sum_{t=0}^{\infty} \beta^t x_{1t}' W x_{1t}, \textit{ where } W = P'K'DKP. \end{aligned} \tag{5.15}$$

Remark 20 *(Trace of a product.) The trace of a matrix is the sum of the element on the principal diagonal. Provided the dimensions fit, we have $tr(ABC) = tr(BCA) = tr(CAB)$, and also $tr(A + B) = tr(A) + tr(B)$.*

We try to find the loss function value by guessing and verifying. Guess that the loss (value) function is

$$J_t = x_{1t}' V x_{1t} + v, \tag{5.16}$$

and note from (5.13) that it must satisfy

$$J_t = x_{1t}' W x_{1t} + \beta \text{E}_t J_{t+1}. \tag{5.17}$$

Use (5.16) for J_{t+1} and (5.4) for x_{1t+1} to get

$$
\begin{aligned}
J_t &= x_{1t}' W x_{1t} + \beta \mathrm{E}_t x_{1t+1}' V x_{1t+1} + \beta v \text{ /*(5.16) for } t+1\text{*/} \\
&= x_{1t}' W x_{1t} + \beta \mathrm{E}_t \left(x_{1t}' M' + \varepsilon_{t+1}' \right) V \left(M x_{1t} + \varepsilon_{t+1} \right) + \beta v \text{ /*(5.4)*/} \\
&= x_{1t}' W x_{1t} + \beta x_{1t}' M' V M x_{1t} + \beta \mathrm{E}_t \varepsilon_{t+1}' V \varepsilon_{t+1} + \beta v. \qquad (5.18)
\end{aligned}
$$

Of course, (5.16) and (5.18) must give the same value for every possible realization of the x_{1t} vector, which requires that

$$
v = \beta \mathrm{E}_t \varepsilon_{t+1}' V \varepsilon_{t+1} + \beta v, \text{ and} \qquad (5.19)
$$

$$
V = W + \beta M' V M. \qquad (5.20)
$$

V can be calculated by iterating on (5.20) by starting at, for instance, $V = \mathbf{0}$. v is a scalar, and can therefore be calculated as

$$
\begin{aligned}
v &= \frac{\beta}{1-\beta} \mathrm{E}_t \varepsilon_{t+1}' V \varepsilon_{t+1} \\
&= \frac{\beta}{1-\beta} \operatorname{trace}(V \Sigma). \qquad (5.21)
\end{aligned}
$$

Together, this gives us the loss (value) function (5.16) with v and V given by (5.21) and (implicitly) (5.20).

5.4 Optimal Simple Rule

Several authors have considered optimal simple rules where they optimize (5.10) with respect to the decision rules F in (5.2) and the restriction that the choice of F should give a unique equilibrium (according to Proposition 1 in the Blanchard-Kahn, see above). This rule is, in general, not the same as the globally optimal rule (under commitment) since there are restrictions on decision rule.

That this is a commitment rule is intuitively easy to understand if we recall how the optimal policy is found: *(i)* guess a policy rule (F); *(ii)* solve the model as above (with the Schur decomposition) and calculate the loss; *(iii)* try another F matrix and see if it gives a lower loss function value. Note that we required the policy rule to be the same for all periods, and for each policy rule we calculate the equilibrium by assuming that the private sectors take F for granted when they form their expectations. This means that the

policy maker acts like a Stackelberg leader.

It can be shown that the optimal simple rule is not certainty equivalent, so the decision rule will depend on the covariance matrix Σ, and the initial state vector, x_{10}. Only the combination of a quadratic loss function, linear transition equations, and no restrictions on the form of decision rule gives certainty equivalence.

5.5 Singular Dynamic Equations*

Instead of (5.1), let the dynamic equations be like

$$
\begin{bmatrix} x_{1t+1} \\ H \mathrm{E}_t x_{2t+1} \end{bmatrix} = A \begin{bmatrix} x_{1t} \\ x_{2t} \end{bmatrix} + B u_t + \begin{bmatrix} \varepsilon_{t+1} \\ \mathbf{0}_{n_2 \times 1} \end{bmatrix}, \qquad (5.22)
$$

where H can be singular (if not, premultiply by H^{-1} to get the system on standard form). This type of model can often be solved by using the Generalized Schur decomposition. This is discussed in Chaper 7.

6 Optimal Policy under Commitment

6.1 Model

Reference: Currie and Levine (1993), Backus and Driffil (1986), Svensson (1994), and Söderlind (1999).

Let

$$
x_t = \begin{bmatrix} x_{1t} \\ x_{2t} \end{bmatrix}, \qquad (6.1)
$$

where x_{1t} is an $n_1 \times 1$ vector of "backward looking" variables and x_{2t} an $n_2 \times 1$ vector of "forward looking" variables. Let $n = n_1 + n_2$, so x_t is an $n \times 1$ vector.

The problem is to minimize the loss function

$$
J_0 = \mathrm{E}_0 \sum_{t=0}^{\infty} \beta^t \begin{bmatrix} x_t' & u_t' \end{bmatrix} \begin{bmatrix} Q & U \\ U' & R \end{bmatrix} \begin{bmatrix} x_t \\ u_t \end{bmatrix}, \qquad (6.2)
$$

by choosing an optimal sequence of the $k \times 1$ vector of policy instrument, u_t. The constraints are

$$\begin{bmatrix} x_{1t+1} \\ \mathrm{E}_t x_{2t+1} \end{bmatrix} = A \begin{bmatrix} x_{1t} \\ x_{2t} \end{bmatrix} + Bu_t + \begin{bmatrix} \varepsilon_{t+1} \\ \mathbf{0}_{n_2 \times 1} \end{bmatrix}, \text{ or} \tag{6.3}$$

$$x_{t+1} = Ax_t + Bu_t + \xi_{t+1}, \text{ and } x_{10} \text{ given,} \tag{6.4}$$

where $\xi_{t+1} = (\varepsilon_{t+1}, x_{2t+1} - E_t x_{2t+1})$. The second part of ξ_{t+1} are the innovations in x_{2t+1}, which have to be functions of ε_{t+1}, but that is something we will return to later.

Form the Lagrangian

$$L_0 = \min_{\{u_t\}_{t=0}^{\infty}} \mathrm{E}_0 \sum_{t=0}^{\infty} \beta^t \left[x_t' Q x_t + 2x_t' U u_t + u_t' R u_t + 2\rho_{t+1}' \left(Ax_t + Bu_t + \xi_{t+1} - x_{t+1} \right) \right]. \tag{6.5}$$

The k first order conditions for u_t are

$$-B' \mathrm{E}_t \rho_{t+1} = U' x_t + R u_t. \tag{6.6}$$

The n first order conditions for x_t are

$$\beta A' \mathrm{E}_t \rho_{t+1} = \rho_t - \beta Q x_t - \beta U u_t. \tag{6.7}$$

We can write (6.4), (6.7), and (6.6) as

$$\begin{bmatrix} I_n & \mathbf{0}_{n\times k} & \mathbf{0}_{n\times n} \\ \mathbf{0}_{n\times n} & \mathbf{0}_{n\times k} & \beta A' \\ \mathbf{0}_{k\times n} & \mathbf{0}_{k\times k} & -B' \end{bmatrix} \begin{bmatrix} x_{t+1} \\ u_{t+1} \\ \mathrm{E}_t \rho_{t+1} \end{bmatrix} = \begin{bmatrix} A & B & \mathbf{0}_{n\times n} \\ -\beta Q & -\beta U & I_n \\ U' & R & \mathbf{0}_{k\times n} \end{bmatrix} \begin{bmatrix} x_t \\ u_t \\ \rho_t \end{bmatrix} + \begin{bmatrix} \xi_{t+1} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix}. \tag{6.8}$$

6.1.1 Initial Conditions

(See Currie and Levine (1993) p 171.)

We have n_1 initial conditions from the predetermined x_{10}, and n_2 from $p_{20} = \mathbf{0}_{n_2 \times 1}$. The control variables in u_t (a $k \times 1$ vector) should belong to the forward looking variables, since we have no initial value for them. Note that p_{2t} will typically be non-zero, except in the initial period. This can be interpreted as if the policy maker in $t = 0$ exploits the fact that private sector expectations formed in $t < 0$ (which still influence today's economy, for instance, through capital stocks and prices determined in previous periods).

In fact, there is always a temptation to exploit this, that is, to set policy in such a way that $\rho_{2t} = \mathbf{0}$, but the commitment rules out this—expect for the initial period. The "timeless perspective," advocated by Woodford (1999), is essentially to use the policy rule that comes out from solving (6.8), but only from some period $t > 0$ where ρ_{2t} is set to some non-zero value such that the policy is stationary.

6.2 Solving

Reference: Klein (2000) (see also Sims (2001)).

Remark 21 *(Generalized Schur Decomposition.) The decomposition of the $n \times n$ matrices G and D gives the matrices Q, S, T, and Z such that Q and Z are unitary and S and T upper triangular. They satisfy*

$$G = QSZ^H \text{ and } D = QTZ^H. \tag{6.9}$$

The generalized Schur decomposition solves the generalized eigenvalue problem $Dx = \lambda Gx$, where λ are the generalized eigenvalues (which will equal the diagonal elements in T divided by the corresponding diagonal element in S). Note that (6.9) implies

$$Q^H GZ = S \text{ and } Q^H DZ = T. \tag{6.10}$$

Note: the notation Q has been recycled here.

Example 22 *If $G = I$ in generalized eigenvalue problem $Dx = \lambda Gx$, then we are back to the standard eigenvalue problem. Clearly, we can pick $S = I$ and $Q = Z$ in this case, so (6.9) becomes $G = I$ and $D = ZTZ^H$, as in the standard Schur decomposition.*

6.2.1 Writing the System on a Convenient Form

We now write the system on a form which allows us to apply standard formulas (see Klein (1997) and Sims (1996)) to solve the system.

Partition (6.8) as

$$\begin{bmatrix} \tilde{G}_{11} & \tilde{G}_{12} & \tilde{G}_{13} & \tilde{G}_{14} & \tilde{G}_{15} \end{bmatrix} \begin{bmatrix} x_{1t+1} \\ x_{2t+1} \\ u_{t+1} \\ \mathrm{E}_t \rho_{1t+1} \\ \mathrm{E}_t \rho_{2t+1} \end{bmatrix} = \begin{bmatrix} \tilde{D}_{11} & \tilde{D}_{12} & \tilde{D}_{13} & \tilde{D}_{14} & \tilde{D}_{15} \end{bmatrix} \begin{bmatrix} x_{1t} \\ x_{2t} \\ u_t \\ \rho_{1t} \\ \rho_{2t} \end{bmatrix} + \begin{bmatrix} \varepsilon_{t+1} \\ \epsilon_{t+1} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \tag{6.11}$$

and reshuffle so x_{1t} and ρ_{2t} come first (since we have initial values for these)

$$\begin{bmatrix} \tilde{G}_{11} & \tilde{G}_{15} & \tilde{G}_{12} & \tilde{G}_{13} & \tilde{G}_{14} \end{bmatrix} \begin{bmatrix} x_{1t+1} \\ \mathrm{E}_t \rho_{2t+1} \\ x_{2t+1} \\ u_{t+1} \\ \mathrm{E}_t \rho_{1t+1} \end{bmatrix} = \begin{bmatrix} \tilde{D}_{11} & \tilde{D}_{15} & \tilde{D}_{12} & \tilde{D}_{13} & \tilde{D}_{14} \end{bmatrix} \begin{bmatrix} x_{1t} \\ \rho_{2t} \\ x_{2t} \\ u_t \\ \rho_{1t} \end{bmatrix} + \begin{bmatrix} \varepsilon_{t+1} \\ \mathbf{0} \\ \epsilon_{t+1} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix}. \tag{6.12}$$

Let

$$k_t = \begin{bmatrix} x_{1t} \\ \rho_{2t} \end{bmatrix} \text{ and } \lambda_t = \begin{bmatrix} x_{2t} \\ u_t \\ \rho_{1t} \end{bmatrix}. \tag{6.13}$$

(Warning: k is the number of elements in the control vector, u_t, but k_t is the vector defined in (6.13)).

Take expectations, based on the information in t, of (6.33), and use the notation in (6.34) to write the result as

$$G\mathrm{E}_t \begin{bmatrix} k_{t+1} \\ \lambda_{t+1} \end{bmatrix} = D \begin{bmatrix} k_t \\ \lambda_t \end{bmatrix}, \tag{6.14}$$

where G and D are the matrices in (6.33).

6.2.2 "Decoupling" the System

Calculate the generalized Schur decomposition and reorder so the block corresponding to the stable generalized eigenvalues, $|T_{ii}/S_{ii}| < 1$, come first (this involves a bit of tricky "book keeping," but there is software available for it). Define the auxiliary variables

$$\begin{bmatrix} \theta_t \\ \delta_t \end{bmatrix} = Z^H \begin{bmatrix} k_t \\ \lambda_t \end{bmatrix}, \tag{6.15}$$

where we will associate the stable roots with θ, and the unstable with δ.

Use the generalized Schur decomposition (6.9), $G = QSZ^H$ and $D = QTZ^H$, in (6.14). Premultiply with the non-singular matrix Q^H ("no information is lost," that is, we get an equivalent system) from the generalized Schur decomposition to get

$$\begin{aligned} Q^H QSZ^H \mathrm{E}_t \begin{bmatrix} k_{t+1} \\ \lambda_{t+1} \end{bmatrix} &= Q^H QTZ^H \begin{bmatrix} k_t \\ \lambda_t \end{bmatrix} \\ SZ^H \mathrm{E}_t \begin{bmatrix} k_{t+1} \\ \lambda_{t+1} \end{bmatrix} &= TZ^H \begin{bmatrix} k_t \\ \lambda_t \end{bmatrix} \\ S\mathrm{E}_t \begin{bmatrix} \theta_{t+1} \\ \delta_{t+1} \end{bmatrix} &= T \begin{bmatrix} \theta_t \\ \delta_t \end{bmatrix}. \text{ /*from (6.15)*/} \end{aligned} \tag{6.16}$$

6.2.3 Solving the System of $\mathrm{E}_t\, \theta_{t+1}$ and $\mathrm{E}_t\, \delta_{t+1}$

Since both S and T are upper triangular (6.16) can be written

$$\begin{bmatrix} S_{\theta\theta} & S_{\theta\delta} \\ \mathbf{0} & S_{\delta\delta} \end{bmatrix} \mathrm{E}_t \begin{bmatrix} \theta_{t+1} \\ \delta_{t+1} \end{bmatrix} = \begin{bmatrix} T_{\theta\theta} & T_{\theta\delta} \\ \mathbf{0} & T_{\delta\delta} \end{bmatrix} \begin{bmatrix} \theta_t \\ \delta_t \end{bmatrix}. \tag{6.17}$$

The lower right block contains the unstable roots, so a stable solution requires that $\delta_t = \mathbf{0}$ for all t. The remaining equations are then $S_{\theta\theta}\mathrm{E}_t\theta_{t+1} = T_{\theta\theta}\theta_t$, which we solve

$$\mathrm{E}_t\theta_{t+1} = S_{\theta\theta}^{-1} T_{\theta\theta}\theta_t, \tag{6.18}$$

since $S_{\theta\theta}$ is invertible. The reason is that $\det(S_{\theta\theta})$ equals the product of the diagonal elements since $S_{\theta\theta}$ is triangular, and that all diagonal elements are non-zero since $|T_{ii}/S_{ii}| < 1$ are sorted first so there cannot be any zeros in the diagonal of $S_{\theta\theta}$; $\det(S_{\theta\theta})$ is therefore non-zero and $S_{\theta\theta}$ is invertible.

6.2.4 Initial Values of θ_t

Invert (6.15) (recall that $Z^{-1} = Z^H$), partition,

$$\begin{bmatrix} k_t \\ \lambda_t \end{bmatrix} = \begin{bmatrix} Z_{k\theta} & Z_{k\delta} \\ Z_{\lambda\theta} & Z_{\lambda\delta} \end{bmatrix} \begin{bmatrix} \theta_t \\ \delta_t \end{bmatrix} = \begin{bmatrix} Z_{k\theta} \\ Z_{\lambda\theta} \end{bmatrix} \theta_t, \tag{6.19}$$

since $\delta_t = \mathbf{0}$. The boundary conditions give us a given value of x_{10} and $\rho_{20} = \mathbf{0}$, which together with the relevant blocks in (6.19) give

$$k_0 = \begin{bmatrix} x_{10} \\ \mathbf{0}_{n_2\times 1} \end{bmatrix} = Z_{k\theta}\theta_0. \tag{6.20}$$

We can solve for θ_0 if $Z_{k\theta}$ is invertible. It has n rows (n_1 in x_{1t} and n_2 in ρ_{2t}) and as many columns as there are stable roots. A necessary condition is therefore that (6.14) has the "saddle path" property: the number of stable roots equals number of backward looking variables (see Blanchard and Kahn (1980), Proposition 1). Suppose $Z_{k\theta}^{-1}$ exists, then

$$\theta_0 = Z_{k\theta}^{-1} k_0. \tag{6.21}$$

6.2.5 Putting the Innovations Back

From (6.3) we have $x_{1t+1} - \mathrm{E}_t x_{1t+1} = \varepsilon_{t+1}$, and Backus and Driffil (1986) show that $\rho_{2t+1} - \mathrm{E}_t \rho_{2t+1} = \mathbf{0}_{n_2\times 1}$. Use (6.13) and (6.19) to write these expressions in terms of θ_{t+1}

$$k_{t+1} - \mathrm{E}_t\, k_{t+1} = Z_{k\theta}\left(\theta_{t+1} - \mathrm{E}_t\theta_{t+1}\right) = \begin{bmatrix} x_{1t+1} - \mathrm{E}_t\, x_{1t+1} \\ \rho_{2t+1} - \mathrm{E}_t\, \rho_{2t+1} \end{bmatrix} = \begin{bmatrix} \varepsilon_{t+1} \\ \mathbf{0}_{n_2\times 1} \end{bmatrix}, \tag{6.22}$$

which, under the same conditions are above, can be inverted

$$\theta_{t+1} = \mathrm{E}_t\theta_{t+1} + Z_{k\theta}^{-1} \begin{bmatrix} \varepsilon_{t+1} \\ \mathbf{0}_{n_2\times 1} \end{bmatrix}. \tag{6.23}$$

Substitute for $\mathrm{E}_t\theta_{t+1}$ from (6.18) to get

$$\theta_{t+1} = S_{\theta\theta}^{-1} T_{\theta\theta}\theta_t + Z_{k\theta}^{-1} \begin{bmatrix} \varepsilon_{t+1} \\ \mathbf{0}_{n_2\times 1} \end{bmatrix}, \tag{6.24}$$

which together with (6.21) and (6.19) is a complete description of how the model variables evolve.

6.2.6 Dynamics in Terms of x_{1t}, ρ_{2t}, x_{2t}, and ρ_{1t}

The relation between x_{1t}, ρ_{2t}, and θ_t in (6.19) can be inverted

$$\theta_t = Z_{k\theta}^{-1} k_t = Z_{k\theta}^{-1} \begin{bmatrix} x_{1t} \\ \rho_{2t} \end{bmatrix} \tag{6.25}$$

(see above) and used in (6.24) to get

$$\begin{bmatrix} x_{1t+1} \\ \rho_{2t+1} \end{bmatrix} = Z_{k\theta} S_{\theta\theta}^{-1} T_{\theta\theta} Z_{k\theta}^{-1} \begin{bmatrix} x_{1t} \\ \rho_{2t} \end{bmatrix} + \begin{bmatrix} \varepsilon_{t+1} \\ \mathbf{0}_{n_2\times 1} \end{bmatrix} \tag{6.26}$$

Similarly, combine (6.25) with the relation between x_{2t}, u_t, ρ_{1t}, and θ_t in (6.19)

$$\begin{bmatrix} x_{2t} \\ u_t \\ \rho_{1t} \end{bmatrix} = Z_{\lambda\theta} Z_{k\theta}^{-1} \begin{bmatrix} x_{1t} \\ \rho_{2t} \end{bmatrix}. \tag{6.27}$$

Equations (6.26) and (6.27), together with the initial values of x_{10} and $\rho_{20} = \mathbf{0}_{n_2\times 1}$ give a complete description of the evolution of the system.

6.3 Alternative Expression when R is Invertible*

This section summarizes an alternative way of setting up the first order conditions, which can be used when R is invertible. It is sometimes found in the literature.

Invert (6.6)

$$u_t = -R^{-1}\left(U' x_t + B' \mathrm{E}_t \rho_{t+1}\right) \text{ if } R^{-1} \text{ exists.} \tag{6.28}$$

Use (6.6) in (6.7)

$$\beta A'\mathrm{E}_t\rho_{t+1} = \rho_t - \beta Q x_t + \beta U R^{-1}\left(U' x_t + B'\mathrm{E}_t\rho_{t+1}\right), \text{ or}$$

$$\left(\beta A' - \beta U R^{-1} B'\right)\mathrm{E}_t\rho_{t+1} = \rho_t + \left(\beta U R^{-1} U' - \beta Q\right) x_t. \tag{6.29}$$

Similarly, use (6.6) in (6.4)

$$x_{t+1} = A x_t - B R^{-1}\left(U' x_t + B'\mathrm{E}_t\rho_{t+1}\right) + \xi_{t+1}, \text{ or}$$

$$x_{t+1} + \left(B R^{-1} B'\right)\mathrm{E}_t\rho_{t+1} = \left(A - B R^{-1} U'\right) x_t + \xi_{t+1} \tag{6.30}$$

Therefore, if R^{-1} exists, then we can write the first order conditions as

$$\begin{bmatrix} I_n & B R^{-1} B' \\ \mathbf{0}_{n\times n} & \beta A' - \beta U R^{-1} B' \end{bmatrix}\begin{bmatrix} x_{t+1} \\ \mathrm{E}_t\rho_{t+1} \end{bmatrix} = \begin{bmatrix} A - B R^{-1} U' & \mathbf{0}_{n\times n} \\ \beta U R^{-1} U' - \beta Q & I_n \end{bmatrix}\begin{bmatrix} x_t \\ \rho_t \end{bmatrix} + \begin{bmatrix} \xi_{t+1} \\ \mathbf{0} \end{bmatrix}. \tag{6.31}$$

Partition (6.31) as

$$\begin{bmatrix} \tilde{G}_{11} & \tilde{G}_{12} & \tilde{G}_{13} & \tilde{G}_{14} \end{bmatrix}\begin{bmatrix} x_{1t+1} \\ x_{2t+1} \\ \mathrm{E}_t\rho_{1t+1} \\ \mathrm{E}_t\rho_{2t+1} \end{bmatrix} = \begin{bmatrix} \tilde{D}_{11} & \tilde{D}_{12} & \tilde{D}_{13} & \tilde{D}_{14} \end{bmatrix}\begin{bmatrix} x_{1t} \\ x_{2t} \\ \rho_{1t} \\ \rho_{2t} \end{bmatrix} + \begin{bmatrix} \varepsilon_{t+1} \\ \epsilon_{t+1} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \tag{6.32}$$

and reshuffle so x_{1t} and ρ_{2t} come first (since we have initial conditions on these)

$$\begin{bmatrix} \tilde{G}_{11} & \tilde{G}_{14} & \tilde{G}_{12} & \tilde{G}_{13} \end{bmatrix}\begin{bmatrix} x_{1t+1} \\ \mathrm{E}_t\rho_{2t+1} \\ x_{2t+1} \\ \mathrm{E}_t\rho_{1t+1} \end{bmatrix} = \begin{bmatrix} \tilde{D}_{11} & \tilde{D}_{14} & \tilde{D}_{12} & \tilde{D}_{13} \end{bmatrix}\begin{bmatrix} x_{1t} \\ \rho_{2t} \\ x_{2t} \\ \rho_{1t} \end{bmatrix} + \begin{bmatrix} \varepsilon_{t+1} \\ \mathbf{0} \\ \epsilon_{t+1} \\ \mathbf{0} \end{bmatrix}. \tag{6.33}$$

Let k_t represent the predetermined variables (x_{1t} and ρ_{2t}), and λ_t the forward looking variables (x_{2t} and ρ_{1t})

$$k_t = \begin{bmatrix} x_{1t} \\ \rho_{2t} \end{bmatrix} \text{ and } \lambda_t = \begin{bmatrix} x_{2t} \\ \rho_{1t} \end{bmatrix}. \tag{6.34}$$

The rest of the analysis is then as in the general case.

6.4 Singular Dynamic Equations*

Instead of (6.3) and (6.4), let the dynamic equations be

$$\begin{bmatrix} x_{1t+1} \\ H\mathrm{E}_t x_{2t+1} \end{bmatrix} = A\begin{bmatrix} x_{1t} \\ x_{2t} \end{bmatrix} + B u_t + \begin{bmatrix} \varepsilon_{t+1} \\ \mathbf{0}_{n_2\times 1} \end{bmatrix}, \text{ or} \tag{6.35}$$

$$\tilde{H} x_{t+1} = A x_t + B u_t + \xi_{t+1}, \text{ and } x_{10} \text{ given, where } \tilde{H} = \begin{bmatrix} I_{n_1} & \mathbf{0}_{n_1\times n_2} \\ \mathbf{0}_{n_2\times n_1} & H \end{bmatrix}, \tag{6.36}$$

where H can be singular (if not, premultiply by H^{-1} to get the system on standard form).

The first order conditions for x_t, corresponding to (6.7), are then

$$\beta A'\mathrm{E}_t\rho_{t+1} = \tilde{H}'\rho_t - \beta Q x_t - \beta U u_t, \tag{6.37}$$

so we can write (6.35), (6.37), and (6.6) as , corresponding to (6.8), as

$$\begin{bmatrix} \tilde{H} & \mathbf{0}_{n\times k} & \mathbf{0}_{n\times n} \\ \mathbf{0}_{n\times n} & \mathbf{0}_{n\times k} & \beta A' \\ \mathbf{0}_{k\times n} & \mathbf{0}_{k\times k} & -B' \end{bmatrix}\begin{bmatrix} x_{t+1} \\ u_{t+1} \\ \mathrm{E}_t\rho_{t+1} \end{bmatrix} = \begin{bmatrix} A & B & \mathbf{0}_{n\times n} \\ -\beta Q & -\beta U & \tilde{H}' \\ U' & R & \mathbf{0}_{k\times n} \end{bmatrix}\begin{bmatrix} x_t \\ u_t \\ \rho_t \end{bmatrix} + \begin{bmatrix} \xi_{t+1} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix}. \tag{6.38}$$

We can then apply the same solution algorithm as for to this system.

7 Simple Rules with Singular Dynamic Equations*

We now return to the issue of how to solve a model like (5.22), where the forward looking equations are singular. In this case,

$$\tilde{H}\begin{bmatrix} x_{1t+1} \\ \mathrm{E}_t x_{2t+1} \end{bmatrix} = A\begin{bmatrix} x_{1t} \\ x_{2t} \end{bmatrix} + B u_t + \begin{bmatrix} \varepsilon_{t+1} \\ \mathbf{0}_{n_2\times 1} \end{bmatrix}, \text{ where } \tilde{H} = \begin{bmatrix} I_{n_1} & \mathbf{0}_{n_1\times n_2} \\ \mathbf{0}_{n_2\times n_1} & H \end{bmatrix}. \tag{7.1}$$

where H can be singular (if not, premultiply by H^{-1} to get the system on standard form). First, use the policy rule (5.2) to substitute for u_t to get

$$\tilde{H}\begin{bmatrix} x_{1t+1} \\ \mathrm{E}_t x_{2t+1} \end{bmatrix} = (A - BF)\begin{bmatrix} x_{1t} \\ x_{2t} \end{bmatrix} + \begin{bmatrix} \varepsilon_{t+1} \\ \mathbf{0}_{n_2\times 1} \end{bmatrix}. \tag{7.2}$$

Clearly, this has the same form as (4.20), but with $B = \mathbf{0}$ (there is no control variable in that model), so the approach here can be used also for that model.

Note that (7.2) has the same structure as (6.14) with $k_t = x_{1t}$ and $\lambda_t = x_{2t}$, so we can apply the same solution algorithm as in Section 6.2—with appropriate changes of dimensions and notation. In practice, (6.14)-(6.19) still hold, but in (6.20)-(6.27) we should just leave out ρ_{1t}, ρ_{2t}, and u_t (think of them as zero dimension (empty) vectors). In particular, this means that a requirement for an invertible $Z_{k\theta}$ is that there are n_1 (number of elements in $k_t = x_{1t}$) stable roots.

8 Discretionary Solution

8.1 Summary

References: Currie and Levine (1993), Backus and Driffil (1986), Oudiz and Sachs (1985), Svensson (1994), and Söderlind (1999).

The main features are the following.

1. The policy maker reoptimizes every period, but we can find a stationary policy rule if we let the time horizon go to infinity.

2. The state of the economy is given by the predetermined variables, x_{1t}. As a consequence, the decision rule and the non-predetermined variables, x_{2t}, must be linear functions of x_{1t} ($u_t = -Fx_{1t}$, and $x_{2t} = Cx_{1t}$, respectively in the stationary equilibrium).

3. The policy maker takes the expectations of private agents as given ("Nash equilibrium" - not like in commitment case where the policy maker is a "Stackelberg leader"). From above it is $\mathrm{E}_t x_{2t+1} = C\mathrm{E}_t x_{1t+1}$.

4. No closed form solution exists—not even a proof (except in the scalar case) of convergence of the solution algorithm.

8.2 The Model

Let the

$$x_t = \begin{bmatrix} x_{1t} \\ x_{2t} \end{bmatrix}, \tag{8.1}$$

where x_{1t} is an $n_1 \times 1$ vector of "backward looking" variables and x_{2t} an $n_2 \times 1$ vector of "forward looking" variables. Let $n = n_1 + n_2$, so x_t is an $n \times 1$ vector.

The problem is to minimize the loss function

$$J_t = \mathrm{E}_t \sum_{s=0}^{\infty} \beta^s \begin{bmatrix} x'_{t+s} & u'_{t+s} \end{bmatrix} \begin{bmatrix} Q & U \\ U' & R \end{bmatrix} \begin{bmatrix} x_{t+s} \\ u_{t+s} \end{bmatrix}. \tag{8.2}$$

The constraints are

$$\begin{bmatrix} x_{1t+1} \\ \mathrm{E}_t x_{2t+1} \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} x_{1t} \\ x_{2t} \end{bmatrix} + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} u_t + \begin{bmatrix} \varepsilon_{t+1} \\ \mathbf{0}_{n_2 \times 1} \end{bmatrix}, \tag{8.3}$$

where u_t is $k \times 1$, and where x_{10} is given.

8.3 Optimization in Period t

The policy maker optimizes in every period, taking into account that he will be able to reoptimize next period. The state of the economy is summarized by x_{1t} and the period return is a quadratic form, so we know that we can write the value of loss function as

$$J_t = r_t + \beta \mathrm{E}_t \left\{ x'_{1t+1} V_{t+1} x_{1t+1} + v_{t+1} \right\}, \text{ where} \tag{8.4}$$

$$r_t = \begin{bmatrix} x_{1t} \\ x_{2t} \end{bmatrix}' \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix} \begin{bmatrix} x_{1t} \\ x_{2t} \end{bmatrix} + 2 \begin{bmatrix} x_{1t} \\ x_{2t} \end{bmatrix}' \begin{bmatrix} U_1 \\ U_2 \end{bmatrix} u_t + u'_t R u_t. \tag{8.5}$$

We have not yet specified what the $n_1 \times n_1$ matrix V_{t+1} and the scalar v_{t+1} are. They are assumed to be known, and that V_{t+1} is symmetric (no loss of generality). We also assume, once again without loss of generality, that R is symmetric. The tricky aspect of this optimization problem is that the objective function depends on the forward looking variables, x_{2t}, which are determined endogenously - and depend on expected future values of x_{1t} and x_{2t}.

The approach to solve this problem discussed below is to express x_{2t} in (8.5) in terms of x_{1t} and u_t. Since x_{1t} is predetermined, optimizing (8.4) is then a standard problem. The value of x_{2t} depends on $\mathrm{E}_t x_{2t+1}$, however, so the first step is to use the guess on how expectations are formed, that is, that $\mathrm{E}_t x_{2t+1} = C_{t+1}\mathrm{E}_t x_{1t+1}$.

8.3.1 Rewriting the System by Using $\mathbf{E}_t x_{2t+1} = C_{t+1}\mathbf{E}_t x_{1t+1}$

Expectations of the non-predetermined variables are

$$\mathrm{E}_t x_{2t+1} = C_{t+1}\mathrm{E}_t x_{1t+1}, \tag{8.6}$$

where C_{t+1} is some $n_2 \times n_1$ matrix which we know (unspecified so far).

Use in (8.6) in the constraints (8.3), and take expectations

$$\begin{bmatrix} I \\ C_{t+1} \end{bmatrix} \mathrm{E}_t x_{1t+1} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} x_{1t} \\ x_{2t} \end{bmatrix} + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} u_t. \tag{8.7}$$

Note that x_{1t} is given, and suppose we consider what happens if a certain u_t is chosen. Then (8.7) specifies n equations for n unknown (n_1 in $\mathrm{E}_t x_{1t+1}$ and n_2 in x_{2t}). We therefore rewrite the system (8.7) as

$$\begin{aligned} \begin{bmatrix} I & -A_{12} \\ C_{t+1} & -A_{22} \end{bmatrix} \begin{bmatrix} \mathrm{E}_t x_{1t+1} \\ x_{2t} \end{bmatrix} &= \begin{bmatrix} A_{11} \\ A_{21} \end{bmatrix} x_{1t} + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} u_t, \text{ or} \\ \begin{bmatrix} \mathrm{E}_t x_{1t+1} \\ x_{2t} \end{bmatrix} &= \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} \left(\begin{bmatrix} A_{11} \\ A_{21} \end{bmatrix} x_{1t} + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} u_t \right), \end{aligned} \tag{8.8}$$

where

$$\begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} = \begin{bmatrix} I & -A_{12} \\ C_{t+1} & -A_{22} \end{bmatrix}^{-1}.$$

From the rules of inverses of partitioned matrices we can note that

$$\begin{aligned} P_{21} &= (A_{22} - C_{t+1}A_{12})^{-1} C_{t+1}, \text{ and} \\ P_{22} &= -(A_{22} - C_{t+1}A_{12})^{-1}. \end{aligned} \tag{8.9}$$

The last n_2 equations in (8.8) can therefore be rewritten as

$$\begin{aligned} x_{2t} &= (P_{21}A_{11} + P_{22}A_{21}) x_{1t} + (P_{21}B_1 + P_{22}B_2) u_t \\ &= \left[(A_{22} - C_{t+1}A_{12})^{-1} C_{t+1}A_{11} - (A_{22} - C_{t+1}A_{12})^{-1} A_{21} \right] x_{1t} \\ &+ \left[(A_{22} - C_{t+1}A_{12})^{-1} C_{t+1}B_1 - (A_{22} - C_{t+1}A_{12})^{-1} B_2 \right] u_t \\ &= \underbrace{(A_{22} - C_{t+1}A_{12})^{-1} (C_{t+1}A_{11} - A_{21})}_{D_t} x_{1t} + \underbrace{(A_{22} - C_{t+1}A_{12})^{-1} (C_{t+1}B_1 - B_2)}_{G_t} u_t \\ &= D_t x_{1t} + G_t u_t, \end{aligned} \tag{8.10}$$

where D_t is $n_2 \times n_1$ and G_t is $n_2 \times k$.

Evolution of x_{1t}

Use (8.10) in the first n_1 equations in the constraints (8.3) to get the "reduced form" evolution of the predetermined variables

$$\begin{aligned} x_{1t+1} &= (A_{11} + A_{12}D_t) x_{1t} + (B_1 + A_{12}G_t) u_t + \varepsilon_{t+1} \\ &= A_t^* x_{1t} + B_t^* u_t + \varepsilon_{t+1}, \end{aligned} \tag{8.11}$$

where A_t^* is $n_1 \times n_1$ and B_t^* is $n_1 \times k$.

Value of r_t

Use (8.10) in the period t loss function (8.5)

$$r_t = \begin{bmatrix} x_{1t} \\ D_t x_{1t} + G_t u_t \end{bmatrix}' \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix} \begin{bmatrix} x_{1t} \\ D_t x_{1t} + G_t u_t \end{bmatrix} + 2 \begin{bmatrix} x_{1t} \\ D_t x_{1t} + G_t u_t \end{bmatrix}' \begin{bmatrix} U_1 \\ U_2 \end{bmatrix} u_t + u_t' R u_t. \tag{8.12}$$

The second term can be written (after straightforward multiplication and rearrangement)

$$\begin{aligned} &2x_{1t}' \left(U_1 + D_t' U_2 \right) u_t + \\ &u_t' \left(G_t' U_2 + U_2' G_t \right) u_t. \end{aligned} \tag{8.13}$$

(We write the middle matrix in the second line $G_t'U_2 + U_2'G_t$ instead of $2G_t'U_2$ since it will guarantee that R^* below is symmetric, which is convenient.)

Similarly, the first term can be written

$$\begin{aligned}
&x_{1t}' \left(Q_{11} + Q_{12}D_t + D_t'Q_{21} + D_t'Q_{22}D_t\right) x_{1t} + \\
&x_{1t}' \left(Q_{12}G_t + D_t'Q_{22}G_t\right) u_t + \\
&u_t' \left(G_t'Q_{21} + G_t'Q_{22}D_t\right) x_{1t} + \\
&u_t'G_t'Q_{22}G_t u_t.
\end{aligned} \tag{8.14}$$

Use (8.13) and (8.14) in (8.12), and collect terms

$$\begin{aligned}
r_t = {} & x_{1t}' \underbrace{\left(Q_{11} + Q_{12}D_t + D_t'Q_{21} + D_t'Q_{22}D_t\right)}_{Q_t^*} x_{1t} + \\
& x_{1t}' \underbrace{\left(Q_{12}G_t + D_t'Q_{22}G_t + U_1 + D_t'U_2\right)}_{U_t^*} u_t + \\
& u_t' \underbrace{\left(G_t'Q_{21} + G_t'Q_{22}D_t + U_1' + U_2'D_t\right)}_{U_t^{*\prime}} x_{1t} + \\
& u_t' \underbrace{\left(R + G_t'Q_{22}G_t + G_t'U_2 + U_2'G_t\right)}_{R_t^*} u_t,
\end{aligned}$$

which we write as

$$r_t = x_{1t}'Q_t^*x_{1t} + 2x_{1t}'U_t^*u_t + u_t'R_t^*u_t, \tag{8.15}$$

where Q_t^* is $n_1 \times n_1$, U_t^* is $n_1 \times k$, and R_t^* is $k \times k$.

8.3.2 Reformulated Optimization Problem

Use the new expression for r_t, (8.15), in the loss function (8.4) to rewrite it as

$$J_t = x_{1t}'Q_t^*x_{1t} + 2x_{1t}'U_t^*u_t + u_t'R_t^*u_t + \beta \mathrm{E}_t \left\{x_{1t+1}'V_{t+1}x_{1t+1} + v_{t+1}\right\}. \tag{8.16}$$

Substitute for x_{1t+1} by using the "reduced form" evolution, (8.11), to get

$$\begin{aligned}
J_t = {} & x_{1t}'Q_t^*x_{1t} + 2x_{1t}'U_t^*u_t + u_t'R_t^*u_t + \\
& \beta \mathrm{E}_t \left[\left(A_t^*x_{1t} + B_t^*u_t + \varepsilon_{t+1}\right)' V_{t+1} \left(A_t^*x_{1t} + B_t^*u_t + \varepsilon_{t+1}\right) + v_{t+1}\right],
\end{aligned} \tag{8.17}$$

which should be minimized with respect to the $k \times 1$ vector u_t. This is (since x_{1t} is not a choice variable and $\mathrm{E}_t\varepsilon_{t+1} = \mathbf{0}_{n_1\times 1}$) equivalent to minimizing

$$\begin{aligned}
\tilde{J}_t &= 2x_{1t}'U^*u_t + u_t'R^*u_t + 2\left(\beta x_{1t}'A_t^{*\prime}V_{t+1}B_t^*u_t\right) + \beta u_t'B_t^{*\prime}V_{t+1}B_t^*u_t \\
&= u_t'\left(R_t^* + \beta B_t^{*\prime}V_{t+1}B_t^*\right)u_t + 2x_{1t}'\left(U_t^* + \beta A_t^{*\prime}V_{t+1}B_t^*\right)u_t.
\end{aligned} \tag{8.18}$$

The first order conditions are (if R_t^* and V_{t+1} are symmetric)

$$2\left(R_t^* + \beta B_t^{*\prime}V_{t+1}B_t^*\right)u_t + 2\left(U_t^{*\prime} + \beta B_t^{*\prime}V_{t+1}A_t^*\right)x_{1t} = 0, \tag{8.19}$$

which we solve as

$$\begin{aligned}
u_t &= -\left(R_t^* + \beta B_t^{*\prime}V_{t+1}B_t^*\right)^{-1}\left(U_t^{*\prime} + \beta B_t^{*\prime}V_{t+1}A_t^*\right)x_{1t} \\
&= -F_t x_{1t},
\end{aligned} \tag{8.20}$$

where F_t is $k \times n_1$. This is the decision rule in t.

8.3.3 Finding the Implied C_t and V_t

By using the decision rule (8.20) we can substitute for u_t in (8.10) in order to relate x_{2t} to x_{1t} only

$$\begin{aligned}
x_{2t} &= \left(D_t - G_tF_t\right)x_{1t}, \\
&= C_t x_{1t},
\end{aligned} \tag{8.21}$$

where C_t is $n_2 \times n_1$.

We now rewrite the value function in terms of x_{1t} only. Use decision rule (8.20) in (8.17)

$$\begin{aligned}
J_t^{opt} = {} & x_{1t}'Q_t^*x_{1t} - x_{1t}'\left(U_t^*F_t + F_t'U_t^{*\prime}\right)x_{1t} + x_{1t}'F_t'R_t^*F_tx_{1t} \\
& + \beta \mathrm{E}_t\left\{\left[\left(A_t^* - B_t^*F_t\right)x_{1t} + \varepsilon_{t+1}\right]' V_{t+1}\left[\left(A_t^* - B_t^*F_t\right)x_{1t} + \varepsilon_{t+1}\right] + v_{t+1}\right\}, \text{ or}
\end{aligned}$$

$$\begin{aligned}
J_t^{opt} = {} & x_{1t}'\left[Q_t^* - U_t^*F_t - F_t'U_t^{*\prime} + F_t'R_t^*F_t + \beta\left(A_t^* - B_t^*F_t\right)'V_{t+1}\left(A_t^* - B_t^*F_t\right)\right]x_{1t} \\
& + \mathrm{E}_t\varepsilon_{t+1}'\beta V_{t+1}\varepsilon_{t+1} + \beta \mathrm{E}_t v_{t+1}.
\end{aligned} \tag{8.22}$$

Recall the following useful fact.

Remark 23 *(Cyclical permutation of trace.) Trace(ABC) =Trace(BCA) =Trace(CAB), if the dimensions allow the products.*

Since the second term in (8.22) is

$$\begin{aligned} \mathrm{E}_t \varepsilon'_{t+1} \beta V_{t+1} \varepsilon_{t+1} &= \beta \mathrm{trace}(\mathrm{E}_t V_{t+1} \varepsilon_{t+1} \varepsilon'_{t+1}) \\ &= \beta \mathrm{trace}(V_{t+1} \Sigma), \end{aligned}$$

we can rewrite (8.22) as

$$J_t^{opt} = x'_{1t} V_t x_{1t} + v_t, \text{ where } v_t = \beta \mathrm{trace}(V_{t+1}\Sigma) + \beta \mathrm{E}_t v_{t+1}. \quad (8.23)$$

Note that this equation is of the same form as the value function in $t+1$ in (8.4), which suggests a recursive algorithm.

8.4 A Recursive Algorithm

1. Start with some guesses of C_{t+1} and a symmetric and positive definite V_{t+1}.
2. Find F_t, C_t, and V_t according to the previous section.
3. Keep on iterating until F_t, C_t, and V_t converge.

Note that the also the matrices D_t, G_t, A_t^*, and B_t^*, change during this iterative process.

To sum up, the equations used in the iterations are

$$\begin{aligned}
D_t &= (A_{22} - C_{t+1}A_{12})^{-1}(C_{t+1}A_{11} - A_{21}), \text{ (which is } n_2 \times n_1) \\
G_t &= (A_{22} - C_{t+1}A_{12})^{-1}(C_{t+1}B_1 - B_2), \ (n_2 \times k) \\
A_t^* &= A_{11} + A_{12}D_t, \ (n_1 \times n_1) \\
B_t^* &= B_1 + A_{12}G_t, \ (n_1 \times k) \\
Q_t^* &= Q_{11} + Q_{12}D_t + D_t'Q_{21} + D_t'Q_{22}D_t, \ (n_1 \times n_1) \\
U_t^* &= Q_{12}G_t + D_t'Q_{22}G_t + U_1 + D_t'U_2, \ (n_1 \times k) \\
R_t^* &= R + G_t'Q_{22}G_t + G_t'U_2 + U_2'G_t, \ (k \times k) \\
F_t &= \left(R_t^* + \beta B_t^{*\prime} V_{t+1} B_t^*\right)^{-1} \left(U_t^{*\prime} + \beta B_t^{*\prime} V_{t+1} A_t^*\right), \ (k \times n_1) \\
C_t &= D_t - G_t F_t, \ (n_2 \times n_1) \\
V_t &= Q_t^* - U_t^* F_t - F_t' U_t^{*\prime} + F_t' R_t^* F_t + \beta \left(A_t^* - B_t^* F_t\right)' V_{t+1} \left(A_t^* - B_t^* F_t\right), \ (n_1 \times n_1)
\end{aligned} \quad (8.24)$$

8.5 The Time Invariant Solution

When the decision rules have converged we have

$$u_t = -F x_{1t}, \quad (8.25)$$

$$x_{2t} = C x_{1t}, \quad (8.26)$$

and the loss function value is

$$J_t^{opt} = x'_{1t} V x_{1t} + \frac{\beta}{1-\beta} \mathrm{trace}(V\Sigma). \quad (8.27)$$

Note that it should be possible to use this F matrix (padded with $\mathbf{0}_{k \times n_2}$ at the right) as a simple rule. Provided the decision rule puts the system on a unique stable path, the simple rule equilibrium (including the C matrix) should be the same as in the discretionary case.

8.6 Dynamics in Terms of x_{1t} and x_{2t}

From the constraints (8.3) we have

$$x_{1t+1} = A_{11}x_{1t} + A_{12}x_{2t} + B_1 u_t + \varepsilon_{t+1}. \quad (8.28)$$

Combining this with the converged decision rule, (8.25), and the converged relation between x_{1t} and x_{2t}, (8.26), gives

$$x_{1t+1} = (A_{11} + A_{12}C - B_1 F)\, x_{1t} + \varepsilon_{t+1}. \tag{8.29}$$

8.7 Singular Dynamic Equations*

Instead of (8.3), let the constraints be

$$\begin{bmatrix} x_{1t+1} \\ H\mathrm{E}_t x_{2t+1} \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} x_{1t} \\ x_{2t} \end{bmatrix} + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} u_t + \begin{bmatrix} \varepsilon_{t+1} \\ \mathbf{0}_{n_2 \times 1} \end{bmatrix}, \tag{8.30}$$

where H can be singular (if not, premultiply by H^{-1} to get the system on standard form).

It is straightforward to see that most of the previous derivation of the discretionary equilibrium still holds. However, in (8.7) we should substitute HC_{t+1} for C_{t+1}—and this carries through in all equations up to and including (8.10). As a consequence, we have to do the same in the definitions of D_t and G_t in (8.24).

Bibliography

Backus, D., and J. Driffil, 1986, "The Consistency of Optimal Policy in Stochastic Rational Expectations Models," CEPR Discussion Paper 124.

Blanchard, O. J., and C. M. Kahn, 1980, "The Solution to Linear Difference Models under Rational Expectations," *Econometrica*, 5, 1305–1311.

Currie, D., and P. Levine, 1993, *Rules, Reputation and Macroeconomic Policy Coordination*, Cambridge University Press.

Golub, G. H., and C. F. van Loan, 1989, *Matrix Computations*, The John Hopkins University Press, Baltimore, MD, 2nd edn.

King, R. G., and M. W. Watson, 1995, "The Solution of Singular Linear Difference Equations under Rational Expectations," mimeo University of Virginia.

Klein, P., 2000, "Using the Generalized Schur Form to Solve a Multivariate Linear Rational Expectations Model," *Journal of Economic Dynamics and Control*, 24, 1405–1423.

Oudiz, G., and J. Sachs, 1985, "International Policy Coordination in Dynamic Macroeconomic Models," in Willem H. Buiter, and Richard C. Marston (ed.), *International Economic Policy Coordination*, Cambridge University Press, Cambridge.

Sims, C. A., 2001, "Solving Linear Rational Expectations Models," *Journal of Computational Economics*, 20, 1–20.

Söderlind, P., 1999, "Solution and Estimation of RE Macromodels with Optimal Policy," *European Economic Review*, 43, 813–823.

Svensson, L. E. O., 1994, "Why Exchange Rate Bands? Monetary Independence in Spite of Fixed Exchange Rates," *Journal of Monetary Economics*, 33, 157–199.

Woodford, M., 1999, "Commentary: How Should Monetary Policy Be Conducted in an Era of Price Stability?," in *New Challenges for Monetary Policy*, Federal Reserve Bank of Kansas City.

9 Monetary Policy in VAR Systems

9.1 VAR System, Structural Form, and Impulse Response Function

Reference: Walsh (1998) 1.3. For a statistical background, see Hamilton (1994) and Greene (2000).

Let y_t be an $n \times 1$ vector of macro variables, including the policy instrument (usually a short interest rate or a narrow money aggregate). The VAR system, that is, the *reduced form* is

$$y_t = \mu + A_1 y_{t-1} + ... + A_p y_{t-p} + \varepsilon_t, \; \varepsilon_t \text{ is white noise, } \mathrm{Cov}(\varepsilon_t) = \Omega. \tag{9.1}$$

The underlying *structural form* is assumed to be

$$F y_t = \alpha + B_1 y_{t-1} + ... + B_p y_{t-p} + u_t, \; u_t \text{ is white noise, } \mathrm{Cov}(u_t) = D. \tag{9.2}$$

We are, in most cases, interested in understanding the effect of the structural shocks, u_t. This essentially requires an estimate of the structural form, but that can be achieved by imposing identifying restrictions on the VAR. As an example, the impulse response function of the VAR in (9.1) is

$$y_t = \varepsilon_t + C_1 \varepsilon_{t-1} + C_2 \varepsilon_{t-2} + ... \tag{9.3}$$

By comparing (9.1) and (9.2) we see that $\varepsilon_t = F^{-1} u_t$ (or $u_t = F \varepsilon_t$). We can then rewrite the impulse response function (9.3) in terms of the structural shocks

$$y_t = F^{-1} u_t + C_1 F^{-1} u_{t-1} + C_2 F^{-1} u_{t-2} + ... \tag{9.4}$$

A VAR estimation gives us C_i, $i = 1, 2, \ldots$, but not F, so we need to impose restrictions in order to identify the impulse responses to structural shocks.

Remark 24 *The easiest way to calculate this representation is by first finding F^{-1} (see*

below), then using $\varepsilon_t = F^{-1} u_t$ to write (9.1) as

$$y_t = \mu + A_1 y_{t-1} + ... + A_p y_{t-p} + F^{-1} u_t. \tag{9.5}$$

To calculate the impulse responses to the first element in u_t, set $y_{t-1}, ..., y_{t-p}$ equal to the long-run average, $(I - A_1 - ... - Ap)^{-1} \mu$, make the first element in u_t unity and all other elements zero. Calculate the response by iterating forward on (9.5), but putting all elements in $u_{t+1}, u_{t+2}, ...$ to zero. This procedure can be repeated for the other elements of u_t.

To see the mapping between the reduced form and the structural form, premultiply (9.2) by F^{-1}. This shows that the relation between the VAR parameters and the structural parameters is

$$\begin{array}{lcl} \underline{\text{VAR}} & & \underline{\text{in terms of structural form parameters}} \\ \Omega & = & F^{-1} D \left(F^{-1}\right)' \\ A_s & = & F^{-1} B_s \text{ for } s = 1, ..., p. \end{array} \tag{9.6}$$

In the VAR, there are pn^2 elements in $A_1, ..., A_p$ and $n(n+1)/2$ (unique) elements in Ω. In the structural form, there are $(1 + p)\, n^2$ elements in $F, ..., B_p$ and $n(n+1)/2$ (unique) elements in D. We therefore have to impose at least n^2 (non-trivial) restrictions on the structural form in order to back out the structural form parameters from the reduced form.

9.2 Fully Recursive Structural Form

9.2.1 Identification

Remark 25 *(Cholesky decomposition) Let Ω be an $n \times n$ symmetric positive definite matrix. The Cholesky decomposition gives the unique lower triangular P such that $\Omega = PP'$ (some software returns an upper triangular matrix, that is, Q in $\Omega = Q'Q$ instead). Note that each column of P is only identified up to a sign transformation; they can be reversed at will.*

Remark 26 *(Changing sign of column and inverting.) Suppose the square matrix A_2 is the same as A_1 except that the i^{th} and j^{th} columns have the reverse signs. Then A_2^{-1} is the same as A_1^{-1} except that the i^{th} and j^{th} rows have the reverse sign.*

The most common set of restrictions is to assume that F is lower triangular and that $D = I$, which gives exact identification. The Cholesky decomposition is useful in this case.

A Cholesky decomposition of the covariance matrix of the VAR residuals, Ω, gives a lower triangular matrix, which by (9.6) can be taken to represent F^{-1}, since a lower triangular F (as assumed) implies a lower triangular F^{-1} and $D = I$. Note however, that the signs of each column of F^{-1} are arbitrary. Therefore, we have

$$\text{chol}(\Omega) = F^{-1}, \tag{9.7}$$

up to a sign transformation of each column of F^{-1}, which implies a sign transformation of each row of F. With F identified, B_1, ..., B_p can be calculated from (9.6).

Expression (9.2) with a lower triangular F and $D = I$ is, in fact, a fully recursive system of simultaneous equations (Greene (2000) 16.3). Using (9.6) and (9.7) is just a way to recover the fully recusive system from the VAR.[1]

9.2.2 Monetary Policy

We now consider monetary policy in a fully recusrive structural model. Partition the vector of endogenous variables, y_t, into the (scalar) policy instrument, s_t, variables which come before s_t, x_{1t}, and those which come after s_t, x_{2t},

$$y_t = \begin{bmatrix} x_{1t} \\ s_t \\ x_{2t} \end{bmatrix}. \tag{9.8}$$

[1]We would asymptotically get the same structural parameters by equation-by-equation LS of (9.2) LS is FIML in this case (assuming normally distributed shocks), since the structural shocks are assumed to be uncorrelated. The reason why the two estimates are not identical in small samples is that the VAR approach imposes that also the small sample estimate of D is an identify matrix, while the equation-by-equation LS does not.

Rewrite (9.2) as (assuming $\alpha = \mathbf{0}$)

$$\begin{bmatrix} F^{11} & \mathbf{0} & \mathbf{0} \\ F^{21} & F^{22} & \mathbf{0} \\ F^{31} & F^{32} & F^{33} \end{bmatrix} \begin{bmatrix} x_{1t} \\ s_t \\ x_{2t} \end{bmatrix} = \begin{bmatrix} B_1^{11} & B_1^{12} & B_1^{13} \\ B_1^{21} & B_1^{22} & B_1^{23} \\ B_1^{31} & B_1^{32} & B_1^{33} \end{bmatrix} \begin{bmatrix} x_{1t-1} \\ s_{t-1} \\ x_{2t-1} \end{bmatrix} + \dots$$
$$+ \begin{bmatrix} B_p^{11} & B_p^{12} & B_p^{13} \\ B_p^{21} & B_p^{22} & B_p^{23} \\ B_p^{31} & B_p^{32} & B_p^{33} \end{bmatrix} \begin{bmatrix} x_{1t-p} \\ s_{t-p} \\ x_{2t-p} \end{bmatrix} + \begin{bmatrix} u_{1t} \\ u_{st} \\ u_{2t} \end{bmatrix}, \tag{9.9}$$

where F^{22} is a scalar, and F^{11} and F^{33} are lower-triangular matrices (not necessarily with diagonal elements equal to unity). The covariance matrix of the shocks is the identity matrix. This model has $D = I$ and a lower triangular F.

The equation for s_t in (9.9) is

$$F^{22}s_t = -F^{21}x_{1t} + \begin{bmatrix} B_1^{21} & B_1^{22} & B_1^{23} \end{bmatrix} \begin{bmatrix} x_{1t-1} \\ s_{t-1} \\ x_{2t-1} \end{bmatrix} + \dots + \begin{bmatrix} B_p^{21} & B_p^{22} & B_p^{23} \end{bmatrix} \begin{bmatrix} x_{1t-p} \\ s_{t-p} \\ x_{2t-p} \end{bmatrix} + u_{st}. \tag{9.10}$$

If we divide by the scalar F^{22}, then we get a traditional *reaction function*. Policy in t is determined by *(i)* a rule which depends on the contemporaneous x_{1t} (but not x_{2t}); *(ii)* all lagged variables; and *(iii)* a monetary policy shock, u_{st}.[2]

Suppose s_t is the j^{th} element in y_t. The *impulse response with respect to the monetary policy shock* is then found from the j^{th} columns of the matrices in (9.4), that is, the j^{th} columns of F^{-1}, C_1F^{-1}, C_2F^{-1},... Since F^{-1} is lower triangular, a policy shock in period t, u_{st}, has a contemporaneous effect on x_{2t}, but not on x_{1t}.

9.2.3 Importance of the Ordering of the VAR

Suppose the our objective is to analyze the effects of monetary policy shocks on the other variables in the VAR system, for instance, output and prices. The identification rests on the ordering of the VAR, that is, on the structure of the contemporaneous correlations as captured by F. It is therefore important to understand how the results on the monetary

[2]Note also that since $\text{Std}(u_{st}) = 1$, $\text{Std}(u_{st}/F^{22}) = 1/\left|F^{22}\right|$. This clarifies the relation to the traditional normalization in systems of simulataneous equations (diagonal elements of F equal to unity and D diagonal but not restricted to be an identity matrix); the absolute values of the diagonal elements in F here corresponds to the inverses of the standard deviations of the shocks in the traditional normalization.

policy shock are changed if the variables are reordered.

My conjecture is summarized below. (I have not been able to locate any proof in the literature and my own proof is only half-baked.)

1. The partitioning of y_t into variables which come before, x_{1t}, and after, x_{2t}, the policy instrument is important for u_{st} and the impulse response function of all variables with respect to u_{st}.

2. The order *within* x_{1t} and x_{2t} does not matter for u_{st} or the impulse response function of any variable with respect to u_{st}.

This suggests that we can settle for *partial identification* in the sense that we must take a stand on which variables that come before and after the policy instrument, but the ordering within those blocks are unimportant for understanding the effects of monetary policy shocks.

The typical identifying assumption in much of Sims' work (see for instance, Sims (1980)) is that the monetary policy variable is unaffected by contemporaneous innovations in the other variables, that is, it is put "first" in the VAR. In later work, by Sims and others, monetary policy is instead put last (so monetary policy is potentially affected by, but does not affect, contemporaneous macro variables).

9.2.4 On Variance Decompositions

It is sometimes found in VAR studies that policy surprises explains only a small part of the variance of y_t (a typical result for US studies for the period after 1982, see for instance, Leeper, Sims, and Zha (1996)). Two comments are warranted (see also Bernanke (1996)). *First*, this does not mean that all monetary policy has been unimportant. For instance, it could be the case that anticipated monetary policy, or more generally, the systematic monetary policy, decreases the variance of output and inflation. *Second*, the variance decomposition does not tell us about the potential effects of monetary policy surprises (the impulse response function does, however), only about the combination of the potential effect with the actual monetary policy shocks for that particular sample.

9.3 Some Controversies

9.3.1 Choice of Policy Instrument: i_t or m_t?

Sims (1980) showed that the fraction of the forecasting error variance in US output that can be attributed to money stock innovations is much lower when an interest rate is added to a VAR of money, price, and output. (The typical identifying assumption in much of Sims' work is that the monetary policy variable is unaffected by contemporaneous innovations in the other variables, that is, it is put "first" in the VAR.)

McCallum (1983) argues that the policy instrument of the Fed is a short interest rate and that the correct measure of the monetary policy shocks is the residual in a reaction function for this interest rate

$$\text{short interest rate} = f(\text{lagged macro data}) + \text{policy shock}. \tag{9.11}$$

The crucial assumption is that the policy instrument does not depend on contemporaneous macro data. In contrast, the money stock does, since money demand is probably affected by shocks to income and prices (as well as the policy shock to the interest rate). The innovation in the money stock, $m_t - \mathrm{E}_{t-1} m_t$, is therefore a mixture of the policy shock and other shocks.

Bernanke and Blinder (1992) argue for using the federal funds rate as the policy instrument. (The federal funds rate is the market interest rate for over-night US dollar loans. It is usually loans of reserves between banks, called "federal funds loans" since they have typically been used to meet the reserve requirements. It is not directly controlled by Fed; only the discount rate is.) *First*, a policy instrument, if effective, should be able to predict macro economic variables. They find that the federal funds rate produces better forecasts of output, employment, and consumption than M1, M2, T-bill rates, or long bonds. *Second*, they notice that the federal funds rate was raised at all cyclical peaks (NBER) and at most of the "Romer dates." Estimates of reaction functions like (9.11) produce reasonable responses to inflation and unemployment shocks. *Third*, the estimated supply curve of non-borrowed reserves is extremely elastic at the target funds rate between FOMC meetings. This suggests that the federal funds rate is predetermined within the month (and presumably set by policy), and not driven by demand for reserves which changes continually as the economy is hit by shocks.

Sims (1992) argues that expansionary shocks to monetary policy should drive output up and lead to opposite movements in money stock and interest rates. Eichenbaum (1992) comments that this makes Sims choose the interest rate rather than M1 as the policy instrument since positive shocks to M1 lead (in typical VAR of the US economy) to an increase in the federal funds rate and a decline in output! He also notes that M0, but not non-borrowed reserves, has the same property. Since the former has a less pronounced price puzzle (see below), Eichenbaum (1992) argues that it is a better measure of monetary policy than the federal funds rate.

In short (and for the US), the policy instrument is now usually taken to be the federal funds rate, or in some cases, some narrow money aggregate.

9.3.2 The "Price Puzzle"

The price puzzle is that in a VAR of output, prices, money, interest rate and perhaps some more variables, contractionary shocks to monetary policy leads to persistent price increases! This seems to hold not just in the US, but also in several other countries, and is more pronounced if the policy instrument is taken to be a short interest rate rather than a money aggregate. It is often not statistically significant, but is so common that it signals that the VAR might be misspecified.

Sims (1992) discusses how this could be due to a missing element in the reaction function of the central bank. Commodity prices may signal inflation expectations, so the central bank may react now by raising interest rates which makes the inflation somewhat lower than it would otherwise have been (but still positive). If commodity prices are excluded from the VAR, this may appear as monetary policy shocks having a positive effect on inflation.

9.4 Interpretation of the VAR Results

Reference: Cochrane (1998).

9.4.1 Setup and Important VAR Results

Suppose we have estimated an output-money (plus whatever) VAR, imposed identifying restrictions, and calculated the impulse responses to the structural shocks

$$\begin{bmatrix} m_t \\ y_t \end{bmatrix} = \begin{bmatrix} R_{mm}(\mathrm{L}) & R_{my}(\mathrm{L}) \\ R_{ym}(\mathrm{L}) & R_{yy}(\mathrm{L}) \end{bmatrix} \begin{bmatrix} u_{mt} \\ u_{yt} \end{bmatrix}, \mathrm{Cov}\left(\begin{bmatrix} u_{mt} \\ u_{yt} \end{bmatrix} \right) = 1. \quad (9.12)$$

A typical result is that monetary policy shocks have a hump-shaped effect on output, but also a long lasting effect on monetary policy. In fact, $R_{mm}(\mathrm{L})$ and $R_{ym}(\mathrm{L})$ are quite similar.

The issue in Cochrane's paper is whether the impulse response of output, $R_{ym}(\mathrm{L})$, depends on the monetary policy or not and how this affects the interpretation of the impulse response functions obtained from an identified VAR model.

9.4.2 If Only Unanticipated Policy Matters

Suppose the true model is that output depends on policy surprises (current and lagged) as well as other shocks

$$y_t = a^*(\mathrm{L})(m_t - \mathrm{E}_{t-1}m_t) + b^*(\mathrm{L})\,\delta_t, \quad (9.13)$$

where $\mathrm{L}^s(m_t - \mathrm{E}_{t-1}m_t) = m_{t-s} - \mathrm{E}_{t-s-1}m_t$, and where the non-monetary shock, δ_t, is uncorrelated with the monetary shock.

The coefficients in (9.13) are supposed to be unaffected by monetary policy. (This rules out, among other things, that a change in the volatility of money supply affects the coefficient of the money supply surprise, as in Lucas' model.)

Attaching a lag polynomial to policy surprises is ad hoc, but was done almost immediately after the Lucas model was published. The original Lucas model could not explain the business cycles or the long responses of output to monetary shocks. (The first motivation for these lags was in terms of capital accumulation, but this can hardly be a plausible explanation given the stability of capital stock.)

From (9.12) we have

$$m_t - \mathrm{E}_{t-1}m_t = R_{mm}(0)\,u_{mt} + R_{my}(0)\,u_{yt},$$

which can be used to rewrite (9.13) as

$$y_t = a^* (\mathrm{L})\, R_{mm} (0)\, u_{mt} + a^* (\mathrm{L})\, R_{my} (0)\, u_{yt} + b^* (\mathrm{L})\, \delta_t. \tag{9.14}$$

Equation (9.14) and the second line in (9.12) must be identical. This implies that

$$a^* (\mathrm{L})\, R_{mm} (0) = R_{ym} (\mathrm{L}), \tag{9.15}$$

so the VAR impulse response of output to policy shocks, R_{ym} (L), is proportional to the true propagation mechanism, a^* (L), which is invariant to actual monetary policy. In this case, the VAR is useful tool for understanding the effect of monetary policy surprises on output.

This means that the hump-shaped and long-lasting effect of monetary policy shocks found in VAR studies , R_{ym} (L), reflects the ad-hoc dynamics, a^*(L), attached to the Lucas' model.

The similarly between R_{ym} (L) and R_{mm} (L) found in VAR studies should, in this setting, be interpreted as a coincidence.

9.4.3 If No Distinction Between Anticipated and Unanticipated Policy

Consider the extreme case where anticipated policy has the same effect as unanticipated policy, so the true model is

$$y_t = a^* (\mathrm{L})\, m_t + b^* (\mathrm{L})\, \delta_t. \tag{9.16}$$

By using (9.12) this can be written

$$y_t = a^* (\mathrm{L})\, R_{mm} (\mathrm{L})\, u_{mt} + a^* (\mathrm{L})\, R_{my} (\mathrm{L})\, u_{yt} + b^* (\mathrm{L})\, \delta_t. \tag{9.17}$$

This should be equal to the second line in (9.12), so

$$a^* (\mathrm{L})\, R_{mm} (\mathrm{L}) = R_{ym} (\mathrm{L}). \tag{9.18}$$

The VAR impulse response function of output to policy shocks, R_{ym}(L), is no longer invariant to the policy rule—rather the opposite.

In the extreme case when the propagation mechanism is such that output only depends on the current monetary policy shock, $a^* (\mathrm{L}) = a^*$, then the typical hump-shaped pattern

of R_{ym} (L) is a reflection of a hump-shaped pattern of how policy shocks affect future policy, that is of the hump-shaped pattern of R_{mm} (L).

Since estimates of R_{mm} (L) are typically hump-shaped in data, and fairly similar to the estimates of R_{ym} (L), this setting suggests that $a^* (\mathrm{L}) \approx a^*$, that is, the effect of money on output is almost contemporaneous.

In this case the VAR is a not a useful tool for understanding the effect of monetary policy surprises on output. Instead, a direct estimation of (9.16) should work well.

Note that this does not concern the endogenous part of monetary policy, that is, how monetary policy affects R_{my} (L) and thereby how output reacts to output shocks.

9.4.4 Federal Funds Rate

In this case the monetary policy is measured by the innovations in the Federal funds rate (ordered last in a VAR including, among other things, commodity prices to deal with the "price puzzle"). The results are similar to those discussed above.

9.4.5 Sticky Price Models

Sticky price model (for instance the Taylor model) has built-in dynamics, where both anticipated and unanticipated policy matters, but where the latter is usually more powerful. The built-in dynamics decreases the need for ad hoc dynamics, as captured by a^* (L) above, in order to explain the observed VAR impulse response of output to money supply surprises.

9.5 "The Federal Funds Rate and the Channels of Monetary Transmission" by Bernanke and Blinder

Reference: Bernanke and Blinder (1992).

- Monthly US data 1959:1-1978:12.
- VAR of federal funds rate, unemployment rate, log of CPI, deposits/securities/loans.
- Identifying assumption: monetary policy is predetermined (does not depend on other contemporaneous shocks, as in much of Sims' work).

- Results: *(i)* policy shocks (higher federal funds rate) increases the unemployment rate after a year; *(ii)* bank deposits fall; *(iii)* banks initially sell off securities to balance the drop in deposits, but this is later undone and the volume of loans is reduced instead.

- Interpretation: adjustment of the stock of loans takes time, so the fall in deposits is initially met by selling of liquid securities. Unemployment starts to rise at the same time as stock of loans is reduced. More than a coincidence (decreased supply of credit - the "credit channel")? Or is it that the demand for loans decrease as the interest rate increase creates a recession by the standard IS-LM mechanism?

9.6 "The Effects of Monetary Policy Shocks: Evidence from the Flow of Funds" by Christiano, Eichenbaum, and Evans

Reference: Christiano, Eichenbaum, and Evans (1996).

- A study of the effect of monetary policy shocks on, for instance, "net funds raised in the financial markets" by nonfinancial business or households.

- Quarterly US data 1960:Q1-1992:Q4.

- VAR (in levels?) of log real GDP, log of GDP deflator, log commodity prices, *federal funds rate*, *minus log of non-borrowed reserves*, total reserves, and net funds raised in the financial markets. This is also the ordering in the identification.

  - Choice of policy instrument: federal funds rate or minus log of non-borrowed reserves (in the latter case the order of the federal funds rate and the non-borrowed reserves is reversed in the VAR). The results are not particularly sensitive to this choice.
  - Commodity prices included to avoid the prize puzzle.

- Inspection of the estimated shocks to the federal funds rate.

  - The policy shocks are relatively high before each NBER recession, et vice versa. Causality?

  - Persistent effects on the federal funds rate (lasts almost two years).
  - A positive shock decreases Fed's holdings of US government securities. Are the subsequent increases in the interest rate accomplished by selling bills and bonds (open market operations)?
  - A positive shock decreases M1 and output (the latter with a lag of two quarters).

- Result: the initial effect of a positive shock to the federal funds rate is to increase net funds raised by the business sector for almost a year, and it is only thereafter that we observe a decline! This is quite contrary to most models, including the "credit channel " interpretation of Bernanke and Blinder (1992). Why? Interest rate shocks create recessions, firm revenues decrease, but costs take time to change?

9.7 "Do Measures of Monetary Policy in a VAR Make Sense" by Rudebusch

Reference: Rudebusch (1998) and Sims (1998).

- VAR interest rate equations.

  - Can be interpreted as a reaction function, see (9.10).
  - Time invariant, linear structure: tests of parameter stability in the reaction function often rejects stability (monthly US data 1960-:1-1995:3.)
  - Small information set: traditional reaction functions typically use a much larger information set (trade deficit, stance of fiscal policy, measures of political pressure). The official records indicate that different types of data has been of interest at different times.
  - Use of final data, Y^F,while the true reaction function can only include preliminary data, Y^p, where $Y_t^F = Y_t^P + w_t$. Suppose the true reaction function is $s_t = \alpha Y_t^P + u_{st}$, but we estimate $s_t = \alpha Y_t^F + e_{st} = \alpha Y_t^P + \alpha w_t + u_{st}$. If the statistical agency produces inefficient preliminary estimates, then w_t and Y_t^P will be correlated and the estimator that produces $\hat{\alpha}$ is inconsistent. Important?

- long distributed lags: VAR estimates often show significant coefficients at lags of many months, which indicate that there is some variation in the federal funds rate which can be predicted many months in advance. This is at variance with other evidence.

- VAR interest rate shocks.
  - Comparison of the VAR innovations, that is, the j^{th} element in ε_t from (9.1) (rather than the structural shocks u_{st}) with the difference between forward federal funds rate and the realized federal funds rate (short sample: 1988:10-1995:3).
  - This is done for the forecast/forward price of the federal funds rate average over a month which be realized one-, two-, and three-months in the future.
  - The VAR shocks are much more volatile than the surprise according to financial data, and the correlation between them is low.

- Other observations.
  - Different VARs in the literature have produced different time series of policy shocks, but fairly similar impulse response functions. Strange! Data mining to get a reasonable impulse response function?

Sims (1998) has two important comments on Rudebusch's paper. *First*, excluding an exogenous variable from a regression (in a system of equations) does not necessarily lead to bad estimates of the coefficients (this is clearly the case if the excluded explanatory variables are uncorrelated with the other explanatory variables), but will obviously change the fitted residual. *Second*, the Federal funds rate is often changed quickly as new information about the state of the economy arrives. This means that innovations to the federal funds rate contain both policy surprises (what we want to measure) and reactions to innovations in the state of the economy. The difference between the federal funds futures and the actual federal funds rate is such an innovation.. In contrast, a VAR where the policy instrument is allowed to depend on current values of the state of the economy, may potentially be able to separate the components of the innovation. Note that this is an argument for *not* making the monetary policy instrument predetermined, as is the case in much of Sims' own work.

9.8 "What Does Monetary Policy Do?" by Leeper, Sims. and Zha

Reference: Leeper, Sims, and Zha (1996) and Bernanke (1996).

This paper estimates large VAR systems of monetary policy, sometimes with split of the sample. The degrees of freedom problem is handled with a Bayesian approach.

In his comment on the paper, Bernanke finds that the most important conclusions from the paper are the following.

1. The estimated effects of monetary policy seems plausible: the VAR approach might work.
2. Empirically, short interest rates are better indicators of monetary policy than monetary aggregates. This finding is not surprising given the amount of interest rate smoothing that most central bank pursue: the supply curve of reserves is almost flat between infrequent interest rate changes; most innovations in monetary aggregates reflect money demand innovations.
3. Monetary policy surprises have been relatively unimportant for US macroeconomic fluctuations since 1960.
4. Monetary policy reacts strongly to the macroeconomic situation (the "feedback" or "systematic" part of monetary policy is an important part of monetary policy).

Bernanke points out that the third point does not prove that monetary policy shocks cannot have large effects. In order to assess this possibility, the impulse response function is more useful than the forecasting error variance decomposition. He also notes that the VAR approach has little to say about the effects of anticipated monetary policy.

9.9 "Identifying Monetary Policy in a Small Open Economy under Flexible Exchange Rates" by Cushman and Zha

Reference: Cushman and Zha (1997).

The authors argue that a VAR study of a small open economy cannot be done in the same way as for the US. They try to incorporate the following aspects in a VAR of monthly Canadian macro data.

1. Interest rates movements are likely to react contemporaneously to foreign interest rates. This is an argument against assuming that the monetary policy instrument is predetermined (like in much of Sims' work).

2. Under a flexible exchange rate regime, the exchange rates should be allowed to react to all contemporaneous shocks. It is, after all, a forward looking asset price.

3. Trade flows are interesting and important.

4. Foreign variables can be treated as a separate block, which is (block) exogenous for the domestic (small open) economy. In practice, this means that domestic variables are not allowed to affect foreign variables - not even with a lag.

5. Data: monthly 1974-1993 data for Canada. US is taken to be the "the rest of the world."

6. A (Sims style) identification gives the strange result that a monetary contraction leads to price increases (the "price puzzle" once again) and a depreciation of the exchange rate (an "exchange rate puzzle").

7. Their identification implies a traditional money demand equation ($M1$, P, y, i), and a money supply equation which may depend on the foreign interest rate and commodity prices, but not on contemporaneous output. (Plus a few more things.)

8. Their results indicate that a monetary contraction leads to an appreciation of the Canadian dollar, an increase in the interest rate and a decrease in the money stock, a prolonged negative effect on the price level, and a small but negative effect on output. The variance decomposition indicate that monetary policy shocks account for only a small fraction of forecast error variance of output.

9.10 "What Does the Bundesbank Target?" by Bernanke and Mihov

Reference: Bernanke and Mihov (1997). See also Walsh 9.4 and Bernanke and Mihov (1998).

Remark 27 *(Bundesbank's interest rates. See, for instance, Burda and Wyplosz (1997) 9.3) The floor of the interest rate tunnel was the discount rate. Access to this was limited by quotas. The ceiling was the Lombard rate at which banks borrowed in emergencies. The repo rate was in between. The instruments of the Bundesbank were these three rates, the quotas, and the reserve requirements.*

Bernanke and Mihov study what the Bundesbank has actually done over the period 1969:01 to 1990:12. They find that inflation forecasts explain much more of the variance in the Lombard rate than does money growth. The conclude that the Bundesbank has, in fact, been running an inflation target.

Their structural model looks like

$$\begin{bmatrix} I-F & \mathbf{0} \\ -D_0 & I-G_0 \end{bmatrix} \begin{bmatrix} Y_t \\ P_t \end{bmatrix} = \sum_{s=1}^{k} \begin{bmatrix} B_s & C_s \\ D_s & G_s \end{bmatrix} \begin{bmatrix} Y_{t-s} \\ P_{t-s} \end{bmatrix} + \begin{bmatrix} A^y & \mathbf{0} \\ \mathbf{0} & A^p \end{bmatrix} \begin{bmatrix} v_t^y \\ v_t^p \end{bmatrix}, \tag{9.19}$$

where shocks are uncorrelated with each other. P_t is a vector of "policy variables" with the associated structural shocks v_t^p, where the innovations to the variables in P_t may be correlated through a non-diagonal matrix A^p. Y_t is a vector of non-policy variables, like output and inflation.

Policy variables have no contemporaneous effect on the non-policy variables (the opposite approach to what Sims typically use). If P_t was a scalar, this would be enough for identifying the policy shock, v_t^p. It would simply be the residual in a regression of P_t on contemporaneous Y_t and lags of both P_t and Y_t.

In this paper, P_t contains total reserves, tr, non-borrowed reserves, nbr, the call rate (a market rate of reserves, similar to the federal funds rate), cr, and the Lombard rate, lr. It is therefore necessary to put extra restrictions on the system in order to extract a *scalar policy shock*, v_t^s.

The reduced form, VAR, of (9.19) is

$$\begin{bmatrix} Y_t \\ P_t \end{bmatrix} = \sum_{s=1}^{k} A_s \begin{bmatrix} Y_{t-s} \\ P_{t-s} \end{bmatrix} + \begin{bmatrix} u_t^y \\ u_t^p + u_t^q \end{bmatrix}. \tag{9.20}$$

The VAR residuals for the policy block has been split up into two components. To see what they are, note that by the rules of partitioned matrices, we have that the inverse of the leading matrix in (9.19) is.

$$\begin{bmatrix} I-F & \mathbf{0} \\ -D_0 & I-G_0 \end{bmatrix}^{-1} = \begin{bmatrix} (I-F)^{-1} & \mathbf{0} \\ (I-G_0)^{-1} D_0 (I+F)^{-1} & (I-G_0)^{-1} \end{bmatrix}. \tag{9.21}$$

The VAR shocks of the policy variables must have the following relation to the structural shocks

$$u_t^p + u_t^q = (I - G_0)^{-1} A^p v_t^p + (I - G_0)^{-1} D_0 (I + F)^{-1} A^y v_t^y. \tag{9.22}$$

Define u_t^p to be $(I - G_0)^{-1} A^p v_t^p$, that is, the part of the VAR shock (of the policy variables) which is uncorrelated with structural non-policy shocks, v_t^y. (Recall that v_t^p and v_t^y are uncorrelated). We must therefore have

$$(I - G_0) u_t^p = A^p v_t^p. \tag{9.23}$$

The model for the policy innovations are (corresponding to equations (2.7)-(2.10) in the paper)

$$\begin{aligned}
\text{Total reseves demand}: &\ u_{tr} = -\alpha u_{cr} + v^d \\
\text{Lombard loans demand}: &\ u_{ll} = \beta (u_{cr} - u_{lr}) + v^b \\
\text{Nonborrowed reserves supply}: &\ u_{nbr} = \phi^d v^d + \phi^b v^b + \phi^s v^s + v^n \\
\text{Lombard rate}: &\ u_{lr} = \gamma^d v^d + \gamma^b v^b + \gamma^n v^n + v^s,
\end{aligned}$$

and v^s is taken to be the "policy shock," which we want to identify. Using the identity $u_{ll} = u_{tr} - u_{nbr}$, the model implies the following restrictions on (9.23)

$$\underbrace{\begin{bmatrix} 1 & 0 & \alpha & 0 \\ 1 & -1 & -\beta & \beta \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}}_{I-G_0} \begin{bmatrix} u_{tr} \\ u_{nbr} \\ u_{cr} \\ u_{lr} \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \phi^d & \phi^b & 1 & \phi^s \\ \gamma^d & \gamma^b & \gamma^n & 1 \end{bmatrix}}_{A^p} \begin{bmatrix} v^d \\ v^b \\ v^n \\ v^s \end{bmatrix}. \tag{9.24}$$

The model is estimated in a two-step procedure. First, each equation in the VAR (9.20) is estimated separately with least squares. Second, the policy shock is identified by matching the covariance matrix of the VAR residuals with the covariance matrix implied by the theoretical model (9.24). The idea is that if we had the $\text{Cov}(u_t^p)$ (has $4(4+1)/2 = 10$ unique elements), then we could solve for the parameters in (9.24) (there are 8 parameters) plus the variances of the structural shocks (4). With at least additional two restrictions (on top of all the zeros and cross restrictions already assumed), the parameters could be identified, that is, we solve (typically a non-linear problem) for the parameters and the variances from $\text{Cov}(u_t^p)$.

They therefore put additional restrictions. For instance, in case of "Lombard rate targeting" they set $\gamma^d = \gamma^b = \gamma^n = 0$ (gives overidentification), which means that $v^s = u_{lr}$. Alternatively, with "nonborrowed-reserves targeting," they impose $\phi^d = \phi^b = \phi^s = 0$ (gives overidentification).

The previous discussion supposed that we could observe u_t^p and calculate $\text{Cov}(u_t^p)$. However, the VAR shocks for the policy blocks are $u_t^p + u_t^q$ as given in (9.22). One way of dealing is as follows (I do not know how Bernanke and Mihov did). By (9.19)-(9.21) the VAR shock for the non-policy block must be

$$u_t^y = (I - F)^{-1} A^y v_t^y. \tag{9.25}$$

From (9.21) we then get

$$u_t^p + u_t^q = (I - G_0)^{-1} A^p v_t^p + (I - G_0)^{-1} D_0 u_t^y. \tag{9.26}$$

This gives

$$\begin{aligned}
\text{Cov}\left(u_t^p + u_t^q\right) = (I - G_0)^{-1} A^p \text{Cov}\left(v_t^p\right) A^{p\prime} \left[(I - G_0)^{-1}\right]' + \\
(I - G_0)^{-1} D_0 \text{Cov}\left(u_t^y\right) D_0' \left[(I - G_0)^{-1}\right]',
\end{aligned} \tag{9.27}$$

since v_t^y, and therefore u_t^y, and v_t^p are uncorrelated. From (9.26) we also get

$$\text{Cov}\left(u_t^p + u_t^q, u_t^y\right) = (I - G_0)^{-1} D_0 \text{Cov}\left(u_t^y\right) \left[(I - G_0)^{-1}\right]' D_0', \tag{9.28}$$

The matrices $\text{Cov}(u_t^p + u_t^q)$, $\text{Cov}(u_t^p + u_t^q, u_t^y)$ and $\text{Cov}(u_t^y)$ can be estimated from the VAR residuals. If A^p and $I - G_0$ are identified as discussed in conjunction with (9.24), then the equations in (9.28) are sufficient to identify D_0, since it has as many unique elements as there are elements in D_0.

Bibliography

Bernanke, B. S., 1996, "Comment to 'What Does Monetary Policy Do?' by Leeper, Sims, and Zha," *Brookings Papers on Economic Activity*, 2, 69–73.

Bernanke, B. S., and A. S. Blinder, 1992, "The Federal Funds Rate and the Channels of Monetary Transmission," *American Economic Review*, 82, 901–921.

Bernanke, B. S., and I. Mihov, 1997, "What Does the Bundesbank Target?," *European Economic review*, 41, 1025–1053.

Bernanke, B. S., and I. Mihov, 1998, "Measuring Monetary Policy," *Quarterly Journal of Economics*, 113, 869–902.

Burda, M., and C. Wyplosz, 1997, *Macroeconomics - A European Text*, Oxford University Press, 2nd edn.

Christiano, L. J., M. Eichenbaum, and C. Evans, 1996, "The Effects of Monetary Policy Shocks: Evidence from the Flow of Funds," *Review of Economics and Statistics*, 78, 16–34.

Cochrane, J. H., 1998, "What Do the VARs Mean? Measuring the Output Effects of Monetary Policy," *Journal of Monetary Economics*, 41, 277–300.

Cushman, D. O., and T. Zha, 1997, "Identifying Monetary Policy in a Small Open Economy under Flexible Exchange Rates," *Journal of Monetary Economics*, 39, 433–448.

Eichenbaum, M., 1992, "Comment on 'Interpreting the Macroeconomic Time Series Facts: The Effects of Monetary Policy' by Christopher Sims," *European Economic Review*, 36, 1001–1011.

Greene, W. H., 2000, *Econometric Analysis*, Prentice-Hall, Upper Saddle River, New Jersey, 4th edn.

Hamilton, J. D., 1994, *Time Series Analysis*, Princeton University Press, Princeton.

Leeper, E. M., C. A. Sims, and T. Zha, 1996, "What Does Monetary Policy Do?," *Brookings Papers on Economic Activity*, 2, 1–63.

McCallum, B. T., 1983, "A Reconsideration of Sim's Evidence Regarding Monetarism," *Economic Letters*, 13, 167–171.

Rudebusch, G. D., 1998, "Do Measures of Monetary Policy in a VAR Make Sense?," *International Economic Review*, 39, 907–931.

Sims, C. A., 1980, "Comparison of Interwar and Postwar Business Cycles: Monetarism Reconsidered," *American Economic Review*, 70, 250–257.

Sims, C. A., 1992, "Interpreting the Macroeconomic Time Series Facts: The Effects of Monetary Policy," *European Economic Review*, 36, 975–1000.

Sims, C. A., 1998, "Comment on Glenn Rudebusch's 'Do Measures of Monetary Policy in a VAR Make Sense?'," *International Economic Review*, 39, 933–942.

Walsh, C. E., 1998, *Monetary Theory and Policy*, MIT Press, Cambridge, Massachusetts.