The material on these notes is borrowed from qed.econ.queensu.ca/pub/faculty/lloyd-ellis/econ320

RATIONAL EXPECTATIONS AND THE LUCAS CRITIQUE

Adaptive Expectations

A common specification for the adjustment of expectations, developed in 1960's by Milton Friedman and Edmund Phelps, is the adaptive expectations model. This model specifies the change in an agent's forecast of inflation as being equal to a fraction of the previous forecast error:

$$\pi_{t+1}^e - \pi_t^e = \lambda(\pi_t - \pi_t^e),$$

where $0 < \lambda < 1$. This can be written as

$$\pi_{t+1}^e = \lambda\pi_t + (1-\lambda)\pi_t^e$$

$$\pi_{t+1}^e = \lambda\pi_t + (1-\lambda)(\lambda\pi_{t-1} + (1-\lambda)\pi_{t-1}^e)$$

$$\pi_{t+1}^e = \lambda\pi_t + (1-\lambda)\lambda\pi_{t-1} + (1-\lambda)^2\pi_{t-1}^e$$

$$\pi_{t+1}^e = \lambda\pi_t + (1-\lambda)\lambda\pi_{t-1} + (1-\lambda)^2(\lambda\pi_{t-2} + (1-\lambda)\pi_{t-2}^e)$$

$$\pi_{t+1}^e = \lambda\pi_t + (1-\lambda)\lambda\pi_{t-1} + (1-\lambda)^2\lambda\pi_{t-2} + (1-\lambda)^3\pi_{t-2}^e$$

....

$$\pi_{t+1}^e = \sum_{i=0}^{T}\lambda(1-\lambda)^i\pi_{t-i} + (1-\lambda)^{T+1}\pi_{t-T}^e.$$

If T is large enough, the last term becomes close to zero. Thus, according to this hypothesis, expected inflation can be expressed as a weighted sum of past inflation rates, with the weights being greater for recent inflation and declining as we go further back. This lends certain plausibility to the adaptive expectations hypothesis. Moreover if, in the long—run, the actual inflation rate settles down to a constant value, then inflationary expectations will converge to that value.

However, the adaptive expectations hypothesis has some unappealing properties. In general, adaptive expectations involve systematic forecast errors – if inflation doubles, expected inflation will always lie below actual inflation. Another problem is that only past observations matter, so that announcements concerning the future values of key variables, such as the money supply, are precluded from having any effect. The appeal of the rational expectations hypothesis (which we will discuss shortly) is that it does not involve these shortcomings. Nevertheless, many economists (especially those involved in policy) continue to rely on the adaptive scheme, although modern versions are somewhat more sophisticated.

In a series of papers in the early 1970s, Robert E. Lucas Jr. developed an alternative theory of expectations formation, under the assumption of rational expectations. He showed that a positive relationship between output and inflation could arise in the data because of imperfect information regarding the aggregate price level. His work emphasized the crucial distinction between anticipated and unanticipated movements in the money supply, and the limits imposed on monetary policy. However, the broader impact of his work had not so much to do with this particular theory, but rather with the implications of rational expectations for the use of traditional macroeconomic models in evaluating policy – the so-called Lucas Critique.

Rational Expectations

The hypothesis of rational expectations means that economic agents forecast in such a way as to minimize forecast errors, subject to the information constraints that confront them. It does not mean that they make no forecast errors; it simply means that such errors have no serial correlation, no systematic component. Most economists would not object to this general proposition. The hypothesis becomes controversial however, when strong assumptions are made about the constraints. In much of the macroeconomics literature it is assumed (mainly for simplicity) that these constraints are negligible, so that agents have an almost perfect knowledge of the structure of the economy. Under these conditions, rational expectations regarding a future variable π_{t+1} can be defined as

$$\pi^e_{t+1} = E[\pi_{t+1} | \Omega_t] = E_t[\pi_{t+1}],$$

where E is the statistical expectation of π_{t+1}, which may be based on both past observations of π_t, current observations of other relevant data and information about the structure of the economy. The symbol Ω_t is used to represent this information set at time t, and the symbol | is short-hand for “conditional upon”. Thus $E[\pi_{t+1}|\Omega_t]$ should be read as the expectation of π_{t+1} conditional upon the information available at time t. This is often abbreviated by putting a time subscript on the expectations sign as in the last term above. The notion of a statistical expectation may be understood by imagining that there is a distribution of possible “states of the world” indexed by s. Each state involves a different level of inflation, $\pi(s)$, and the probability of state s occurring at time t+1, given the information available at time t, is given by $\mu_{t+1}(s|\Omega_t)$. Then we would write the time t expected value of inflations as

$$E_t[\pi_{t+1}] = \sum_{s=1}^{S} \pi(s)\mu_{t+1}(s|\Omega_t).$$

Example: Suppose people know that inflation evolves according to the first-order autoregressive (AR1) process

$$\pi_{t+1} = \alpha + \beta\pi_t + \varepsilon_{t+1},$$

where ε_{t+1} is a normally distributed random variable with mean zero and variance σ^2. This just says that inflation today depends linearly on past inflation plus a random shock which represents uncertainty. Then the expected value of π_{t+1} is given by

$$E_t[\pi_{t+1}] = \alpha + \beta\pi_t.$$

This is the rational expectations estimate of π_{t+1} conditional upon π_t and the structure of this simple economy (which is the information set a time t). While everyone in the economy does not usually form their expectations in such a mechanical way, the assumption is often justified by arguing that on average people act “as if” they do. For example, when you drive a car down the street, you do not compute the probabilities of every possible event and fine-tune your driving in response. However, people typically drive “as if” they had done this, by implicitly assigning a very low (subjective) probability to a deer running out on to the 401, and therefore driving at 120kms per hour. Sometimes an individual’s behavior is determined by what they see other people or firms (or the Fed) doing. If some of these economic actors have computed the probabilities correctly, others may act “as if” they have, even if they are simply following suit.

1.1 Prior Expectations and Posterior Expectations

A crucial aspect of the rational expectations hypothesis is the assumptions it makes about the way in which people update their beliefs regarding the future in response to new information. Suppose a person has some prior about the distribution of the aggregate price level. For example, by looking at past aggregate price levels she could have determined the mean and variance of that distribution. Her prior expectation of the price level would then equal the mean. However, suppose that while she does not actually know the aggregate price level, she can observe directly some prices in the economy. Then she might use that information to update her priors and form a new posterior distribution of the aggregate price level. For example, if the prices that she can observe are unexpectedly high then she might place some weight on the probability that the general price level has risen. Suppose all prices in the economy are such that their mean is E[p] based on past observation. Then the rational expectations estimate of the (posterior) expected price level conditional on the observed price p_i is given by

$$E[p|p_i] = \theta p_i + (1-\theta)E[p],$$

where θ is a parameter than depends on the variance of the underlying price distributions. The intuition behind this is simply that if you observe prices above what you expect ($p_i >$ E[p]), then you should rationally place greater probability on the overall price level having increased relative to your prior expectation. However, you should not revise your expectations upward by the full extent p_i-E[p] because it is possible that other prices have fallen and the average price has actually remained the same or even fallen. The parameter θ depends on how correlated price increases are across markets and how variable the prices are for each good relative to the average price level. For example, if the prices of individual goods are generally very variable then θ will tend to be low because observing some price increases does not tell you that much about the overall average.

2 Lucas' Imperfect Information Model

This model shows how a positive relationship between output and inflation might arise in an economy where goods and labor markets clear, but where information regarding the average price level is imperfect. As described above, competitive firms are assumed to be able to observe changes in the nominal price of their own goods, but cannot tell whether this is due to changes in tastes for the good or changes in aggregate demand due to increases in the money supply. The latter implies overall inflation, so that the real price of the good does not change. In what follows I will use lower case letters to denote the logarithm of the associated aggregate. So, for example, m = lnM. Also, I will drop the time subscript until later, so it should be understood that all variables are time t variables. To simplify things further, assume that the LM curve is vertical so that aggregate demand depends only on real money balances (fluctuations in the IS curve then determine the interest rate). The simplified log-linear aggregate demand curve is then given by

$$y = \alpha + m - p, \qquad \text{(AD)}$$

where α is a constant parameter. Note that to further simplify things I have assumed that the elasticity of output w.r.t. real money balances is 1. It is assumed, for simplicity, that the share of expenditure allocated to each good is the same. Thus, since there are n goods, the share spent on each is 1/n. This implies that the aggregate price index here, is the geometric average across n goods indexed by i:

$$p = \frac{1}{n}\sum_{i=1}^{n} p_i.$$

The demand curve for an individual good i is assumed to be given by

$$q_i^d = y + z_i - \eta(p_i - p).$$

Here, the demand for a good depends positively on aggregate income and negatively on the real price of the good (if the price of the good rises in proportion to average prices p, the demand does not change). The parameter η is the price elasticity of demand. The term z_i is supposed to represent a "taste" shock that may cause the demand curve to shift. It is assumed to be normally distributed with mean 0. Note that the geometric average of the individual demands yields aggregate demand:

$$\frac{1}{n}\sum_{i=1}^{n} q_i^d = y.$$

Finally, the supply of good i is assumed to be given by

$$q_i^s = \gamma + \beta(p_i - p),$$

where β is the price elasticity of supply.

2.1 Equilibrium with Perfect Information

Suppose firms have perfect information regarding the aggregate price level. The equilibrium in the market for each good i would be given by

$$\gamma + \beta(p_i - p) = y + z_i - \eta(p_i - p).$$

Solving for $\ln P_i$ we get

$$p_i = \frac{1}{\beta + \eta}(y + z_i - \gamma) + p$$

Aggregating over all markets we get

$$p = \frac{1}{\beta + \eta}(y - \gamma) + p \Rightarrow \bar{y} = \gamma. \qquad \text{(AS1)}$$

Thus, in a market clearing equilibrium, aggregate supply is fixed and independent of the money supply. Setting aggregate supply (AS1) equal to aggregate demand (AD) yields

$$\gamma = \alpha + m - p,$$

so that the average price level is

$$p = \alpha - \gamma + m.$$

2.2 Equilibrium with Imperfect Information

Now suppose that at any point in time firms can only observe prices of their own goods, and cannot determine with fluctuations in demand are resulting from real demand shocks, z_i, or aggregate demand fluctuations, y. The supply curve of firm i is given by

$$q_i^s = \gamma + \beta(p_i - E[p|p_i]),$$

They must therefore form a rational expectation of the aggregate price index conditional on the information they have. As described in the previous section, this posterior expectation is given by

$$E[p|p_i] = E[p] + \theta(p_i - E[p]).$$

Substituting this into the supply function we get

$$q_i^s = \gamma + \beta(1-\theta)(p_i - E[p]).$$

Averaging across all firms yields the short-run aggregate supply (or Lucas supply) function

$$y = \gamma + \beta(1-\theta)(p - E[p]). \qquad \text{(AS2)}$$

Since the full information output is $\bar{y} = \gamma$, we could re-write this equation, after adding time subscripts, as

$$y_t - \bar{y} = \beta(1-\theta)(p_t - p_{t-1} - E_{t-1}[p_t - p_{t-1}]),$$
$$y_t - \bar{y} = \beta(1-\theta)(\pi_t - \pi_t^e),$$

which looks like an inverted version of the short-run (expectations augmented) Phillips curve. Note here, however, that goods market clear and the relationship does not come from sluggish price setting, but from “monetary misperceptions”. To simplify the algebra let b = β(1 − θ). Then, in equilibrium, set aggregate demand (AD) equal to aggregate supply (AS2), we get

$$\alpha + m - p = \gamma + b(p - E[p]).$$

It is convenient to write this as

$$\alpha - \gamma + (m - E[p]) = (1+b)(p - E[p]). \qquad (*)$$

Taking expectations of this equation, we have

$$\alpha - \gamma + E[m - p] = 0.$$

And so

$$E[p] = \alpha - \gamma + E[m].$$

Using this to substitute out E[p] from (*) gives

$$m - E[m] = (1+b)(p - \alpha - \gamma - E[m]).$$

Solving for p, we get

$$p = \alpha - \gamma + E[m] + \frac{1}{1+b}(m - E[m]). \qquad (P)$$

Now substituting this into the aggregate demand equation (AD) yields

$$y = \alpha + m - p$$
$$y = \bar{y} + \frac{b}{1+b}(m - E[m]). \qquad (Y)$$

Thus, deviations in output from the full—information level depend only on unanticipated movements in the money supply. Anticipated movements affect only the price level.

The Lucas model is a Classical model in that all markets clear and prices adjust immediately. However, it shows how short—run deviations from trend output may arise in response to unexpected changes in the money supply. Thus, it may be viewed as providing a money-driven theory of business cycle fluctuations (unlike the traditional classical model). Its most important implications, however are:

• It shows how a positive statistical relationship can arise between output and inflation even though prices adjust instantaneously and markets clear. Thus, it provides an alternative theory for Phillips' empirical observations. To see this, suppose that the money supply grows at a constant rate, g_m, except for random fluctuations denoted by the variable ε_t, which has mean zero. That is

$$m_t = m_{t-1} + g_m + \varepsilon_t.$$

The time t−1 expectation of m is $E_{t-1}[m_t] = m_{t-1} + g_m$ and so

$$m_t - E_{t-1}[m_t] = \varepsilon_t.$$

From the equilibrium output equation (Y) it follows that

$$y_t - \bar{y} = \frac{b}{1+b}\varepsilon_t. \qquad (**)$$

Using the equilibrium price equation (P) we have

$$p_t = \alpha - \gamma + E_{t-1}[m_t] + \frac{1}{1+b}(m_t - E_{t-1}[m_t]),$$

$$p_t = \alpha - \gamma + m_{t-1} + g_m + \frac{1}{1+b}\varepsilon_t.$$

Lagging one period yields

$$p_{t-1} = \alpha - \gamma + m_{t-2} + g_m + \frac{1}{1+b}\varepsilon_{t-1}.$$

It follows that inflation is given by

$$\pi_t = p_t - p_{t-1} = m_{t-1} - m_{t-2} + \frac{1}{1+b}(\varepsilon_t - \varepsilon_{t-1}),$$

$$\pi_t = g_m + \varepsilon_{t-1} + \frac{1}{1+b}(\varepsilon_t - \varepsilon_{t-1}),$$

$$\pi_t = g_m + \frac{1}{1+b}\varepsilon_t + \frac{b}{1+b}\varepsilon_{t-1}.$$

Now we can use (**) to substitute out the ε_t's and get

$$\pi_t = g_m + \frac{1}{b}(y_t - \bar{y}) + (y_{t-1} - \bar{y}).$$

Thus, the model predicts that inflation and output are positively correlated over time in a statistical sense. However, this correlation is not due to sluggish price adjustment, but to the effects of unanticipated movements in the money supply which initially affect output temporarily, but eventually only cause inflation.

• Although there is a statistical relationship, the trade—off cannot be exploited by the central bank so as to maintain output above its “natural” (full information) level for long. For example, suppose the central bank raises money supply growth, gm. If this is not observed, there may be a period during which output rises due to the unexpected increase in the money supply. However, once agents in the economy determine that gm has risen, output will return to its long run level.

3 The Lucas Critique

On its own, the Lucas model may not have had much impact outside a small group of theoretical macroeconomists. However, almost immediately after that paper was published, its predictions appeared to become eerily accurate. Following the first oil price shock of 1973, policy-makers attempted to offset the effects of the ensuing downturn through a monetary expansion. Implicitly they were trying to exploit a presumed trade off along the short—run Phillips curve. If inflationary expectations had adjusted in a sluggish fashion then this may have worked. However, it appeared that within a short period of time firms and households figured out what was happening and adjusted their expectations of inflation upwards. These expectations were then built into employment contracts and supply relationships, thereby fuelling high inflation with little impact on output or employment. This combination of low output and high inflation became known as stagflation.

Although not all economists accept this exact interpretation of what happened during the 1970s, most agree on the broader implications of the Lucas Critique. The above example is a special case of the general problem that if policy-makers attempt to exploit a statistical relationship based on past data, effects operating through expectations may cause the relationship to break down. The problem arises in the example above because a short—run Phillips curve estimated using past data takes the policy regime as given. However, when money growth gm increases the short-run Phillips curve shifts up as firms and households rationally adjust their expectations of inflation upwards (as predicted by the model). Policy changes that do not take into account the endogenous response of the private sector's reaction function will not work as predicted or may even fail.

Other examples of the Lucas Critique abound in the static Keynesian model and the neoclassical synthesis. For example, in considering the impacts of a deficit—financed increase in government spending, these models do not allow for the impacts of changing expectations regarding the future implications. Although there is no tax increase today, if the government must maintain a sustainable debt load, there most be implications for future taxes. An anticipated increase in future taxes, reduces future expected disposable income. In anticipation of this how might we expect households to react? One effect maybe that they borrow less or save more today than they otherwise would have, for a given level of current income and the interest rate. In other words, the Keynesian consumption function would shift down because change in expected wealth that results from the policy. If we based policy evaluation on a model which included a consumption function estimated before the increase in spending, the predictions of the model would be wrong.

Thus, the main implication of the Lucas Critique is that as far as possible, dynamic macroeconomic models need to be based on sound microeconomic foundations. Consumer and firm behavior must be derived from dynamic microeconomic models in which their response to changes in policy and other phenomena is determined by policy-invariant structural parameters. Moreover, policy evaluation must take the reaction of the private sector into account. As we shall see in the second part of this course, the Lucas Critique has had far reaching consequences for the nature of mainstream macroeconomic thinking in the last 25 years and also for macroeconomic policy, although the impact on the latter has been much slower.