Lectures in Quantitative Monetary Economics

Martin Ellison
University of Warwick and CEPR
martin.ellison@warwick.ac.uk

2003

Contents

Chapter 1

Stylised monetary facts

1.1 Motivation

In this lecture we outline some simple summary statistics regarding the cyclical behaviour of real and nominal variables and their interactions. It provides an empirical reference point to help guide us in our subsequent analysis of how important money is for the business cycle. While informative, we interpret the following results as preliminary. Before they could be considered as definitive we would need to do substantial extra work in establishing their robustness across data periods, data definitions and empirical techniques.

1.2 Key readings

Walsh (1999), *Monetary Theory and Practice*, Chapter 1. This is an excellent graduate textbook.

Cooley (ed.) (1995), *Frontiers of Business Cycle Research*, Chapter 7, Sections 1 and 2. This is also a very good reference.

1.3 Related reading

Cogley and Nason (1994), "Effects of the Hodrick-Prescott filter on trend and difference stationary time series: Implications for business cycle research", Journal of Economic Dynamics and Control, 19, 253-278.

Harvey and Jaeger (1993), "Detrending, stylised facts and the business cycle", Journal of Applied Econometrics, 8, 231-247

Hodrick and Prescott (1980), "Postwar US business cycles: An empirical investigation", Discussion Paper 451, Carnegie Mellon University. Unpublished.

Holland and Scott (1998), "The determinants of UK business cycles", Economic Journal, 108, 1067-92

King and Plosser (1993), "Low frequency filtering and real business cycles", Journal of Economic Dynamics and Control, 17, 207-232

Nelson and Kang (1981), "Spurious periodicity in inappropriately detrended time series", Econometrica, 49, 741-751

1.4 Tools for analysis

Business cycle theorists face a problem - namely what does the business cycle component of the data look like. Observed macroeconomic data reflects many different components - an underlying trend, the business cycle component, seasonality, as well as purely random fluctuations - but business cycle theorists are interested in only one component. Effectively what is required is a filter, something which ignores everything other than the business cycle component.

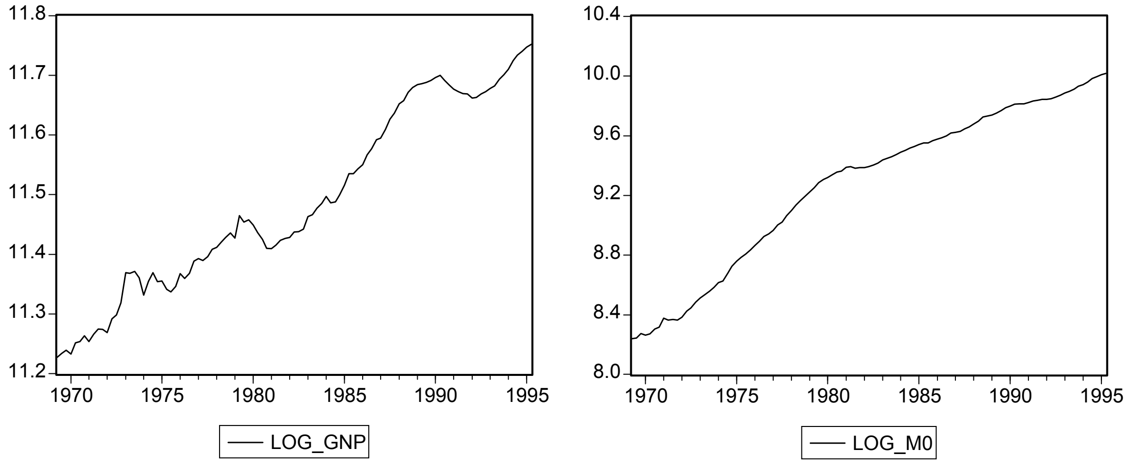


Figure 1.1: GNP and M0 in the UK

Figure 1.1 shows the logarithm of GNP and narrow money M0 for 1969q2 to 1994q4[1]. At present the figure is not very informative and it is difficult to see the interactions between real and nominal variables. To make things clearer there are an infinite number of different filters that can be used. Unfortunately because we can never know what the true business cycle component actually looks like we can never with complete certainty claim that one filter is better than another. Instead we have to choose a particular filter because it is plausible given certain beliefs about what a business cycle filter should do. Ideally, the cyclical component extracted from any dataset should not vary greatly between different filters. However, as we shall see this is not the case.

In the modern macroeconomics literature one filter has been used almost exclusively. This filter is based on an unpublished paper by Hodrick and Prescott (1980). The idea behind this filter is that the data consists of a trend and a business cycle. Hodrick and Prescott's starting point is that the trend must be a smooth time series - in other words, it does not make sense to think of a trend which fluctuates wildly

[1] This dataset was used in Ellison and Scott (JME, 2000). You can download the data as an Eviews file from the course homepage. It would be a very useful exercise for you to load this into Eviews and replicate the results seen in the lecture.

from quarter to quarter. However, the trend (which we will denote by τ_t) must follow the observed data (y_t) closely. They therefore infer the trend from the following minimisation problem.

$$\min_{\{\tau_t\}_{t=1}^{T}} \sum_{t=1}^{T}(y_t - \tau_t)^2 + \lambda \sum_{t=2}^{T-1} \left[(\tau_{t+1} - \tau_t) - (\tau_t - \tau_{t-1})\right]^2$$

The first part of the minimisation problem ensures that the trend component tracks the data fairly well. The constraint however prevents the change in the trend being too volatile. The larger the value of λ the smoother the changes in the growth of the trend have to be. Hodrick and Prescott suggest a value of $\lambda = 1600$ for quarterly data. Although, as pointed out many critics many critics, the choice of λ is likely to vary from variable to variable. Figure 1.2 shows UK GNP and Hodrick-Prescott trends for different smoothing parameters.

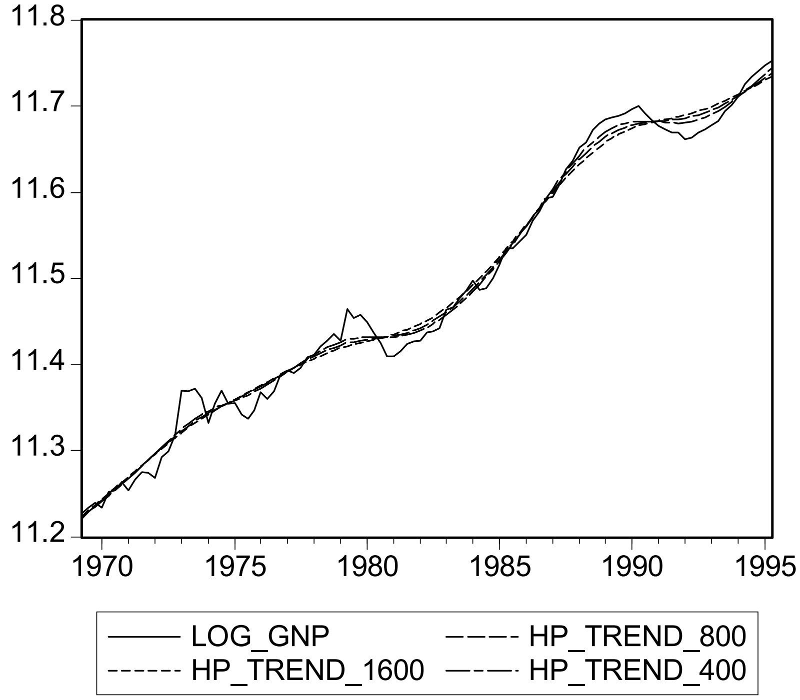


Figure 1.2: UK GNP and Hodrick-Prescott trends

Having solved this minimisation problem to arrive at an estimate of the trend, the cyclical component is defined as $y_t - \tau_t$. As shown in King and Plosser (1993) the Hodrick-Prescott filter has the ability to successfully detrend any series of an order of integration less than or equal to I(4). As mentioned above, the Hodrick-Prescott filter is now the industry standard approach towards detrending. However, there is widespread concern about its use.

(i) Business cycle facts are not invariant to the detrending filter used.

(ii) Other filters may be more optimal (see for instance Harvey and Jaeger (1993)). A little bit of thought will reveal that if variables have different stochastic properties then a different detrending filter should be applied. Therefore we cannot expect any one technique to be optimal for all variables. The crucial thing is whether one particular technique is a better approximation across a wider range of variables than another.

(iii) The Hodrick-Prescott filter may produce spurious cycles. This is the subject of Cogley and Nason (1994). A well-known result in the econometrics literature is the "spurious cycle" result of Nelson and Kang (1981). They show that if a linear time trend is fitted to a series which follows a random walk then the detrended data will display spurious cycles. In other words, if the researcher mistakenly thinks the trend is deterministic then the cycles derived will be misspecified. Cogley and Nason (1994) show that under incorrect assumptions about the stochastic behaviour of a variable, the HP filter will exaggerate the pattern of long term growth cycles at cyclical frequencies and depress the influence of cycles at other frequencies. The result is that the HP filter may exaggerate the importance of business cycles.

Even more striking they show that in the context of the Frisch-Slutsky paradigm the HP filter can be dramatically misleading. Observed stylised facts about the business cycle reflect three factors: (i) an impulse (ii) a propagation mechanism and (iii) the data being detrended by the HP filter and the certain statistics reported. Cogley and Nason (1994) show that for a typical macroeconomic model (ii) is unnecessary - merely assuming a process for the shock and applying the HP filter will be enough to generate business cycle patterns **even if they are not there in the model**. In other words, so called "stylised facts" are nothing more than artifacts. This is why some call the HP filter the Hocus Pocus filter - it simply creates business cycles from nothing. However, some words of caution are necessary here. The reason why the HP filter goes wrong is that the researcher makes the wrong assumption about the trend behaviour of a series and so applies the wrong filter. It is not so much the HP filter that produces the wrong result but the misspecification of the trend by the researcher. Using another detrending filter aside from the HP does not remove the risk of misspecification.

These are clearly serious criticisms of simple stylised facts calculated from the HP filter. Some efforts have been made to put this literature on a more secure statistical footing. However, it still remains the case that nearly every quatitative macroeconomic paper attempts to justify itself by using stylised facts constructed using the HP filter. Be warned.

1.5 Stylised facts

Figures 1.3 and 1.4 plot detrended real GNP alongside M0 and M4 (all variables are in logarithms) for the period 1969q3 - 1994q4. Figure 1.3 shows a strong positive relationship between GNP and M0, whilst Figure 1.4 has a less strong but nonetheless pronounced relationship between GNP and M4. Interpreting detrended data should always be done with caution, but visual inspection strongly suggests that M4 *lags* GDP by 18-24 months while Figure 1.3 suggests that M0 is more contemporaneous. There would appear to be little evidence that money *leads* output, consistent with the empirical findings in Holland and Scott (1998).

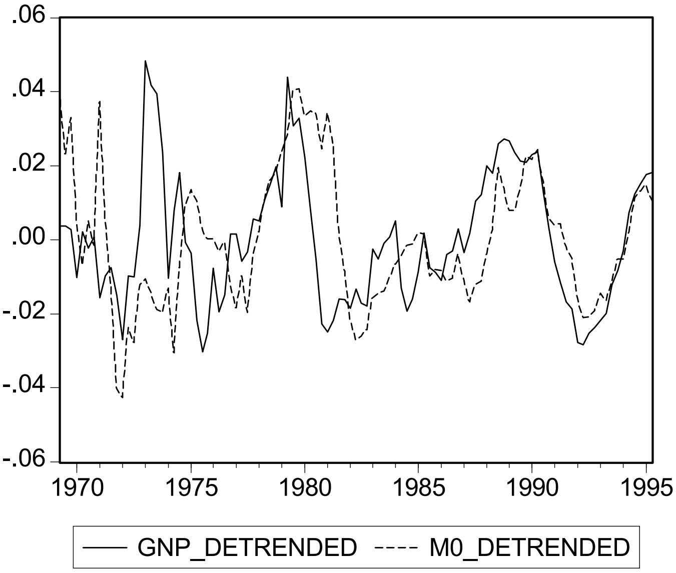


Figure 1.3: Detrended GNP and M0

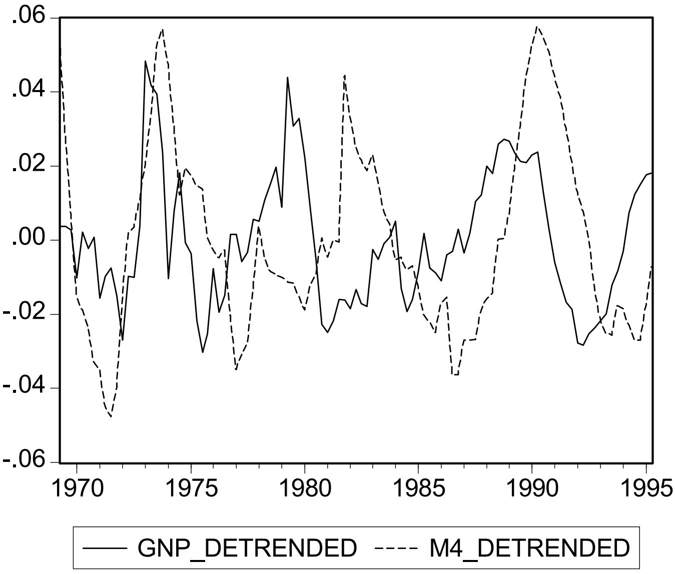


Figure 1.4: Detrended GNP and M4

Table 1.1 shows a more complete description of the relationships amongst real and nominal variables. We report the standard deviations (SD%), the first order autocorrelation coefficient on each variable (rho) and cross correlations at various leads and lags with real GNP. The positive correlations in Table 1.1 show M0 and M4 to be strongly procyclical. As Figures 1.3 and 1.4 suggest, M0 has a stronger contemporaneous correlation with output than M4. In addition, M4 is a more volatile series than output, whose standard deviation is similar to M0. The cross correlations of money and output show strong evidence that both

M0 and M4 lags output (the correlations in the right half of the table being larger). This contrasts with US evidence surveyed by Cooley and Hansen (1995) which suggests that money is strongly correlated with output but with a lead. As commented earlier, these results would need to be robustly established using different detrending filters before they could be considered as stylised facts of the UK business cycle. However, Figures 1.3 and 1.4 and Table 1.1 are strongly suggestive that variations in money are **not** an important cause of UK business cycles.

Table 1.1 suggests that the price level is countercyclical with negative contemporaneous and cross correlations with output. This is a similar result as for the US, and suggests the preponderance of supply shocks during this period, see Holland and Scott (1998) for support of this proposition using UK data. The contemporaneous relationship between output and inflation is very weak, becoming positive for leading output, negative for lagged. This differs from the US, where there is a strong positive correlation between cyclical output and inflation. This lack of a well defined correlation for the UK questions the existence of strong cyclical movements along a Phillips curve.

Nominal interest rates show a varied correlation with output. Contemporaneous short term nominal rates are positively related to output, whilst long term rates have virtually no correlation. The larger absolute values in the left half of the table suggest it is higher long term interest rates and not short term rates that are followed by lower cyclical output. The velocity of money is strongly counter-cyclical.

There is a positive correlation between M0 growth and output, also for hours, investment and consumption. There is a negative correlation between prices and money growth and a smaller negative correlation between money growth and nominal interest rates. The relatively small correlation coefficient between M0 growth and interest rates suggests that the liquidity effect in the UK is small.

1.6 Summary

We can tentatively conclude 5 stylised facts for the UK monetary economy.

1. Examining the cyclical component of UK data we find evidence which suggests that M4 *lags* GDP by around 2 years and M0 either lags or is contemporaneous with GDP.

2. We find little evidence that short term nominal interest rates are negatively correlated with future output but we do find such a correlation for long term rates.

3. There is a small negative correlation between contemporaneous money growth and short term interest rates.

4. The velocity of money is strongly counter-cyclical.

5. We find (a) little evidence in favour of a short run Phillips curve and (b) the price level is counter-cyclical, suggesting the post-war predominance of supply side shocks.

Variable	SD%	rho	VariableCross-correlation of output with:							Correlation with M0 growth
			t-3	t-2	t-1	t	t+1	t+2	t+3	
Output	1.66	0.81	0.434	0.624	0.810	1.000	0.810	0.624	0.434	0.233
Consumption	1.95	0.80	0.425	0.564	0.671	0.763	0.637	0.558	0.458	0.177
Investment	4.00	0.74	0.382	0.486	0.587	0.695	0.665	0.573	0.454	0.235
Labour input	1.11	0.57	0.297	0.371	0.479	0.605	0.441	0.239	0.054	0.157
Capital stock	0.23	0.97	-0.505	-0.410	-0.291	-0.146	0.024	0.186	0.333	-0.094
Prices										
GDP deflator	2.32	0.90	-0.439	-0.544	-0.605	-0.611	-0.524	-0.391	-0.216	-0.214
RPI	2.37	0.86	-0.560	-0.648	-0.667	-0.587	-0.420	-0.291	-0.120	-0.262
Inflation	1.03	0.36	-0.244	-0.240	-0.144	-0.019	0.171	0.266	0.360	-0.064
Money										
M0	1.84	0.84	-0.076	0.093	0.246	0.343	0.433	0.463	0.489	0.022
M4	2.73	0.94	-0.063	0.020	0.082	0.136	0.216	0.273	0.284	-0.016
Velocity	1.89	0.89	-0.486	-0.627	-0.717	-0.734	-0.668	-0.609	-0.500	-0.407
Interest rates										
10 year	0.96	0.75	-0.489	-0.333	-0.164	-0.016	0.114	0.227	0.294	-0.095
3 month	1.63	0.79	-0.330	-0.140	0.063	0.230	0.346	0.444	0.472	-0.123

Table 1.1: Cyclical behaviour of the UK economy 1965-1994

Chapter 2

Empirical monetary economics

2.1 Motivation

In the previous lecture we examined the relationship between money and output using simple bivariate measures. While these are easy to understand, they allow for only a limited view of the way in which variables may interact. In the next two lectures we turn our attention to more sophisticated techniques which can easily be extended to multivariate analysis. We also introduce more structure to our empirical analysis so later we can more clearly see the links between empirical and theoretical monetary economics.

2.2 Key readings

Walsh (1999) *Monetary Theory and Policy*, MIT Press, Section 1.3

Balke and Emery (1994), "Understanding the Price Puzzle", Federal Reserve Bank of Dallas Economic Review, 4th Quarter, 15-26

Sims (1972) "Money, Income and Causality", *American Economic Review*, 62, no. 4, 540-542

Sims (1980), "Macroeconomics and Reality", *Econometrica*, 48, 1-48

2.3 Related reading

Bernanke and Blinder (1992) "The Federal Funds Rate and the Channels of Monetary Transmission", *American Economic Review*, 82, 901-21

Blanchard and Quah (1989) "The Dynamic Effects of Aggregate Supply and Demand Disturbances", *American Economic Review*, 79, 4, 655-673

Canova and De Nicolo' (1999) "Monetary Disturbances Matter for Business Cycle Fluctuations in the G-7", Federal Reserve Board, International Finance Division working paper 660

Canova and Pina (1998) "Monetary Policy Misspecification in VAR Models", *CEPR Discussion Paper*

Ehrmann and Ellison (2001), "The changing nature of the response of US industry to monetary policy", *mimeo*, European Central Bank and University of Warwick. Unpublished.

Friedman (1961) "The Lag in the Effect of Monetary Policy", *Journal of Political Economy*, 69, 447-466.

Garratt and Scott (1998) "Money and Output: Post hoc, propter hoc or ad hoc?", *mimeo*, University of Cambridge. Unpublished

Huh (1993) "Causality and correlations of output and nominal variables in a real business cycle model", *Journal of Monetary Economics*, 32, 147-168

Sims (1992) "Interpreting the macroeconomic time series facts: The effects of monetary policy", *European Economic Review*, 36, 976-1011.

Sims (1996) "Macroeconomics and Methodology", *Journal of Economic Perspectives*, Vol. 10, No. 1, 105-20

Strongin (1995) "The Identification of Monetary Policy Disturbances: Explaining the Liquidity Puzzle", *Journal of Monetary Economics*, 35, 463-497

Uhlig (1999) What are the effects of monetary policy on output? : results from an agnostic identification procedure", CentER for Economic Research Discussion Papers No.28

2.4 Vector autoregressions (VAR)

Vector autoregression models have rapidly established themselves as the dominant research methodology in empirical monetary economics. The flexibility of the autoregressive formulation not only allows a statistical description of a wide range of real data sets but also provides a unifying framework in which to analyse alternative theories and hypotheses. Such models do not represent the truth in economics but are a useful tool for gaining insight into the interactions between different variables. In practice, the models can often provide a quite adequate description of the data.

The basic form of a p-th order VAR in money m_t and output y_t is shown in Equation (2.1). The order of the VAR refers to the number of lags in each equation. e_{1t} and e_{2t} are Gaussian residuals which may well be correlated. Equation (2.1) can be estimated using standard OLS regression techniques.

$$\begin{aligned} y_t &= y_0 + a_{11}m_{t-1} + \cdots + a_{p1}m_{t-p} + b_{11}y_{t-1} + \cdots + b_{p1}y_{t-p} + e_{1t} \\ m_t &= m_0 + a_{12}m_{t-1} + \cdots + a_{p2}m_{t-p} + b_{12}y_{t-1} + \cdots + b_{p2}y_{t-p} + e_{2t} \end{aligned} \tag{2.1}$$

In estimating a VAR we are faced with two difficulties. Firstly, we have to decide which variables should be included. Economic theory can often guide us to make sensible choices but as we will see the results are often sensitive to exactly which variables we choose. The second difficulty relates to deciding on the order of the VAR. If too few lags are included then we have an omitted variables problem: too many lags and estimates will be imprecise. The standard solution to this problem is to rely on information criteria. The Akaike Information Criterion (AIC) and Schwarz-Bayesian (BIC) are amongst the most-widely used. However, they should be treated carefully and often give conflicting advice. As a rule of thumb, it is important that the order of the VAR is at least sufficient to eliminate serial correlation in the residuals.

2.5 Granger causality tests

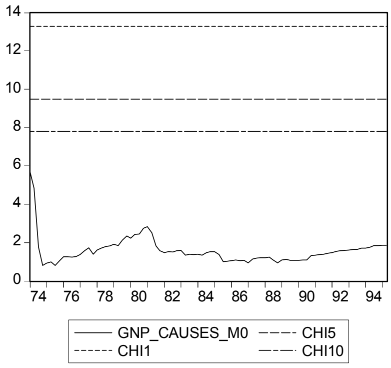
The VAR framework is also popular because of significant theoretical developments which enhance the basic model. The first of these is the Granger causality test introduced by Sims (1972). This enables us to say something about the causality links between two variables. It is always difficult to establish causality in econometrics so this is especially nowadays a fairly weak test. We say that money Granger-causes output if and only if lagged values of money have predictive power in forecasting output. In other words, if knowing past values of money enables us to predict output better then money Granger-causes output. In practice, a test of whether money Granger-causes output involves testing whether the coefficients on lagged money in the output equation in (2.1) are significant, i.e. we test

$$\text{equivalent} \begin{cases} H_0: \text{ Money does not Granger-cause output} \\ H_0: a_{11} = \cdots = a_{p1} = 0 \end{cases}$$

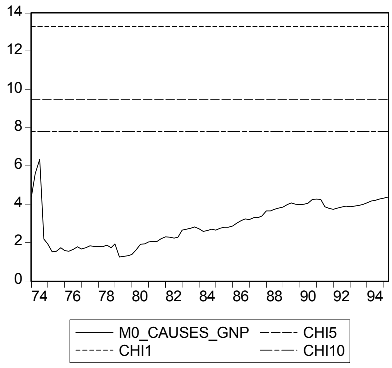
This can easily be tested with an F-test or a Wald test. Figure 2.1 shows recursive Granger-causality tests for the UK economy using the Ellison-Scott data. We use Δm_t and Δy_t and plot critical χ^2 values for the test at 1%, 5% and 10% levels respectively. The hypothesis is rejected if the test statistic exceeds the critical value. The results of this simple analysis do not suggest a great deal of Granger-causality between money and output. The only hypothesis that is rejected is for the non-causality of M4 by GNP in the early part of the sample. This is not surprising given that we saw M4 lagging GNP in the last lecture.

In a more robust analysis, Garratt and Scott (1998) perform Granger causality tests on a multivariate system of output, money, prices and short-term interest rates. They focus on three different sets of Granger causality tests: whether money predicts output, whether interest rates predict output and whether money and interest rates jointly predict output. The most striking result is the strong significance that M0 Granger-causes output, which we did not pick up in our simple bivariate analysis. By contrast, M4 only Granger-causes after 1994. For the US, the general result is that money does Granger-cause output, see Strongin (1995) and the references therein.

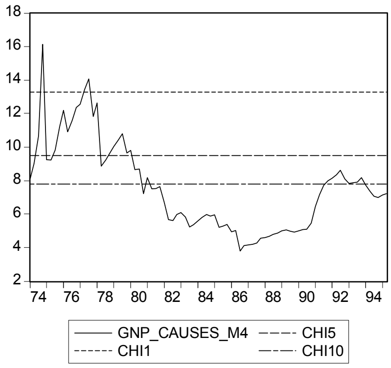
The usefulness of Granger causality tests has been questioned in an interesting paper by Huh (1993). He simulates a dynamic general equilibrium model in which output is exogenous with respect to money. Even in such a model with “reverse causation” from output to money, repeated applications of the Granger causality test to simulated data lead to the finding that money causes output more often than expected. This casts serious doubt on the validity of using Granger causality tests.



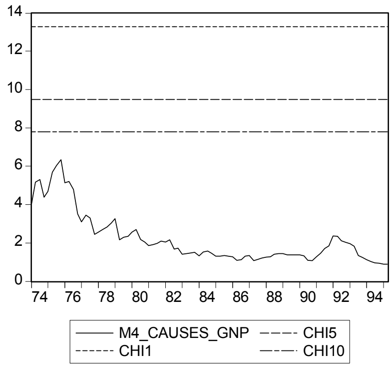
H_0 : GNP does not Granger-cause M0



H_0 : M0 does not Granger-cause GNP



H_0 : GNP does not Granger-cause M4



H_0 : M4 does not Granger-cause GNP

Figure 2.1: Granger causality tests for GNP and money in the UK economy

2.6 Structural VARs

To interpret the estimation results of an unrestricted VAR is still difficult. We are unable to say anything about how the economy reacts to different shocks. All we have are two Gaussian residuals which can be correlated. The general form of their variance-covariance matrix is given in (2.2).

$$\begin{pmatrix} e_{1t} \\ e_{2t} \end{pmatrix} \sim \mathbf{N} \left[\mathbf{0}; \begin{pmatrix} \sigma_1 & \sigma_{1,2} \\ \sigma_{1,2} & \sigma_2 \end{pmatrix} \right] \tag{2.2}$$

Suppose that theory suggests that really there are two fundamental shocks u_{1t} and u_{2t} driving the system: output shocks and monetary policy shocks. These shocks are assumed to have unit variance and to be uncorrelated at all leads and lags. The problem is that both these shocks may affect both the variables. Output shocks could affect output and money, whereas money shocks could in turn affect both output and money. In other words, there is the identification problem shown in Figure 2.2.

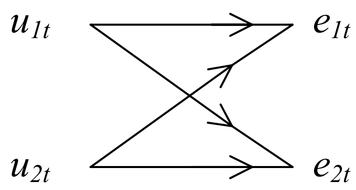


Figure 2.2: The identification problem

The idea of a structural VAR is to impose identifying restrictions on how the fundamental disturbances u_{1t} and u_{2t} affect the residuals e_{1t} and e_{2t}. The simplest idea and grandfather of all structural VAR models is by Sims (1980), with a nice application in Sims (1992). The idea is to order the disturbances in the model. In the bivariate system we would assume that output is causally prior to money. This means that output shocks have contemporaneous effects on both output and money but money only has contemporaneous effects on money itself. Sims justifies this by saying that monetary policy only affects output with a lag. Figure 2.3 shows the general idea. An identification scheme of this type is referred to as a recursive ordering or a Wold causal chain.

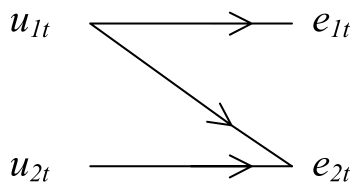


Figure 2.3: Bivariate identification by recursive ordering

Mathematically, the identification implies that there is a well-defined relationship between the fundamental disturbances and the residuals of the unrestricted VAR. Assuming linearity, equations (2.3) and (2.4) show that the residual in the output equation is only composed of output shocks but the residual in the money equation contains both output and monetary policy shocks.

$$e_{1t} = \theta_1 u_{1t} \tag{2.3}$$

$$e_{2t} = \theta_2 u_{1t} + \theta_3 u_{2t} \tag{2.4}$$

Since the fundamental disturbances have unit variance and are uncorrelated at all leads and lags the identity (2.5) must hold. This is obtained by looking at the variance-covariance of the residuals from the VAR.

$$\begin{pmatrix} \sigma_1 & \sigma_{1,2} \\ \sigma_{1,2} & \sigma_2 \end{pmatrix} \equiv \begin{pmatrix} \theta_1^2 & \theta_1\theta_2 \\ \theta_1\theta_2 & \theta_2^2 + \theta_3^2 \end{pmatrix} \tag{2.5}$$

This identity gives unique values for θ_1, θ_2 and θ_3 according to equations (2.6)-(2.8). It is then easy to recover the fundamental disturbances from the VAR residuals using equations (2.3) and (2.4).

$$\theta_1 = \sqrt{\sigma_1} \tag{2.6}$$

$$\theta_2 = \frac{\sigma_{1,2}}{\sqrt{\sigma_1}} \tag{2.7}$$

$$\theta_3 = \sqrt{\sigma_2 - \frac{\sigma_{1,2}^2}{\sigma_1}} \tag{2.8}$$

There is nothing unique about this causal ordering. Bernanke and Blinder (1992) take an opposite view and assume that the fundamental money shock affects both output and money whereas output shocks only affect output. The rational behind this is that the money supply is set by the central bank and output data is only available with a lag. Quite simply, money cannot react to output shocks because the information about output shocks is not available to the central bank. Not surprisingly, the results are somewhat sensitive to which ordering is used. Several other competing identification schemes exist, such as long-term restrictions (the recursive ordering essentially restricts the short-term behaviour of the economy) by Blanchard and Quah (1989) and sign restrictions by Uhlig (2000) and Canova and Nicolo' (2000). The recursive causal structure is the easiest to handle and can easily be extended to many dimensions. In the trivariate ordering in Figure 2.4 the first disturbance u_{1t} affects all three variables contemporaneously through e_{1t}, e_{2t} and e_{3t}. The second disturbance u_{2t} only affects contemporaneously the first two variables through e_{2t} and e_{3t}. The last disturbance u_{3t} only contemporaneously affects the last variable through e_{3t}. The multivariate system has a similar set of unique equations as the bivariate system (2.6)-(2.8). However, a neat mathematical trick involving a Choleski decomposition makes calculating the restrictions easier. For this reason, many papers talk about imposing structure through a Choleski decomposition. It is the same as imposing a Wold causal ordering.

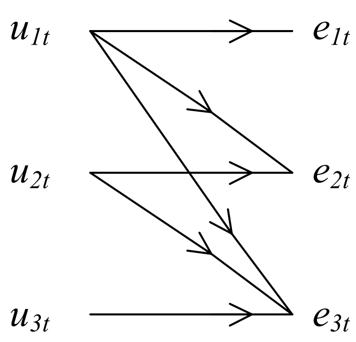


Figure 2.4: Trivariate identification by recursive ordering

Whilst traditionally Granger causality tests looked at the relationship between output and money, it has become increasingly common in structural VAR studies to use interest rates as the indicator of monetary policy. This partly reflects the stronger causal relationship between interest rates and the economy, but also a shift by central banks to using interest rates as the instrument of monetary policy. The rest of this section therefore studies systems in interest rates rather than money. As an example, we take monthly data from a study by Ellison and Ehrmann (2001), which looks at the relationship between capacity utilisation in US manufacturing and the federal funds rate - the key monetary policy interest rate in the US economy.[1] The causal ordering is such that interest rate shocks only have a delayed effect on capacity utilisation. Figure 2.5 shows the residuals of the unrestricted VAR and the shocks that are identified by the Wold causal ordering. In this simple bivariate model the residuals and shocks are somewhat similar, reflecting the low covariance between the two residuals. In models of higher dimension this may not be the case.

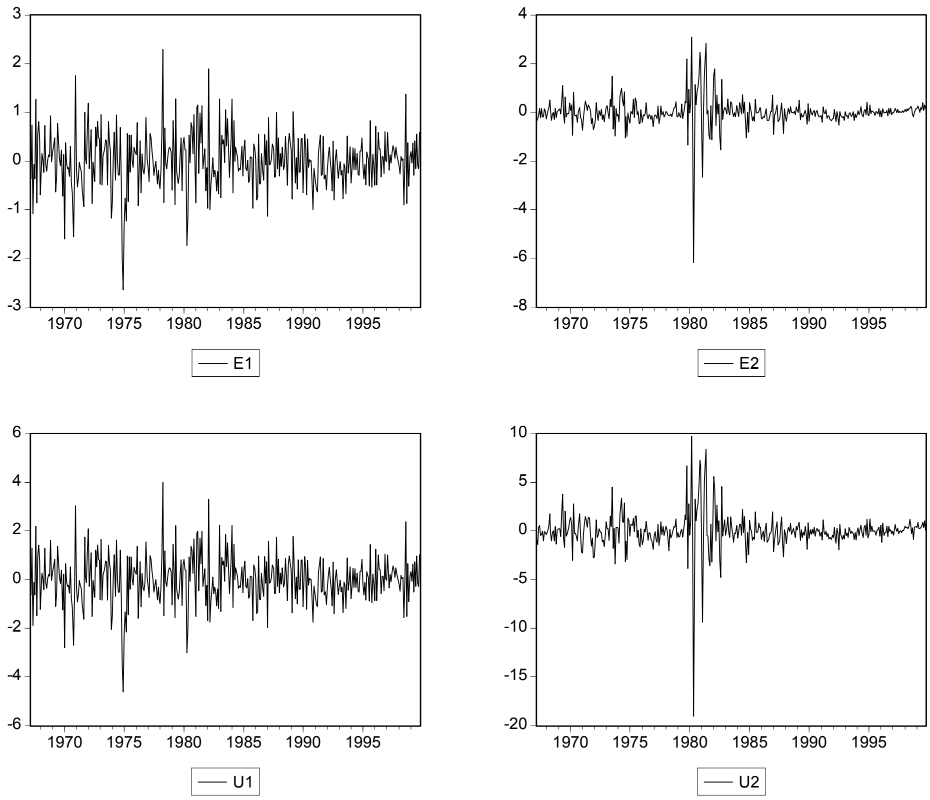


Figure 2.5: Residuals (e_{1t} and e_{2t}) and shocks (u_{1t} and u_{2t}) in the SVAR model

2.7 Impulse response functions

The structural VAR is a powerful framework in which to see how the variables in the system interact. We can use the technique of impulse response functions to show how the fundamental shocks we have identified feed through into the economy. As an example, suppose we are interested in the effect of interest

[1] The data is available for download on the course homepage.

rate shocks and ask what would be the effect of an interest rate shock in the absence of a simultaneous output shock. To see this we can set $u_{1t} = 0$ and $u_{2t} = 1$ and suppose the economy is in steady-state at time $t-1$, with $y_t = R_t = 0$. The size of the interest rate shock is normalised to one standard deviation to enable easy interpretation. The effects of this shock each period are given algebraically in Table 2.1. The first line shows the impact of the shock using the equations (2.3) and (2.4) - not surprisingly given our identification scheme the interest rate shock does not contemporaneously affect output. The remaining rows are obtained by feeding the impact effect through the autoregressive structure of the model.

$$
\begin{array}{ll}
t=0 & \begin{array}{l} y_t = 0 \\ R_t = \theta_3 \end{array} \\
t=1 & \begin{array}{l} y_t = a_{11}\theta_3 \\ R_t = a_{12}\theta_3 \end{array} \\
t=2 & \begin{array}{l} y_t = a_{11}a_{12}\theta_3 + a_{21}\theta_3 + b_{11}a_{11}\theta_3 \\ R_t = a_{12}a_{12}\theta_3 + a_{22}\theta_3 + b_{12}a_{11}\theta_3 \end{array} \\
\vdots &
\end{array}
$$

Table 2.1: Response to an interest rate shock

Calculating impulse response functions numerically is much easier than analytically. Eviews and most other econometric packages have built-in routines to estimate and identify systems quickly. Impulse response functions estimated from the Ehrmann-Ellison (2001) data are shown in Figure 2.6. The dashed lines surrounding each impulse response are error bands, giving an idea of the significance of the response. They can be obtained either analytically or by standard bootstrapping techniques.

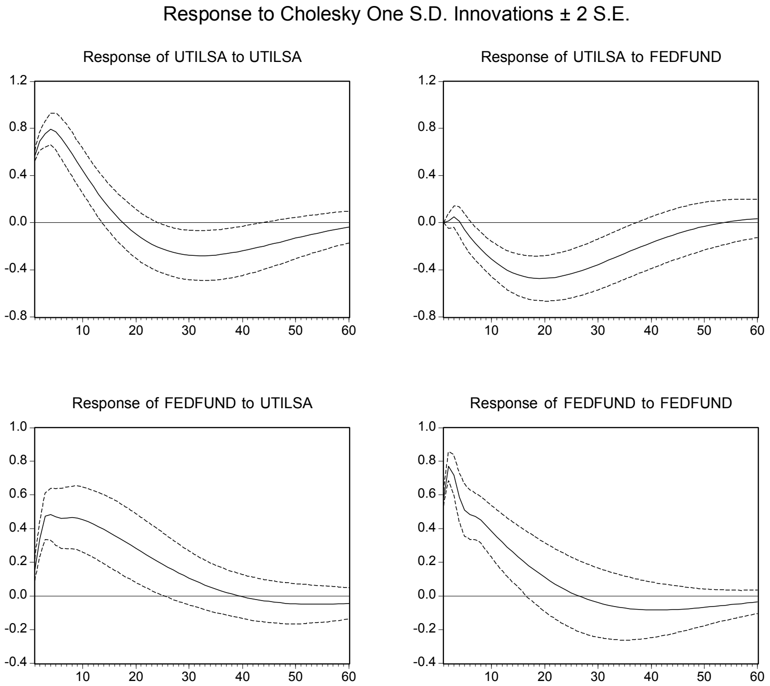


Figure 2.6: Impulse response functions with the Ehrmann-Ellison (2001) data

Now, at last, we can start to see some monetary economics. The right column in Figure 2.6 shows the response of the economy to an interest rate shock. In the lower right-hand panel, interest rates obviously rise but also stay above baseline for a long period, reflecting a high degree of persistence in interest rates. The upper right-hand panel shows how output (proxied by capacity utilisation) reacts. We see that a positive interest rate shock depresses output. However, the effect is not immediate and it takes about 2 years for the output effect to hit a maximum. This is a robust result across many countries and data sets. It was originally referred to by Friedman (1961) as due to the "long and variable" lags of monetary policy.

The left-hand column in Figure 2.6 shows the effect of an output shock on the economy. In the upper left-hand panel the output shock has a persistent effect on output - probably the easiest way to think of this is as a demand shock. The lower left-hand panel shows that interest rates rise in response already at time $t = 0$. This is the response of the central bank: After a demand shock they increase interest rates to stabilise the economy. Again there is an effect on interest rates for a long period.

2.8 Problems with SVARs

Structural VARs are not without their problems and critics. In making the identifying assumptions, we are imposing *a priori* beliefs about how the economy behaves.[2] These cannot be tested - Granger causality

[2] A paper by Canova and Pina (1998) shows that stochastic dynamic general equilibrium models typically do not imply a recursive structure on the model. In these models, both interest rate and output shocks impinge upon both variables

tests do not help because they say nothing about contemporaneous causality - so many econometricians consider SVARs as more art than science. An insightful debate into this was published in a volume of the *Journal of Economic Perspectives* in 1996. Structural VARs do though form a useful baseline against which to assess the theoretical models in the next lectures.

One way to assess the robustness of the results is to see whether the impulse responses match our economic intuition. In the previous example, we found that a positive interest shock caused a contraction in output, which accords well with our expectations from economic theory. This is an example of an over-identifying assumption - we expect high interest rates to cause recessions - which can be tested to check our results. The bivariate case creates no problems but in higher order systems there are many more things that can go wrong. Consider a trivariate system in output, prices and the interest rate. Figure 2.7 concentrates on one of the estimated impulse responses: the response of prices to a monetary policy shock.

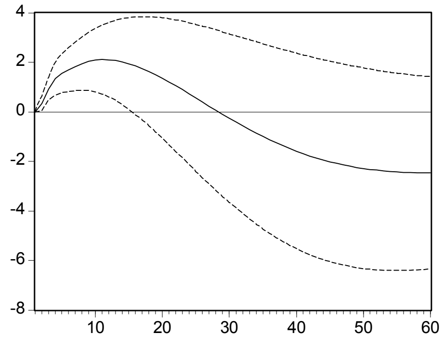


Figure 2.7: The estimated response of prices to an interest rate shock - the price puzzle

This picture is worrying. The interest rate shock actually leads to an increase in the price level for 2-3 years. This again is a very robust result across many countries and datasets and is referred to in the literature as the *price puzzle*. Whilst this is not impossible, it does not fit well with established wisdom concerning the deflationary effect of positive interest rate shocks. What is wrong then? One answer would be to conclude that our theory is wrong and set out to write a theory which predicts the sort of response we estimate. More likely though, there is something wrong in our econometric approach. The shock we have identified as an interest rate shock does not really look like an interest rate shock at all and this seems to be the heart of the problem.

One explanation of the price puzzle that has become standard in the literature is that the shock identified as an interest rate shock is actually a shock to inflation expectations. If people expect higher prices in the future then prices are already driven up today. Simultaneously, the central bank increases interest rates to control future price information. In other words, the expectations shock increases both prices and interest rates together. The solution to this is to control for the inflation shocks in the estimation. The standard way to do this is by including an index of world commodity prices as an exogenous variable in the model. This acts as a proxy for changes in expectations. Balke and Emery (1994) have a clear explanation of

simultaneously.

the price puzzle problem and how to solve it. Figure 2.8 shows the response of prices to an interest rate shock in the Ehrmann-Ellison (2001) data once world commodity prices have been included as a proxy for expectations shocks. Now the price puzzle is much briefer and less dramatic. Including commodity prices has not completely eradicated the price puzzle but has certainly reduced it significantly.

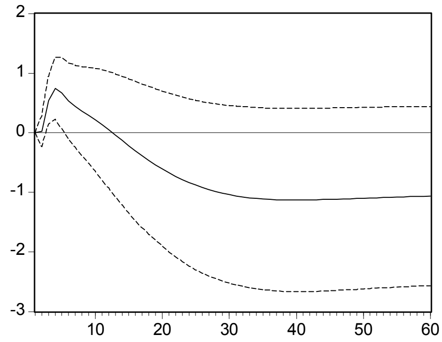


Figure 2.8: The estimated response of prices to an interest rate shock after controlling for commodity prices

This concludes our brief review of empirical monetary economics. We will return to these results in various parts of the remainder of the course to assess the validity of the theoretical models we develop. There are thousands of structural VAR studies already done but this is still very much an active area of research. Contributions are still very welcome.

Chapter 3

Lucas islands model

3.1 Motivation

The papers by Lucas (1972), (1973) and (1975) sought to explain the short term procyclicality of output and inflation (the "Phillips curve") while still maintaining the neoclassical assumption that in the long run money is a veil and does not affect output. The impulse in his model is money supply shocks and the propagation mechanism consists of (a) competitive markets (b) Rational Expectations (c) imperfect information, in the sense that agents have better information regarding the market they operate in than the economy as a whole.

These papers were enormously influential and generated a substantial literature investigating the relative impact of expected and unexpected money. However, the more long-lasting impact of this paper has been methodological. Firstly, it illustrates the power of Rational Expectations as a modelling technique. Secondly, it shows how to use simple general equilibrium models to arrive at structural models of economic fluctuations. Therefore, the best way to view these lectures is less as *the* model of the business cycle and more as a useful way of illustrating modelling techniques which are now standard in macroeconomics.

3.2 Key reading

Lucas (1973) "Some International Evidence on Output-Inflation Trade-offs", *American Economic Review*, 63, 326-334

3.3 Related reading

Lucas (1972) "Expectations and the neutrality of money", *Journal of Economic Theory*, 4, 103-24

Lucas (1975) "An Equilibrium Model of the Business Cycle", *Journal of Political Economy*, 83, 1113-1144

Barro (1978) "Unanticipated money growth, output and the price level", *Journal of Political Economy*, 86, 549-80

3.4 The model

There exist N islands. On each island is a producer who charges price $p_t(z)$, where z denotes a particular island. We shall denote the aggregate price as p_t, which is simply the average of $p(z)$ across all Z islands.

The main focus in the paper is on the supply side where the following "Lucas supply" function is assumed

$$y_t(z) = \gamma(p_t(z) - p_t) \tag{3.1}$$

so that a producer increases his output when his own price is greater than the aggregate price in the economy. Thus output is given by a standard upward sloping supply schedule. $y(z)$ should be interpreted as deviations in output from trend, so that when $y(z) = 0$ output is equal to its trend value. The innovation in this model is to assume imperfect information so that a producer on an island knows at time t his own price $p(z)$ but does not know the economy-wide price level p. Instead, they have to form a guess of p based on the information at their disposal. Let $I_t(z)$ denote the information available at time t to the producer in market z. Then $E(p_t\,|I_t(z)\,)$ denotes the expectation of the aggregate price p_t given the information available to the producer in island z. Because of imperfect competition, we have the incomplete information supply curve

$$y_t(z) = \gamma(p_t(z) - E(p_t\,|I_t(z)\,)) \tag{3.2}$$

However, the crucial question is how do agents form $E(p_t\,|I_t(z)\,)$? Without an expression for this guess of aggregate prices, we can do very little with our model.

Lucas' trick is to use Rational Expectations. There are two different but related interpretations of Rational Expectations. The first is a statistical one and implies that, when agents have to make a forecast, errors are unpredictable. In other words, agents use all the information at their disposal. The second interpretation is more economic and is due to John Muth. According to this definition, Rational Expectations is when agents use the economic model to form their price predictions.

Assuming rational expectations, it must be the case that $p_t = E(p_t\,|I_{t-1}\,) + \epsilon_t$ where ϵ is a forecast error which is on average zero and which cannot be predicted from I_{t-1}. We shall denote the variance of ϵ ($E(\epsilon^2)$) as σ^2.

Lucas also assumes that the price in each island, $p(z)$, differs only randomly from the aggregate price level, p. In other words, $p_t(z) = p_t + z_t$ where z is on average zero and has a variance equal to τ^2.

If the producer has perfect information about the aggregate price level then $y(z)$ would respond only to z, the relative price shock. However, due to imperfect information agents only observe the gap between $p(z)$ and their expectation of the aggregate price level, that is they observe the composite error $z + \epsilon$. The producers problem is to decide how much of this composite error is due to mistakes in forecasting the aggregate price level (ϵ) and how much is the relative price shock (z) and to only alter output in response to the latter.

How does the producer decide how much of the composite shock is due to ϵ and how much is due to z? Technically this problem is called "signal extraction". The answer is to look at historical data. Our model only assumes that agents do not know the current value of ϵ, but they do observe historical data on z and ϵ. Running a regression of z on $(z+\epsilon)$ will provide a guess of what proportion of $(z+\epsilon)$ is due to z. Using standard OLS formulae we have

$$\theta = \frac{\tau^2}{\sigma^2+\tau^2}, \text{ where } z_t = \theta(z_t+\epsilon_t)+u_t \tag{3.3}$$

Therefore an agents best guess of z is $\theta(z+\epsilon)$. By definition $p(z) = p+z$ so that an agent's best guess of the aggregate price level given that they observe $p_t(z)$ is

$$\begin{aligned} E(p_t\,|I_{t-1}(z), p_t(z)\,) &= p_t(z) - E(z_t\,|I_{t-1}(z), p_t(z)\,) \\ &= p_t(z) - \theta(p_t(z) - E(p_t\,|I_{t-1}\,)) \\ &= (1-\theta)p_t(z) + \theta E(p_t\,|I_{t-1}\,) \end{aligned} \tag{3.4}$$

Equation (3.4) gives an expression for $E(p_t\,|I_t(z)\,)$ which we can insert into (3.2) to get a supply curve. Aggregating across all islands gives an aggregate supply curve.

$$y_t = \gamma\theta(p_t - E(p_t\,|I_{t-1}\,)) \tag{3.5}$$

so that output only responds to unexpected aggregate price shocks. Note that by adding and subtracting p_{t-1} to the right-hand of (3.5) and re-arranging we have

$$(p_t - p_{t-1}) = \frac{1}{\gamma\theta}y_t + (E(p_t\,|I_{t-1}\,) - p_{t-1}) \tag{3.6}$$

so that inflation depends positively on output and expected inflation (y should be interpreted throughout this note as deviation from trend). Assuming a negative relationship between output and unemployment, (3.6) defines a Phillips curve. However, there only exists a short term trade-off between output and inflation. Due to the assumption of Rational Expectations, on average forecasts of inflation equal inflation so that for (3.6) to hold $y = 0$. Output only differs from its trend value due to unexpected forecast errors.

So far we have only discussed the supply side of the model, we now turn to the demand side. We assume that demand in each island, $y^d(z)$, depends on nominal money and the price level in each island, e.g. $y^d = m(z) - p(z)$ and that the money supply in each island equals the aggregate money supply plus a random error, e.g. $m(z) = m + \eta(z)$ where $\eta(z)$ has a zero mean and variance δ^2. Finally, we assume that the aggregate money supply evolves according to

$$m_t = m_{t-1} + \mu + \zeta_t \tag{3.7}$$

where μ is the expected growth in the money supply and ζ is the unexpected part of money growth with a zero mean and a variance equal to λ^2.

We now have expressions for demand and supply in each island. Equilibrium requires that supply equals demand in each market so

$$\gamma\theta(p_t(z) - E(p_t\,|I_{t-1})) = m_t(z) - p_t(z) \tag{3.8}$$

It is here that we use our assumption of Rational Expectations. Our supply functions all require $E(p_t\,|I_{t-1})$. Different assumptions regarding this term will lead to very different implications. According to Rational Expectations, we should use the model's predictions for p_t as the best possible forecast. To see how this works aggregate (3.8) over all markets to give

$$\gamma\theta(p_t - E(p_t\,|I_{t-1})) = m_t - p_t \tag{3.9}$$

The left-hand side of (3.9) depends on the forecast error made by agents in predicting the economy-wide price level. However, by the definition of Rational Expectations these forecasts must be unpredictable at $t-1$. If this is the case then the expected value of the left and right hand side of (3.9) at $t-1$ must equal zero. From this we can therefore deduce that $E(p_t\,|I_{t-1}) = E(m_t\,|I_{t-1})$ so that the forecast of next period's aggregate price is equal to the forecast of next period's money supply, which from (3.7) is $m_{t-1}+\mu$ (unanticipated money is forecast to be zero or else it wouldn't be unanticipated). Therefore we have used the implication of the model (3.8) to derive agent's expectations.

Using this price expectations term, we can now return to our equilibrium condition (3.8) and solve our model. Replacing $E(p_t\,|I_{t-1})$ with $m_{t-1}+\mu$ and rearranging we have

$$p_t(z) = m_{t-1} + \mu + \frac{\zeta_t + \eta_t(z)}{1+\theta\gamma} \tag{3.10}$$

$$p_t = m_{t-1} + \mu + \frac{\zeta_t}{1+\theta\gamma} \tag{3.11}$$

for the market clearing price in each market and for the economy as a whole respectively. Notice our assumption that agent's price expectations are rational. We assumed that $E(p_t\,|I_{t-1}) = m_{t-1}+\mu$ and having made this assumption we have derived (3.11) as our expression for prices and (3.11) implies $E(p_t\,|I_{t-1}) = m_{t-1}+\mu$. Therefore agent's expectations are consistent with the model, this is the essence of Rational Expectations.

We earlier assumed the existence of a forecasting error ϵ_t for the aggregate price level and a relative price shock for each island z_t. Using (3.10) and (3.11) we can now deduce that

$$\epsilon_t = \tfrac{\zeta_t}{1+\theta\gamma} \quad z_t = \tfrac{\eta_t(z)}{1+\theta\gamma} \tag{3.12}$$

and we can also write the crucial signal extraction parameter θ as $\delta^2/(\lambda^2+\delta^2)$, in other words it depends on the relative variability of aggregate and island specific money supply.

Inserting our price expectations (from (3.10) and (3.11)) into the aggregate and island specific supply curve gives

$$y_t(z) = \tfrac{\gamma\theta}{1+\gamma\theta}[\zeta_t + \eta_t(z)] \quad y_t = \tfrac{\gamma\theta}{1+\gamma\theta}\zeta_t \tag{3.13}$$

Equation (3.13) gives expressions for the equilibrium sequence of output in each island and in the economy as a whole. As is the case in Rational Expectations models, these equations are highly structural. Every parameter can be interpreted in terms of the basic parameters of the model.

3.5 Implications

i) anticipated money does not influence output, i.e. y_t is independent of μ. From (3.13) the only thing that matters is shocks to the money supply. Agents in each island will observe an increase in $p(z)$ but will be unsure whether this is a relative price shock or an inflation shock. Even though they respond rationally, they are partly fooled into increasing output in response to aggregate price shocks. However, they only respond to unanticipated shocks and so this cannot be exploited systematically by the monetary authorities.

ii) unanticipated money can influence the current value of y_t but not the average level of output. From (3.13) the average deviation of output from trend is 0. Therefore monetary policy cannot shift the so-called natural rate of output.

3.6 Empirical evidence

There exists some support for the key idea, namely that output responds more to unexpected money than it does to expected money although there is some evidence that in the short run both types of money influence output. The idea that money does not affect the long run level of output is widely believed.

However, even though there is this empirical evidence the Lucas model is not seen as a good model of the cycle. The key equation is the expression for aggregate output in (3.13). There are three drawbacks for this equation as a model of the business cycle:

i) ϵ_t is white noise whereas output fluctuations are serially correlated. In other words, the model does not generate any persistence in output. It has as an impulse white noise shocks and produces as output white noise fluctuations, This is a fairly standard failing of neoclassical models. Neoclassical propagation mechanisms do not appear to increase by much the persistence of shocks assumed as impulses to the model. Lucas (1975) tries to overcome this problem by introducing capital but with limited success.

ii) $\gamma\theta/(1+\gamma\theta)$ is less than one. In other words, the model not only does not increase the persistence of shocks assumed as the impulse to the model but it also serves to dampen down their effect. Money supply shocks influence output by less than one for one.

iii) ϵ is very small. Monthly innovations to the money supply are not very large and certainly not large enough given (ii) to generate observed business cycle movements. Further, given the welfare costs of business cycles it is relatively costless for the government every period to notify all agents what they have done with the aggregate money supply. In other words, the imperfect information could be easily rectified (although this raises important issues of credibility).

To see the failings of the model more clearly consider the following table taken from Cooley and Hansen "Money and the Business Cycle" in Cooley (ed.) "Frontiers of Business Cycle Research".

	SD%	t-4	t-3	t-2	t-1	t	t+1	t+2	t+3	t+4
U.S. data	1.72	0.16	0.38	0.63	0.85	1.00	0.85	0.63	0.38	0.16
Simulations	0.30	-0.06	-0.06	-0.08	-0.09	1.00	-0.09	-0.08	-0.06	-0.06

Tables like this are now a standard way of examining the properties of economic models. The first row shows the information on the cyclical part of US data. The first column quotes the standard deviation of output growth. The second row shows the results of simulating a version of Lucas' model which has been "calibrated" using US data. The first column shows that the Lucas model fails to generate anything like the right amount of volatility in output fluctuations, as suggested by (ii) and (iii). In other words, the combination of monetary shocks and imperfect information cannot account for business cycle volatility. the remaining columns quote the autocorrelation coefficients of output at different lags. In other words, the $t-2$ column shows how the cyclical component of output is correlated with its own value two periods ago. The value of 0.63 for US data shows that there is a strong correlation. This reveals that cyclical output is positively serially correlated, that is that output growth tends to be high for several consecutive periods (booms) and then low for several consecutive periods (recessions). However, looking at the results from the simulated Lucas economy we see no sign of such serial correlation. Cyclical output seems uncorrelated with either its recent past or recent future. In other words, the Lucas island model does not generate business cycles of booms and recessions but occasional random high and low values.

Chapter 4

Cash in advance model

4.1 Motivation

In this lecture we will look at ways of introducing money into a neoclassical model and how these methods can be developed in an effort to try and explain certain facts. As in previous lectures, we shall find that while we can develop models to improve our understanding of the business cycle we still remain some distance from a reliable model. While we introduce money in ways which break the neoclassical dichotomy, the models we look at are still partly neoclassical in nature. They are often referred to as monetary RBC models because they essentially keep the same structure of the real economy as RBC models but superimpose a monetary sector.

4.2 Key readings

Once again the origins of the modern literature rest with Lucas (1982) "Interest Rates and Currency Prices in a Two Country Model", *Journal of Monetary Economics*, and Lucas and Stokey (1987) "Money and Interest in a Cash in Advance Economy", *Econometrica*. However, I would recommend instead Svensson (1985) "Money and Asset Prices in a Cash in Advance Economy", *Journal of Political Economy*. Less theoretical papers (which are also easier to read) which focus on the business cycle implications of cash in advance are Cooley and Hansen (1989) "The inflation tax in a real business cycle model", *American Economic Review* and by the same authors "Money and the Business cycle" in Cooley (ed) "Frontiers of Business Cycle Research".

4.3 Introduction

The first problem with any neoclassical general equilibrium approach to business cycles when it comes to modelling monetary phenomena is how to explain why consumers need to hold money. Without a justification of why there is a demand for money it is obviously impossible to model the impact that variations in money supply will have on the economy. There are three broad approaches:

(i) Money in the Utility Function - if utility depends upon real money balances then money is like any other good and will be demanded by consumers. However, most people are reluctant to start with this assumption as it is rather *ad hoc*. It is a far better modelling strategy to try and point to a reason why money is held by consumers other than it is a direct source of utility. However, that said it is well known that under certain conditions there exists an equivalence between putting money in the utility function or specifying a transactions technology which involves money.

(ii) Cash in Advance Models - this is the route which has been most thoroughly explored in the literature. the assumption here is that before a consumer can buy goods they must pay for them in cash. Therefore money is demanded because it is the only means of purchasing some goods.

(iii) Transactions Cost (Shopping Time Technology) - in these models consumers have a choice (unlike in (ii)). They can obtain goods on credit or barter or they can purchase goods with cash. However, purchasing goods consumes resources and the more cash that an individual holds the lower these shopping costs are (e.g. they can avoid very costly bartering processes). By holding money, consumers lose any interest they would otherwise have gained on their savings but they economise on their transactions costs.

4.4 Cash in advance models

In what follows we shall focus entirely on cash in advance models. The basic cash in advance model is due to Lucas. Every period a consumer has to choose (a) their consumption (denoted c) (b) their money balances (denoted m) and (c) their savings (denoted a, assets). However, all consumption goods have to be paid for by cash so there is a constraint the consumer faces, $P_t c_t \leq m_t$. Assets deposited in the bank earn an interest rate $R > 0$ but no interest is earned on assets held in the form of money. Instead, money earns a rate of return equal to P_{t-1}/P_t, so if there is inflation money earns a negative return (it loses value). Consumers choose their consumption, assets and money balances once they observe the state of the world (i.e. after seeing what today's money supply growth is, what the value of the current productivity shock is, etc.) Because consumers earn interest on deposits but not on money they will always prefer to keep assets on deposit. Therefore they will hold only just enough cash to finance their consumption, e.g. $P_t c_t = m_t$. This has a rather unfortunate consequence that the velocity of money is constant. The velocity of money (V) is defined by the identity $MV = PY$, where M is the money supply, P is the price level and Y is the volume of transactions in the economy. Assuming no capital, the volume of transactions in this economy is just c, and because $m = Pc$ it must be that the velocity of money is always equal to 1. In reality, the velocity of money shows considerable variation and depends in particular on the interest rate. These are features which the basic cash in advance (CIA) model cannot account for.

Svensson (1985) proposes a simple amendment to Lucas' basic model. Like Lucas' article, Svensson's main concern is how to price assets when you have a cash in advance constraint. Svensson assumes that consumers have to choose how much cash to hold before they know the current state of the world (i.e. they are ignorant of the current money supply or productivity shock). As a result of this uncertainty the velocity of money is no longer constant. Agents will usually choose to hold $m > Pc$ for precautionary reasons. In a very good state of the world, agents know they would like their consumption to be high and they can only achieve this if they have high money balances. Therefore, agents tend to hold more money than they otherwise would need as a precaution in case they find themselves wanting to consume large amounts in a surprisingly good state of the world. The greater the uncertainty facing the consumer (e.g. the higher the probability of wanting to spend a lot on average) the larger these precautionary balances. However, the higher is the interest rate the lower the level of precautionary balances held by the consumer. Consumers have to trade the benefits of higher money balances (increased insurance against a good state of the world) against the costs (loss of interest). As a result the velocity of money becomes time-varying and depends on the interest rate.

4.5 Cash-credit models

Another version of the CIA model is the so called cash-credit model of Lucas and Stokey (1987). In this model agents gain utility from two goods, c_1 and c_2, where c_1 can only be purchased using cash but c_2 can be purchased on credit. The timing of the model is as follows. Agents observe the state of the world, decide on c_1 and c_2 and m, they then go and purchase cash goods paying for them with their money balances and also purchase credit goods, and then at the end of the period all credit bills are settled. This is another way of making the velocity of money variable. In this model, agents get utility from two goods, but on one good they have to pay cash and so lose R on any assets held in the form of cash. Therefore, when the interest rate is high they will tend to lower c_1 and increase c_2 to compensate, because they consume less of the cash good they also hold fewer money. Therefore the velocity of money $((c_1 + c_2)/m)$ varies positively with the interest rate - the higher the interest rate, the lower are money balances and the harder money has to work.

4.6 Simulation evidence

How well do these cash in advance models work in generating plausible patterns for money holdings in simulated economies? This is an issue examined in Hodrick, Kocherlakota and D. Lucas (1992) "The variability of velocity in cash in advance models", *Journal of Political Economy*. The authors choose a variety of values for certain key parameters, i.e. variability of money supply growth, persistence of money supply growth, risk aversion etc. They find that if they use the basic cash in advance model, whether it be that of Lucas or Svensson, they cannot generate under any plausible scenario enough variation in the velocity of money. For instance, over the last 100 years the velocity of money has had a standard deviation of around 4.5% but with the basic cash in advance model they cannot generate a standard deviation any greater than 0.09%. However, the cash-credit CIA model can generate more plausible numbers for the variability of velocity, with the numbers ranging between 0.6 % and 5.1%. However, to generate these more plausible numbers the authors have to assume very high levels of risk aversion. As a consequence, while they can explain observed volatility in the velocity of money they are unable to account for several other features of the data, such as the low level of actual interest rates. With the levels of risk aversion required to explain the velocity variability they need to have very large interest rates (20%) to explain consumption growth of around 2%, whereas in the data real interest rates tend to be around 3%. This is exactly the same type of failure others have identified as the equity premium puzzle. To account for one feature in the data we need to assume high risk aversion but in doing so we cannot explain why the risk-free rate of interest is so low. The core problem with these CIA models is that if interest rates are positive (i.e. consumers want to minimise cash holdings) and if consumption (or more precisely the marginal utility of consumption) is not very volatile (so that agents hold lower precautionary balances as they are less likely to have the need for very high levels of consumption) then because precautionary balances are very low, money is nearly equal to consumption of cash goods and so the velocity of money is low. Only by increasing risk aversion so that agents are concerned about even low levels of variability in consumption can you rescue the model. Not surprisingly, because the CIA models suffer from the same problems that generate the equity premium puzzle these basic CIA models are also not very successful in generating plausible asset price and interest rate data, an issue examined in Giovannini and Labadie (1991) "Asset Prices and Interest Rates in Cash in Advance Models", *Journal of Political Economy.*

4.7 Business cycle implications

The above suggests that CIA models have difficulty in explaining certain facts about nominal variables, but how do they fare in explaining business cycle fluctuations? This is investigated in Cooley and Hansen (1989) and also in their chapter in the volume edited by Cooley "Frontiers of Business Cycle Research". Before examining the outcome from simulations using these models it is important to analyse the propagation mechanism at work in the CIA model.

CIA models do not satisfy the conditions of the neoclassical dichotomy and so nominal variables such as money can have an impact on real variables such as consumption. The means by which this happens in a CIA model is through the "inflation tax". When inflation is high, the real value of money declines sharply so agents will hold less money. However, if cash is required to finance consumption purchases then high inflation will cause lower consumption and so a nominal variables (inflation/money supply growth) will affect a real variable (consumption).[1] Effectively, inflation acts as a tax on goods which require money to be purchased and as a subsidy on credit goods which do not require cash. Notice that what matters here is anticipated inflation and not unanticipated inflation. When agents choose their money holdings they do so on the basis of their expectations of inflation (which are assumed rational) and so unexpected inflation does not alter their decision.

Table 4.1 shows the results of simulating a CIA cash-credit model which has been calibrated for the US economy. An important feature of this simulation is that money growth is positively serially correlated, i.e. money growth tends to be above (below) average for several consecutive periods. Without this assumption CIA models tend to perform very poorly.

[1] Inflation acts as a distorting tax in this model. As a consequence, the fundamental welfare theorems do not hold and so the problem can no longer be analysed as the social planner's problem.

Variable	SD%	VariableCross-correlation of output with:							Correlation with M0 growth
		t-3	t-2	t-1	t	t+1	t+2	t+3	
Output	1.69	0.240	0.444	0.0692	1.000	0.692	0.444	0.240	-0.01
Consumption	0.53	0.397	0.488	0.582	0.676	0.378	0.153	-0.013	-0.60
Investment	5.90	0.169	0.381	0.643	0.975	0.699	0.473	0.282	0.16
Hours	0.35	0.145	0.362	0.637	0.987	0.715	0.492	0.305	-0.03
Prices	1.88	-0.135	-0.161	-0.190	-0.218	-0.126	-0.064	-0.021	0.43
Inflation	1.23	0.040	0.045	0.042	-0.138	-0.093	-0.065	-0.049	0.92
Nominal interest rate	0.58	0.010	0.011	-0.002	-0.014	0.003	0.011	0.007	0.72
ΔMoney	0.87	0.001	0.005	-0.000	-0.008	0.008	0.011	0.002	1.00
Velocity	1.40	0.141	0.351	0.613	0.948	0.691	0.477	0.295	0.27

Table 4.1: Simulated Cash-Credit RBC Model
(from Cooley and Hansen (ed) "Frontiers of Business Cycle Research")

The RBC model used in this simulation is essentially the basic Stochastic Growth model but with Hansen's assumption regarding indivisible labour markets. Comparison of the rows corresponding with the real variables in Table 4.1 with those from non-monetary RBC simulations suggest that little has altered except that consumption is slightly more volatile in Table 4.1. The effects of the inflation tax are immediately evident from the final column of Table 4.1, with consumption showing a strong negative correlation with money growth and investment a positive correlation, exactly what our intuition would expect. Unfortunately, comparing this with the earlier stylised facts we see that these correlations are not observed in the data. There are also a number of additional problems with Table 4.1. Firstly, money displays very little correlation with output either contemporaneously or in advance. The data suggest that high monetary growth is followed by higher output, which is not revealed in the simulations in Table 4.1. Further, the nominal interest rate displays no correlation with output at any leads or lags whereas the data shows strongly that high interest rates tend to be associated with lower output. Finally, the model is capable of generating a negative correlation between prices and output (because productivity shocks are the main source of fluctuations) but does not generate a positive correlation between inflation and output. In other words, adding a CIA model to as simple RBC structure cannot account either for the observed cyclical behaviour of nominal variables or the interaction between real and nominal variables. This suggests that in order to successfully account for the interaction between real and nominal variables in the data we need to introduce more sources of non-neutrality than just the inflation tax.

4.8 Exercise

Let the consumer's utility function be $U(c,l) = \alpha \ln c_{1t} + (1-\alpha)\ln c_{2t} - \gamma h_t$, where h denotes hours worked, c_1 is a good which can only be purchased with cash and c_2 denotes a good which can be purchased using credit. Households can hold two assets: money (m) or government bonds (b), the latter earn the return R. The household seeks to maximise utility subject to two constraints:

(i) a cash in advance constraint

$$p_t c_{1t} \leq m_t + (1 + R_{t-1})b_t - b_{t+1}$$

(ii) a resource constraint

$$c_{1t} + c_{2t} + \frac{m_{t+1}}{p_t} + \frac{b_{t+1}}{p_t} \leq w_t h_t + \frac{m_t}{p_t} + \frac{(1 + R_{t-1})b_t}{p_t}$$

(a) Write down an expression for the share of cash goods in total consumption as a function of the interest rate

(b) Write down an expression for the velocity of money. Is this a plausible model?

4.9 Solution

(a) The consumer maximises the present discounted value of utility subject to the two constraints, i.e.

$$\max E_0 \sum_{j=0}^{\infty} \beta^{t+j} \left[\begin{array}{l} \alpha \ln c_{1t+j} + (1-\alpha) \ln c_{2t+j} - \gamma h_{t+j} \\ +\lambda_{1t+j}(m_{t+j} + (1 + R_{t+j-1})b_{t+j} - b_{t+j+1} - p_{t+j} c_{1t+j}) \\ +\lambda_{2t+j}(w_{t+j} h_{t+j} + \frac{m_{t+j}}{p_{t+j}} + \frac{(1+R_{t+j-1})b_{t+j}}{p_{t+j}} - c_{1t+j} - c_{2t+j} - \frac{m_{t+j+1}}{p_{t+j}} - \frac{b_{t+j+1}}{p_{t+j}} \end{array} \right]$$

The first order conditions are:

$$w.r.t.\ c_1 \quad \tfrac{\alpha}{c_{1t}} - \lambda_{1t} p_t - \lambda_{2t} = 0 \tag{4.1}$$

$$w.r.t.\ c_2 \quad \tfrac{1-\alpha}{c_{2t}} - \lambda_{2t} = 0 \tag{4.2}$$

$$w.r.t.\ m \quad \lambda_{1t} + \tfrac{\lambda_{2t}}{p_t} - \tfrac{1}{\beta} \tfrac{\lambda_{2t-1}}{p_{t-1}} = 0 \tag{4.3}$$

$$w.r.t.\ b \quad \lambda_{1t}(1 + R_{t-1}) + \lambda_{2t} \tfrac{(1+R_{t-1})}{p_t} - \tfrac{1}{\beta}\lambda_{1t-1} - \tfrac{1}{\beta} \tfrac{\lambda_{2t-1}}{p_{t-1}} = 0 \tag{4.4}$$

$$w.r.t.\ h \quad \lambda_{2t} w_t = \gamma \tag{4.5}$$

When finding the share S_t of cash goods in total consumption we are interested in

$$S_t = \frac{c_{1t}}{c_{1t} + c_{2t}}$$

We can substitute out for consumption using (4.1) and (4.2).

$$\begin{aligned} S_t &= \left(\frac{\alpha}{\lambda_{1t} p_t + \lambda_{2t}} \right) \Big/ \left(\frac{\alpha}{\lambda_{1t} p_t + \lambda_{2t}} + \frac{1-\alpha}{\lambda_{2t}} \right) \\ &= \frac{\alpha \lambda_{2t}}{\alpha \lambda_{2t} + (1-\alpha)\lambda_{1t} p_t} \end{aligned}$$

Now substitute for λ_{1t} from (4.3) into (4.4).

$$\left(\frac{1}{\beta}\frac{\lambda_{2t-1}}{p_{t-1}} - \frac{\lambda_{2t}}{p_t}\right)(1+R_{t-1}) + \lambda_{2t}\frac{(1+R_{t-1})}{p_t} - \frac{1}{\beta}\lambda_{1t-1} - \frac{1}{\beta}\frac{\lambda_{2t-1}}{p_{t-1}} = 0$$

$$\lambda_{1t-1} = R_{t-1}\frac{\lambda_{2t-1}}{p_{t-1}}$$

And the share of cash goods is given by

$$S_t = \frac{\alpha}{1+(1-\alpha)R_t}$$

and the share of cash goods is inversely related to the interest rate, $S_t = S_t(\overset{-}{R_t})$.

(b) Because there is an opportunity costs to holding money, the only money held will be to finance cash goods, i.e.

$$m_t = c_{1t}p_t$$

Total income at any time t is equal to total consumption spending on cash and credit goods.

$$y_t = p_t c_{1t} + p_t c_{2t}$$

Therefore, velocity of circulation is

$$v_t = \frac{y_t}{m_t} = \frac{p_t c_{1t} + p_t c_{2t}}{p_t c_{1t}} = \frac{1}{S_t}$$

$$v_t = \frac{1}{\alpha} + \frac{(1-\alpha)}{\alpha}R_t$$

This varies positively with the rate of interest, $v_t = v_t(\overset{+}{R_t})$, and so potentially the model is quite realistic.

Chapter 5

Sticky wages and prices

5.1 Motivation

The previous lectures have highlighted the apparent inability of monetary RBC models to replicate key features of the data. Simply adding a monetary sector to a standard model does not give a very realistic model of the business cycle and so is inappropriate for serious monetary policy analysis. In this lecture we introduce an additional propagation mechanism in the economy by allowing for stickiness in wages and prices. Such models begin to be more Keynesian in spirit but still remain some way from being useful.

5.2 Key readings

Cooley and Hansen (1994) "Money and the Business Cycle", in Cooley (ed) (1994) "Frontiers of Business Cycle Research". Especially Sections 5 and 6.

5.3 Related reading

Calvo (1983) "Staggered Prices in a Utility-Maximising Framework", *Journal of Monetary Economics*, 12, 383-398

Fischer (1977) "Long-Term Contracts, Rational Expectations, and the Optimal Money Supply Rule", *Journal of Political Economy*, 85, 191-206

Gray (1976) "Wage Indexation: A Macroeconomic Approach", *Journal of Monetary Economics*, 2, 221-235

Taylor (1980) "Aggregate Dynamics and Staggered Contracts", *Journal of Political Economy*, 88, 1-24

Yun (1996) "Nominal Price Rigidity, Money Supply Endogeneity and Business Cycles", *Journal of Monetary Economics*, 37, 175-391.

5.4 Fixed wage contracts

In their chapter in the Cooley volume, Cooley and Hansen suggest that it may be worth investigating what the effects of fixed wage contracts are in monetary RBC models. They consider a very simple example. Imagine that, for whatever reason, the nominal wage has to be set one period in advance using the expectations of both firms and workers. During the period in which the preset wage holds, firms are then free to determine their employment level such that marginal product of labour equals the preset nominal wage. The wage that is set one period in advance is based on rational expectations of next period's inflation and productivity. However, if the money supply is unexpectedly high next period the nominal wage will have been set too low and so inflation will be higher than expected and the real wage will have fallen. Because the real wage will have fallen, firms recruit more labour and so produce more output. Therefore positive money supply shocks are associated with higher output. However, because of rational expectations it is only *unexpected* money supply shocks which have this effect on output. Essentially, a price has been fixed in advance and cannot adjust within the period so that money can affect real variables. However, after the period has elapsed wages are reset at the appropriate new inflation level. Therefore the real wage rises to its "true" value and employment declines. In other words, this model generates little persistence as money supply shocks only have a positive effect while wages cannot be adjusted, which in this case is one year. If wages are set for several periods in advance, however, money can be non-neutral for longer periods. This idea that fixed nominal contracts may explain business cycle features is an old ideas and was the initial Keynesian response to the policy ineffectiveness claims of Rational Expectations, see Gray (1976), Fischer (1977) and Taylor (1980).

Variable	SD%	Cross-correlation of output with:							Correlation with M0 growth
		t-3	t-2	t-1	t	t+1	t+2	t+3	
Output	2.37	0.098	0.202	0.320	1.000	0.320	0.202	0.098	0.61
Consumption	0.49	0.363	0.392	0.346	0.261	0.304	0.131	-0.004	-0.42
Investment	9.00	0.047	0.154	0.289	0.986	0.298	0.198	0.108	0.67
Hours	3.10	0.001	0.066	0.142	0.937	0.189	0.128	0.070	0.77
Prices	1.79	0.053	0.055	0.041	-0.011	-0.361	-0.266	-0.191	0.42
Inflation	1.15	-0.003	0.022	0.078	0.561	-0.145	-0.117	-0.093	0.92
Nominal interest rate	1.04	-0.019	0.012	0.079	0.475	-0.069	-0.053	-0.039	0.71
ΔMoney	0.87	-0.015	0.064	0.236	0.608	-0.091	-0.076	-0.069	1.00
Velocity	2.39	0.022	0.122	0.254	0.975	0.254	0.172	0.097	0.74

Table 5.1: Simulated Monetary Economy with Nominal Contracts
(from Cooley and Hansen (ed) "Frontiers of Business Cycle Research")

There are a number of problems with this approach. Firstly, it requires a strong justification of why nominal wages may be fixed for several periods in advance and why they cannot be altered when the state of the world is revealed (i.e. why can they not be indexed?). Secondly, to explain the fact that the money supply one or two years ago predicts current output we need to assume that wages are fixed for long periods of time. In reality, wage contracts do not last this long. Thirdly, even if we make these assumptions we can still not explain some of the features we observe in the data. Table 5.1 from Cooley and Hansen (1994) (in Cooley (ed)) shows some simulation results from a CIA-RBC model with one period ahead fixed nominal wages. An important failing of this model (aside from the short lead time money has over output, which could be corrected by assuming longer contracts) is that high (low) interest rates do not predict low (high) future output. This is a very strong feature of the data and one which most monetary economists feel is a defining feature of the transmission mechanism of monetary policy. In the next section we turn to constructing models which try and explain this correlation in the data.

5.5 Sticky prices

Yun (1996) investigates how the assumption of sticky prices impacts on the ability of RBC monetary models to explain the cyclical behaviour of inflation. He uses a cash in advance RBC model and shows that by introducing sticky prices he can substantially improve the model's ability to

(a) explain procyclical inflation

(b) explain why output and inflation rise in response to temporary shocks.

To introduce sticky prices into the analysis it is necessary to introduce market power - if firms do not have some monopoly power then prices cannot be different across competitors. Yun uses standard assumptions regarding preferences to derive the demand for good i as

$$D_t(i) = \left(\frac{P_t(i)}{P_t}\right)^{-\epsilon} D_t$$

where D_t is aggregate demand, $P_t(i)/P_t$ is the relative price for good i and ϵ denotes the elasticity of demand. To model price stickiness he follows the formulation of Calvo (1983). This formulation assumes that each period a fraction of firms, $1-\alpha$, charges a new price and the remaining fraction α simply index their price in line with inflation. Under these assumptions firms expect to keep their prices unchanged for $1/\alpha$ periods. However, note this is an expectation - some firms will be changing prices frequently whilst others will go for long periods before changing their prices. We should stress that this is done for analytical convenience rather than empirical plausibility.

Table 5.2a shows the correlation of inflation (at various leads and lags) with the changes in US GNP along with the correlation of inflation with the components of US GNP due to permanent productivity shocks and temporary monetary disturbances. Tables 5.2b - 5.2e show the same set of numbers but for different assumptions about price flexibility. Comparing Tables 5.2a and 5.2b shows that assuming flexible prices e.g. $\alpha = 0$ means that the model cannot explain why inflation responds in a strong way to temporary money supply shocks. However, notice that only by assuming $\alpha = 0.75$, that is that only 25% of firms change their prices each period, can we replicate findings based on US data. This then begs the question what can justify such large and important price rigidites and whether this value of $\alpha = 0.75$ is plausible. The second thing to note about the Yun paper is that it *only* focuses on the inflation-output correlation.

In particular, it does not try to focus on other problematic correlations - namely money-output-interest rates.

j	-3	-2	-1	0	1	2	3
π_{t+j} and output growth at t	-0.200	-0.104	-0.048	-0.025	0.026	0.014	-0.028
π_{t+j} and Δtrend output at t	0	0	0	-0.451	-0.177	-0.151	-0.169
π_{t+j} and transitory output at t	0.744	0.814	0.850	0.867	0.649	0.550	0.469

Table 5.2a: Inflation and Output Correlations for US data

j	-3	-2	-1	0	1	2	3
π_{t+j} and output growth at t	-0.019	-0.020	-0.031	-0.378	0.035	0.034	-0.033
π_{t+j} and Δtrend output at t	0	0	0	-0.378	-0.033	-0.035	-0.024
π_{t+j} and transitory output at t	0.172	0.179	0.187	0.199	0.098	0.094	0.091

Table 5.2b: Inflation and Output Correlations for Flexible Price Model

j	-3	-2	-1	0	1	2	3
π_{t+j} and output growth at t	-0.042	-0.076	-0.168	-0.209	0.012	0.021	-0.026
π_{t+j} and Δtrend output at t	0	0	0	-0.371	-0.025	-0.025	-0.0.025
π_{t+j} and transitory output at t	0.187	0.212	0.288	0.289	0.122	0.109	0.104

Table 5.2c: Inflation and Output Correlations for $\alpha = 0.25$

j	-3	-2	-1	0	1	2	3
π_{t+j} and output growth at t	-0.091	-0.187	-0.477	-0.306	0.053	0.021	-0.001
π_{t+j} and Δtrend output at t	0	0	0	-0.374	-0.025	-0.026	-0.026
π_{t+j} and transitory output at t	0.220	0.256	0.363	0.522	0.184	0.148	0.125

Table 5.2d: Inflation and Output Correlations for $\alpha = 0.5$

j	-3	-2	-1	0	1	2	3
π_{t+j} and output growth at t	-0.088	-0.196	-0.553	-0.601	0.093	0.055	-0.031
π_{t+j} and Δtrend output at t	0	0	0	-0.370	-0.025	-0.032	-0.033
π_{t+j} and transitory output at t	0.243	0.308	0.455	0.867	0.305	0.228	0.177

Table 5.2e: Inflation and Output Correlations for $\alpha = 0.75$

Chapter 6

The persistence puzzle

6.1 Motivation

At the beginning of the lectures we derived evidence using structural vector autoregressions that monetary policy shocks have a delayed yet persistent effect on output. All of the models we have examined so far (Lucas islands, cash in advance, nominal wage rigidity, sticky prices) have failed to replicate this key feature of the data. Whilst they are able to generate real-nominal interactions, the effect is not very persistent and output quickly returns to baseline. In the literature this is referred to as the "persistence puzzle". In this lecture we show how it is a general failing of monetary RBC models.

6.2 Key readings

The origin of the persistence puzzle is Chari, Kehoe and McGrattan (2000) "Sticky Price Models of the Business Cycle: Can the Contract Multiplier Solve the Persistence Problem?", *Econometrica*, 68(5), 1151-1180. This is a difficult paper to read. Our discussion is a simplified version of Section 5 of the paper.

6.3 Related reading

Chari, Kehoe and McGrattan (2000) "Can Sticky Price Models Generate Volatile and Persistent Real Exchange Rates?", *NBER working paper,* no. 7869

Ellison and Scott (2000) "Sticky Prices and Volatile Output", *Journal of Monetary Economics*, 46, 621-632

Taylor (1980) "Aggregate Dynamics and Staggered Contracts", *Journal of Political Economy*, 88, 1-24

Watson (1993) "Measures of Fit for Calibrated Models", *Journal of Political Economy*, 101, 1011-1041

6.4 Introduction

The basic intuition behind the persistence puzzle can be seen by considering the case with a static money demand equation, $y_t + p_t = m_t$, where y_t is output, p_t is the aggregate price level and m_t is the aggregate money supply. This equation enables us to see how money supply shocks are divided between output and prices - to have a large output effect it must be the case that prices do not respond too much to money supply shocks. However, prices do respond to costs in the form of the wage rate w_t. In other words, if prices are not to respond too much to money supply shocks then wages must not respond too much either.

To see the response of wages we assume a general form for the labour supply equation, $w_t = \gamma y_t$, where $\gamma > 0$ is the elasticity of the real wage with respect to output. According to this equation, wages will not react to the money supply shock as long as γ is small. We will see in what follows that in all our models γ is by definition large so there is too much pass-through of money supply shocks to wages and prices and not enough to output.

6.5 Staggered price setting and persistence

The paper by Chari, Kehoe and McGrattan (2000), CKM, formally analyses whether staggered price setting can solve the persistence problem. To demonstrate their argument we will assume that there are two types of firms: one changes its price in even periods and one changes its price in odd periods. We denote the price set by firms able to change their price in period t as x_t, so the aggregate price at time t is $p_t = (x_t + x_{t-1})/2$. This is potentially a strong propagation mechanism. Because of the staggered and overlapping price contracts, if the money supply shock affects the price set at time t, x_t, then this affects the aggregate prices p_t and p_{t+1}. However, the aggregate price p_{t+1} affects the price set at $t+1$, x_{t+1}, and so also the aggregate price p_{t+2} and so on. CKM refer to this as a "contract multiplier" which could pass the shock on through time.

CKM derive an equation (6.1) for the optimal price-setting behaviour of a firm setting its price at time t. Using microfoundations, they show that the price set depends on (i) the price x_{t-1} set by other firms in the previous period, (ii) the price x_{t+1} expected to be set by other firms in the next period, (iii) expected wages in the current period and (iv) expected wages in the next period. This formulation is very close to Taylor (1980).

$$x_t = \frac{1}{2}x_{t-1} + \frac{1}{2}E_{t-1}x_{t+1} + E_{t-1}w_t + E_{t-1}w_{t+1} \tag{6.1}$$

With a static money demand equation, $y_t + p_t = m_t$, a labour supply function of the form $w_t = \gamma y_t$, and the definition of aggregate prices $p_t = (x_t + x_{t-1})/2$, this can be re-written as equation (6.2).

$$E_{t-1}x_{t+1} - \left[\frac{2\,(1+\gamma)}{1-\gamma}\right] x_t + x_{t-1} + \left[\frac{2\gamma\,(1+\gamma)}{(1-\gamma)^2}\right] E_{t-1}\left(m_t + m_{t+1}\right) = 0 \tag{6.2}$$

This is a standard second order difference equation. We make the further simplifying assumption that the money supply is a random walk $E_{t-1}m_{t+j} = m_{t-1}\ \forall\ j$, so the solution to the equation is of the form

$x_t = ax_{t-1} + bm_{t-1}$. Comparing coefficients on x_{t-1} we obtain the quadratic equation (6.3) in a and select the stable solution (6.4).

$$a^2 - \left[\frac{2(1+\gamma)}{1-\gamma}\right] a + 1 = 0 \tag{6.3}$$

$$a = \frac{1-\sqrt{\gamma}}{1+\sqrt{\gamma}} \tag{6.4}$$

The full solution to the model is as follows:

$$\begin{aligned}
x_t &= ax_{t-1} + (1-a)m_{t-1} \\
p_t &= ap_{t-1} + \frac{1}{2}(1-a)(m_{t-1} - m_{t-2}) \\
y_t &= ay_{t-1} + (m_t - m_{t-1}) + \frac{1}{2}(1-a)(m_{t-1} - m_{t-2})
\end{aligned}$$

The degree of persistence in the model is directly related to the parameter a. If $a = 0$ ($\gamma = 1$), a shock Δ to the money supply at time t leads to an increase Δ in output at time t, an increase of $\Delta/2$ at time $t+1$ and no increase for all subsequent periods. There is no persistence beyond the length of price stickiness. In contrast, if $a = 1$ ($\gamma = 0$), a shock Δ to the money supply at time t has a permanent effect on output. The value of a depends directly on the elasticity of the real wage with respect to output by equation (6.4). To generate the require degree of persistence needs a value of γ which is not too small (in which case there is too much persistence) but not too large either (in which case there is very little persistence).

6.6 Labour supply in classical models.

The elasticity of the real wage with respect to output, γ, is not a free parameter in classical RBC models. Rather, the supply of labour is determined by the tastes of workers in competitive labour markets. In such models, workers deciding their labour supply typically maximise their utility function (6.5) subject to a budget constraint (6.6).

$$\begin{aligned}
u(c_t, l_t) &= \ln c_t + \psi \ln(1 - l_t) \qquad (6.5)\\
c_t &= w_t l_t \qquad (6.6)
\end{aligned}$$

c_t is consumption, l_t is hours worked and $\psi > 0$ is a parameter measuring the weight of leisure time in the utility function. The utility function (6.5) is popular because it is consistent with balanced growth in the economy. In the budget constraint (6.6), w_t is the wage rate. Assuming no capital in the economy so $c_t = y_t \forall t$, the first order condition can be written as equation (6.7).

$$w_t = (1 + \psi) y_t \tag{6.7}$$

This is the aggregate supply curve of labour in a competitive labour market. Compared with the general form $w_t = \gamma y_t$, we see that $\gamma = 1 + \psi > 1$, where ψ is a true structural parameter. This illustrates why there is a "persistence puzzle" in the models we have analysed. All our models have competitive labour markets so $\gamma > 1$ and *by definition* the money supply shocks do not have a persistent effect on output. The wage rate and prices react too strongly to money supply shocks to have persistent output effects. The CKM result is that price stickiness alone cannot solve the persistence problem. It is worth noting that Taylor (1980) treats γ itself as a structural parameter and calibrates it at 0.05 to generate persistence. However, CKM derive this parameter from first principles to arrive at a much larger calibration of 1.22.

6.7 Extensions

The paper by Ellison and Scott (2000) looks at nominal price rigidity in a fully-specified dynamic general equilibrium model. They calibrate ψ to 1.56 on the basis of an 8 hour working day in competitive labour markets. The elasticity of the real wage with respect to output is then $\gamma = 2.56$ so there are persistence puzzles in the model simulations. The impulse response in Figure 6.1 from the model is much less than that estimated from the data using a structural vector autoregression.

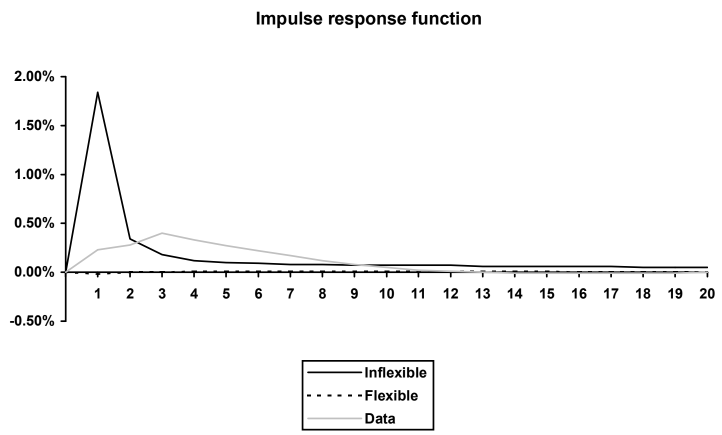


Figure 6.1: Impulse response function of monetary shocks on output

Ellison and Scott (2000) also identify other key failings of the model. Rather than rely on a qualitative assessment of stylised facts, they look at a quantitative measure of the goodness of fit of the model using the approach of Watson (1993). This measures how well the model fits the data in four dimensions: output, inflation, employment and consumption. The conclusion of the paper is that overall the presence of price rigidity *worsens* the ability of the model to match the data. A flexible price model actually performs better as a representation of the data than a sticky price model. Most of the deterioration in the performance is in the form of excessively high output volatility.

A second paper by Chari, Kehoe and McGrattan (2000) looks at the analogous problem of volatility and persistence of real exchange rates in an open economy. The conclusion is somewhat similar to the closed economy case - stickiness can only account for some properties of the data.

Chapter 7

IS-LM is back

7.1 Motivation

The IS-LM model originally developed by Hicks (1937) fell out of favour with macroeconomists as the Rational Expectations Revolution forced researchers to work harder on the microfoundations of their models. However, the recent trend has been to recast modern stochastic dynamic general equilibrium models in the IS-LM framework. In this lecture, we see how this is possible and re-examine the persistence problem analysed in the previous lecture.

7.2 Key readings

Goodfriend and King (1998) "The New Neoclassical Synthesis and the Role of Monetary Policy" *Federal Reserve Bank of Richmond Working Paper*, No. 98-5

Jeanne (1997) "Generating Real Persistent Effects of Monetary Shocks: How Much Nominal Rigidity Do We Really Need?", *NBER Working Paper*, No. 6258

7.3 Related reading

Calvo (1983) "Staggered Prices in a Utility-Maximising Framework", *Journal of Monetary Economics*, 12, 383-398

Clarida, R., J. Galí and M. Gertler, 1999, "The Science of Monetary Policy: A New Keynesian Perspective", *Journal of Economic Literature* 27, 1661-701

Mankiw (2000) "The Inexorable and Mysterious Trade-Off Between Inflation and Unemployment", *Economic Journal*, 111, C45-C61

Mankiw and Reis (2001) "Sticky Information Versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve", *NBER Working Paper*, No. 8290

7.4 Log-linearisation

To derive the dynamic IS-LM model we will follow the paper and notation of Jeanne (1997). To make the algebra tractable we will describe a log-linearised version of the model, i.e. all the key equations are log-linearised using the formula (7.1).

$$\ln(f(x_t)) \simeq \ln(f(x_0)) + \frac{f'(x_0)}{f(x_0)} x_0 \ln\left(\frac{x_t}{x_0}\right) \tag{7.1}$$

After applying the approximation, the solution of the model can always be written in the form $X_t = \beta X_{t-1}$. This equation describes the evolution of an $n \times 1$ vector of variables X_t, where X_t is covariance stationary and β is a matrix of coefficients.

7.5 The IS curve

The IS curve is derived from intertemporal maximisation of the representative agents, who maximises a utility function subject to a resource constraint. The agent equates the marginal utility of consumption in adjacent periods after allowing for the real rate of return on savings (the nominal interest rate minus the rate of inflation). In other words, there is an Euler equation for consumption in the form of

$$u'(c_t) = \beta E_t \left(\frac{1+i_t}{1+\pi_{t+1}} u'(c_{t+1}) \right)$$

$u'(c_t)$ is the marginal utility of consumption c_t at time t. This equals the marginal utility of consumption at time $t+1$ after allowance for nominal interest i_t paid on savings and inflation π_{t+1}. To log-linearise this equation first take logarithms.

$$\ln u'(c_t) = \ln \beta + E_t \ln(1+i_t) - E_t \ln(1+\pi_{t+1}) + E_t \ln u'(c_{t+1})$$

Using the CRRA utility function, $u(c_t) = c_t^{1-\frac{1}{\sigma_u}}/(1-\frac{1}{\sigma_u})$, $\ln u'(c_t) = -\frac{1}{\sigma_u} \ln c_t$. The above equation also holds at steady-state c_0, i_0, π_0 so we can write

$$\begin{aligned} -\frac{1}{\sigma_u}(\ln c_t - \ln c_0) &= E_t \ln(1+i_t) - \ln(1+i_0) \\ &\quad - E_t \ln(1+\pi_{t+1}) + \ln(1+\pi_0) \\ &\quad - \frac{1}{\sigma_u}(E_t \ln c_{t+1} - \ln c_0) \end{aligned}$$

We introduce the hat notation $\hat{x}_t = (x_t - x_0)/x_0 \approx \ln x_t - \ln x_0$, where $\hat{x}_t$ is the percentage deviation from steady-state. Finally, assume no capital so $\hat{y}_t = \hat{c}_t$ and we have the dynamic IS curve (7.2).

$$\hat{y}_t = -\sigma_u[\hat{\imath}_t - E_t \hat{\pi}_{t+1}] + E_t \hat{y}_{t+1} \tag{7.2}$$

This is a dynamic relationship between output and the real interest rate, as was true statically in Hick's (1937) analysis. When the real interest rate is expected to be high their either current output is low or future output is high.

7.6 The LM curve

We now turn our attention to the monetary side of the model. Money non-neutralities are introduced by assuming that the representative consumer faces a cash-in-advance constraint.

$$p_t c_t = M_t$$

M_t is nominal money holdings. The log-linearisation is simply (7.3), where $\hat{m}_t$ is real money holdings, $\hat{y}_t = \hat{c}_t$ because of no capital, and the inequality is assumed to always bind.

$$\hat{y}_t = \hat{m}_t \tag{7.3}$$

Equation (7.3) is the dynamic LM curve, linking output to the money supply. Because of our assumption of cash-in-advance, the velocity of circulation of money is always constant at unity and interest rates do not have a separate effect. A more general specification such as a cash-credit model would lead to an additional interest rate term in the LM equation.

7.7 The AS curve

To complete the model we have to specify an equation for aggregate supply in the economy. The supply curve in Jeanne (1997) is an example of the New Keynesian Supply Curve made popular by Clarida, Galí and Gertler (1999). They begin with the Calvo (1983) approach to price stickiness in which a fraction $1-\phi$ of firms are allowed to change their price each period. Under such a scheme, the aggregate price level (in percentage deviations from steady state) follows a simple linear expression.

$$\hat{p}_t = (1-\phi)\hat{p}_{it} + \phi\hat{p}_{t-1}$$

$\hat{p}_{it}$ is the price set by the fraction $1-\phi$ of firms able to change their price in the current period. The remaining fraction ϕ retain the previous average price $\hat{p}_{t-1}$. Using standard assumptions, it is possible to derive from microfoundations the pricing behaviour of a firm able to set its price in period t.

$$\hat{p}_{it} = (1-\beta\phi)\hat{p}^*_t + \beta\phi E_t\hat{p}_{it+1}$$

This expression shows that the price the firm sets is a weighted average of what it would set purely looking at the current period ($\hat{p}^*_t$) and the price it anticipates setting if able to in the next period ($E_t\hat{p}_{it+1}$). $\beta < 1$ is the subjective discount factor. Note that if $\phi = 0$ then there is perfect price flexibility, the probability of being able to change price each period is 1 and the firm sets its price according to short-run aims, $\hat{p}_{it} = \hat{p}^*_t$. In contrast, as $\phi \to 1$ and the degree of price stickiness increases, the firm pays less and less attention to current market conditions and focuses on the future expected price, $\hat{p}_{it} \simeq E_t\hat{p}_{it+1}$.

An expression for $\hat{p}^*_t$ can be obtained by considering the equilibrium that would prevail when the firm maximises its short-run profit. Typically, in such a situation relative price is set as a mark-up over marginal cost, where the magnitude of the mark-up depends on the degree of market power held by the firm. This relationship, $p_t^*/p_t = k \times mc_t$ when log-linearised becomes

$$\hat{p}_t^* = \hat{p}_t + \hat{mc}_t$$

Since there is no capital in the model, marginal costs arise only from wages and, assuming log-linearity between wages and marginal costs, we have that $\hat{mc}_t = \hat{w}_t$. To complete the aggregate supply curve we assume a general form for the labour supply function as in the previous lecture.

$$\hat{w}_t = \frac{1}{\alpha}\hat{y}_t$$

Note again that this is consistent with a competitive labour for a suitable choice of the α parameter. Combing the various behavioural relationships (the evolution of aggregate prices, optimal price setting, definition of short-run desired price, and the labour supply function) gives the New Keynesian aggregate supply curve (7.4).

$$\hat{y}_t = \frac{\alpha\phi}{(1-\phi)(1-\beta\phi)}(\hat{\pi}_t - \beta E_t\hat{\pi}_{t+1}) \tag{7.4}$$

The intuition behind this supply function has some appeal. If inflation is expected to increase then output falls: firms able to change their price in the current period set high prices (thereby choking demand) in anticipation of future high prices. Similarly, if inflation is expected to fall then firms set lower prices (stimulating demand) ready for the future. This supply function is part of what McCallum calls the nearest thing we have to a standard model in monetary economics. However, recently there has been some unease about certain features of the model, see Mankiw (2000) and Mankiw and Reis (2001).

7.8 The money supply

The money supply is assumed to be exogenous with growth rate $\mu_t = M_t/M_{t-1}$ following an AR(1) process $\mu_t - \bar{\mu} = \rho_m(\mu_t - \bar{\mu}) + \varepsilon_t$. A log-linearisation of the real money balances is (7.5).

$$\hat{\mu}_t = \hat{\pi}_t + \hat{m}_t - \hat{m}_{t-1} \tag{7.5}$$

7.9 The persistence puzzle revisited

Equations (7.2) - (7.5) describe an IS-LM-AS framework that is dynamic and derived from optimising behaviour of individual agents. To look at persistence we can solve the system of log-linear equations and find an expression for any endogenous variable in terms of exogenous variables. We have four equations in four unknowns ($\hat{y}_t, \hat{\imath}_t, \hat{\pi}_t$ and $\hat{m}_t$) and so have the solution

$$\hat{y}_{t-1} - \left[1 + \beta + \frac{(1-\phi)(1-\beta\phi)}{\alpha\phi}\right]\hat{y}_t + \beta E_t\hat{y}_{t+1} = \beta E_t\hat{\mu}_{t+1} - \hat{\mu}_t$$

This is a second order stochastic difference equation which satisfies standard assumptions. Using standard techniques, the solution of the equation may be written as

$$\hat{y}_t = \rho \hat{y}_{t-1} + \rho \frac{1-\beta\rho_m}{1-\beta\rho\rho_m} \hat{\mu}_t$$

where ρ is the stable root satisfying

$$1 - \left[1 + \beta + \frac{(1-\phi)(1-\beta\phi)}{\alpha\phi}\right]\rho + \beta\rho^2 = 0$$

Once again, ρ is the crucial parameter for the persistence of monetary policy shocks. Jeanne (1997) asks how much rigidity do we need for a certain level of persistence. He shows that both nominal price rigidity (ϕ) and real wage rigidity (α) increase the level of persistence, i.e. $\partial\rho/\partial\phi > 0$ and $\partial\rho/\partial\alpha > 0$. To make this clearer we can ask which combinations of nominal and real rigidity give rise to a certain level of persistence. This is equivalent to holding ρ constant in the quadratic equation and searching for suitable values of ϕ and α. Re-arranging the quadratic equation gives a simple expression for the level of wage rigidity needed for persistence ρ when nominal rigidity is ϕ.

$$\alpha = \frac{\rho(1-\phi)(1-\beta\phi)}{\phi(1-\rho)(1-\beta\rho)}$$

We can plot this curve in (ϕ, α) space to see which combinations of nominal and real rigidity give a desired level of persistence. The iso-persistence curves are shown in Figure 7.1.

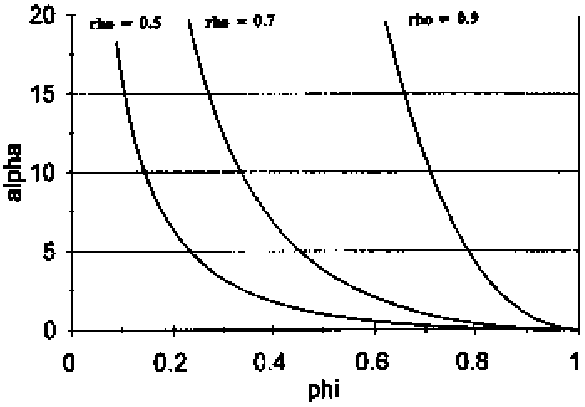


Figure 7.1: The iso-persistence curves in (ϕ,α) space

The figure reveals important non-linearities in the relationship. The curve are flat around $\phi = 1$ so a small increase in real rigidity is sufficient to offset a large amount of nominal rigidity. This suggests that it is not necessary to assume implausibly large amounts of wage stickiness to generate persistence. Jeanne (1997) suggests a calibration of $\phi = 0.5, \alpha = 3, \rho_m = 0.5$ as realistic. The impulse response function of output to a money supply shock for this calibration is shown in Figure 7.2. The response comes very close to that estimated from the data. Note, though that $\alpha = 3$ implies a considerable amount of wage rigidity compared with competitive labour markets.

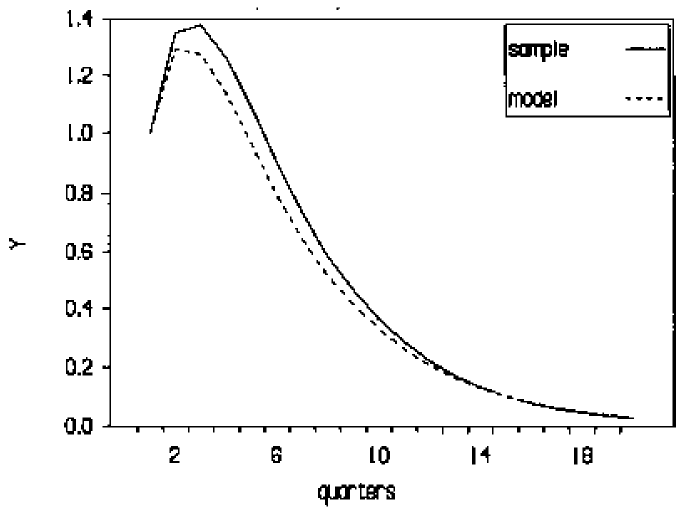


Figure 7.2: Impulse response functions

Chapter 8

The liquidity effect

8.1 Motivation

In all the models we have examined so far there has been a very strong positive correlation between money supply shocks and the nominal rate of interest. This is a major failing of the models since in the data there is a negative correlation - in other words money supply shocks should reduce the nominal interest rate but in our models they increases it. What is missing in the models is a liquidity effect. Positive money supply shocks increase liquidity and so should reduce the price of money (the nominal interest rate). In this lecture we show how the basic cash-in-advance model can be amended to generate a liquidity effect.

8.2 Key readings

On the liquidity effect the best reference is Christiano “Modelling the Liquidity Effect of a Monetary Shock” which is printed in Miller (ed) ”The Rational Expectations Revolution” and also Christiano and Eichenbaum (1992) *American Economic Review Papers and Proceedings.* Fuerst (1992) “Liquidity, Loanable Funds and Real Activity” *Journal of Monetary Economics*, should be readable enough in places to be worth a look. Once again Lucas (1990) “Liquidity and Interest Rates” *Journal of Economic Theory* is the seminal reference but it is a difficult read for those unfamiliar with dynamic programming.

8.3 Related reading

Alvarez, Atkenson and Kehoe (1999) “Money and Interest Rates with Endogeneously Segmented Markets”, *NBER Working Paper*, No. 7060

Cogley and Nason (1995) “Output Dynamics in Real-Business-Cycle Models”, *American Economic Review*, 85, 492-511

Strongin (1995) “The Identification of Monetary Policy Disturbances: Explaining the Liquidity Puzzle”, *Journal of Monetary Economics*, 35, 463-497

8.4 Introduction

Before we can develop a model which explains how monetary policy affects the real economy we obviously need empirical evidence on whether monetary policy is important and if so through which channels it operates. Not surprisingly this is a long and controversial literature. Some more recent references have been given in the key readings sections and the enthusiast can follow up those references from earlier literature. For our purposes we shall assume that the results of Strongin (1995) are broadly correct: money supply innovations (correctly measured) are important and account for a large proportion of output fluctuations, a positive money supply shock tends to lower nominal interest rates and then lead to higher output.

The crucial thing here is the way in which positive money supply shocks lower the nominal interest rate. If we can explain how that happens then we can explain relatively easily how interest rates then affect output. However, explaining why the **nominal** interest rate falls is not a trivial exercise. In a neoclassical model the standard way of viewing nominal interest rates is through the Fisher hypothesis. This says that the nominal interest rate equals the real interest rate plus the expected inflation rate. In other words, the nominal interest rate must compensate depositors not just with the real interest rate but also for any purchasing power they are expected to lose during the period arising from inflation. An increase in the money supply can have two effects: (i) it can reduce the real interest rate (this is called the "liquidity effect", more money, i.e. more liquidity, tends to lower the price of money which is equivalent to lowering the interest rate) (ii) it forecasts higher future inflation (called the expected inflation or Fisher effect). Therefore to generate a falling nominal interest rate in response to a positive money supply shock we require the liquidity effect to outweigh the Fisher effect. However, in neoclassical models money does not influence real variables such as the real interest rate. Therefore increases in the money supply just forecast higher inflation and so the nominal interest rate rises as there is no liquidity effect only a Fisher effect (higher expected inflation). Therefore to generate the liquidity effect (lower nominal interest rates for higher money supply) we require enough non-neutralities that inflation expectations do not outweigh the liquidity effect.

Examining the stylised facts seen in previous lectures we can see that this was not the case for the basic CIA models. In these models nominal interest rates have little predictive power for output. This is because increases in the nominal interest rate are entirely due to the Fisher effect. In the CIA model when agents have higher expectations of inflation the nominal interest rate rises accordingly and there is no liquidity effect. We therefore need to amend the CIA model if we are to explain the liquidity effect.

8.5 Modelling the liquidity effect

Lucas (1992) sketched out a way this could be achieved by subtly altering the structure of CIA models and this idea was developed more fully in Fuerst (1992). Accessible treatments are to be found in Christiano and Eichenbaum (1992) (which has the merit of being very short) and Christiano which is in Miller (ed) "The Rational Expectations Revolution". The trick in these models is not dissimilar to some we have seen earlier, and it essentially involves altering the information at agents' disposal. Agents are to be thought of as a family who at the beginning of each period split up into (i) a worker (ii) a shopper (iii) a firm (iv) a financial intermediary. What is crucial to the model is that before the household separates at the beginning of the period they have to decide how much cash to hold for consumption purposes (equals $m_t - n_t$ where m denotes the money holdings inherited from last period and n is cash deposited with the financial intermediary). All consumption goods have to be purchased with cash so that $m - n > pc$. Only after a decision has been made on n does the household separate and at this point the state of the world is revealed, e.g. the current productivity shock and money supply shock. After separating at the beginning of the period, agents cannot come into touch with each other again and have to respond to shocks subsequently in an uncoordinated way. Once they have separated the worker in the household decides how many hours to work (l), the member of the household who acts as a firm hires h hours of work at the wage rate w (we ignore capital for the moment) and produces $f(h)$ which sells at price p. However, the wage bill has to be paid in advance and to finance this wage bill (wh) the firm borrows the amount b, where $b \geq wh$, on which it has to pay the nominal interest rate R. At the end of the period the new level of money balances equals what was inherited from last period, interest earned on n this period plus wage income less consumption plus the profits made by the household firm. We can, using Lagrange multipliers, outline the consumer's maximisation problem as

$$\max \left\{ E_0 \sum_{t=0}^{\infty} \left[\begin{array}{c} \beta^t u(c_t, l_t) + \lambda_{1t}\beta^t(m_t - n_t - p_t c_t) + \lambda_{2t}\beta^t(b_t - w_t h_t) \\ +\lambda_{3t}\beta^t(m_{t+1} - m_t - n_t R_t - w_t l_t + p_t c_t - p_t f(h_t) + w_t h_t + b_t R_t) \end{array} \right] \right\}$$

The consumer has to maximise lifetime utility by choosing for each period n (savings), m (money holdings), c (consumption), h (labour demand), l (labour supply) and b (borrowing). This maximisation has three constraints (i) the first reflects the cash in advance constraint faced by the consumer (ii) the second reflects the cash in advance constraint faced by the firm and (iii) the final way defines how money grows over time. For a given choice of n, made in the absence of any knowledge on the current productivity shock or money supply shock, the household's first order conditions are:

$$\begin{array}{lr} c_t & \beta^t u'_{ct} - \lambda_{1t}\beta^t p_t + \lambda_{3t}\beta^t p_t = 0 \\ m_t & \lambda_{1t}\beta^t - \lambda_{3t}\beta^t + \lambda_{3t-1}\beta^{t-1} = 0 \\ h_t & -\lambda_{2t}\beta^t w_t - \lambda_{3t}\beta^t p_t f'(h_t) + \lambda_{3t}\beta^t w_t = 0 \\ b_t & \lambda_{2t}\beta^t + \lambda_{3t}\beta^t R_t = 0 \end{array}$$

Using the first order conditions for c and m we have that

$$\begin{aligned} \frac{u'_{ct}}{p_t} &= \lambda_{1t} - \lambda_{3t} \\ \beta E_t \frac{u'_{ct+1}}{p_{t+1}} &= -\lambda_{3t} \end{aligned}$$

and using the first order condition for b we have that

$$\beta E_t \frac{u'_{ct+1}}{p_{t+1}} R_t = \lambda_{2t}$$

and combining them all together we have

$$[1+R_t] = \frac{\Lambda_t + u'_{ct}/p_t}{\beta E_t[u'_{ct+1}/p_{t+1}]}$$

$$\text{where } \Lambda_t = \lambda_{2t} - \lambda_{1t}$$

Consider first the case when $\Lambda = 0$. In this case the equation collapses to the standard Euler equation for consumption except that now we make an allowance for inflation. This equation simply says that the nominal interest rate equals the real interest rate (the ratio of the marginal utility of consumption) times a term reflecting inflation in the goods market. In other words we have the standard Fisherian equation that the nominal interest rate equals the real interest rate plus the inflation rate.

However, if $\Lambda < 0$, which implies $\lambda_1 > \lambda_2$, then the nominal interest rate is less than the standard Fisherian formula. What does this mean? If $\Lambda < 0$ then the value of the multiplier attached to the firm's cash in advance constraint is worth less than that attached to the consumer. In other words, the firm has more cash available than the consumer. Ideally, the household would want to switch these excess funds from the firm to the consumer. However, the structure of the model is that consumers cannot rapidly change their consumption and savings decisions, only firms can alter their borrowing/saving in response to current shocks. Therefore when cash is more readily available to the firm than the household then $\Lambda < 0$ and the nominal interest rate is low. Exactly the opposite reason suggests that when cash is scarce amongst firms then $\Lambda > 0$ and the nominal interest rate will rise.

What is happening here? At the beginning of the period consumers make their decisions regarding consumption and savings and they cannot change these decisions. Once they have made their decision the state of the world is revealed. Let us assume that there is unexpectedly high money supply growth. Banks therefore have more cash than they need and as long as $R > 0$ they will want to lend this money out. However, because consumers have made their decisions the only part of the household that can respond and increase their borrowing is the firm sector. However, to encourage the firm to borrow more the nominal interest rate must fall. Returning to the consumer's first order condition we can see that it is possible for an increase in the money supply to have the desired liquidity effect, i.e. for R to fall. However, this will not always occur as it requires Λ to change by more than inflation expectations. Otherwise the Fisher effect dominates and there is no liquidity effect. Fuerst shows that under some parameter values the liquidity effect dominates and the effect of an increase in the money supply is to lower interest rates and increase output (although see Christiano for some sceptical comments on this).

How does the fall in nominal interest rates increase output? Using the first order conditions for b and h it can be shown that

$$\frac{f'(h)}{1+R_t} = \frac{w_t}{p_t}$$

so that the marginal product of labour divided by the nominal interest rate equals the real wage rate. The reason the nominal interest rate enters into the labour demand condition is that labour is a cash in advance good, and so a borrowing cost is incurred for every person hired. Therefore when the nominal interest rate falls, hiring costs decline and so the firm is prepared to recruit more workers, i.e. the labour demand curve shifts to the right. If instead nominal interest rates rise then the labour demand curve shifts left, lowering employment, wages and output.

Therefore this extended version of CIA has the capacity to explain the liquidity effect (i.e. falling nominal interest rates in response to an increase in the money supply). It does so by combining the standard inflation tax that CIA models involve with an additional non-neutrality. This additional non-neutrality is essentially to assume that there are not complete markets. Ideally the consumer would like to be able to take out an insurance policy such that if the money supply were unexpectedly high more cash would be given to the consumer and if the money supply were low then cash would be taken away from the consumer. If such an insurance policy existed it could be used to ensure that the value of cash was the same for both the shopper and the firm in the household, i.e. $\Lambda = 0$. However, in the absence of such an insurance scheme the household has to precommit to a given level of cash holdings to purchase consumption goods with. Notice however that because of Rational Expectations the household on average chooses the correct amount of cash. In other words, we have the standard result that it is only unexpected money supply shocks that lead to the liquidity effect, not anticipated money. The household always chooses its cash so that on average it expects the value of cash to be the same for both firm and shopper. This leads to a similar problem that we encountered for the fixed nominal contracts model. As we discussed (see the reference to Cogley and Nason), the standard RBC model does not generate much persistence via capital accumulation. Therefore, even though money supply shocks affect output this period they have little effect on output in future periods. What happens in the Fuerst model is that next period households realise there is more money in the economy and reallocate it accordingly between shopper and firm. The rebalancing of the portfolio means that on average both shopper and firm will value money the same and real interest rates return to their previous level. In other words, the Fuerst model only generates the liquidity effect for as long as it takes consumers to rebalance their portfolio in the optimal way. In the Fuerst model this only takes one period and so even though this model can explain the liquidity effect it cannot explain why in the data it operates with a one to two year lead time over output. Christiano and Eichenbaum (1992) show that by assuming costs of adjustment we can get more persistent liquidity effects but it seems unlikely that they can explain a liquidity effect over a two year period.

8.6 Conclusion

We have outlined in the lectures a number of variants of cash in advance models. Whilst each variant gets closer to explaining certain observed facts it is clear that they are still someway from having a convincing and fully articulated model of how money affects output over the business cycle. The shopping time technology approach may hold some hope for the neoclassical monetary economists as it allows for greater flexibility on the part of the consumer than the CIA model. However, it will still prove difficult in that model to explain why the liquidity effect operates with such a long lead time. However, these notes suggest that the appropriate model will have to involve greater complexity than simply assuming certain prices are fixed. Such an approach clearly does not account for several features of the data. It is worth noting here that considerable interest has recently been placed, invariably by Keynesian economists, in the role of credit rather than money in determining business cycle fluctuations. However, it is still undecided whether credit or money will hold the key to understanding why nominal variables affect the economy.

Chapter 9

Monetary policy rules

9.1 Motivation

So far in the course the focus has been on constructing stochastic dynamic general equilibrium models which match various characteristics observed in the data. To some extent this has been successfully achieved so we now proceed to consider how monetary policy itself should be designed. We are able to do this because the models that have been developed are derived from fairly solid microfoundations, in other words we do not fall foul of the Lucas critique. In this lecture we assess different rules for setting monetary policy.

9.2 Key readings

Rotemberg and Woodford (1998) "Interest-Rate Rules in an Estimated Sticky Price Model" *NBER Working Paper*, No. 6618

Taylor (1993) "Discretion Versus Policy Rules in Practice", *Carnegie-Rochester Conference Series on Public Policy*, 39, 195-214

9.3 Related reading

Chow (1980) *Analysis and Control of Dynamic Economic Systems*, Wiley

Hansen and Sargent (2001) "Robust Control and Model Uncertainty", *mimeo*

Levin, Wieland and Williams (1998) "Robustness of Simple Monetary Policy Rules under Model Uncertainty", *NBER Working Paper*, No. 6570

Rotemberg and Woodford (1998) "An Optimization-Based Econometric Framework for the Evaluation of Monetary Policy: Expanded Version", *NBER Technical Working Paper*, No. 233

Tetlow and von zur Muehlen (1999) "Simplicity Versus Optimality: the choice of monetary policy rules when agents must learn", *Board of Governors of the Federal Reserve System Finance and Economics Discussion Series*, No. 1999-10

9.4 An optimising framework

The paper by Rotemberg and Woodford (1998) pulls together many of the theoretical issues we have studied so far. At the heart of their framework is a dynamic IS curve based on intertemporal maximisation and an aggregate supply curve based on the sticky prices in the New Keynesian Phillips curve. Rather than assuming a cash-in-advance constraint and facing the problems of generating substantial liquidity effects, they jump directly to a formulation in which the instrument of monetary policy is the interest rate itself. The central bank is assumed to set interest rates as a function of past rates, current and past inflation and current and past values of the output gap. In other words, interest rates are set according to a rule of the form in equation (9.1).

$$r_t = r^* + \sum_{k=1}^{n_r} \mu_k (r_{t-k} - r^*) + \sum_{k=0}^{n_\pi} \phi_k (\pi_{t-k} - \pi^*) + \sum_{k=0}^{n_y} \theta_k (y_{t-k} - y^*) \tag{9.1}$$

In this equation the central bank is constrained to react linearly to current and past economic variables. However, as we will see later this is not a binding constraint because rules such as (9.1) (subject to including a sufficient number of lags) encompass the optimal monetary policy.

Rotemberg and Woodford (1998) have a neat way of calibrating their model. They wish to calibrate/estimate their theoretical model to match the variances of the three endogenous variables: interest rates, inflation and output. The trick is to recognise that these variances can be completely described by (i) the variances of the fundamental disturbances and (ii) the impulse responses of the three variables to the fundamental disturbances. Hence, instead of matching variances directly, they match the impulse response functions of the structural model to those estimated using a simple structural vector autoregression. This approach has the considerable advantage that shocks in the model have effects similar to those observed in the data, for example Figure 9.1 shows the estimated and theoretical responses of output to a monetary policy shock.

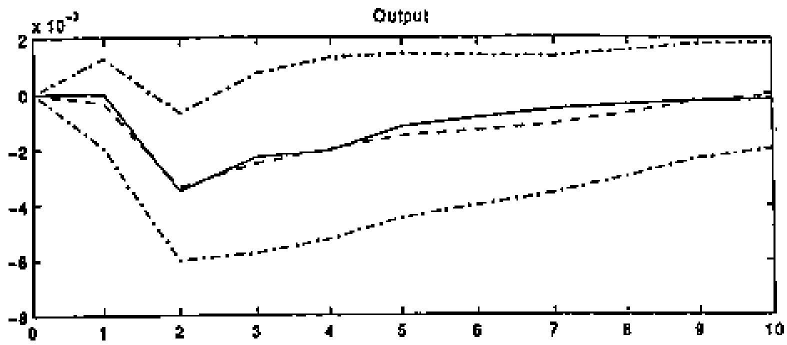


Figure 9.1: Responses of output to a monetary policy shock in the model and data.

The dynamic IS curve, the aggregate supply curve and the interest rate rule (9.1) completely describe the structure of the economy. However, in order to decide on what is the best monetary policy, i.e. what the μ, ϕ and θ coefficients in the monetary policy rule should be, we need to define some target for what monetary policy is trying to achieve. Traditionally, it has been assumed that monetary policy should minimise some quadratic loss function (9.2), where losses are caused by inflation or output being away from their respective targets. χ is the relative weight placed on inflation variability relative to output variability.

$$\mathcal{L} = \sum_{t=0}^{\infty} \beta^t \left[(y_t - y^*)^2 + \chi (\pi_t - \pi^*)^2 \right] \tag{9.2}$$

This is not a completely satisfactory objective for a central bank. If the model is micro-founded on the basis of a representative optimising agent then maximising the welfare of that agent is the appropriate target for central bank policy. In an important innovation, Rotemberg and Woodford (1998) do exactly this and derive the central bank's objective function directly from the welfare of the representative agent. They begin with the utility function (9.3), in which welfare depends on the expected utility of consumption c_t and the expected disutility of the work needed to produce output y_t in each state of the world z.

$$\mathcal{W} = E \left\{ u(c_t) - \int_0^1 v(y_t(z))dz \right\} \tag{9.3}$$

Rotemberg and Woodford (1998) take a second order Taylor series approximation of this to give an objective function of the form (9.4).

$$\mathcal{W} = -\Omega \left(L + \pi^{*^2} \right)$$

$$L = var(\pi_t) + \psi^{-1} var(\pi_t - E_{t-2}\pi_t) + \Lambda var \left\{ E_{t-1}(\hat{y}_t - y_t^s \right\} \tag{9.4}$$

This is similar in structure to the *ad hoc* quadratic loss function (9.2), although the match is not perfect. If the central bank did minimise (9.2) then there is no guarantee that the welfare of the representative agent (9.4) would be maximised. In this framework the problem of the central bank is to choose the monetary policy which maximises the welfare of the representative agent (9.4) subject to the dynamic IS curve and the New Keynesian aggregate supply curve. The objective (9.4) is quadratic and the constraints can be log-linearised so the problem is in the standard form of linear-quadratic control. For this type of problem, it is possible to show that the optimal monetary policy is indeed of the form shown in Equation (9.1), with interest rates reacting linearly to interest rates, inflation and the output gap. In other words, Equation (9.1) does describe the optimal policy and the problem of the central bank reduces to that of finding the correct coefficients μ, ϕ and θ and the correct lags n_r, n_π and n_y for the monetary policy rule.[1]

9.5 Optimal monetary policy

The full optimal rule as calculated by Rotemberg and Woodford (1998) is shown in Equation (9.5).

[1] For more details of how this is done using the Matrix Riccati equation see Chow (1980), pp. 156-160.

$$
\begin{aligned}
\hat{r}_t &= 0.29\hat{r}_{t-1} + 0.42\hat{r}_{t-2} + 1.28\hat{r}_{t-3} + 0.22\hat{y}_{t-1} - 0.25\hat{y}_{t-2} + \cdots \\
&\quad +0.16\hat{\pi}_t + 1.00\hat{\pi}_{t-1} + 2.45\hat{\pi}_{t-2} - 1.45\hat{\pi}_{t-3} + 0.74\hat{\pi}_{t-4} \\
&\quad -0.08\hat{\pi}_{t-5} + 0.25\hat{\pi}_{t-6} + 0.33\hat{\pi}_{t-7} + 0.23\hat{\pi}_{t-8} + 0.25\hat{\pi}_{t-9} \\
&\quad +0.19\hat{\pi}_{t-10} + 0.17\hat{\pi}_{t-11} + 0.13\hat{\pi}_{t-12} + 0.19\hat{\pi}_{t-13} + 0.06\hat{\pi}_{t-14} + \cdots
\end{aligned}
\tag{9.5}
$$

The omitted terms in $\hat{y}_{t-j}$ are all of size smaller than 0.01 and the omitted terms in $\hat{\pi}_{t-j}$ are all of size 0.03 or smaller. If coefficients of 0.0001 were to be included then the optimal policy would need 18 lags of output and 49 lags of inflation. Such complexity raises obvious questions about the operational feasibility of fully optimal monetary policy rules such as (9.5).

9.6 Simple monetary policy rules

In contrast to the complex monetary policy rules outlined above, Taylor (1993) suggested a very simple rule for setting interest rates (9.6).

$$
i_t = 0.5y_t + 1.5\pi_t \tag{9.6}
$$

This paper has been enormously influential, generating a whole industry of research. As Taylor himself comments, it is simple enough to put on the back of a business card. Interest rates are set according to the current output gap and inflation. Clearly such a rule is not optimal so why would a central bank ever want to follow this rule? One answer is that it may be a good approximation to the optimal rule. It is certainly a fair approximation to the behaviour of the Federal Reserve Board over the past forty years. Figure 9.2 shows the path of the actual Federal funds rate (the key instrument for monetary policy in the US) and the rate that would have prevailed if monetary policy had followed the Taylor rule. The two series match each other quite well, especially bearing in mind the considerable difficulties involved in obtaining a reasonable estimate of the current output gap. This evidence suggests that the Taylor rule is a good approximation of the behaviour of the Fed. Whether that implies optimality depends on whether the Fed is believed to have acted optimally over the period.

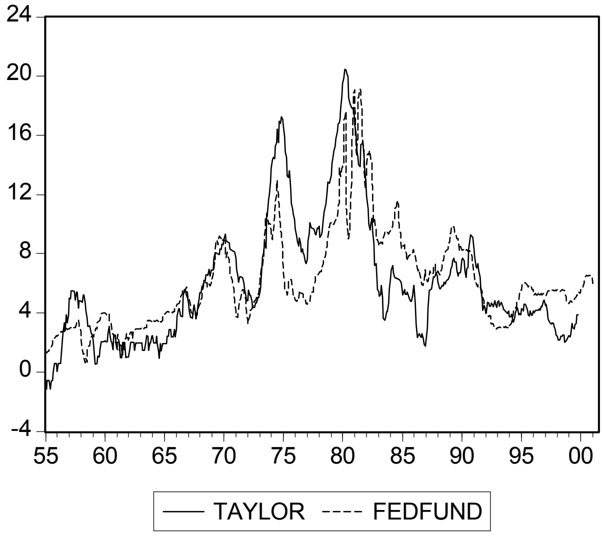


Figure 9.2: The Federal funds rate and Taylor rule based interest rate.

In addition to matching the behaviour of the Federal Reserve Board, simple rules have several other desirable features. We will consider two possibilities: robustness and transparency.

9.7 Robustness

In optimal control theory, policies are distinguished between the good, the bad and the optimal. The complex rule calculated for the Rotemberg-Woodford economy is an example of an optimal rule, the question remains whether it is a good or bad rule. In the context of the Rotemberg-Woodford framework it is a good rule - in fact the best by definition. However, if the economy is not well characterised by the Rotemberg-Woodford model (maybe the economy is not New Keynesian after all) then the optimal policy could turn out to be a very bad policy. What the policy maker would really like is a policy rule which performs well across a wide range of models and in a variety of situations. Such a rule would be robust to errors made in the specification of the underlying model. A recent paper by Levin, Wieland and Williams (1998) assesses whether the Taylor rule may be a good rule in this sense. They analyse the performance of a modified Taylor rule against more complex rules across three different underlying models (the Taylor Multi-Country model (TMCM), the MSR model of Orphanides and Wieland, and the FRB staff model). Their first step is to calculate the optimal weights for a modified Taylor rule in the FRB model, resulting in the rule (9.7).

$$r_t = r_{t-1} + 0.8(\pi_t - \pi^*) + 1.0y_t \tag{9.7}$$

This rule is then applied to each model in turn and compared to what could have been achieved if the Taylor rule had been fine-tuned to each model. The results are shown in Table 9.1.

Model	Loss with Taylor rule fine-tuned to model	Loss with Taylor rule fine-tuned to FRB model	Difference
TMCM	3.61	3.64	0.03
MSR	0.33	0.33	0.00
FRB	2.02	2.02	0.00

Table 9.1: Robustness of Taylor rules

The results indicate that the Taylor rule gives a robust performance in the face of specification error. Even the worst-case scenario, a Taylor rule tuned to the FRB model when the correct model is the Taylor Multi-Country model (TMCM), does not lead to large welfare losses. We now turn our attention to the extent to which more complex rules are similarly robust to model misspecification. The computational burden of calculating such rules is considerable so Levin, Wieland and Williams (1998) restrict themselves to 12-parameter or 8-parameter rules. To give an example, they calculate a 12-parameter rule optimised for the MSR model and apply it to both the FRB and TMCM models. The results are shown in Table 9.2.

Model	Loss with optimal rule fine-tuned to MSR model
TMCM	3.78
MSR	0.16
FRB	1.92

Table 9.2: Robustness of optimal rules

Fine-tuning the monetary policy rule to the MSR model leads to a significant reduction in loss if the MSR model is the correct specification. The loss falls from 0.33 in Table 9.1 to 0.16 in Table 9.2. However, if there is misspecification and the TMCM is really the correct model then this optimal rule performs worse than the Taylor rule. The loss in Table 9.1 under the FRB Taylor rule was only 3.64 but the loss with the optimal MSR rule rises to 3.78 in Table 9.2. In this sense the optimal policy is not robust to model misspecification. It is bad rather than good.

The analysis of Levin, Wieland and Williams (1998) has been criticised because the models it considers are all rather similar. In such an analysis it is not surprising that generally *all* monetary policy rules are quite robust. A more satisfactory approach is to formally search for robust policies across well-defined misspecifications. According to robust control theory, the best policy is a min-max strategy, which effectively entails choosing the policy action which gives the least costly (i.e. leads to the smallest deviation of inflation from target) worst outcome, given an assumption about the range of misspecifications that may occur. Recent work by Hansen and Sargent (2001) and others has developed these ideas more fully.

9.8 Transparency

One further advantage of a simple monetary policy rule is that it is easily understood. Private agents are able to quickly learn the policy of the central bank, an issue examined by Tetlow and von zur Muehlen (1999). They consider rules of the type (9.8) in a Fuhrer-Moore type model.

$$r_t = rr^* + E_{t-1}\pi_t + \beta_R(R_{t-1} - \pi_{t-1}) + \beta_\pi(\pi_{t-1} - \pi^*) + \beta_y y_{t-1} + u_t \tag{9.8}$$

A simple Taylor rule is equal to equation (9.8) with the added restrictions $\beta_R = 0$ and $\beta_\pi = \beta_y = 0.5$. This rule is then compared with optimal 2-parameter rules (with $\beta_R = 0$) and optimal 3-parameter rules. In steady-state there is full information and rational expectations so private agents know the parameters of the rule in force. For this case the expected losses are given in Table 9.3, where losses have been normalised by the loss under the optimal 3-parameter rule.

Rules	Loss
Taylor	1.58
Optimal 2-parameter	1.02
Optimal 3-parameter	1.00

Table 9.3: Steady-state losses under different policy rules

In steady-state it is obvious which rule is the best: the optimal rule is optimal. Now, though, consider a central banker who inherits one of the policy rules. Is it profitable to switch to a new policy rule which gives better steady-state performance? The answer depends on how quickly private agents are able to learn the new coefficients in the policy rule and the costs incurred during the learning process. Learning is not instantaneous since there is always some random noise u_t in the policy rule which disguises the policy maker's true intended actions. Tetlow and von zur Muehlen (1999) report simulation evidence for the welfare cost of switching between different monetary policy rules. Table 9.4 highlights some of their results.

Switch	Percentage change in loss
Taylor $\Rightarrow$ Optimal 2-parameter	- 18 %
Taylor $\Rightarrow$ Optimal 3-parameter	- 13 %
Optimal 2-parameter$\Rightarrow$ Optimal 3-parameter	+ 4 %

Table 9.4: Welfare consequences of changing the monetary policy rule

The simulations show that it is always beneficial to switch away from the Taylor rule. However, there are larger gains in terms of reduced welfare loss in switching to an optimal 2-parameter rule than an optimal 3-parameter rule. This is because the 2-parameter rule is easier for private agents to learn. Even though in steady-state the 3-parameter rule does better, the costs of introducing it (due to private agents having to go through a protracted learning process) are higher than for introducing the 2-parameter rule. Similarly, if the central bank inherits a 2-parameter rule then there are no incentives to switch to the 3-parameter rule, even though it is better in the long run. Such results may explain why simple rules have gained so much favour.

Chapter 10

Monetary policy under uncertainty

10.1 Motivation

In recent times it has become increasingly common for central banks to acknowledge that the do not have perfect information about the structure of they economy they are attempting to control. There may be uncertainty surrounding the precise values of the key parameters of the model or, at a deeper level, there may be fundamental uncertainties regarding which is the correct model. In this lecture we concentrate on parameter uncertainty and discuss how the nature of monetary policy changes when such uncertainty is formally accounted for. The models we use will be static and quite stylised to highlight the precise mechanisms in action.

10.2 Key readings

Brainard (1967) "Uncertainty and the Effectiveness of Policy", *American Economic Review Papers and Proceedings*, 57, 411-425.

Sack (2000) "Does the fed act gradually? A VAR analysis", *Journal of Monetary Economics*, 46, 229-256.

10.3 Related reading

Shellekens (2000) "Caution and Conservatism in the Making of Monetary Policy", *European Central Bank Working Paper*, No. 25

Theil (1958) *Economic Forecasts and Policy*, North-Holland Amsterdam

Tinbergen (1952) *On the Theory of Economic Policy*, North-Holland Amsterdam

10.4 Certainty equivalence

If the uncertainty faced by the central bank takes a particularly simple form then the optimal policy of the central bank is to behave "as if" everything was known with certainty. This will typically be the case if the only source of uncertainty is an additive error term. Equation (10.1) describes a monetary transmission mechanism in which inflation π is determined by the interest rate i through the known coefficient b, where $b < 0$. u is an i.i.d. error term with mean zero and variance σ_u^2.

$$\pi = bi + u \tag{10.1}$$

The central bank is assumed to have a quadratic loss function (10.2), which penalises the deviation of inflation from a target level π^*.

$$\mathcal{L} = (\pi - \pi^*)^2 \tag{10.2}$$

The timing of the model is such that the central bank has to set the interest rate i before the error term u is revealed. In other words, the central bank does not know the true state of the world when it moves and so it has to set interest rates to minimise the expectation of the loss in (10.2). Substituting from (10.1) into (10.2) we can write

$$\begin{aligned} \mathcal{L}^e &= E\left[(bi + u - \pi^*)^2\right] \\ &= b^2 i^2 + E(u^2) + {\pi^*}^2 + 2biE(u) - 2bi\pi^* - 2\pi^* E(u) \end{aligned}$$

Note that the expectation operator only continues to apply to terms in u because b, i and π^* are all known by the central bank at the time the decision is taken. From the definition of u as a random error term, we also have $E(u^2) = \sigma_u^2$ and $E(u) = 0$ so the expected loss can be expressed as in (10.3).

$$\mathcal{L}^e = b^2 i^2 + \sigma_u^2 + {\pi^*}^2 - 2bi\pi^* \tag{10.3}$$

The central bank chooses i to minimise this expected loss and so derives the optimal policy under certainty equivalence (10.4).

$$i = \frac{\pi^*}{b} \tag{10.4}$$

This policy is completely independent of the uncertainty surrounding the error term u. It is as if the central bank has completely ignored the uncertainty and set policy such that the inflation target is met in expectation, i.e. $\pi^e = \pi^*$. In the literature this is known as the "certainty equivalence principle". When uncertainty is additive, as here in the case of a simple additive error term, the central bank can ignore the uncertainty and set policy as if everything was known with certainty. The result was first proposed by Theil (1958) and Tinbergen (1952). However, as we will see, the conditions under which it holds are quite restrictive and it is not really applicable to most real-world situations of interest.

10.5 Brainard conservatism

The paper by Brainard (1967) shows that certainty equivalence no longer holds for more complex specifications of uncertainty. More specifically, if there is uncertainty about the parameters of the model then the central bank should not behave as if the uncertainty does not exist, a result described some thirty years later by Alan Blinder as the "Brainard uncertainty principle". The key difference is that uncertainty about a parameter is multiplicative rather additive uncertainty: the more a policy is used the more that the uncertainty is multiplied into the system. To see how this changes the nature of the optimal policy consider the monetary transmission mechanism with parameter uncertainty (10.5).

$$\begin{aligned} \pi &= bi + u \\ b &\sim (\hat{b}, \sigma_b^2) \end{aligned} \tag{10.5}$$

The first part of equation (10.5) is identical to that in (10.1) in the discussion of certainty equivalence except now there is uncertainty about the parameter b. However, although the central bank does not know the precise value of b, it does know the distribution from which it is drawn, i.e. it knows its mean $\hat{b}$ and variance σ_b^2. There are many reasons why this might be a reasonable description of the central bank's knowledge of the monetary transmission mechanism. It could be that the central bank has poor information about how the transmission mechanism works, for example the current state of the ECB in Euroland. Alternatively, there may be fundamental uncertainties in the transmission of monetary policy which preclude ever being able to predict with certainty what the effect of interest rates on inflation is. The structure of the stylised economy (10.5) is shown in Figure 10.1. The central straight line shows the relationship $\pi = \hat{b}i$, which holds in expectation. The curved lines are confidence bands showing the range of inflation that is expected for given interest rates.

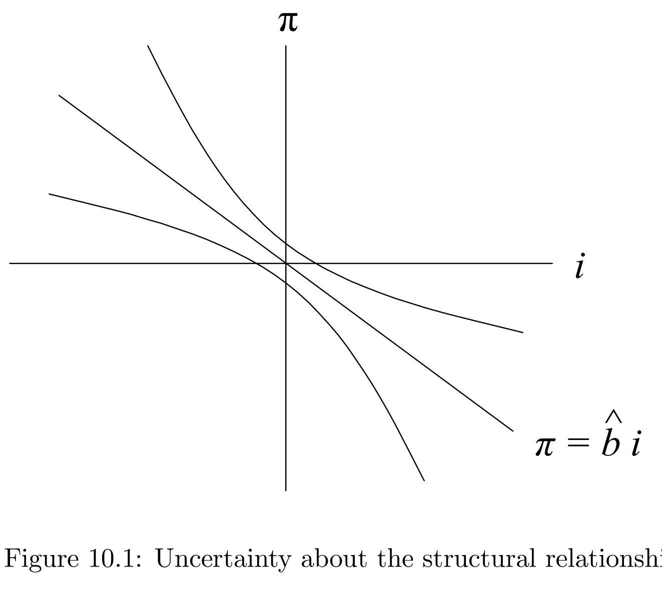


Figure 10.1: Uncertainty about the structural relationship.

Figure 10.1 shows how the parameter uncertainty is multiplicative. As interest rates are moved further away from zero it becomes increasingly more difficult to predict the level of inflation. Uncertainty would be minimised with a zero interest rate, at which the only uncertainty is additive due to the error term, but then expected inflation would be zero and not equal to target. Mathematically, the expected loss with parameter uncertainty is given by

$$\begin{aligned}\mathcal{L}^e &= E\left[(bi + u - \pi^*)^2\right] \\ &= E(b^2)i^2 + E(u^2) + {\pi^*}^2 + 2E(bu)i - 2E(b)i\pi^* - 2\pi^* E(u)\end{aligned}$$

Again, the definition of u as a random error term implies $E(u^2) = \sigma_u^2$ and $E(u) = 0$. The mean of b gives $E(b) = b$ and the variance of $\hat{b}$ can be written as $\sigma_u^2 = E(b - \hat{b})^2 = E(b^2) - E(\hat{b}^2)$, which gives an expression for $E(b^2)$. By making the further simplifying assumption that uncertainties about b and u are unrelated, in other words $E(bu) = 0$, the expected loss can be described by equation (10.6).

$$\mathcal{L}^e = \sigma_b^2 i^2 + \sigma_u^2 + (\hat{b}i - \pi^*)^2 \tag{10.6}$$

The optimal policy (10.7) is derived by differentiating (10.6) with respect to the interest rate i.

$$i = \frac{\hat{b}\pi^*}{\hat{b}^2 + \sigma_b^2} \tag{10.7}$$

This policy differs from the certainty equivalent policy (10.4) by the extra variance term σ_b^2 in the denominator. The presence of parameter uncertainty means that optimal policy is more cautious. $\left|i^{Brainard}\right| \leq \left|i^{Certainty.Equivalence}\right|$ implies that interest rates are closer to zero under the Brainard policy than under certainty equivalence. The reason is that additional caution reduces the amount of uncertainty that policy introduces into the system. At the extreme, as $\sigma_b^2 \to \infty$ and parameter uncertainty becomes infinite, the optimal policy is to do nothing and set interest rates to zero, $i \to 0$. In contrast, parameter uncertainty disappears as $\sigma_b^2 \to 0$ and the interest rate is set equal to that under certainty equivalence. This result is often referred to as Brainard conservatism - parameter uncertainty introduces a motive for caution in optimal policy. Such a policy means that the central bank does not expect to achieve its inflation target, i.e. $\pi^e \neq \pi^*$. The reason for this is that aiming to hit the target exactly involves large potential losses, especially if the parameter b turns out to be high and the monetary transmission mechanism is more potent than expected.

10.6 Is Brainard uncertainty empirically relevant?

Whether Brainard uncertainty is a useful concept to explain the behaviour of the Federal Reserve Board is the subject of a study by Sack (2000). He models the structure of the economy using a five dimensional vector autoregression in industrial production growth, unemployment, consumer price inflation, commodity price inflation (to control for price puzzles) and the federal funds rate. Structure is imposed by a recursive Choleski ordering in which the federal funds rate ordered last. Assuming the model is correctly identified, the first four equations in the model describe the structural form of the economy whilst the last equation is the estimated policy function of the fed.[1] The policy maker is assumed to minimise a present-value quadratic loss function (10.8)

$$\mathcal{L} = \frac{1}{2}E_t\left\{\sum_{i=1}^{\infty}\beta^i\left[(\pi_{t+1}-\pi^*)^2+\lambda_u(u_{t+1}-u^*)^2\right]\right\} \tag{10.8}$$

where λ_u denotes the relative importance of the deviations of unemployment and inflation from their respective targets.

After estimating the model with OLS, Sack (2000) calculates a "certainty equivalent" policy rule by assuming that the point estimates from the VAR are the true values known with certainty. This rule will be linear in the past values of the variables in the system since it is a standard linear-quadratic control problem. The coefficients depend on preferences λ_u and the point estimates of the VAR coefficients. This rule is then compared to a "Brainard" policy rule which takes into account that the point estimates of the VAR are uncertain. The standard OLS errors of the parameter estimates are used as an indicator of the uncertainty surrounding each parameter. The new optimal rule is still a linear function of past values of variables in the system but now the coefficients of the rule depend on both the point estimates of the VAR and the variance-covariance matrix of the coefficient estimates.

Sack (2000) calculates the impulse response functions implied by the optimal rules with and without allowance for parameter uncertainty and compares these to those estimated purely from the data. He finds that the optimal policy rule taking parameter uncertainty into account is closer to the actual behaviour of the federal funds rate than an optimal policy disregarding parameter uncertainty. Figure 10.2 shows how the optimal policy with parameter uncertainty tracks the federal funds rate better, suggesting that caution induced by Brainard uncertainty is quantitatively important.

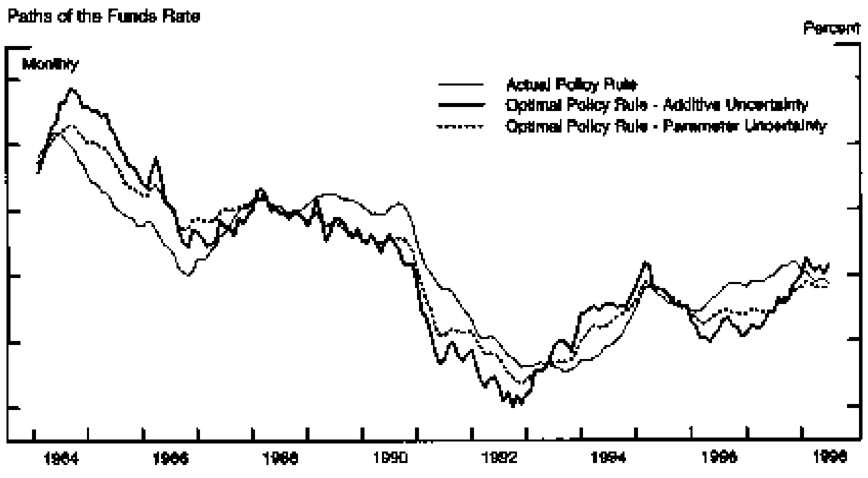


Figure 10.2: The interest rate under actual and optimal policies

[1] Because this is not a true structural model based on microfoundations it is questionable whether the results are robust and not subject to Lucas critique problems.

Both of the optimal policy rules exhibit considerable persistence in interest rates. This is known as "interest rate smoothing" and implies that the fed should adjust interest rates gradually rather than aggressively. Table 10.1 shows simulation evidence for the persistence of interest rate changes under the two optimal policies and compares these to the actual estimated behaviour of the federal funds rate.

Federal funds rate	$\Delta i = -0.02 + 0.46\Delta i_{-1} + 0.01\Delta i_{-2}$
Optimal policy with parameter uncertainty	$\Delta i = -0.02 + 0.10\Delta i_{-1} + 0.22\Delta i_{-2}$
Optimal policy without parameter uncertainty	$\Delta i = -0.04 + 0.15\Delta i_{-1} + 0.09\Delta i_{-2}$

Table 10.1: Persistence structure of interest rates

The optimal rule with parameter uncertainty comes closest to matching the persistence observed in the federal funds rate, giving further support for the claim that Brainard uncertainty and caution are an important feature of the data and do help to explain the behaviour of the Federal Reserve Board. However, the degree of interest rate smoothing in the optimal rule with parameter uncertainty still falls short of that actually observed in the federal funds rate so there must be some additional explanation for the smoothness of interest rates.

Chapter 11

Learning

11.1 Motivation

In all of the lectures so far we have assumed that agents know everything it is possible to know about the structure of the economy. Even in the models of monetary policy under uncertainty, agents knew the final distribution of the uncertain parameters. In this lecture we relax this assumption and look at models where agents are able to learn the key parameters of the model. We begin by examining how the nature of monetary policy is affected by learning considerations and then continue in the second part of the lecture to examine how the equilibrium which emerges in the economy can be different under learning.

11.2 Key readings

There is no required reading for today's lecture. The examples we will discuss are simplified versions of the following two papers.

Ellison and Valla (2001) "Learning, uncertainty and central bank activism in an economy with strategic interactions", *Journal of Monetary Economics*, 48, 153-171

Marcet and Nicolini (1998) "Recurrent hyperinflations and learning", *CEPR Discussion Paper*, No. 1875

11.3 Related reading

Balvers and Cosimano (1994) "Inflation variability and gradualist monetary policy", *Review of Economic Studies*, 61, 721-738

Bertocchi and Spagat (1993) "Learning, experimentation and monetary policy", *Journal of Monetary Economics*, 32, 169-183

Evans and Honkapohja (2001) *Learning and Expectations in Macroeconomics*, Princeton University Press.

Marimon and Sunder (1993) "Indeterminacy of equilibria in a hyperinflationary world: experimental evidence", Econometrica, 61, 1073-1108

Wieland (2000) "Learning by doing and the value of optimal experimentation", *Journal of Economic Dynamics and Control*, 24, 501-534

11.4 Learning and monetary policy

In the previous lecture we saw that the presence of uncertainty lead to a cautious policy by the central bank. This call for Brainard conservatism has been challenged in a number of papers, originally by Bertocchi and Spagat (1993) and Balvers and Cosimano (1994). They argue that a cautious policy is suboptimal because it is very poor from a learning point of view. If the central bank is cautious in its use of policy then it will be very difficult to learn what the effects of monetary policy are. In terms of learning, the central bank should be more aggressive in policy since that way it learns the key parameters about how the policy works. To see how this argument works in practice we will look at a stylised example, which is a simplification of Ellison and Valla (2001).

The structure of the economy is characterised by the Phillips curve (11.1), where output y_t depends on inflation π_t and two supply shocks ε_{1t} and ε_{2t}. Only the first supply shock is observable. Both shocks have the same variance σ_ε.

$$y_t = \beta\pi_t + \varepsilon_{1t} + \varepsilon_{2t} \tag{11.1}$$

The Phillips curve parameter is not known with certainty by the central bank. For simplicity, we will assume it can take one of two possible values, β_H or β_L. The central bank believes that it is high with probability p_t and low with probability $1-p_t$. The loss function of the central bank (11.2) is assumed to punish the expected present discounted value of deviations of output from natural rate and inflation from zero.

$$\mathcal{L} = E_t\sum_{i=0}^{\infty}\delta^n\left[(y_{t+i}-y^*)^2 + \chi\pi_{t+i}^2\right] \tag{11.2}$$

δ is the subjective discount factor and χ is the weight placed upon inflation as against output deviations from target. We begin by solving the model for the case where the central bank does not internalise its learning. In this passive learning case the maximisation problem of the central bank reduces to the static problem (11.3). There are no intertemporal linkages in the model and the central bank just minimises losses period by period.

$$\min_{\pi_t} E_t\left\{(\beta\pi_t + \varepsilon_{1t} + \varepsilon_{2t} - y^*)^2 + \chi\pi_t^2\right\} \tag{11.3}$$

To calculate the expected loss the central bank uses its beliefs about the parameter β so the problem becomes (11.4).

$$\min_{\pi_t}\left\{p_tE_t(\beta_H\pi_t + \varepsilon_{1t} + \varepsilon_{2t} - y^*)^2 + (1-p_t)E_t(\beta_L\pi_t + \varepsilon_{1t} + \varepsilon_{2t} - y^*)^2 + \chi\pi_t^2\right\} \tag{11.4}$$

The associated optimal policy is (11.5).

$$\pi_t = -\frac{p_t\beta_H^2 + (1-p_t)\beta_L^2}{p_t\beta_H^2 + (1-p_t)\beta_L^2 + \chi}\varepsilon_{1t} \tag{11.5}$$

This policy is equivalent to the Brainard policy derived in the previous lecture. The central bank internalises the uncertainty about β in its policy and adjusts its reaction to the observed supply shock accordingly. Figure 11.1 shows this policy for a numerical example based on Ellison and Valla (2001). The upper dashed line shows the inflation choice of the central bank after a one standard deviation negative ε_{1t} supply shock, as a function of beliefs. Whatever the central bank's beliefs about the parameter β it always increases inflation to reflate the economy after a negative supply shock. The lower dashed line shows the corresponding inflation choice for a positive ε_{1t} supply shock.

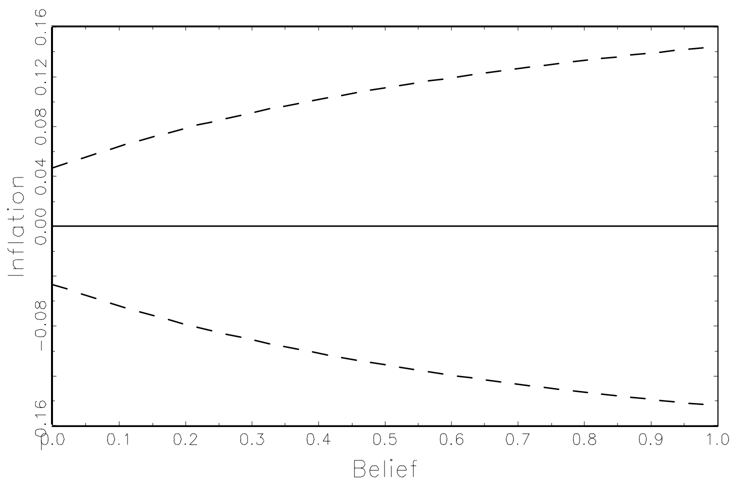


Figure 11.1: Passive learning policy

We now calculate the policy which explicitly takes learning into account. The central bank has several pieces of information which it can use to improve its inference about the actual value of the parameter β. At the end of the period it knows the observable supply shock ε_{1t}, its own inflation choice π_t and the outcome in terms of output y_t. The question is whether y_t has most likely been generated by a model with β_H or β_L. The answer to this lies in comparing the likelihood of outcome y_t in the β_H case (where the distribution of y_t is given by equation (11.6)) to the likelihood in the β_L case (where the distribution of y_t is given by equation (11.7)).

$$y_t\,|\beta_H \sim N\left[\beta_H\pi_t + \varepsilon_{1t}; \sigma_\varepsilon\right] \tag{11.6}$$

$$y_t\,|\beta_L \sim N\left[\beta_L\pi_t + \varepsilon_{1t}; \sigma_\varepsilon\right] \tag{11.7}$$

Comparing the distributions (11.6) and (11.7) to the actual outcome gives information on whether β really is high or low. To see how this feeds into beliefs we use Bayes rule. For this reason the learning in this section is often referred to as Bayesian learning. Equation (11.8) shows how Bayes rule can be applied to our problem.

$$
\begin{aligned}
P(\beta_H \,|y_t\,) &= \frac{P(\beta_H \cap y_t)}{P(y_t)} = \frac{P(\beta_H)P(y_t\,|\beta_H\,)}{P(y_t)} \\
p_{t+1} &= \frac{p_t P(y_t\,|\beta_H\,)}{p_t P(y_t\,|\beta_H\,) + (1-p_t)P(y_t\,|\beta_L\,)} \\
p_{t+1} &= \mathcal{B}(p_t, \pi_t, \varepsilon_{1t}, y_t)
\end{aligned}
\tag{11.8}
$$

$\mathcal{B}$ is the Bayesian operator showing how beliefs p_t are updated on the basis of all relevant information. If the central bank internalises the learning mechanism (11.8) then it is said to be active learning. The problem of the active learning central bank is shown in (11.9).

$$
\begin{aligned}
\underset{\pi_t}{\mathcal{L}} &= E_t \sum_{i=0}^{\infty} \delta^n \left[(y_{t+i} - y^*)^2 + \chi \pi_{t+i}^2\right] \\
& \quad \text{s.t.} \\
y_t &= \beta \pi_t + \varepsilon_{1t} + \varepsilon_{2t} \\
p_{t+1} &= \mathcal{B}(p_t, \pi_t, \varepsilon_{1t}, y_t) \\
p_t &= p_0
\end{aligned}
\tag{11.9}
$$

This is an example of a non-linear-quadratic control problem and so will not have an optimal solution that is linear. For such a problem it is not possible to obtain closed-form analytical solutions but it is possible to use numerical techniques to derive an approximation to the optimal policy. We follow the approach of Wieland (2000) and show our numerical results in Figure 11.2.

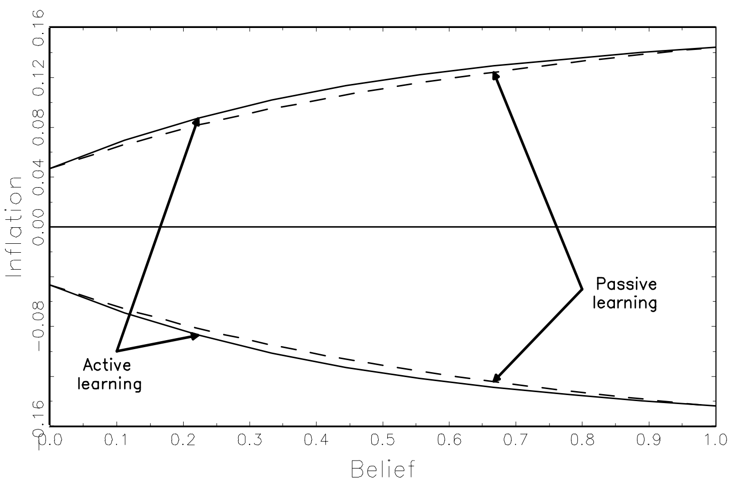


Figure 11.2: Optimal and passive learning policies

Compared with the passive learning policy, the reaction under the active learning policy is stronger, being further away from zero. The active learning policy is more aggressive than the passive learning policy, precisely because this has informational gains which help learning. This result has been used to suggest that central banks should engage in probing or experimentation - adjusting monetary policy to find out how the economy works.

11.5 Learning as an equilibrium selection device

The paper by Marcet and Nicolini (1998) shows how learning has important implications for the equilibrium outcome in the economy. They begin with a simple overlapping generations model in which each generation lives for two periods. When the agent is young (period 0) they receive an endowment of real value w_t^0, save M_{t+1}^d in the form of cash for old age and therefore consume $C_t^0 = w_t^0 - M_{t+1}^d/p_t$. When the agent is old (period 1) they receive endowment w_t^1 and use their money savings to give consumption $C_t^1 = w_t^1 + M_{t+1}^d/p_{t+1}$. With a logarithmic utility function, the decision problem faced by the agent when young is shown by (11.10).

$$\begin{aligned}&\max E\left[\log(C_t^0) + \log(C_t^1)\right] \\ &\text{s.t.} \\ &C_t^0 = w_t^0 - M_{t+1}^d/p_t \\ &C_t^1 = w_t^1 + M_{t+1}^d/p_{t+1}\end{aligned} \tag{11.10}$$

After substituting in for C_t^0 and C_t^1 and taking the first order condition we can derive the demand for money function (11.11), where $a = w_t^0/2$ and $b = w_t^1/2$.

$$\frac{M_{t+1}^d}{p_t} = a - b\pi_{t+1}^e \tag{11.11}$$

The amount of money demanded by the young in this overlapping generations model depends on the pattern of endowments and the expectation of future inflation. In the remainder of this section the focus will be on how the expectation of future inflation is formed - whether everything is already known or does the agent learn things useful for calculating the expectation. We complete the model by specifying a process for the money supply. For simplicity it is assumed that the real government deficit d is completely financed by changes in the money supply so the government budget constraint is given by equation (11.12).

$$M_{t+1}^s - M_t^s = d \tag{11.12}$$

In equilibrium, money demand (11.11) equals money supply (11.12) so we can write the determination of inflation as equation (11.13).

$$\pi_t = \frac{a - b\pi_t^e}{a - b\pi_{t+1}^e - d} \tag{11.13}$$

This is the key equation of the model. It shows how inflation is determined by expectations of current and future inflation. We now assume perfect-foresight rational expectations so $\pi_t^e = \pi_t \forall t$, in which case equation (11.13) becomes an expression showing how inflation evolves. In other words, we see how π_{t+1} depends on π_t as shown in equation (11.14).

$$\pi_{t+1} = R(\pi_t) = \frac{1}{b}\left[a - d + b - \frac{a}{\pi_t}\right] \tag{11.14}$$

There are two things to note about equation (11.14). Firstly, $R'(\pi_t) > 0$ and $R''(\pi_t) < 0$, implying the function is concave. Secondly, the equation has two stationary fixed points at which $\pi^* = R(\pi^*)$. These two properties imply that the mapping $R(\cdot)$ from π_t to π_{t+1} can be described as in Figure 11.3.

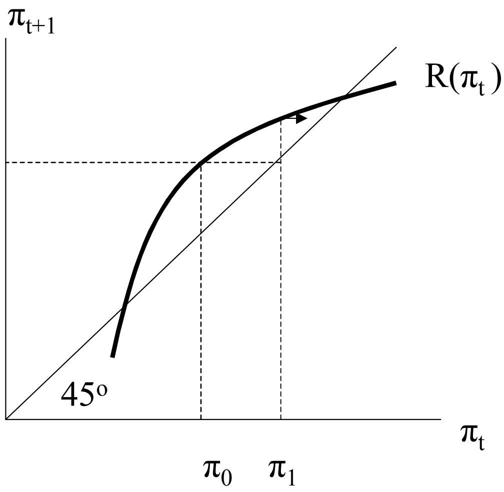


Figure 11.3: Dynamics of inflation under rational expectations

Any point on the $R(\cdot)$ mapping is a rational expectations equilibrium but only two are stationary. To see the dynamic behaviour of this model we can look at the path followed by inflation over time. Suppose that the economy starts out at π_0. Next period's inflation will be given by π_1 and so on. In Figure 11.3 all roads lead to Rome - whatever the level of inflation above the lower equilibrium you start from you will always converge to the high inflation stationary equilibrium.

Convergence to high inflation equilibria is a common feature of rational expectations models. We can say that rational expectations selects the high inflation equilibrium from amongst the set of multiple equilibria available - in this sense it is an equilibrium selection mechanism. However, there are good reasons to believe that in reality we do not see convergence to the high inflation equilibrium. Experimental evidence from Marimon and Sunder (1993) suggests that the economy almost always converges on the low inflation equilibrium. To reconcile these doubts we now turn to learning as an equilibrium selection device. We retain the demand for money function (11.11) and the definition of the money supply (11.12) but relax the assumption of rational expectations and instead propose that agents use an adaptive learning rule (11.15) to form their expectations.

$$\pi^e_{t+1} = \pi^e_t + \alpha(\pi_{t-1} - \pi^e_t) \tag{11.15}$$

Under this rule, agents update their expectations of inflation on the basis of how much inflation in the previous period deviated from that currently expected. If we assume α is small, which will be the case if learning has already proceeded for a reasonable amount of time, $\pi^e_{t-1} \approx \pi^e_t$ in equation (11.13) and we can substitute out for π_{t-1} to arrive at equation (11.16).

$$\pi^e_{t+1} = (1-\alpha)\pi^e_t + \alpha \left[\frac{a - b\pi^e_t}{a - b\pi^e_t - d} \right] \tag{11.16}$$

We now have an equation for the evolution of inflation expectations $\pi^e_{t+1} = G(\pi^e_t)$. Under perfect foresight, we can write this as $\pi_{t+1} = G(\pi_t)$ and note two things about the $G(\cdot)$ mapping (11.17). Firstly, $G'(\pi_t) > 0$ and $G''(\pi_t) > 0$ so the function is convex. Secondly, it has two fixed points $\pi^* = G(\pi^*)$ which coincide with the fixed points of the rational expectations solution. The function is shown in Figure 11.4.

$$\pi_{t+1} = G(\pi_t) = \pi_t + \frac{\alpha d}{a - b\pi_t - d} \tag{11.17}$$

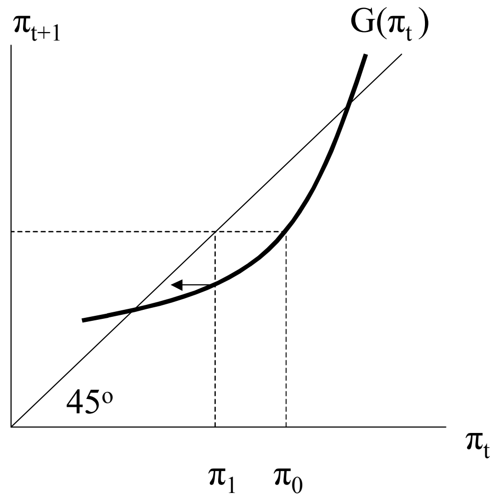


Figure 11.4: Dynamics of inflation under learning

The dynamics of inflation are now quite different from before. Under learning, inflation converges to its low stationary equilibrium instead of the high inflation equilibrium as in rational expectations. The robustness of these type of results has been analysed in many studies. Evans and Honkapohja (2001) is the state of the art textbook on this type of adaptive learning.