Lecture Notes on Macroeconomics

Karl Whelan

University College Dublin

October 2023

Contents

Forward

This is a collection of lecture notes that I have used over a number of years teaching Advanced Macroeconomics to final year undergraduates at University College Dublin. Some of the material has also been used to teach first-term Masters students. The material is not intended to be a comprehensive book on advanced undergraduate macroeconomics but simply reflects my own interests and preferences and is being made available in this format because some people may find it useful.

Over the years, I have made most of these notes available on my own website but they tend to go up and down on the website each year depending on when I am teaching a course. The material also changes from year to year, so some topics that had been previously covered disappear from my site. Because I know that some people are interested in the notes and use them to assist with their own teaching or learning, I have decided to make the full set available permanently in book form.

A few observations on the notes. The first two parts of the book deal with issues related to short-term macroeconomic fluctuations. Part 1 presents an adapted version of the Keynesian IS-LM model featuring monetary policy rules. The approach taken is largely borrowed from Carl Walsh's 2002 paper "Teaching Inflation Targeting: An Analysis for Intermediate Macro" in the *Journal of Economic Education* but the analysis here applies the model to a much wider range of issues. There are also similarities with David Romer's manuscript, *Short-Run*

Fluctuations, particularly in the treatment of the zero bound on interest rates but there are also some important differences.[1]

My intention with this material is to teach what are probably the key insights of New Keynesian economics—that fiscal and monetary policy can be effective in influencing output in the short-run but not the long-run, that the process by which the public formulates inflation expectations is crucial, the advantages to having tough central bankers, that monetary policy rules can be stabilising if they satisfy the Taylor principle, and that some of these key insights are reversed if the economy is in a liquidity trap—without relying on the particular restrictive assumption of rational expectations.

Part 2, in contrast, explores rational expectations in some detail. The unifying theme of this section is the importance of understanding first-order stochastic differences equations; how to solve them for forward-looking solutions and how to relate these solutions to observable data. These methods are used to explore a range of standard macroeconomic topics with some chapters (such as those on consumption and asset prices) more focused on reviewing the empirical evidence than others. In terms of pedagogy, my approach is to explain the important role that rational expectations plays in modern macroeconomic modelling but to encourage students to realise that empirical testing tends to uncover weaknesses in this approach.

The final two parts of the book cover long-run growth theory. The first few topics (growth acccounting, Solow and Romer models) are fairly standard while the remaining topics (on institutions, technology diffusion, Malthusian dynamics and growth and resources) partly reflect my own interests.

Beamer slides are available for almost all the chapters in this book but I do not have time

[1] Romer's manuscript is available at http://eml.berkeley.edu/ dromer/papers/ISMP%20Text%20Graphs%202013.pdf

to adapt my current set of slides to precisely cover all the topics in the book. Anybody who may wish to get teaching slides to accompanying teaching some of the chapters in this book is free to email me at karl.whelan@ucd.ie and I'll help out if I can.

Part I

The IS-MP-PC Model

Chapter 1

Introducing the IS-MP-PC Model

I am guessing that almost everyone reading this has seen the IS-LM and AS-AD models. In the first part of this course, we are going to revisit some of the ideas from those models and expand on them in a number of ways:

- Rather than the traditional LM curve, we will describe monetary policy in a way that is more consistent with how it is now implemented, i.e. we will assume the central bank follows a rule that dictates how it sets nominal interest rates. We will focus on how the properties of the monetary policy rule influence the behaviour of the economy.
- We will have a more careful treatment of the roles played by real interest rates.
- We will focus more on the role of the public's inflation expectations.
- We will model the zero lower bound on interest rates and discuss its implications for policy.

Our model is going to have three elements to it:

- **A Phillips Curve** describing how inflation depends on output.

- **An IS Curve** describing how output depends upon interest rates.

- **A Monetary Policy Rule** describing how the central bank sets interest rates depending on inflation and/or output.

Putting these three elements together, I will call it the IS-MP-PC model (i.e. The Income-Spending/Monetary Policy/Phillips Curve model). I will describe the model with equations. I will also merge together the second two elements (the IS curve and the monetary policy rule) to give a new IS-MP curve that can be combined with the Phillips curve to use graphs to illustrate the model's properties.[1]

Model Element One: The Phillips Curve

Perhaps the most common theme in economics is that you can't have everything you want. Life is full of trade-offs, so that if you get more of one thing, you have to have less of another thing. In macroeconomics, there are important trade-offs facing governments when they implement policy. One of these relates to a trade-off between desired outcomes for inflation and output.

What form does this relationship take? Back when macroeconomics was a relatively young discipline, in 1958, a study by the LSE's A.W. Phillips seemed to provide the answer. Phillips documented a strong negative relationship between wage inflation and unemployment: Low unemployment was associated with high inflation, presumably because tight labour markets stimulated wage inflation. Figure 1.1 shows one of the graphs from Phillips's paper illustrating the kind of relationship he found.

In 1960, a paper by MIT economists Robert Solow and Paul Samuelson (both of whom

[1] The model presented here is basically an adapted version of the model presented in Carl Walsh's 2002 paper "Teaching Inflation Targeting: An Analysis for Intermediate Macro" in the *Journal of Economic Education.*

would go on to win the Nobel prize in economics for other work) replicated these findings for the US and emphasised that the relationship also worked for price inflation. Figure 1.2 reproduces the graph in their paper describing the relationship they found. The Phillips curve quickly became the basis for the discussion of macroeconomic policy decisions. Economists advised that governments faced a tradeoff: Lower unemployment could be achieved, but only at the cost of higher inflation.

However, Milton Friedman's 1968 presidential address to the American Economic Association produced a well-timed and influential critique of the thinking underlying the Phillips curve. Friedman pointed out that it was expected real wages that affected wage bargaining. If low unemployment means workers have a strong bargaining position, then high nominal wage inflation on its own is not good enough: They want nominal wage inflation greater than price inflation.

Friedman argued that if policy-makers tried to exploit an apparent Phillips curve tradeoff, then the public would get used to high inflation and come to expect it. Inflation expectations would move up and the previously-existing tradeoff between inflation and output would disappear. In particular, he put forward the idea that there was a "natural" rate of unemployment and that attempts to keep unemployment below this level could not work in the long run.

Evidence on the Phillips Curve

Monetary and fiscal policies in the 1960s were very expansionary around the world, partly because governments following Phillips curve "recipes" chose booming economies with low unemployment at the expense of somewhat higher inflation.

At first, the Phillips curve seemed to work: Inflation rose and unemployment fell. However, as the public got used to high inflation, the Phillips tradeoff got worse. By the late 1960s inflation was rising even though unemployment had moved up. Figure 1.3 shows the US time series for inflation and unemployment. This *stagflation* combination of high inflation and high unemployment got even worse in the 1970s. This was exactly what Friedman had predicted.

Today, the data no longer show any sign of a negative relationship between inflation and unemployment. If fact, if you look at the scatter plot of US inflation and unemployment data shown in Figure 1.4 , the correlation is positive: The original formulation of the Phillips curve is widely agreed to be wrong.

Figure 1.1: One of A. W. Phillips's Graphs

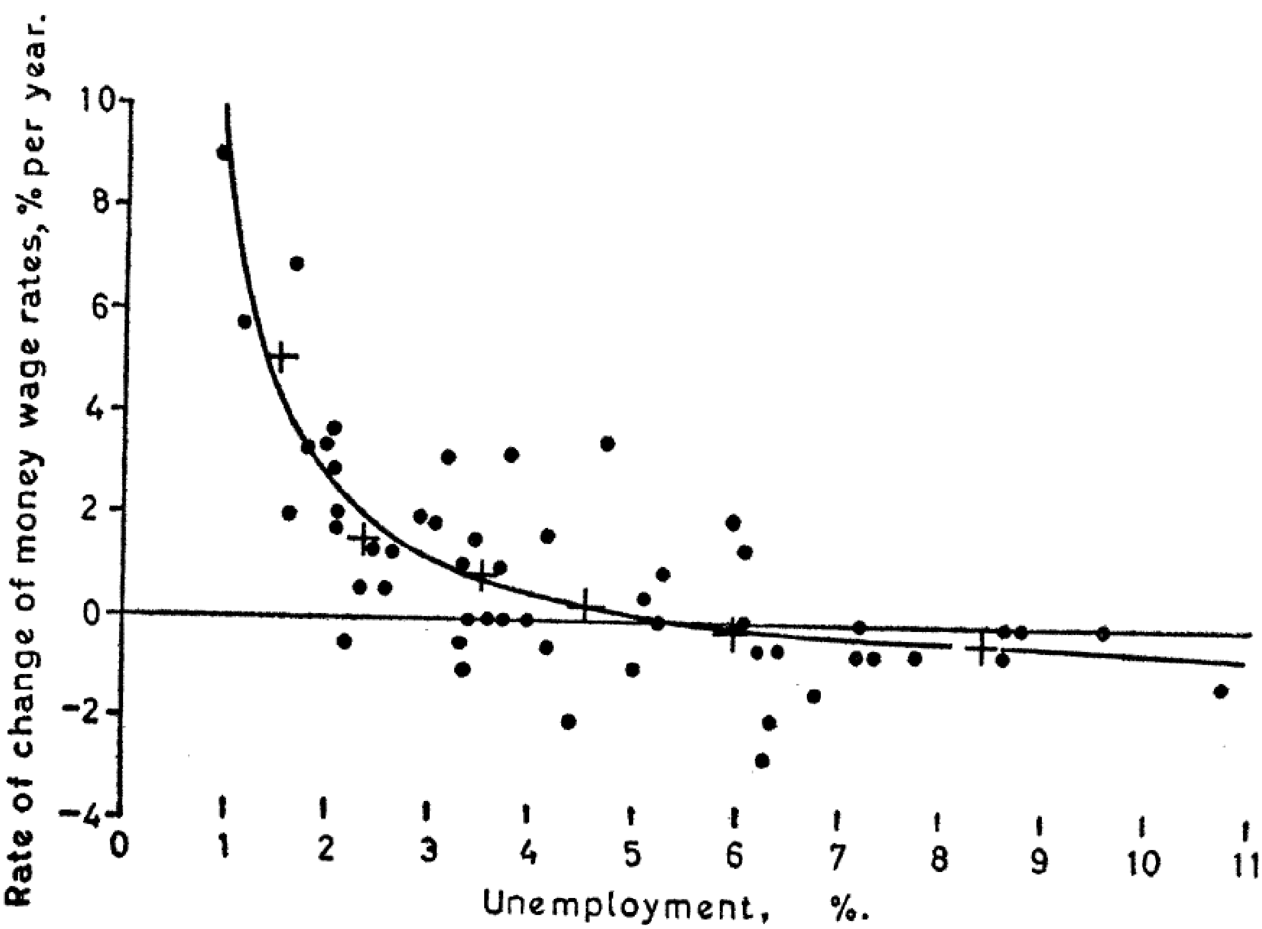


Fig.1. 1861 – 1913

Figure 1.2: Solow and Samuelson's Description of the Phillips Curve

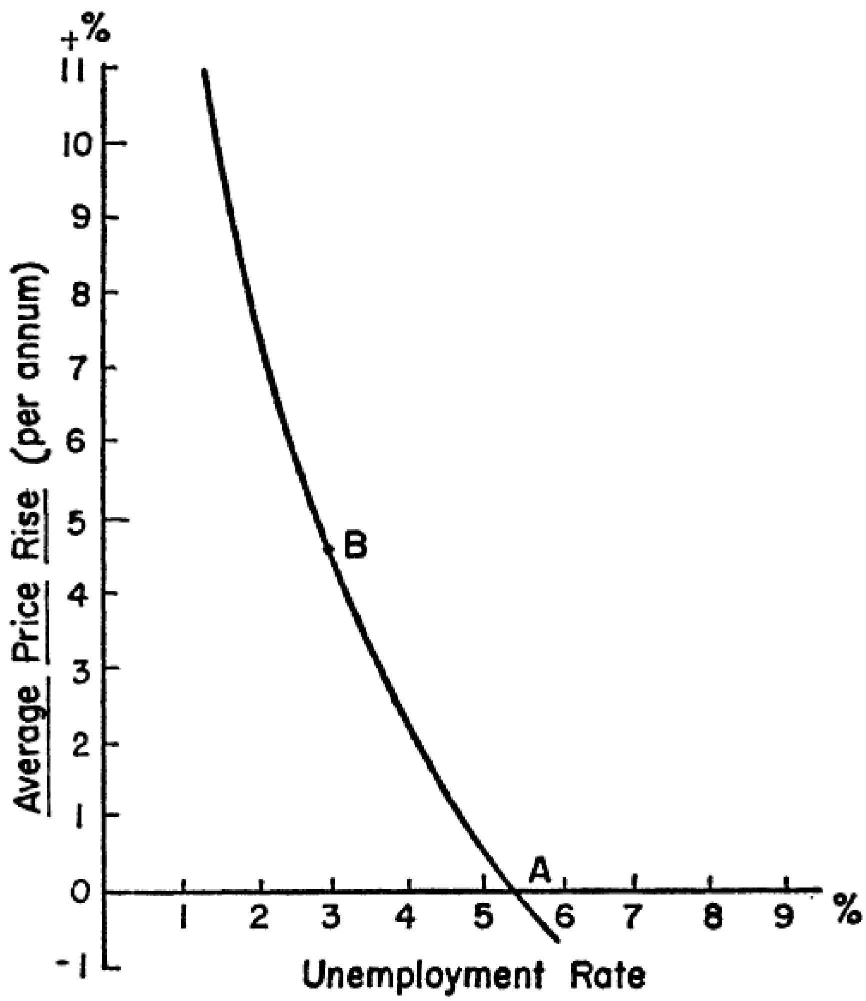


FIGURE 2

MODIFIED PHILLIPS CURVE FOR U.S.

This shows the menu of choice between different degrees of unemployment and price stability, as roughly estimated from last twenty-five years of American data.

Figure 1.3: The Evolution of US Inflation and Unemployment

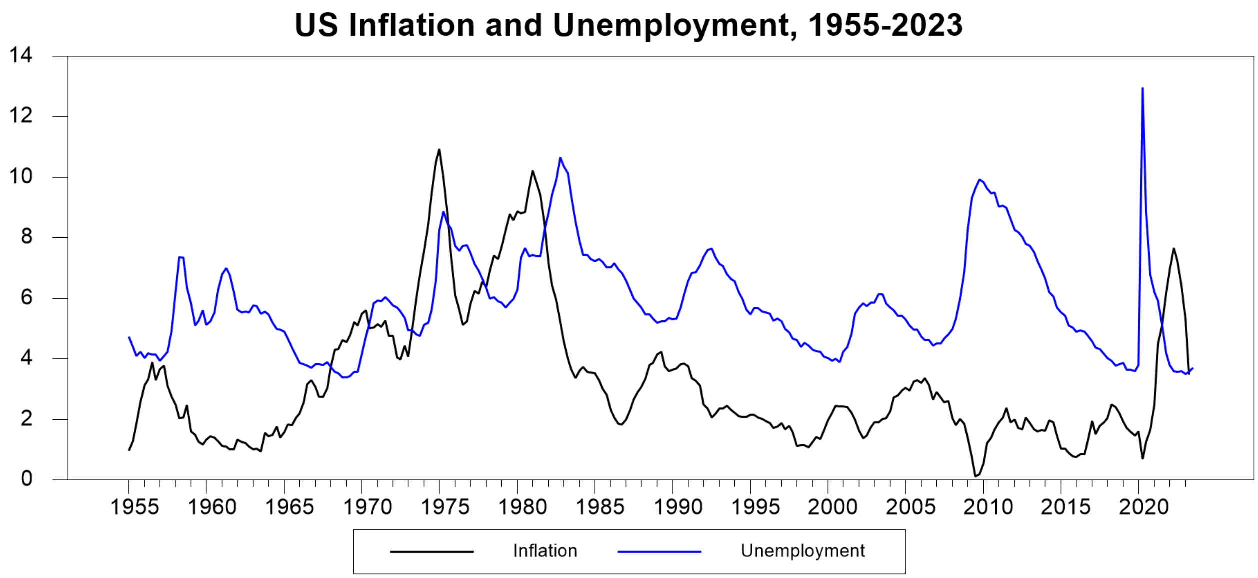
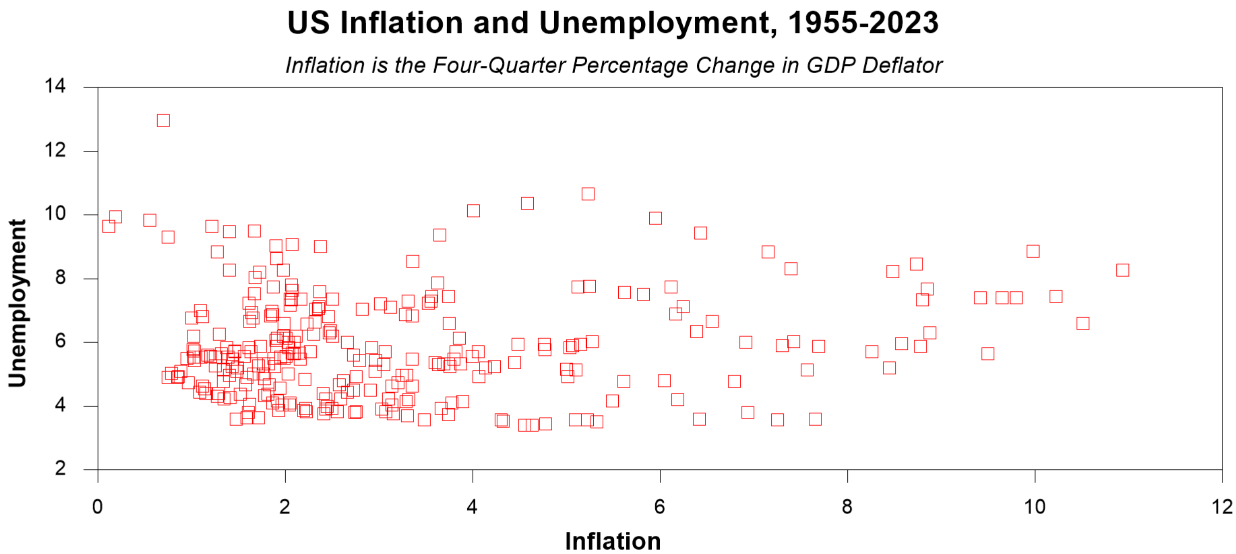


Figure 1.4: The Failure of the Original Phillips Curve



Equations: Variables, Parameters and All That

We will use both graphs and equations to describe the models in this class. I know many students don't like equations but I hope to show that we can learn more about how these model economies behave once we formulate them with equations than we can from just representing them with graphs.

Variables and Coefficients: The equations in this class will generally have a certain format. They will often look a bit like this.

$$y_t = \alpha + \beta x_t \tag{1.1}$$

There are two types of objects in this equation. First, there are the *variables*, y_t and x_t. These will correspond to economic variables that we are interested in (inflation or GDP for example). We interpret y_t as meaning "the value that the variable y takes during the time period t"). For most models in this course, we will treat time as marching forward in discrete intervals, i.e. period 1 is followed by period 2, period t is followed by period $t+1$ and so on.

Second, there are the parameters or *coefficients*. In this example, these are given by α and β. These are assumed to stay fixed over time. There are usually two types of coefficients: Intercept terms like α that describe the value that series like y_t will take when other variables all equal zero and coefficients like β that describe the impact that one variable has on another. In this case, if β is a big number, then a change in the variable x_t has a big impact on y_t while if β is small, it will have a small impact.

Some of you are probably asking what those squiggly shapes— α and β— are. They are Greek letters. While it's not strictly necessary to use these shapes to represent model parameters, it's pretty common in economics. So let me introduce them: α is alpha (Al-Fa), β is beta (Bay-ta), γ is gamma, δ is delta, θ is theta (Thay-ta) and π naturally enough is pi.

Dynamics: One of the things we will be interested in is how the variables we are looking at will change over time. For example, we will have equations along the lines of

$$y_t = \beta y_{t-1} + \gamma x_t \tag{1.2}$$

Reading this equation, it says that the value of y at time t will depend on the value of x at time t and also on the value that y took in the previous period i.e. $t - 1$. By this, we mean that this equation holds in every period. In other words, in period 2, y depends on the value that x takes in period 2 and also on the value that y took in period 1. Similarly, in period 3, y depends on the value that x takes in period 3 and also on the value that y took in period 2. And so on.

Subscripts and Superscripts: When we write y_t, we mean the value that the variable y takes at time t. Note that the t here is a *subscript* – it goes at the bottom of the y. Some students don't realise this is a subscript and will just write yt but this is incorrect (it reads as though the value t is multiplying y which is not what's going on).

We will also sometimes put indicators above certain variables to indicate that they are special variables. For example, in the model we present now, you will see a variable written as π_t^e which will represent the public's expectation of inflation. In the model, π_t is inflation at time t and the e above the π in π_t^e is there to signify that this is not inflation itself but rather it is the public's expectation of it.

Our Version of the Phillips Curve

We will use both graphs and equations to describe the various elements of our model. Our first element is an expectations-augmented Phillips curve which we will formulate as a relationship in which inflation depends on inflation expectations, the gap between output and its "natural" level and a temporary inflationary shock. Our equation for this is the following:

$$\pi_t = \pi_t^e + \gamma \left(y_t - y_t^* \right) + \epsilon_t^\pi \tag{1.3}$$

In this equation π represents inflation and by π_t we mean inflation at time t. The equation states that inflation at time t depends on three factors:

1. **Inflation Expectations,** π_t^e: This term—which puts an e superscript above the π_t—represents the public's inflation expectations at time t. We have put a time subscript on this variable because the public's expectations may change over time. Note that a 1 point increase in inflation expectations raises inflation by exactly 1 point. This is because we are assuming that people bargain over real wages and higher expected inflation translates one-for-one into their wage bargaining, which in turn is passed into price inflation.

2. **The Output Gap,** $(y_t - y_t^*)$: This is the gap between GDP at time t, as represented by y_t, and what we will term the "natural" level of output, which we term y_t^*. This is the level of output at time t that would be consistent with unemployment equalling its natural rate. (Note we are describing inflation as being dependent on the output gap rather than the gap between unemployment and its natural rate because this would require adding an extra element to the model, i.e. the link between unemployment and output). We would expect this natural level of output to gradually increase over time as

productivity levels improve. The coefficient γ (pronounced "gamma") describes exactly how much inflation is generated by a 1 percent increase in the gap between output and its natural rate.

3. **Temporary Inflationary Shocks,** ϵ_t^π: No model in economics is perfect. So while inflation expectations and the output gap may be key drivers of inflation, they won't capture all the factors that influence inflation at any time. For example, "supply shocks" like a temporary increase in the price of imported oil can drive up inflation for a while. To capture these kinds of temporary factors, we include an inflationary "shock" term, ϵ_t^π. (ϵ is a Greek letter pronounced "epsilon"). The superscript π indicates that this is the inflationary shock (this will distinguish it from the output shock that we will also add to the model) and the t subscript indicates that these shocks change over time.

The Phillips Curve Graph

Figure 1.5 shows how to describe our Phillips curve equation in a graph. The graph shows a positive relationship between inflation and output. The key points to notice are the markings on the two axes indicating what happens when output is at its natural rate. This graph illustrates the case when there are no temporary inflationary shocks so $\epsilon_t^\pi = 0$. In this case, the Phillips curve is just

$$\pi_t = \pi_t^e + \gamma \left(y_t - y_t^* \right) \tag{1.4}$$

So when $y_t = y_t^*$ we get $\pi_t = \pi_t^e$.

The curve can move up or down depending on what happens to the inflationary shocks, ϵ_t^π, and with inflation expectations. Figure 1.6 illustrates what happens when there is a positive inflationary shock so that ϵ_t^π goes from being zero to being positive. In this case, even when output equals its natural level (i.e. $y_t = y_t^*$) we still get inflation being above its expected level. Figure 1.7 illustrates how the curve changes when expected inflation rises from π_1 to π_2. The whole curve shifts upwards because of the increase in expected inflation.

Figure 1.5: The Phillips Curve with $\epsilon_t^\pi = 0$

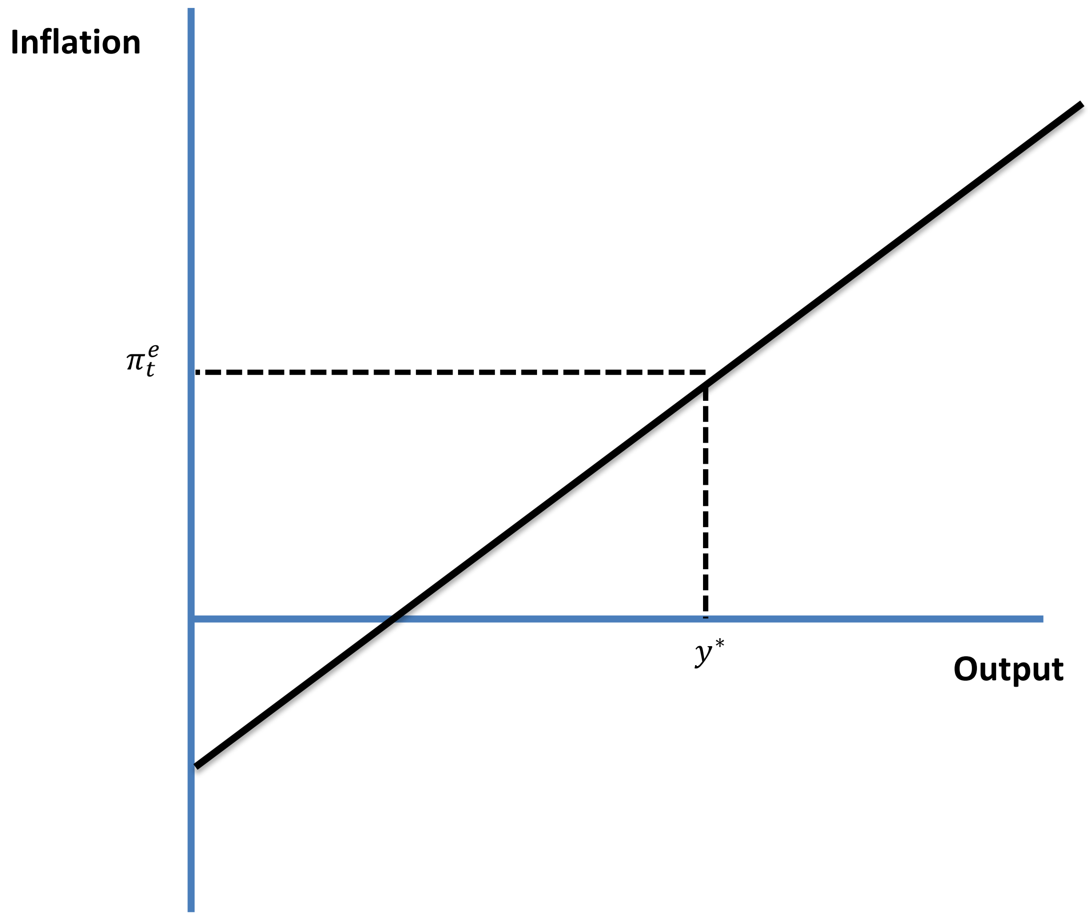


Figure 1.6: The Phillips Curve as we move from $\epsilon_t^\pi = 0$ to $\epsilon_t^\pi > 0$ (An Aggregate Supply Shock)

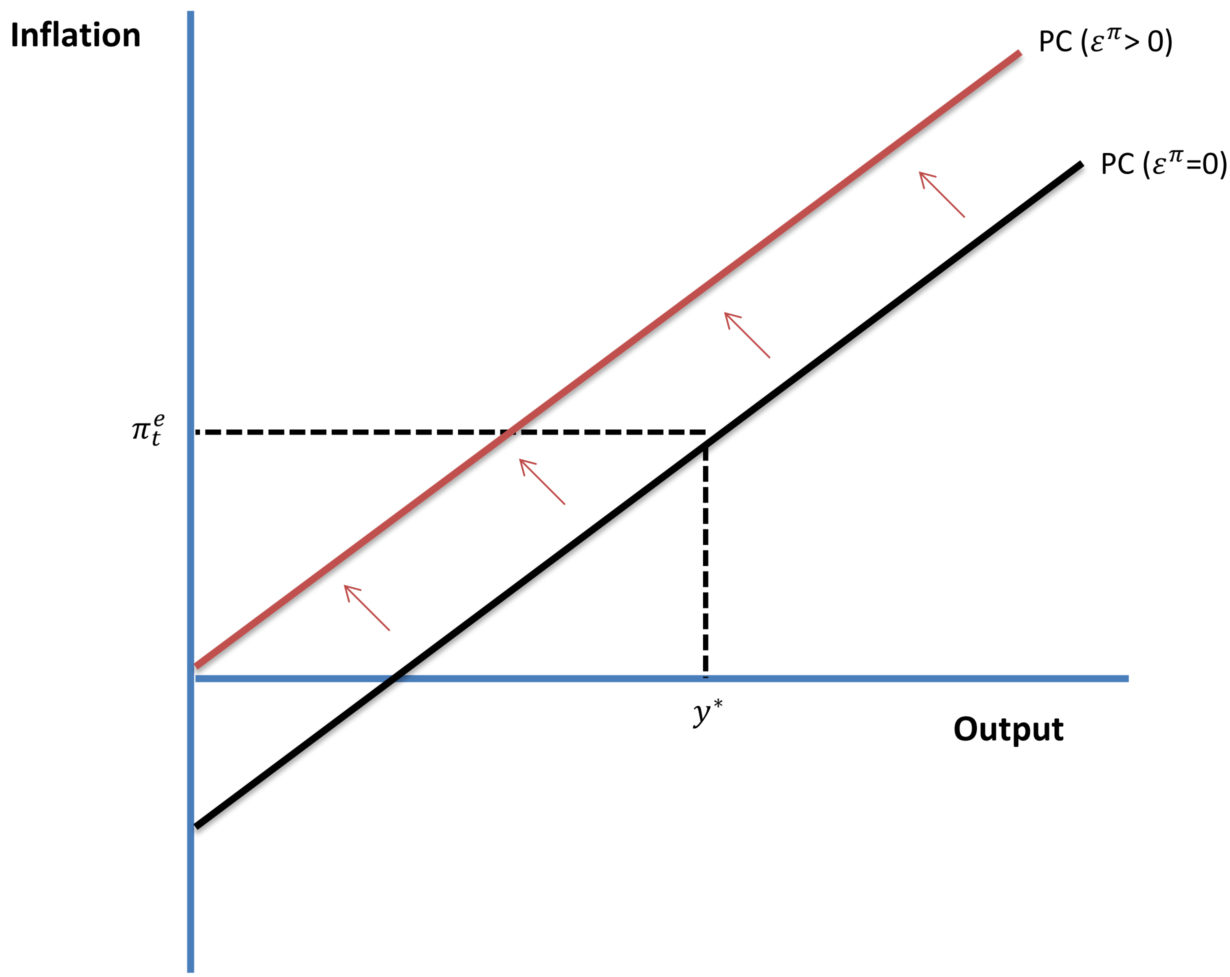
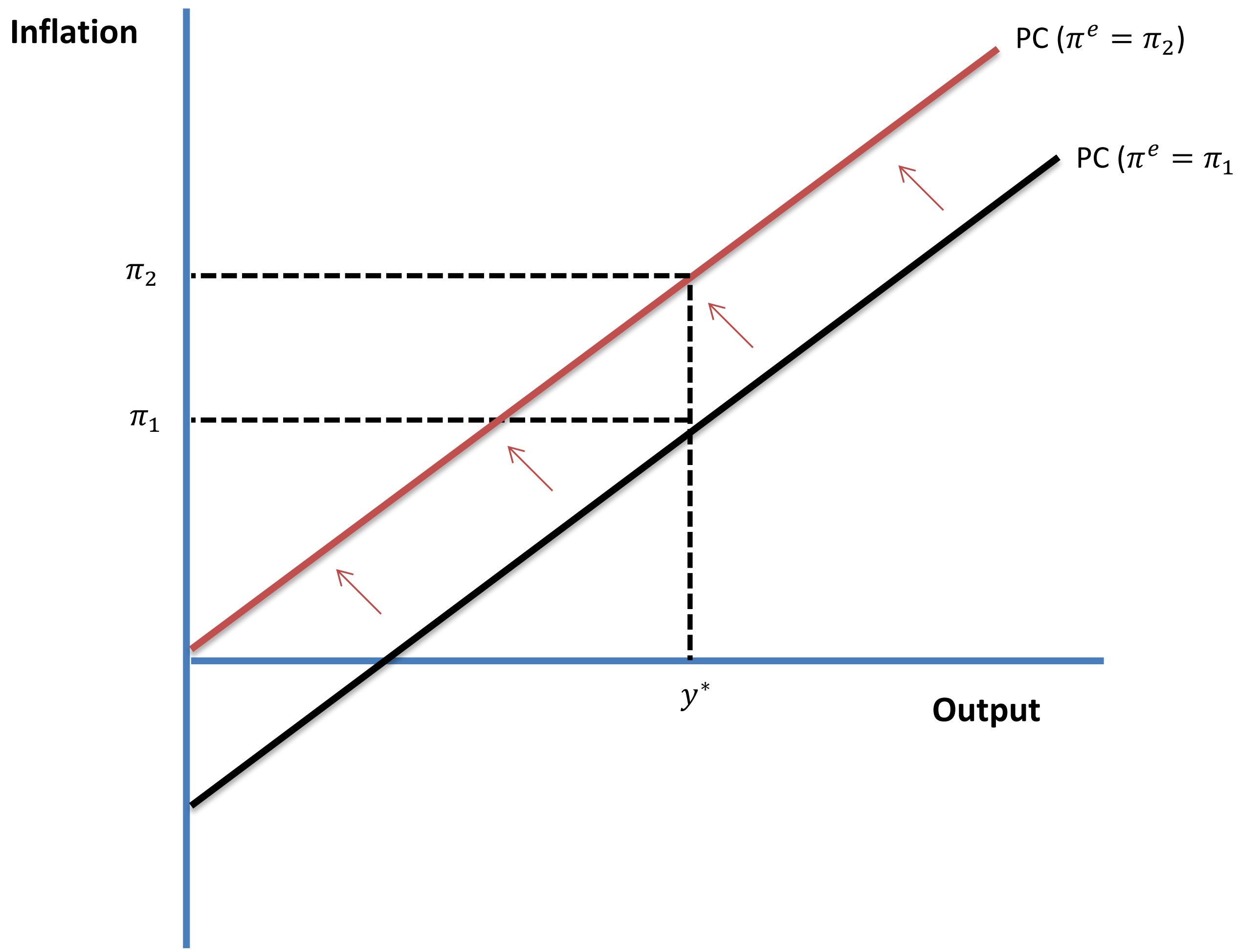


Figure 1.7: The Phillips Curve as we move from $\pi_t^e = \pi_1$ to $\pi_t^e = \pi_2$



Model Element Two: The IS Curve

The second element of the model is one that should be familiar to you: An IS curve relating output to interest rates. The higher interest rates are, the lower output is. However, I want to stress something here that is often not emphasised in introductory treatments of the IS curve. The IS relationship is between output and *real* interest rates, not nominal rates. Real interest rates adjust the headline (nominal) interest rate by subtracting off inflation.

Let's consider why it is that real interest rates are what matters. You know already that high interest rates discourage aggregate demand by reducing consumption and investment spending. But what is a high interest rate? Suppose I told you the interest rate was 10 percent. Is this a high interest rate?

You might be inclined to say, "Yes, this is a high interest rate" but the answer is that it really depends on inflation. Consider the decision to save for tomorrow or spend today. The argument for saving is that it can allow you to consume more tomorrow. If the real interest rate is positive, then this means that you will be able to purchase more goods and services tomorrow with the money that you set aside. For instance, if the interest rate is 5% but inflation is 2%, then you'll be able to buy 3% more stuff next year because you saved. In contrast, if the interest rate is 5% but inflation is 8%, then you'll be able to buy 3% less stuff next year even though you have saved your money and earned interest. For these reasons, it is the real interest rate that determines whether consumers think saving is attractive relative to spending.

The same logic applies to firms thinking about borrowing to make investment purchases. If inflation is 10%, then a firm can expect that its prices (and profits) will be increasing by that much over the next year and a 10% interest rate won't seem so high. But if prices are falling, then a 10% interest rate on borrowings will seem very high.

With these ideas in mind, our version of the IS curve will be the following:

$$y_t = y_t^* - \alpha \left(i_t - \pi_t - r^* \right) + \epsilon_t^y \tag{1.5}$$

Expressed in words, this equation states that the gap between output and its natural rate $(y_t - y_t^*)$ depends on two factors:

1. **The Real Interest Rate**: The nominal interest rate at time t is represented by i_t, so the real interest rate is $i_t - \pi_t$. The coefficient α describes the effect of a one point increase in the real interest rate on output. The equation has been constructed in a particular way so that it explicitly defines the real interest rate at which output will, on average, equal its natural rate. This rate can be termed the "'natural rate of interest" a term first used by early 20th century Swedish economist Knut Wicksell. Specifically, we can see from the equation that if $\epsilon_t^y = 0$ then a real interest rate of r^* will imply $y_t = y_t^*$.

2. **Aggregate Demand Shocks,** ϵ_t^y: The IS curve is an even more threadbare model of output than the Phillips curve model is of inflation. Many other factors beyond the real interest rate influence aggregate spending decisions. These include fiscal policy, asset prices and consumer and business sentiment. We will model all of these factors as temporary deviations from zero of an aggregate demand "shock", ϵ_t^y. Note that this shock has a superscript y to distinguish it from the "aggregate supply" shock ϵ_t^π that moves the Phillips curve up and down.

Model Element Three: A Monetary Policy Rule

Thus far, our model has described how inflation depends on output and how output depends on interest rates. We can complete the model by describing how interest rates are determined.

Traditionally, in the IS-LM model, this is where the LM curve is introduced. The LM curve links the demand for the real money stock with nominal interest rates and output, with a relationship of the form

$$\frac{m_t}{p_t} = \delta - \mu i_t + \theta y_t \tag{1.6}$$

For a given stock of money (m_t) and a given level of prices (p_t), this relationship can be re-arranged to give a positive relationship between output and interest rates of the form

$$y_t = \frac{1}{\theta}\left(\frac{m_t}{p_t} - \delta + \mu i_t\right) \tag{1.7}$$

This positive relationship between output and interest rates is combined with the negative relationship between these variables in the IS curve to determine unique values for output and interest rates, something that can be illustrated in a graph with an upward-sloping LM curve and a downward-sloping IS curve. Monetary policy is then described as taking the form of the central bank adjusting the money supply m_t in a way that sets the position of the LM curve. The determination of prices is usually described separately in an AS-AD model.

We will not be using the LM curve, for three reasons.

1. **Realism 1**: In practice, no modern central bank implements its monetary policy by setting a specified level of the monetary base. Instead, they use their power to supply potentially unlimited amounts of liquidity to set short-term interest rates to equal a target level. This approach — which has been the practice at all the major central banks for at least 30 years — makes the LM curve (and the money supply) of secondary interest

when thinking about core macroeconomic issues. Our approach will be to describe how the central bank sets interest rates and we won't focus on the money supply.

2. **Realism 2**: The traditional approach is for IS-LM to describe the determination of output and interest rates. Then a separate AS-AD model is used to describe the determination of prices (and thus, implicitly, inflation). However, the reality is that rather than being determined independently of inflation, most modern central banks set interest rates with a very close eye on inflationary developments. A model that integrates the determination of inflation with the setting of monetary policy is thus more realistic.

3. **Simplicity**: In simplifying the determination of output, inflation and interest rates down to a single model, this approach is also simpler than one that requires two different sets of graphs (one set for IS-LM and one set for AS-AD).

We will consider two different types of monetary policy rules that may be followed by the central bank in our model. The first one is a simple one in which the central bank adjusts its interest rate in line with inflation with the goal of meeting an inflation target. Specifically, the first rule we will consider is the following one:

$$i_t = r^* + \pi^* + \beta_\pi \left(\pi_t - \pi^*\right) \tag{1.8}$$

In English, the rule can be interpreted as follows: The central bank adjusts the nominal interest rate, i_t, upwards when inflation, π_t, goes up and downwards when inflation goes down (we are assuming that $\beta_\pi > 0$) and it does so in a way that means when inflation equals a target level, π^*, chosen by the central bank, real interest rates will be equal to their natural level.

That's a bit of a mouthful, so let's see that this is the case and then try to understand why the monetary policy rule would take this form. First, note what the nominal interest rate will be if inflation equals its target level (i.e. $\pi_t = \pi^*$). The term after the final plus sign in equation (1.8) will equal zero and the nominal interest rate will be

$$i_t = r^* + \pi^* \tag{1.9}$$

In this case, because $\pi_t = \pi^*$, we can also write this as

$$i_t = r^* + \pi_t \tag{1.10}$$

So the real interest rate will be

$$i_t - \pi_t = r^* \tag{1.11}$$

Now think about why a rule of this form might be a good idea. Suppose the central bank has a target inflation rate of π^* that it wants to achieve. Ideally, it would like the public to understand that this is the likely inflation rate that will occur, so that $\pi_t^e = \pi^*$. If that can be achieved, then the Phillips curve (equation 1.3) tells us that, on average, inflation will equal π^* provided we have $y_t = y_t^*$. And the IS curve tells us that, on average, we will have $y_t = y_t^*$ when $i_t - \pi_t = r^*$. So this is a policy that can help to enforce an average inflation rate equal to the central bank's desired target, provided the public adjusts its inflationary expectations accordingly.

The IS-MP Curve

That's the model. It consists of three equations. Let's pull them together in one place. They are the Phillips curve:

$$\pi_t = \pi_t^e + \gamma \left(y_t - y_t^*\right) + \epsilon_t^{\pi} \tag{1.12}$$

The IS curve:

$$y_t = y_t^* - \alpha \left(i_t - \pi_t - r^*\right) + \epsilon_t^{y} \tag{1.13}$$

And the monetary policy rule:

$$i_t = r^* + \pi^* + \beta_{\pi} \left(\pi_t - \pi^*\right) \tag{1.14}$$

Now you may recall that I had promised a graphical representation of this model. However, this is a system of three variables which makes it hard to express on a graph with two axes. To make the model easier to analyse using graphs, we are going to reduce it down to a system with two main variables (inflation and output). We can do this because the monetary policy rule makes interest rates are a function of inflation, so we can substitute this rule into the IS curve to get a new relationship between output and inflation that we will call the IS-MP curve.

We do this as follows. Replace the term i_t in equation (1.13) with the right-hand-side of equation (1.14) to get

$$y_t = y_t^* - \alpha \left[r^* + \pi^* + \beta_{\pi} \left(\pi_t - \pi^*\right)\right] + \alpha \left(\pi_t + r^*\right) + \epsilon_t^{y} \tag{1.15}$$

Now multiply out the terms in this equation to get

$$y_t = y_t^* - \alpha r^* - \alpha \pi^* - \alpha \beta_{\pi} \left(\pi_t - \pi^*\right) + \alpha \pi_t + \alpha r^* + \epsilon_t^{y} \tag{1.16}$$

We can bring together the two terms being multiplied by α on its own, and cancel out the terms in αr^* to get

$$y_t = y_t^* - \alpha\beta_\pi(\pi_t - \pi^*) + \alpha(\pi_t - \pi^*) + \epsilon_t^y \tag{1.17}$$

which simplifies to

$$y_t = y_t^* - \alpha(\beta_\pi - 1)(\pi_t - \pi^*) + \epsilon_t^y \tag{1.18}$$

This is the IS-MP curve. It combines the information in the IS curve and the MP curve into one relationship.

The IS-MP Curve Graph

What would this curve look like on a graph? It turns out it depends especially on the value of β_π, the parameter that describes how the central bank reacts to inflation. The IS-MP curve says that an extra unit of inflation implies a change of $-\alpha(\beta_\pi - 1)$ in output. Is this positive or negative? Well we are assuming that $\alpha > 0$ (we put a negative sign in front of this when determining the effect of real interest rates on output) so this combined coefficient will be negative if $\beta_\pi - 1 > 0$.

In other words, the IS-MP curve will have a negative slope (slopes downwards) provided that the way the central bank reacts to inflation implies $\beta_\pi > 1$. The explanation for this is as follows. An increase in inflation of x will lead to an increase in nominal interest rates of $\beta_\pi x$ so real interest rates change by $(\beta_\pi - 1)x$. This means that if $\beta_\pi > 1$ an increase in inflation leads to higher real interest rates and, via the IS curve relation, to lower output. So if $\beta_\pi > 1$ then the IS-MP curve can be depicted as a downward-sloping curve. In contrast, if $\beta_\pi < 1$, then an increase in inflation leads to lower real interest rates and higher output, implying an upward-sloping IS-MP curve.

For now, we will assume that $\beta_\pi > 1$ so that we have a downward-sloping IS-MP curve but we will revisit this issue later. Figure 1.8 thus shows what the IS-MP curve looks like when the aggregate demand shock $\epsilon_t^y = 0$. Again, the key thing to notice is the value of inflation that occurs when output equals its natural rate. When $y_t = y_t^*$ we get $\pi_t = \pi^*$. Figure 1.9 shows how the IS-MP curve shifts to the right if there is a positive value of ϵ_t^y corresponding to a positive shock to aggregate demand.

Figure 1.8: The IS-MP Curve with $\epsilon_t^y = 0$

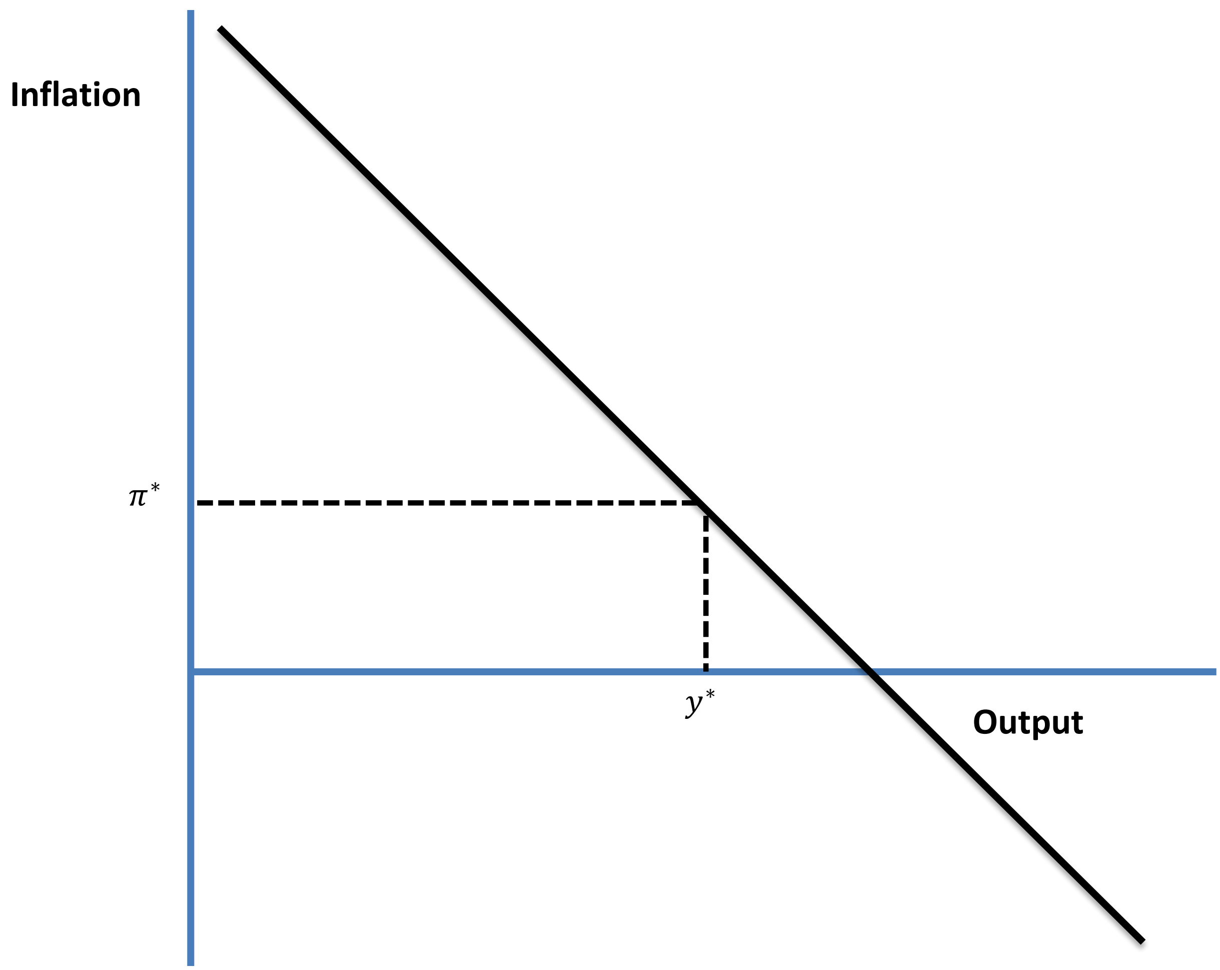
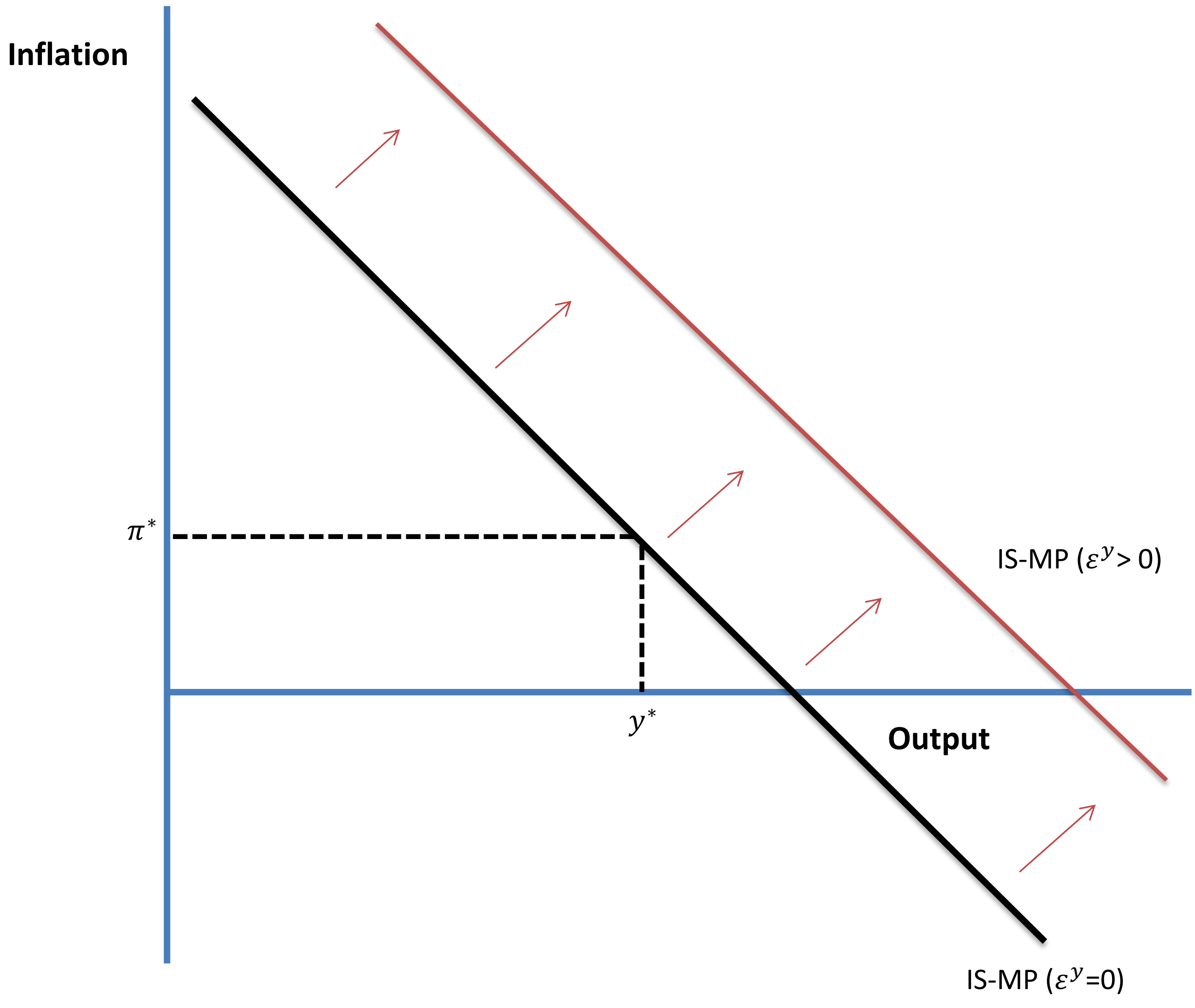


Figure 1.9: The IS-MP curve as we move from $\epsilon_t^y = 0$ to $\epsilon_t^y > 0$

(An Aggregate Demand Shock)

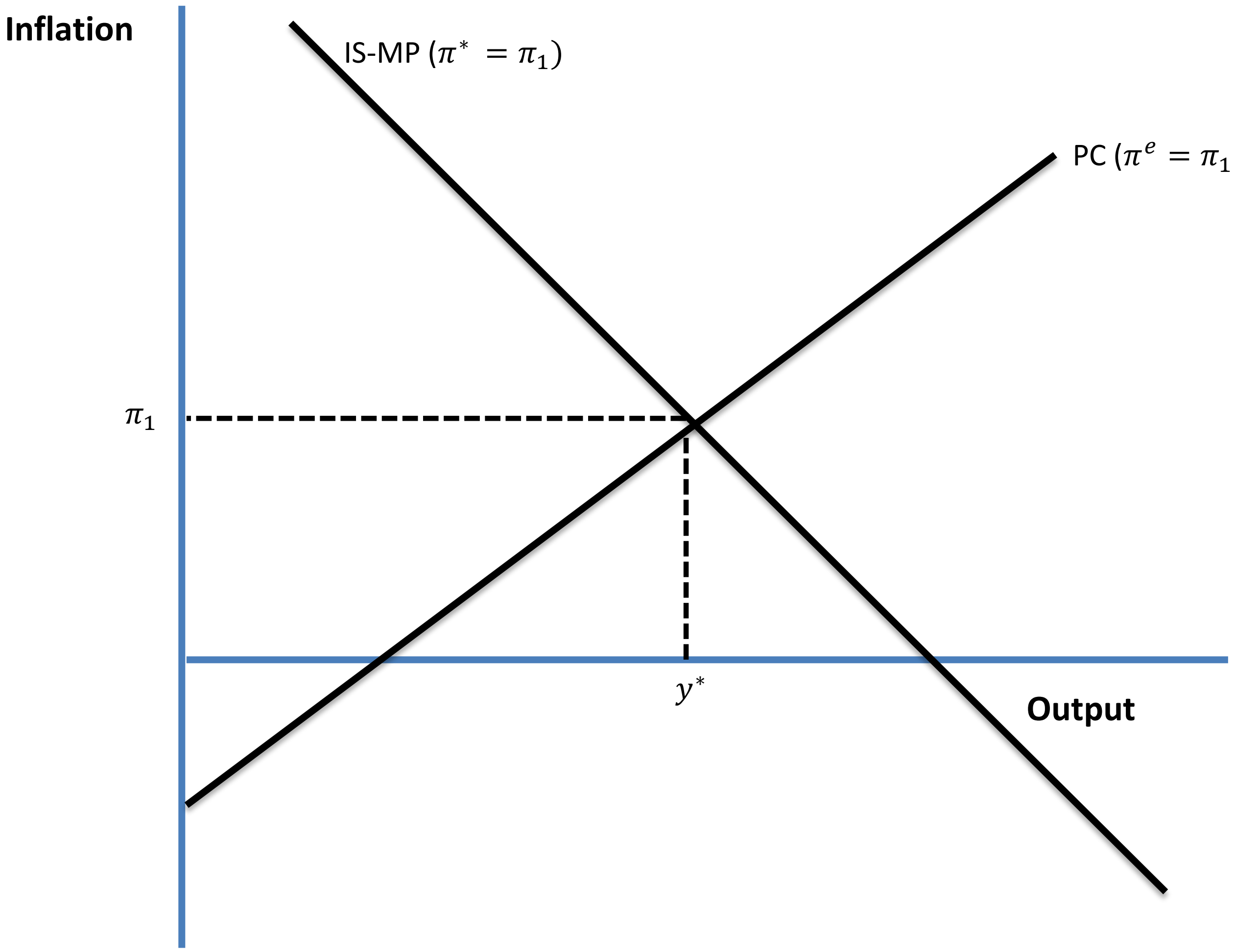


The Full Model

The full IS-MP-PC model can be illustrated in the traditional fashion as a graph with one curve that slopes upwards (the Phillips curve) and one that slopes downwards (the IS-MP curve provided we assume that $\beta_\pi > 1$.) Figure 1.10 provides the simplest possible example of the graph. This is the case where both the temporary shocks, ϵ_t^π and ϵ_t^y equal zero and the public's expectation of inflation is equal to the central bank's inflation target. Note that I have labelled the PC and IS-MP curves to explicitly indicate what the expected and target rates of inflation are and it will be a good idea for you to do the same when answering questions about this model.

In the next chapter, we will analyse this model in depth, examining what happens when various types of events occur and focusing carefully on how inflation expectations change over time.

Figure 1.10: The IS-MP-PC Model When Expected Inflation Equals the Inflation Target



A More Complicated Monetary Policy Rule: The Taylor Rule

Before moving on to analyse the model in more depth, I want to describe the more complex version of the monetary policy rule that I alluded to earlier. This rule takes a form that is associated with Stanford economist John Taylor. In a famous paper published in 1993 called "Discretion Versus Policy Rules in Practice" (you will find a link on the class webpage) Taylor noted that Federal Reserve policy in the few years before his paper was published seemed to be characterised by a rule in which interest rates were adjusted in response to both inflation and the gap between output and an estimated trend.

Within our model structure, we can amend our monetary policy rule to be more like this "Taylor rule" if we make it take the following form:

$$i_t = r^* + \pi^* + \beta_\pi (\pi_t - \pi^*) + \beta_y (y_t - y_t^*) \tag{1.19}$$

It turns out that the properties of the IS-MP-PC model don't really change if we adopt this more complicated monetary policy rule. If we substitute the expression for the nominal interest rate in (1.19) into the IS curve equation (1.5), we get

$$y_t = y_t^* - \alpha \left[r^* + \pi^* + \beta_\pi (\pi_t - \pi^*) + \beta_y (y_t - y_t^*)\right] + \alpha (\pi_t + r^*) + \epsilon_t^y \tag{1.20}$$

This can be re-arranged as follows (canceling out the terms involving r^*):

$$y_t - y_t^* = -\alpha\beta_y (y_t - y_t^*) - \alpha\beta_\pi (\pi_t - \pi^*) - \alpha\pi^* + \alpha\pi_t + \epsilon_t^y \tag{1.21}$$

Bringing together all the terms involving the output gap $y_t - y_t^*$, we get

$$(1 + \alpha\beta_y)(y_t - y_t^*) = -\alpha\beta_\pi (\pi_t - \pi^*) + \alpha (\pi_t - \pi^*) + \epsilon_t^y \tag{1.22}$$

Which can be expressed as

$$y_t - y_t^* = -\frac{\alpha(\beta_\pi - 1)}{1 + \alpha\beta_y}(\pi_t - \pi^*) + \frac{1}{1 + \alpha\beta_y}\epsilon_t^y \tag{1.23}$$

This equation shows us that broadening the monetary policy rule to incorporate interest rates responding to the output gap doesn't change the essential form of the IS-MP curve. As long as $\beta_\pi > 1$, the curve will slope downwards and will feature $\pi_t = \pi^*$ when $y_t = y_t^*$ and there are no inflationary shocks. So while modifying the monetary policy rule so the central bank responds to the output gapchanges the coefficients of the IS-MP model a bit, it doesn't change the underlying economics. In the analysis in the next chapter, we will stick with the model that uses the basic "inflation targeting" monetary policy rule.

Chapter 2

Analysing the IS-MP-PC Model

In our first chapter, we introduced the IS-MP-PC model. We will move on now to examining its properties.

Inflation Expectations and the Inflation Target

Figure 2.1 repeats the simplest possible example of the model. This is the case where both the temporary shocks, ϵ_t^π and ϵ_t^y, equal zero and the public's expectation of inflation equals the central bank's inflation target. Specifically, the graph shows a case where the public's expectation of inflation $\pi_t^e = \pi_1$ and the central bank's inflation target is $\pi^* = \pi_1$. With no temporary shocks, the value of output consistent with $\pi_t = \pi_1$ for the IS-MP curve is y_t^*. Similarly, the value of output consistent with $\pi_t = \pi_1$ for the PC curve is also y_t^*. So the model generates an outcome where $\pi_t = \pi_1$ and $y_t = y_t^*$.

Now consider a case in which the public's inflation expectations shift to being higher than the central bank's target rate. Figure 2.2 illustrates this case. It shows the PC curve shifting upwards to the red line. This position of this red line is determined by the new higher level of expected inflation. Specifically, the public's inflation expectations are now determined by $\pi_t^e = \bar{\pi}$. Note that $\bar{\pi}$ is the higher level of inflation noted on the y-axis and that this level is consistent with $y_t = y_t^*$ in the new Phillips curve described by the red line.

The outcomes for inflation and output of the increase in inflation expectations are described by the intersection of the new red PC line and the old unchanged IS-MP curve. The actual outcome for inflation (denoted as π_2 on the graph) ends up being higher than the central bank's inflation target but lower than the public's inflationary expectations. Output ends up being lower than its natural rate (consistent with a slump or perhaps a full-blown recession) because the higher level of inflation leads the central bank to raise real interest rates which reduces output.

Can We Learn More?

Figure 2.2 is a good example of how we can use graphs to illustrate a model's properties. It gets the basic story across as to what happens when inflation expectations rise above target when the central bank is pursuing a monetary policy rule that increases real rates in response to higher inflation.

Still, one could look to dig a bit deeper. The inflation outcome as drawn in Figure 2.2 is slightly more than halfway towards the public's inflation expectations relative to the central bank's inflation target. But what actually determines this outcome? In other words, what determines how far away from target inflation will move when the public's inflation expectations change? How much does it depend on the monetary policy rule? How much does it depend on other aspects of the model, like the impact of real interest rates on output and the impact of output on inflation? It would be tricky to get these answers from a graph. However, using the equations underlying the model, we can get a full solution that fully answers all these questions.

Figure 2.1: The IS-MP-PC Model When Expected Inflation Equals the Inflation Target

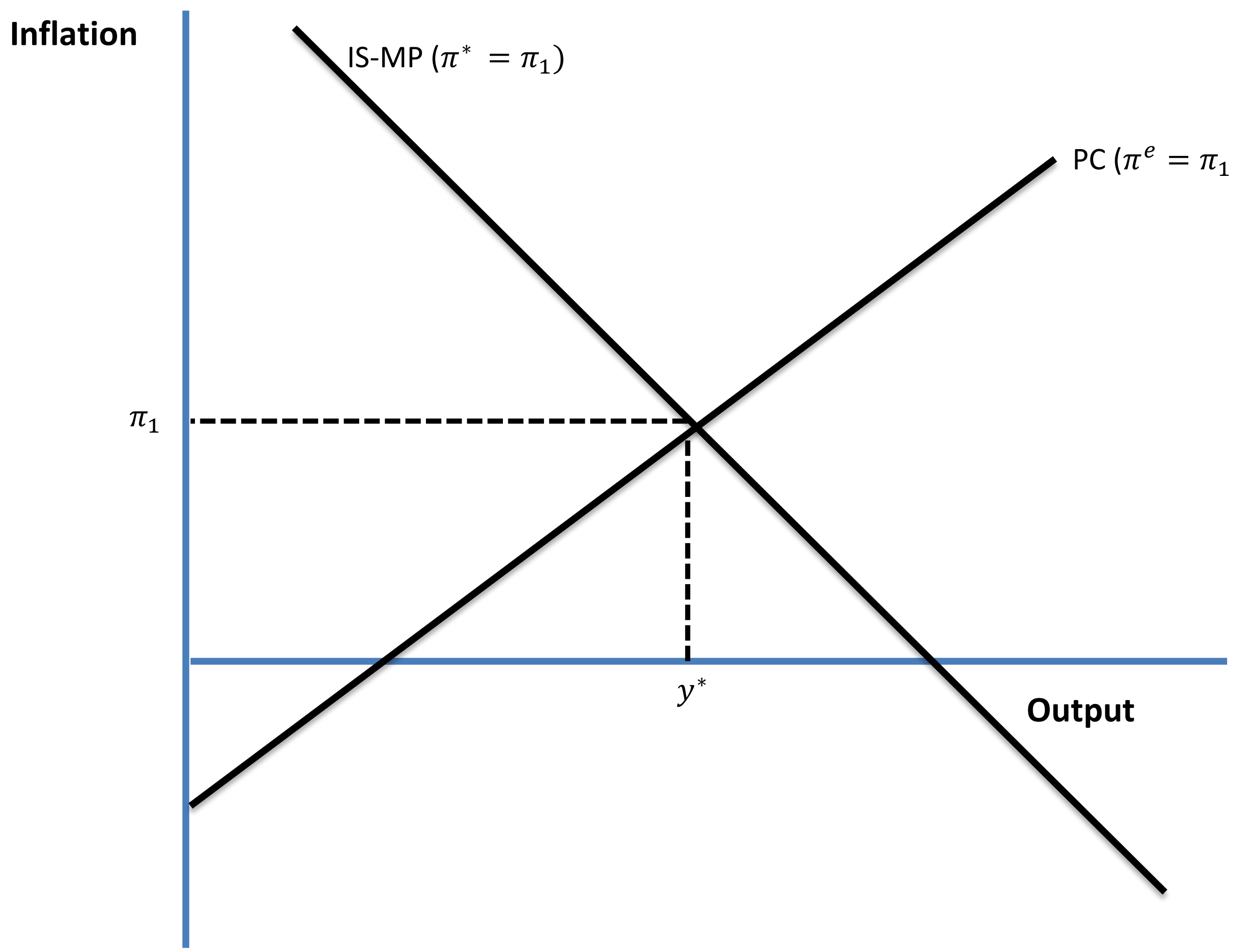
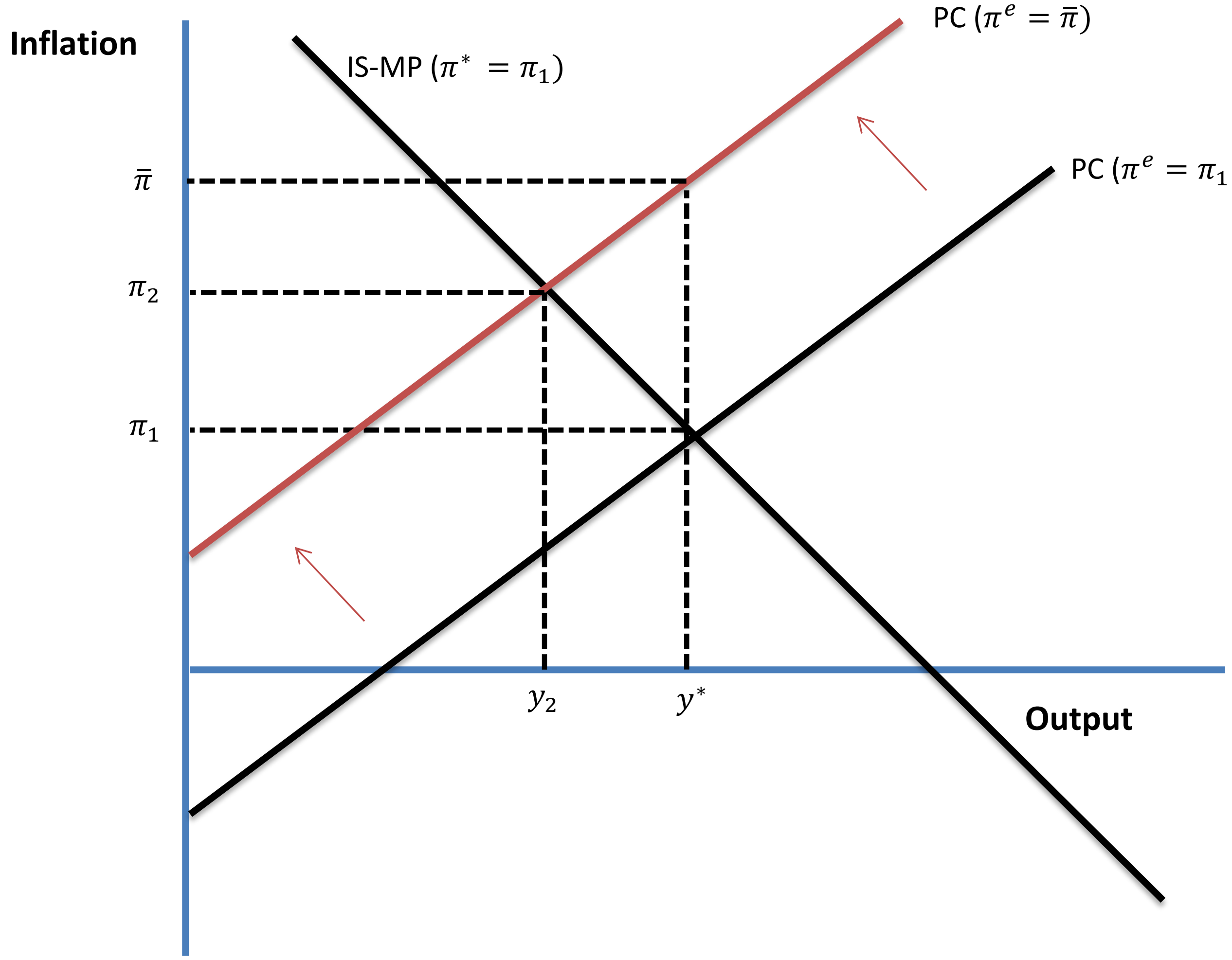


Figure 2.2: The IS-MP-PC Model When Expected Inflation Rises Above the Inflation Target



The IS-MP-PC Model Solution for Inflation

Let's repeat the equations describing our two curves as presented in the previous chapter. The PC curve is

$$\pi_t = \pi_t^e + \gamma \left(y_t - y_t^* \right) + \epsilon_t^\pi \tag{2.1}$$

And the IS-MP curve is

$$y_t = y_t^* - \alpha \left(\beta_\pi - 1 \right) \left(\pi_t - \pi^* \right) + \epsilon_t^y \tag{2.2}$$

Taking all the other elements of the model as given, we can view this as two linear equations in the two variables π_t and y_t. These equations can be solved to give solutions that describe how these two variables depend on all the other elements of the model.

This can be done as follows. First, we will derive a complete expression for the behaviour of inflation and then derive an expression for output. We derive the expression for inflation by starting with the Phillips curve and replacing the output gap term $y_t - y_t^*$ with the variables that the IS-MP curve tells us determines this gap. This gives us the following equation:

$$\pi_t = \pi_t^e + \gamma \left[-\alpha \left(\beta_\pi - 1 \right) \left(\pi_t - \pi^* \right) + \epsilon_t^y \right] + \epsilon_t^\pi \tag{2.3}$$

Adding the term $\alpha\gamma \left(\beta_\pi - 1 \right) \pi_t$ to both sides we get

$$\left[1 + \alpha\gamma \left(\beta_\pi - 1 \right) \right] \pi_t = \pi_t^e + \alpha\gamma \left(\beta_\pi - 1 \right) \pi^* + \gamma\epsilon_t^y + \epsilon_t^\pi \tag{2.4}$$

Now dividing each side by $1 + \alpha\gamma \left(\beta_\pi - 1 \right)$, we get that inflation is determined by

$$\pi_t = \left(\frac{1}{1 + \alpha\gamma \left(\beta_\pi - 1 \right)} \right) \pi_t^e + \left(\frac{\alpha\gamma \left(\beta_\pi - 1 \right)}{1 + \alpha\gamma \left(\beta_\pi - 1 \right)} \right) \pi^* + \frac{\gamma\epsilon_t^y + \epsilon_t^\pi}{1 + \alpha\gamma \left(\beta_\pi - 1 \right)} \tag{2.5}$$

There are a lot of symbols in this equation, which make it a bit hard to read. One way to simplify it is to take the term multiplying inflation expectations and denote it by a single symbol. In this case, we will denote it by the Greek letter θ. So, we define this as

$$\theta = \frac{1}{1 + \alpha\gamma \left(\beta_\pi - 1 \right)} \tag{2.6}$$

Having done this, we can re-write the equation for inflation as

$$\pi_t = \theta \pi_t^e + (1-\theta)\,\pi^* + \theta\left(\gamma \epsilon_t^y + \epsilon_t^\pi\right) \tag{2.7}$$

This equation shows that, apart from the shocks to output and inflation (the $\theta\left(\gamma \epsilon_t^y + \epsilon_t^\pi\right)$ terms) inflation is a weighted average of the public's inflation expectations and the central bank's inflation target i.e. it must lie between these two values as long as $0 < \theta < 1$ (which it should be). What determines whether inflation depends more on the public's expectations or the central bank's target? In other words, what determines the value of θ? Three different factors determine this value.

1. γ: This is the parameter that determines how inflation changes when output changes. The central bank can only influence inflation by influencing output. If the effect of output on inflation gets bigger, then the central bank's inflation target will have more influence on the outcome for inflation.

2. α: This is the parameter that determines how output changes when real interest rates change. If the effect of interest rates on output gets bigger, then the central bank's inflation target will have more influence on the outcome for inflation.

3. β_π: Let's continue to assume $\beta_\pi > 1$ (we'll return to this in the next chapter). Then as β_π gets bigger, the central bank is reacting more to inflation being above its target level. So this parameter getting bigger means less weight on inflation expectations in determining the outcome for inflation and more weight on the central bank's inflation target.

While the calculations here may seem difficult, they illustrate that a formal mathematical solution can sometimes give us a much more complete insight into the properties of a model than graphs. While graphs are often useful at illustrating a particular feature of a model, they also often fall short of explaining its full properties.

The IS-MP-PC Model Solution for Output

Next we provide an expression for output. Looking at the IS-MP curve, we see that the output gap depends on how far inflation is from the central bank's target as well as the "supply shock" term ϵ_t^π. We can use the equation determining inflation, equation (2.7), to get an expression for the gap between inflation and the target level. Subtract π^* from either side of equation (2.7) and you get

$$\pi_t - \pi^* = \theta \left(\pi_t^e - \pi^* + \epsilon_t^\pi + \gamma \epsilon_t^y\right) \tag{2.8}$$

We can now replace the $\pi_t - \pi^*$ on the right-hand-side of the IS-MP curve, equation (1.18), with the right-hand-side of the equation above. This gives

$$y_t = y_t^* - \theta\alpha \left(\beta_\pi - 1\right) \left(\pi_t^e - \pi^* + \epsilon_t^\pi + \gamma \epsilon_t^y\right) + \epsilon_t^y \tag{2.9}$$

which can be simplified to

$$y_t = y_t^* - \theta\alpha \left(\beta_\pi - 1\right) \left(\pi_t^e - \pi^* + \epsilon_t^\pi\right) + \left(1 - \theta\alpha\gamma \left(\beta_\pi - 1\right)\right) \epsilon_t^y \tag{2.10}$$

This equation tells us that whether output is above or below target depends upon the gap between expected inflation and the inflation target as well as on the two temporary shocks ϵ_t^y and ϵ_t^π. Provided we have the usual condition that $\beta_\pi > 1$, the combined coefficient $-\theta\alpha \left(\beta_\pi - 1\right)$ is negative. This means that increases in the public's inflation expectations relative to the inflation target end up having a negative effect on output. Inflationary supply shocks (positive values for ϵ_t^π) also have a negative effect on output while positive aggregate demand shocks ($\epsilon_t^y > 0$) have a positive effect on output.

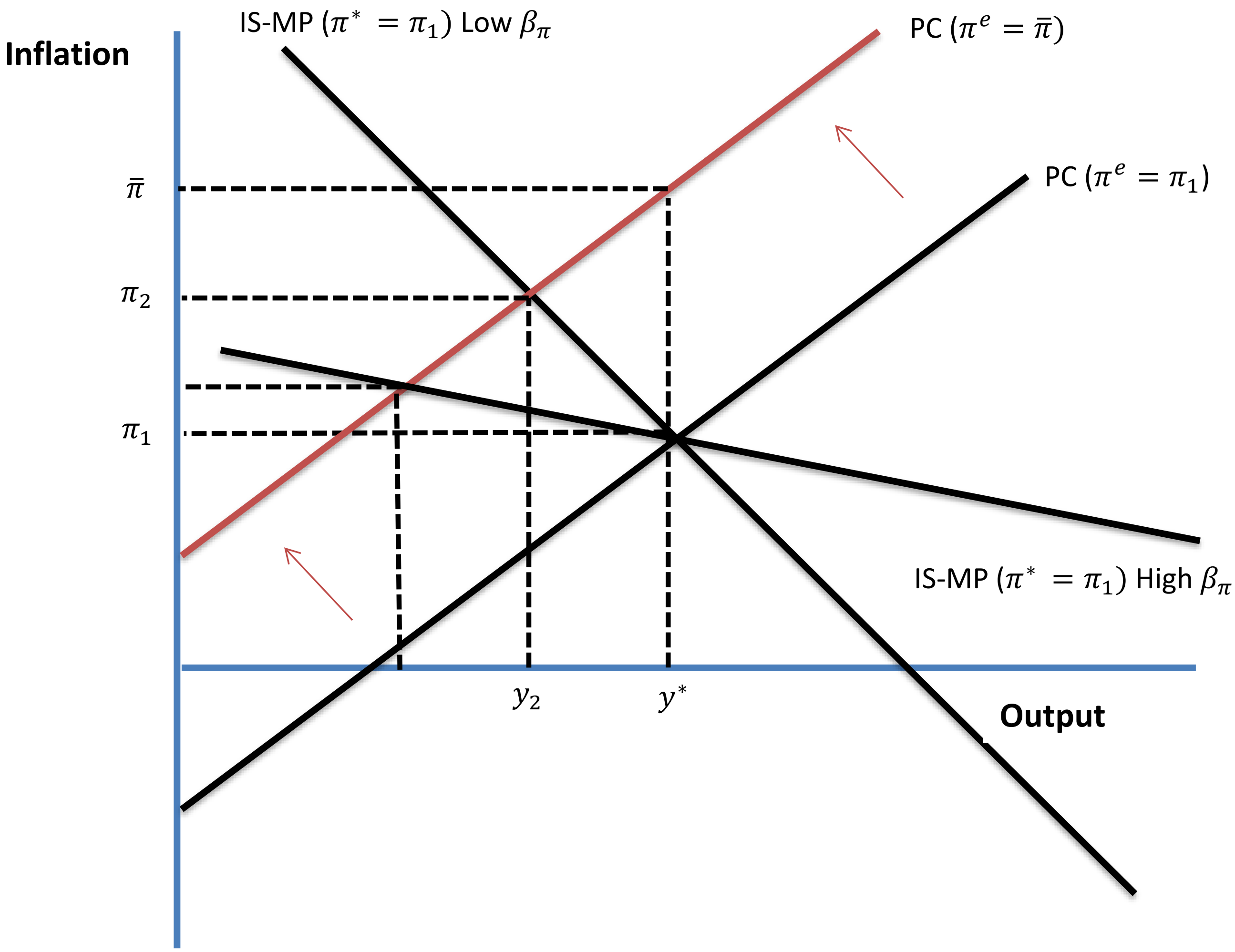
How far does output fall short of its natural rate when inflation expectations rise above the central bank's target? The coefficient determining this is $-\theta\alpha\left(\beta_\pi - 1\right)$. This can be re-expressed as

$$-\theta\alpha\left(\beta_\pi - 1\right) = \frac{\alpha\left(\beta_\pi - 1\right)}{1 + \alpha\gamma\left(\beta_\pi - 1\right)} \tag{2.11}$$

Calculating partial derivatives, you find that the size of the short fall in output depends positively on α and β_π. In other words, the larger the impact of interest rates on output and the larger the central bank's interest rate response to inflation, the larger the shortfall in output will be when inflation expectations rise above the central bank's target. In contrast, the output shortfall depends negatively on γ, the parameter determining the effect of output on inflation: As γ gets bigger, the central bank requires a smaller shortfall in output to implement its policy of getting inflation back to target.

The calculations here tell us that the more aggressive a central bank is in its response to inflation—the higher the value of β_π—then the smaller the rise in inflation will be and the larger the drop in output will be. We can illustrate this graphically by comparing Figure 2.2 with what would have happened if the IS-MP curve had been flatter: A higher value of β_π means a flatter IS-MP curve, meaning each unit increase in inflation is associated with a more aggressive policy response from the central bank and thus a larger fall in output. Figure 2.3 overlays a second, flatter, IS-MP curve on top of Figure 2.2. As with the original IS-MP curve, this curve generated by a higher β_π also intersects with the original curve so that $\pi_t = \pi^*$ and $y_t = y_t^*$ but after the Phillips curve shifts up, it generates a smaller increase in inflation and a larger decrease in output.

Figure 2.3: A Rise in Expected Inflation For Two Values Of $\beta\pi$



How Do Inflation Expectations Change?

Let's go back to Figure 2.2 now. We have seen that after the public's inflation expectations rise, the result is a fall in output below its natural rate and in increase inflation, though this increase is smaller than had been expected by the public. What happens next? How does the public's expectation of inflation change at this point?

Friedman's 1968 paper *The Role of Monetary Policy* suggested that people gradually adapt their expectations based on past outcomes for inflation. Consider now a simple model of this idea of "adaptive expectations" by assuming that, each period, the expected level of inflation is simply equal to the level that prevailed last period. Formally, this can be written as

$$\pi_t^e = \pi_{t-1} \tag{2.12}$$

Under this formulation of expectations, the Phillips curve becomes

$$\pi_t = \pi_{t-1} + \gamma \left(y_t - y_t^*\right) + \epsilon_t^\pi \tag{2.13}$$

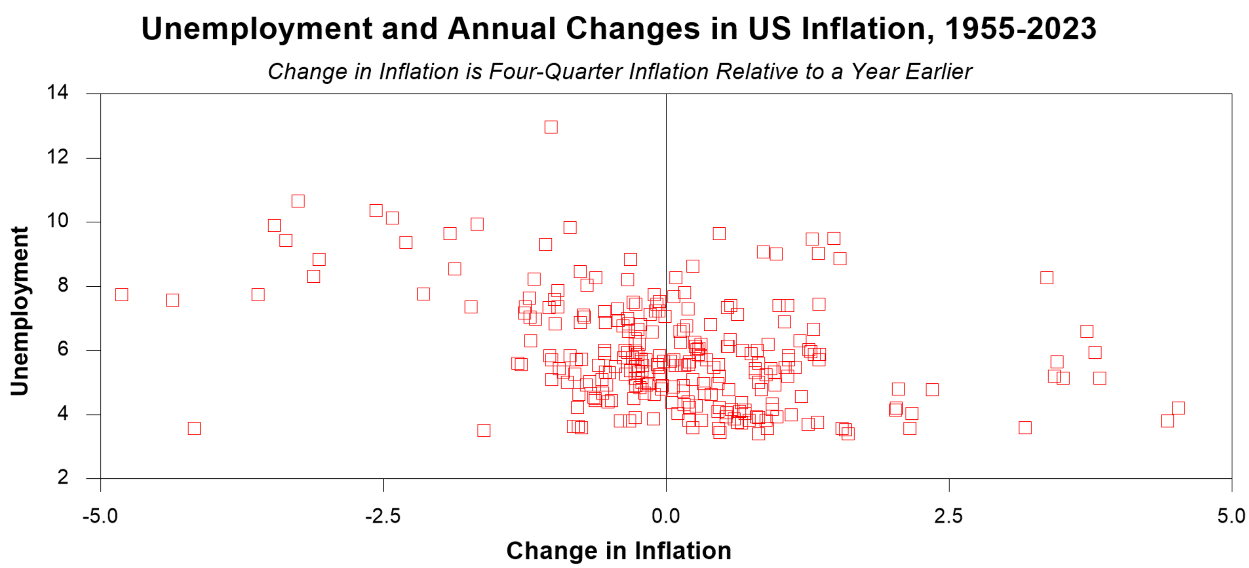
Note that if we subtract π_{t-1} from both sides of this equation, it becomes

$$\pi_t - \pi_{t-1} = \gamma \left(y_t - y_t^*\right) + \epsilon_t^\pi \tag{2.14}$$

In other words, there should be a positive relationship between the *change* in inflation and the output gap. There are various methods for measuring output gaps but one quick and easy method is to use the unemployment rate as an indication of what the output gap might be. If unemployment is high, then output is likely to be below its natural rate so the output gap is negative. In contrast, a low unemployment rate is an indicator that the output gap is likely to be positive. So if the adaptive expectations formulation of the Phillips curve was correct, then we would expect to see a negative relationship in the data between the change in inflation and the unemployment rate.

Figure 2.4 uses the same US quarterly data that we used for Figure 1.4 in the last chapter. That figure showed that there was very little relationship between the level of the unemployment rate and the level of inflation. In contrast, Figure 2.4 shows a scatter plot of datapoints on the change in inflation (measured as the four quarter percentage change in the price level minus the percentage change in the price level over the preceding four quarters) and the unemployment rate. In contrast to the basic Phillips curve, this adaptive-expectations-style Phillips curve shows a clear and strong negative relationship between the change in inflation and the unemployment rate.

Figure 2.4: Evidence for Adaptive Inflation Expectations



These results suggest that the adaptive expectations approach appears to provide a reasonable model of how people formulate inflation expectations. That said, people are unlikely to simply use mechanical formulas to arrive at their expectations and one can imagine conditions in which people's inflation expectations could radically depart from what had happened in the past e.g. the appointment of a new central bank governor with a different approach to inflation, the adoption of a new currency or other major events. Let's examine for now, however, how the IS-MP-PC model behaves when people have adaptive inflationary expectations.

Adjustment of Inflation Expectations

After inflation expectations moved up to $\bar{\pi}$, the outcome was that inflation moved from π_1 (which is the central bank's inflation target) to π_2. If people follow adaptive expectations then the next period, they will set $\pi^e = \pi_2$. Figure 2.5 shows what happens after this. The PC curve moves back downwards and inflation moves down to a lower level, denoted on this graph by π_3. Figure 2.6 indicates how the process plays out. If the public has adaptive expectations, then inflation and output gradually converge back to the point where output is at its natural rate and inflation equals the central bank's target rate.

Here we have illustrated the implications of an increase in inflation expectations away from the central bank's inflation target. But if the public has adaptive expectations, how could inflation expectations just jump upwards? Rather than a random unexplained increase in inflation expectations, the more likely explanation for the Phillips curve shifting upwards is temporary supply shocks, i.e. ϵ_t^π is positive for a number of periods. Under adaptive expectations, the public becomes used to higher inflation and so the Phillips curve will remain above its long-run position even after the temporary supply shock has been reversed.

Figure 2.5: Inflation Expectations Adjusting Back Downwards

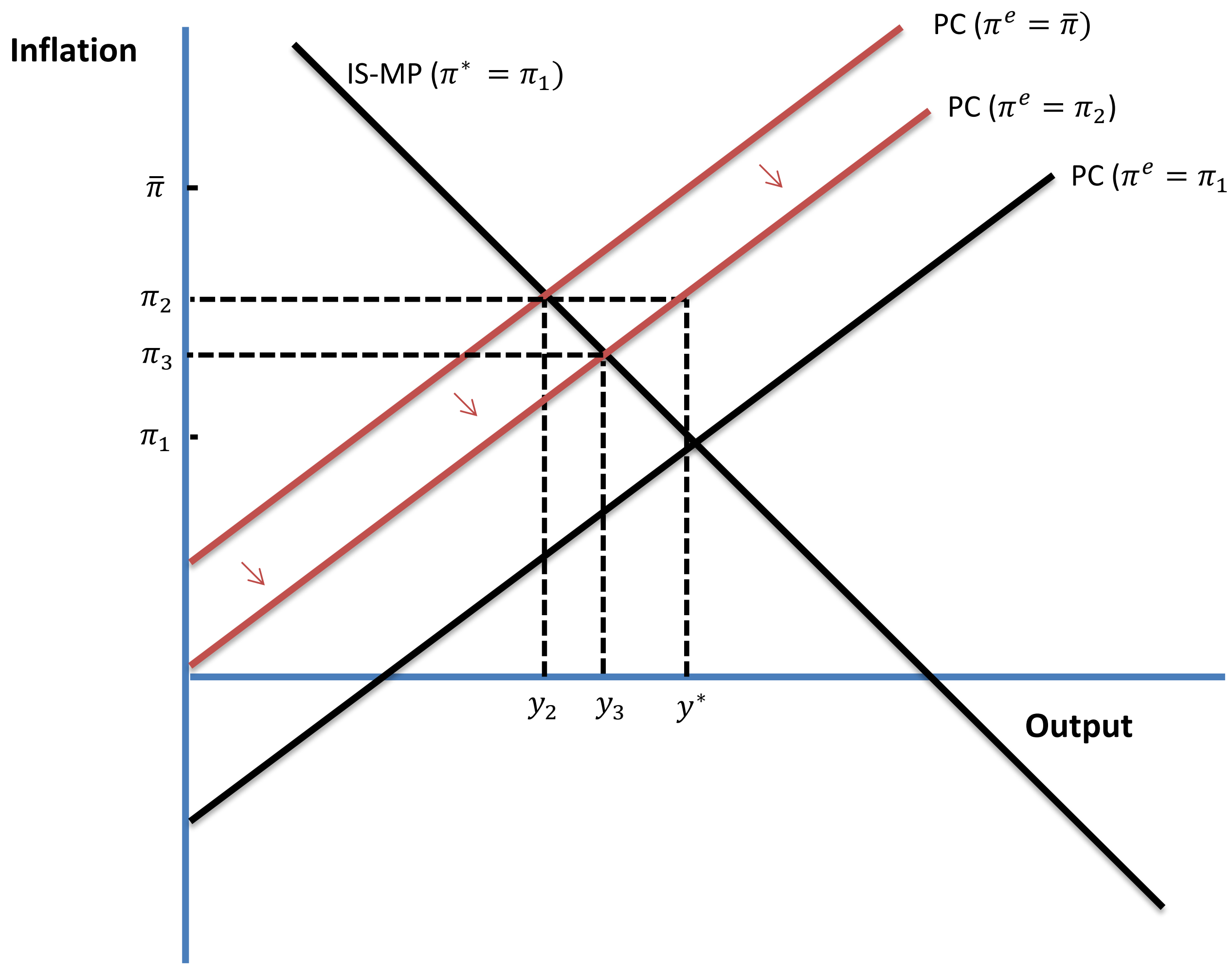
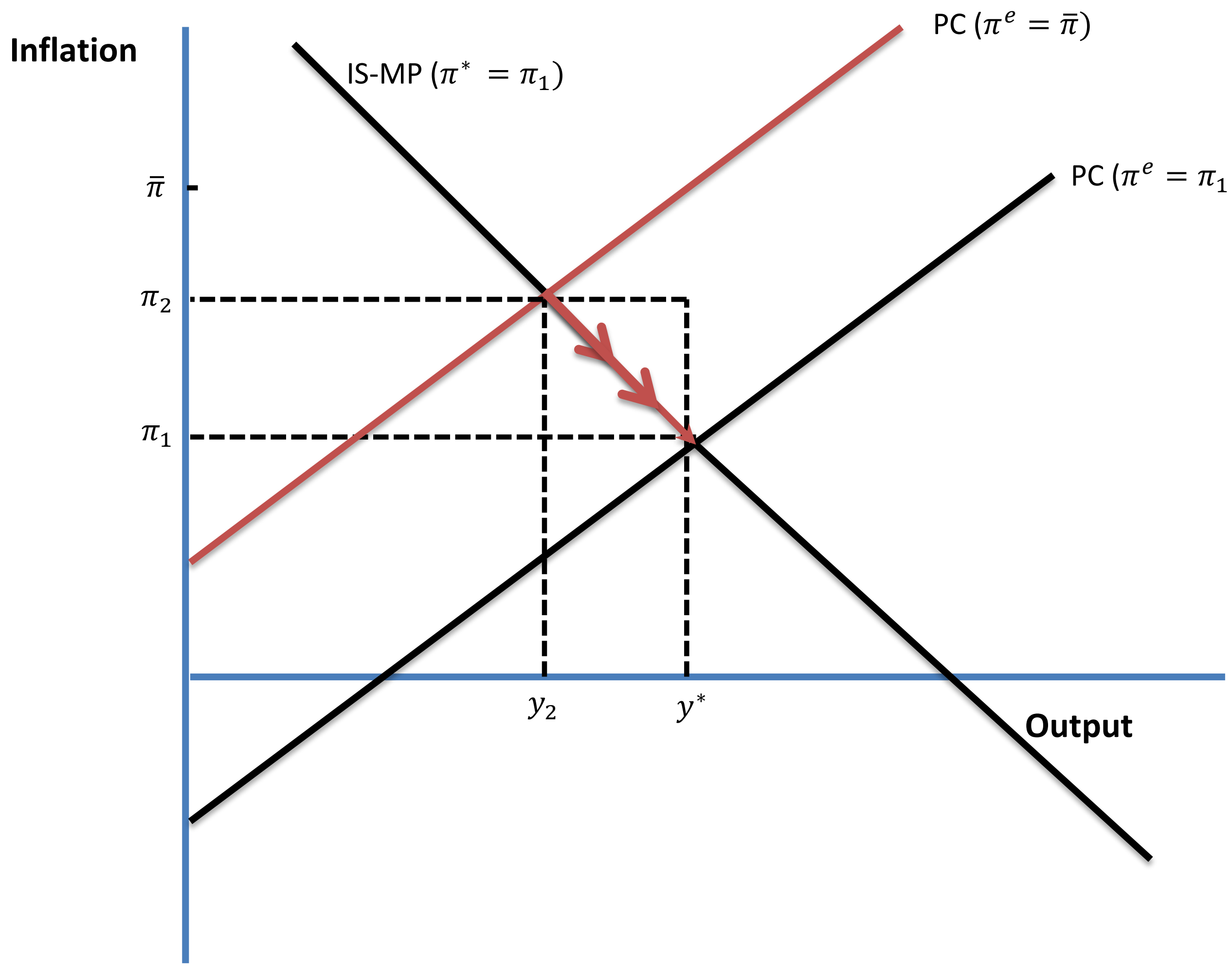


Figure 2.6: Inflation and Output Adjust Back to Starting Position



Inflation and Output Dynamics for Soft and Tough Central Banks

Do we want a "soft" central bank that limits the increase in real interest rates when inflation rises to protect the economy and which isn't too concerned about getting inflation back to target quickly? Or do we want a "tough" central bank that raises interest rates aggressively and is very concerned about getting inflation back to target?

The model doesn't give a clear answer between these two options. Both have positive and negative aspects. If the public's inflation expectations behave in an adaptive fashion, then central banks have a choice between different types of adjustments. We showed above in Figure 2.3 that a central bank that acts more aggressively to inflation—that has a greater β_π—produces a smaller increase in inflation but a larger decline in output. However, with adaptive expectations, this larger reduction in output is short-lasting than when β_π is smaller. This is because the initial increase in inflation is smaller, so the central bank is able to return real interest rates to their natural rate faster.

We can illustrate the differences between the two scenarios by simulating the model on a computer. Figures 2.7 and 2.8 show the results of a computer simulation of the model in which it is assumed that $\pi^* = 2$, that y_t^* is constant at 100 and the other parameters are $\alpha = 1$ and $\gamma = 1$. The model is simulated with two different values for β_π. One version has $\beta_\pi = 1.5$ (this is the "soft" central bank) and other has $\beta_\pi = 3$ (this is the "tough" central bank). Figure 2.7 shows the rise in inflation is smaller and disappears quicker when there is a tough central bank. Figure 2.8 shows that the tough central bank engineers a much larger recession but this ends much quicker. The total average value of output over the whole sample is the same for the two scenarios. This isn't an accident but rather is a feature of the model: A certain amount of cumulative output below its natural rate is required to lower the inflation rate back to the central bank's target.

This suggests central banks face a choice when dealing with high inflation: They can go for the "cold turkey" option and have a sharp but short recession or they can take a softer approach which ends up taking more time to get output and inflation back to target.

Figure 2.7: Inflation Dynamics for High and Low Values of β_π

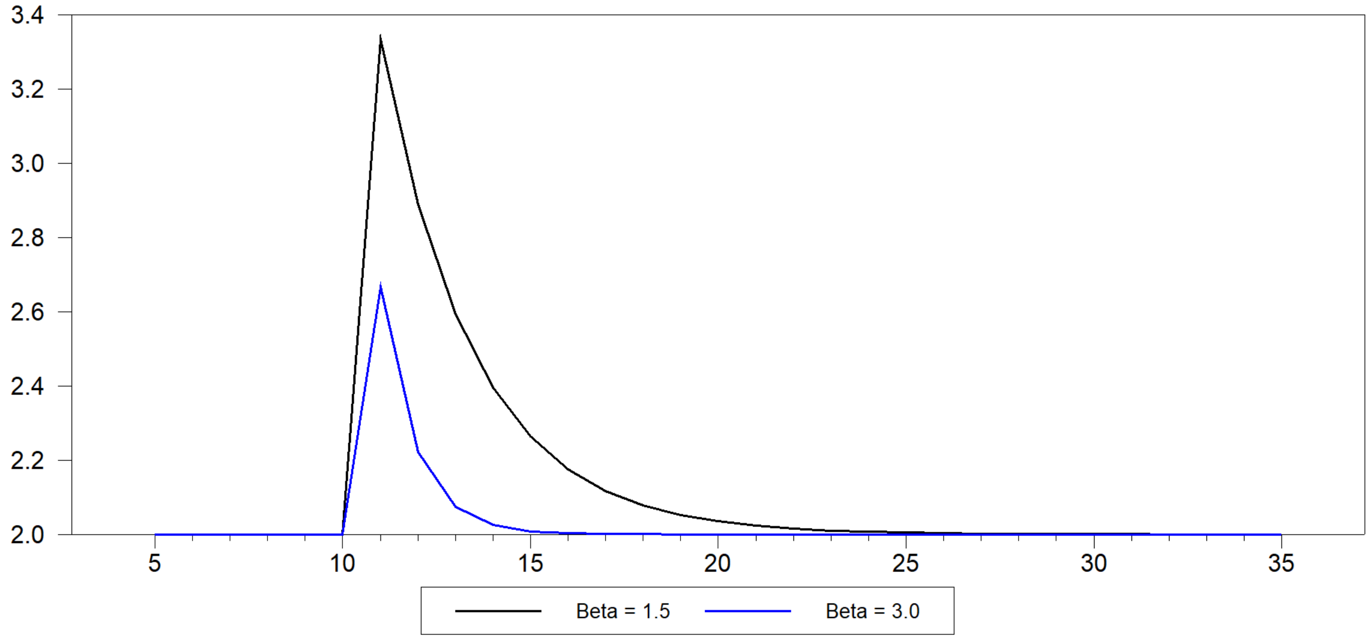
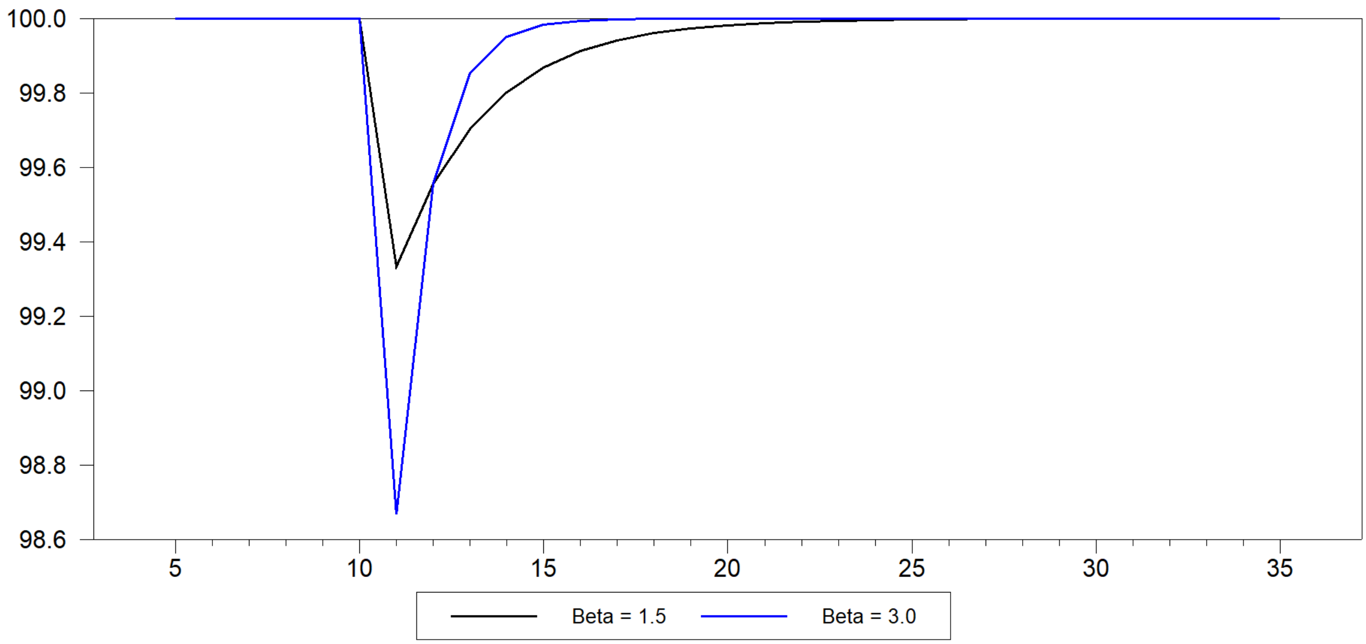


Figure 2.8: Output Dynamics for High and Low Values of β_π



A Temporary Aggregate Demand Shock

Having looked at what happens under adaptive expectations when the Phillips curve shifts, let's consider what happens when we have a temporary shock to aggregate demand, so ϵ_t^y takes a different value from zero, which means a shift in the IS-MP curve. In Figures 2.9 to 2.12, we illustrate a case where there is a shift towards a positive value of ϵ_t^y for a couple of periods but then it shifts back to zero.

Figure 2.9 shows the immediate impact of a positive aggregate demand shock. Output and inflation both go up with inflation reaching the point denoted as π_2 in the figure. If the public has adaptive expectations, then this results in an increase in inflation expectations the following period. Figure 2.10 shows what happens when the aggregate demand shock persists but inflation expectations move up to match the previous period's inflation rate. The inflation rate now rises again to π_3. Figure 2.11 shows how this triggers a further increase in inflation in the next period as inflation expectations move up from π_2 to π_3.

Figure 2.12 shows what happens if the aggregate demand shock then reverses itself in the next period. The IS-MP curve shifts back to its original position but the Phillips curve remains elevated. The result is a nasty combination of high inflation and output below its natural rate. Figure 2.12 contains arrows showing the full set of movements generated by this aggregate demand shock:

- An increase in output and inflation as the shock hits.
- A further increase in inflation as inflation expectations adjust upwards, accompanied by a decline in output.
- A decline in output and inflation as the shock disappears.

- A further decline in inflation accompanied by an increase in output as inflationary expectations gradually return to the central bank's target.

This chart shows that when the public has adaptive inflation expectations, temporary positive aggregate demand shocks lead to counter-clockwise loops on graphs that have output on the x-axis and inflation on the y-axis.

It turns out that much of the data on inflation and output correspond to these kinds of movements. Figure 2.13 is borrowed from Stanford economist Charles I. Jones's website. They show the data from US on inflation and an estimated output gap from 1960 to 1983. The figure shows a number of periods of clear counter-cyclical movements. Figure 2.14 shows the same data from 1983-2009. This figure also shows some evidence counter-cylical loops, thought the movements are smaller than the for the pre-1983 period.

Figure 2.9: A Temporary Aggregate Demand Shock ($\epsilon_t^y > 0$)

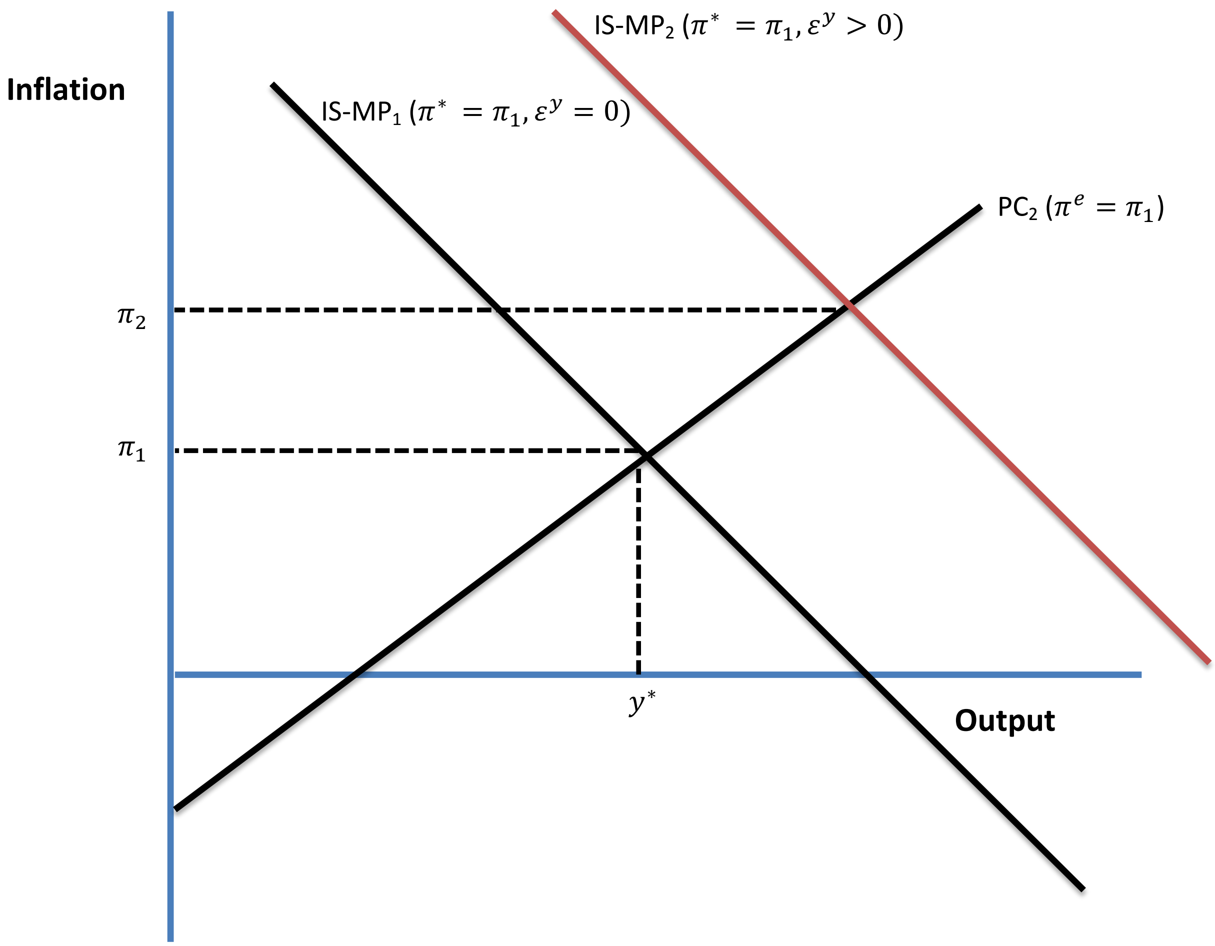


Figure 2.10: Inflation Expectations Adjust Upwards to π_2

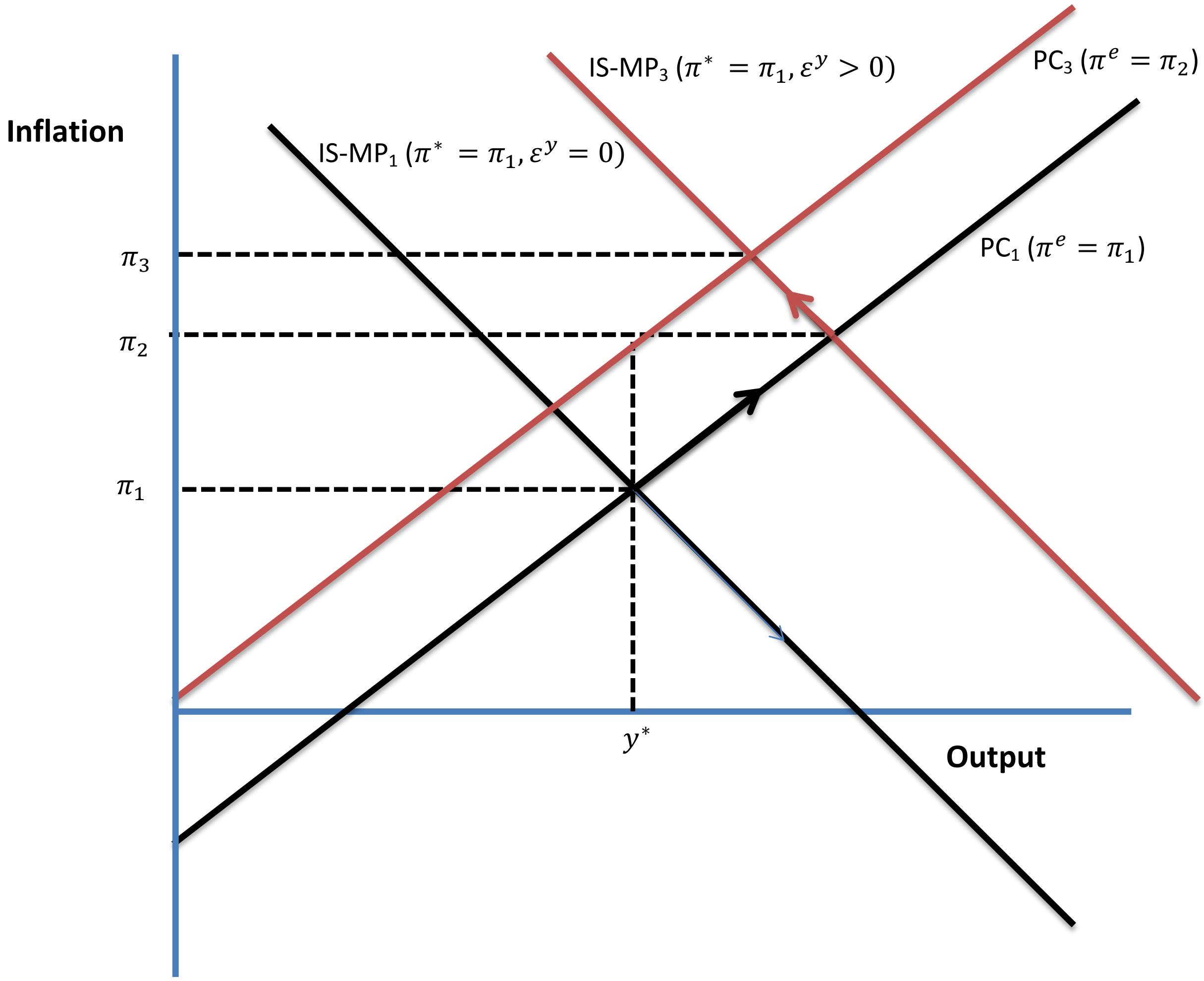


Figure 2.11: Inflation Expectations Adjust Upwards Further to π_3

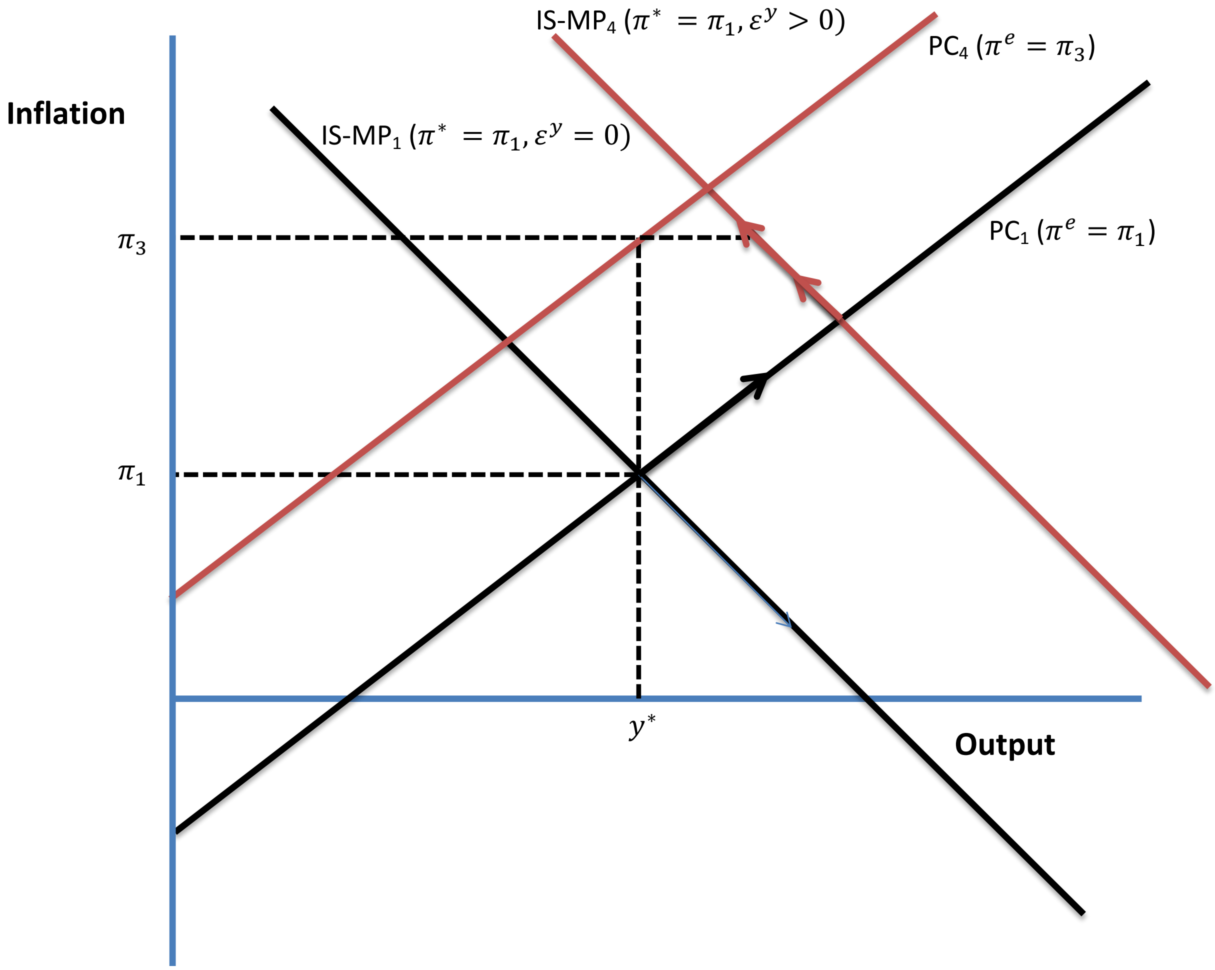


Figure 2.12: Reversal of Aggregate Demand Shock Leads to Recession With High Inflation

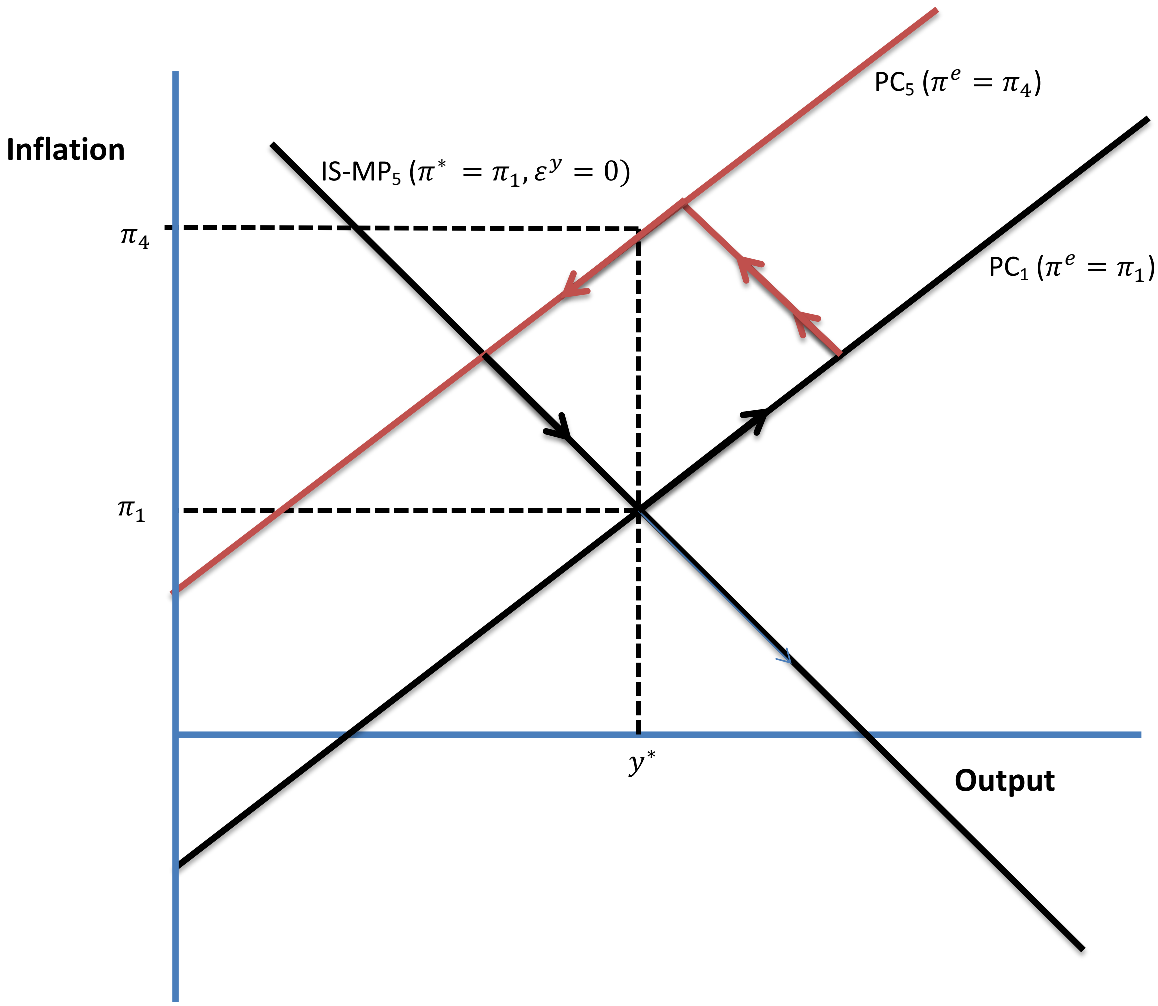


Figure 2.13: From Chad Jones: US Inflation-Output Loops 1960-1983

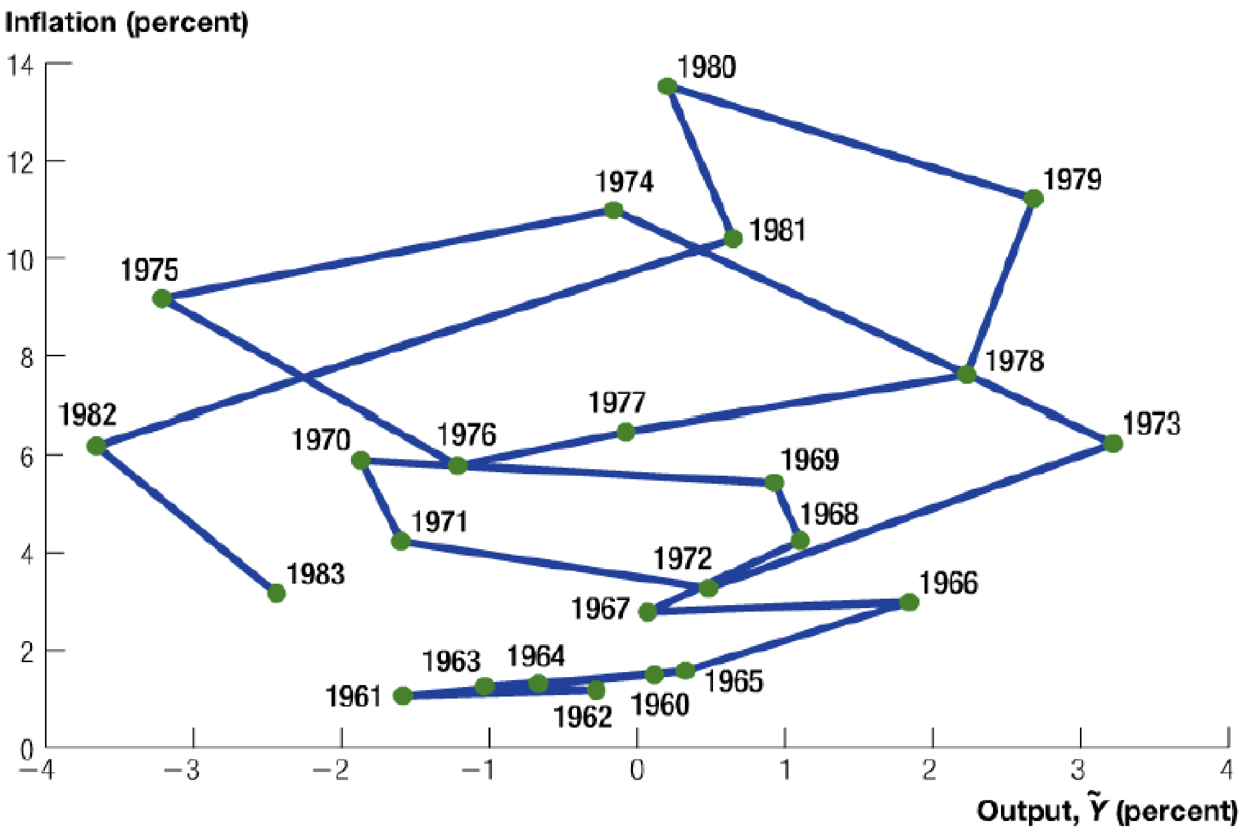
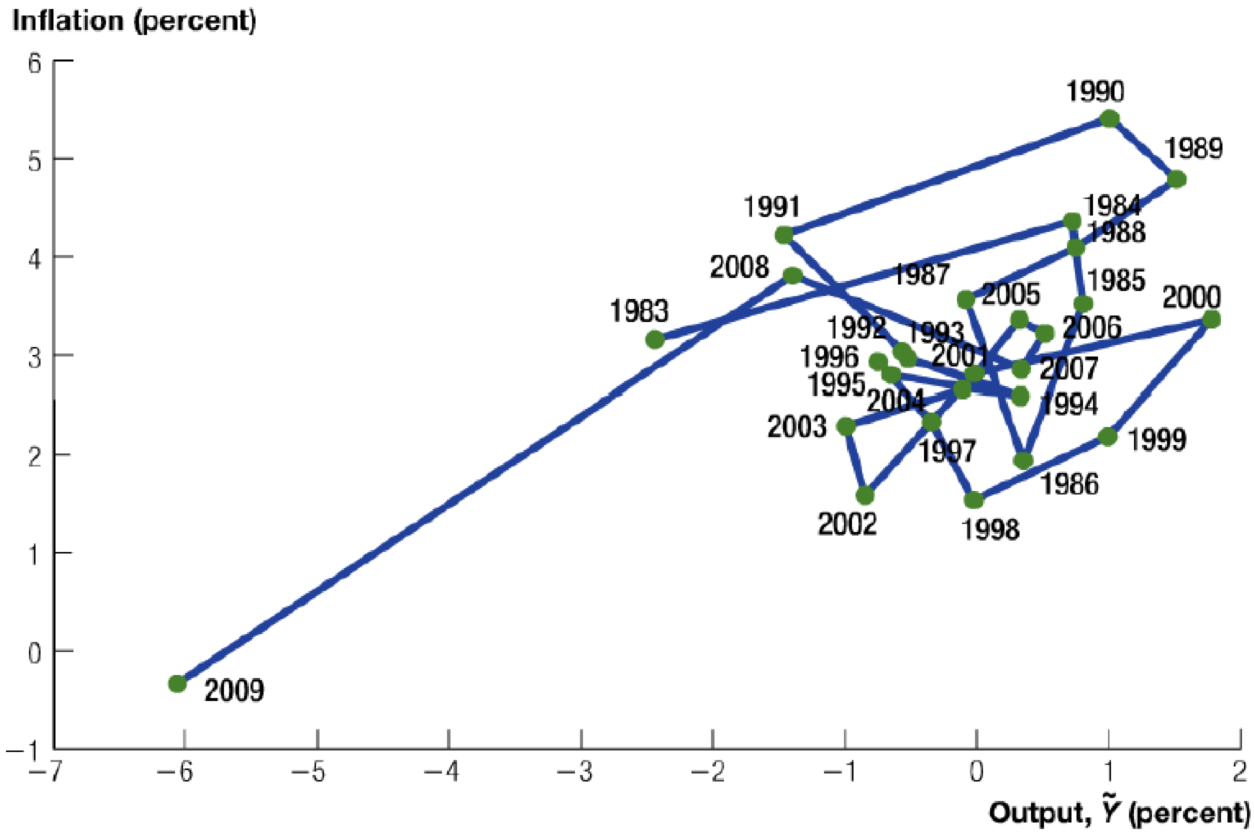


Figure 2.14: From Chad Jones's Notes: US Inflation-Output Loops 1983-2009



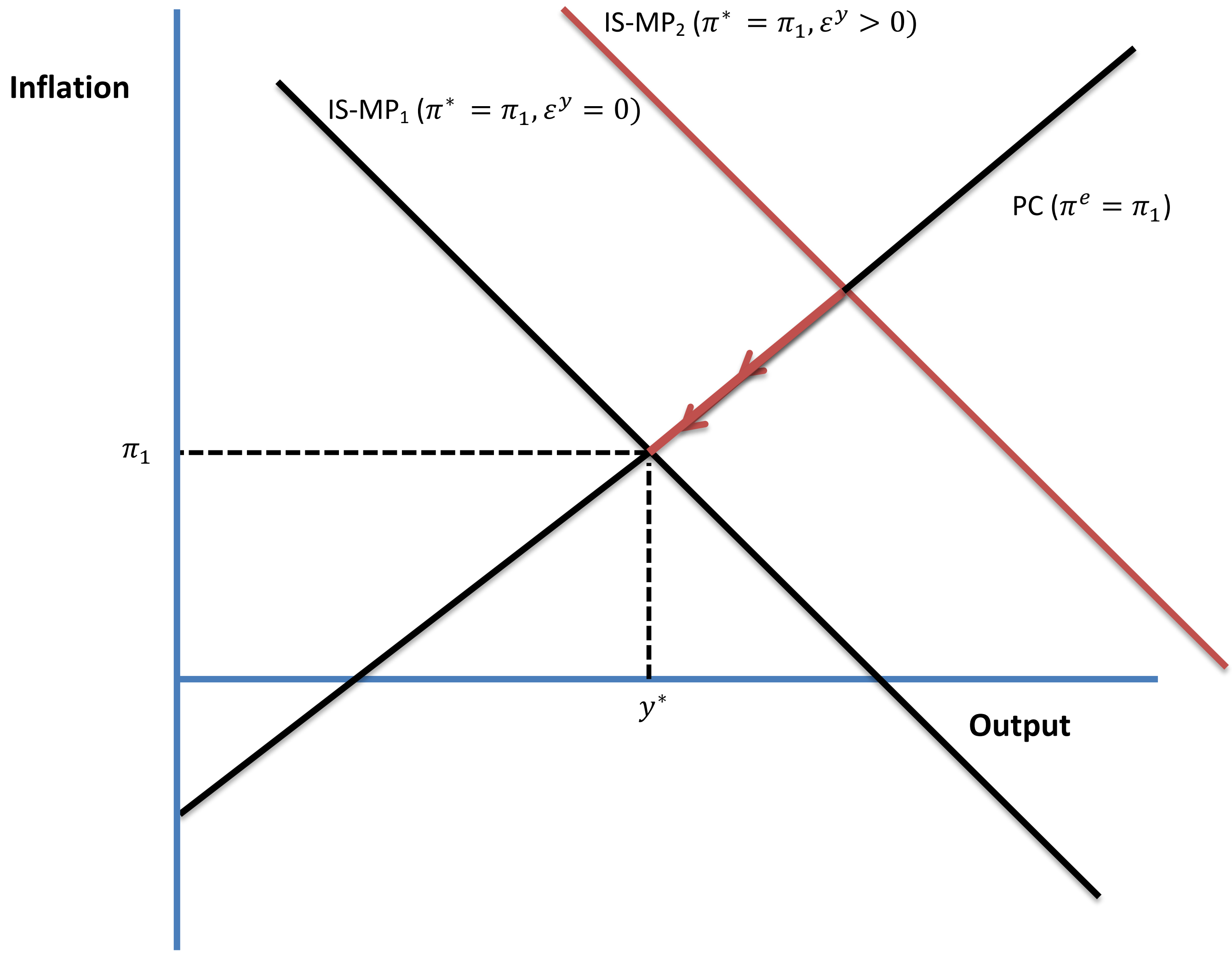
What If Inflation Expectations Don't Adjust?

The evidence presented in Figure 2.4 suggests that adaptive expectations seems to be a reasonable model for how people have formulated their expectations of inflation. And it can be argued that it is a fairly convincing model of how people behave: Most people don't have the time or knowledge to fully understand exactly what's going in the economy and anticipating that last year's conditions provide a guide to what will happen this year probably works well enough for most people. Indeed, if the value of θ is relatively high, then inflation will only change slowly under adaptive expectations, making the adaptive expectations assumption more accurate.

All that said, it is also possible to imagine situations in which the public's inflation expectations are not formed adaptively. For example, if the public believes that the central bank will always act to return inflation quickly towards its target, then they may assume that deviations from the target will be temporary.

Figure 2.15 shows how the economy reacts to a temporary positive demand shock when inflation expectations don't change. The outcome here is much nicer than the counter-clockwise cycle described in Figure 2.12. There is no recession at any point, just a short period of output being above its natural rate and inflation being above its target, followed by a return to their starting levels.

Figure 2.15: Adjustment if Inflation Expectations Don't Change



The Importance of Anchoring Inflation Expectations

The previous examples provide further food for thought about what kind of monetary policy we would like a central bank to implement. The more people believe that a central bank is maintaining its low inflation target, the less likely the economy is to go through boom-bust cycles. We can see this by comparing the dynamics from Figure 2.12 where inflation expectations shift over time (perhaps because the public believes the central bank is willing to be flexible about its target) and Figure 2.15, which shows what happens when inflation expectations do not change after an expansionary shock.

These results predict that we get better outcomes if we have a "tough" central bank which the public believes is committed to keeping the economy near its inflation target. How can this outcome be achieved? The academic literature on this topic has suggested a number of different ways:

1. **Political Independence**: A central bank that plans for the long-term (and does not worry about economic performance during election years) is more likely to stick to a commitment to low inflation. So, independence from political control is an important way to reassure the public about the bank's credibility.

2. **Conservative Central Bankers**: If the central banker is known to really dislike inflation—and the public believes this, the economy gets closer to the ideal low inflation outcome even without commitment. So the government may choose to appoint a central banker who is more inflation-averse than they are.

3. **Consequence for Bad Inflation Outcomes**: Introducing laws so that bad things happen to the central bankers when inflation is high is one way to make the public believe the they will commit to a low inflation rate.

These ideas have had a considerable influence on the legal structure of central banks around the world over the past few decades:

1. **Political Independence**: There has been a substantial move around the world towards making central banks more independent. The Bank of England was made independent in 1997 (previously the Chancellor of the Exchequer had set interest rates) and the ECB/Eurosystem is highly independent from political control.

2. **Conservative Central Bankers**: All around the world, central bankers talk much more now about the evils of inflation and the benefits of price stability. This may be because they believe this to be the case. But there is also a marketing element. Perhaps they can face a better macroeconomic tradeoff if the public believes the central bank's commitment to low inflation.

3. **Consequence for Bad Inflation Outcomes**: In tandem with the move towards increased independence, many central banks now have legally imposed inflation targets and various types of bad things happen to the chief central banker when the inflation target is not met. For instance, the Governor of the Bank of England has to write a letter to the Chancellor explaining why the target was not met.

Appendix: Programme For IS-MP-PC Model Simulations

Figures 2.7 and 2.8 were produced using the programme below. The programme is written for the econometric package RATS but a programme of this sort could be written for lots of different types of software including Excel.

```
allocate 50
set pi = 2.0
set y  = 100
set ystar = 100
set pistarcb = 2.0
set pie = 2.0
comp alpha = 1
comp gamma = 1

*** BETA_PI = 1.5
comp betapi = 1.5
comp kappa  = alpha*gamma*(betapi - 1)
comp theta  = (1 / (1+kappa) )

set pie 11 11 = 4
comp pi(11) = theta*pie(11) + (1-theta)*pistarcb(11)
comp y(11) = ystar(11) - theta*alpha*(betapi - 1)*(pie(11) - pistarcb(11))


do j= 12, 50
comp pie(j) = pi(j-1)
comp pi(j) = theta*pie(j) + (1-theta)*pistarcb(j)
comp y(j) = ystar(j) - theta*alpha*(betapi - 1)*(pie(j) - pistarcb(j))
end do j

print 1 50 pi y pie

set y1 = y
set pi1 = pi


*** BETA_PI = 3
comp betapi = 3
comp kappa  = alpha*gamma*(betapi - 1)
comp theta  = (1 / (1+kappa) )

set pie 11 11 = 4
comp pi(11) = theta*pie(11) + (1-theta)*pistarcb(11)
comp y(11) = ystar(11) - theta*alpha*(betapi - 1)*(pie(11) - pistarcb(11))

do j= 12, 50
comp pie(j) = pi(j-1)
comp pi(j) = theta*pie(j) + (1-theta)*pistarcb(j)
comp y(j) = ystar(j) - theta*alpha*(betapi - 1)*(pie(j) - pistarcb(j))
```

```
end do j


set y2 = y
set pi2 = pi


labels y1 y2
# 'Beta = 1.5' 'Beta = 3.0'

labels pi1 pi2
# 'Beta = 1.5' 'Beta = 3.0'

graph(key=below) 2
# pi1 5 35
# pi2 5 35

graph(key=below)  2
# y1 5 35
# y2 5 35
```

Chapter 3

The Taylor Principle

Up to now, we have maintained the assumption that the central bank reacts to a change in inflation by implementing a bigger change in interest rates. In terms of the equation for our monetary policy rule, this means we are assuming $\beta_\pi > 1$. With this assumption, real interest rates go up when inflation rises and go down when inflation falls. For this reason, our IS-MP curve slopes downwards: Along this curve, higher inflation means lower output. Because John Taylor's original proposed rule had the feature that $\beta_\pi > 1$, the idea that monetary policy rules should have this feature has become known as *the Taylor Principle*. In this chapter, we discuss why policy rules should satisfy the Taylor principle.

Three Different Cases

Recall from our previous chapter that inflation in the IS-MP-PC model is given by

$$\pi_t = \theta \pi_t^e + (1-\theta)\,\pi^* + \theta\left(\gamma \epsilon_t^y + \epsilon_t^\pi\right) \tag{3.1}$$

where

$$\theta = \left(\frac{1}{1+\alpha\gamma\left(\beta_\pi - 1\right)}\right) \tag{3.2}$$

Under adaptive expectations $\pi_t^e = \pi_{t-1}$ and the model can be re-written as

$$\pi_t = \theta \pi_{t-1} + (1-\theta)\,\pi^* + \theta\left(\gamma \epsilon_t^y + \epsilon_t^\pi\right) \tag{3.3}$$

The value of θ turns out to be crucial to the behaviour of inflation and output in this model. We can describe three different cases depending on the value of β_π.

Case 1: $\beta_\pi > 1$

If the Taylor principle is satisfied, then $\alpha\gamma\left(\beta_\pi - 1\right) > 0$. That value being positive means that $1 + \alpha\gamma\left(\beta_\pi - 1\right) > 1$. The parameter θ is calculated by dividing 1 by this amount so this gives us a value of θ that is positive but less than one. So $\beta_\pi > 1$ translates into the case $0 < \theta < 1$.

Case 2: $\left(1 - \frac{1}{\alpha\gamma}\right) < \beta_\pi < 1$

As we reduce β_π below one, $(\beta_\pi - 1)$ becomes negative, meaning $\alpha\gamma\left(\beta_\pi - 1\right) < 0$ and $1 + \alpha\gamma\left(\beta_\pi - 1\right) < 1$. The parameter θ is calculated by dividing 1 by this amount so this gives us a value of θ that is greater than one. As β_π falls farther below one, θ gets bigger and bigger and heads towards infinity as β_π approaches $\left(1 - \frac{1}{\alpha\gamma}\right)$ (this is the value of β_π that makes the denominator in the θ formula equal zero). As long we assume that β_π stays above this level, we will get a value of θ that is positive and greater than one.

Case 3: $0 < \beta_\pi < \left(1 - \frac{1}{\alpha\gamma}\right)$

This produces a "pathological" case in which $1 + \alpha\gamma\left(\beta_\pi - 1\right) < 0$ so the value of θ becomes negative, meaning an increase in inflation expectations actually reduces inflation. We are not going to consider this case.

Macro Dynamics and Difference Equations

These calculations tell us that as long as the Taylor principle is satisfied, the value of θ lies between zero and one but that if β_π slips below one, then θ becomes greater than one. It turns out this is a very important distinction. To understand the difference between these two cases, we need to explain a little bit about difference equations.

A difference equation is a formula that generates a sequence of numbers. In economics, these sequences can be understood as a pattern over time for a variable of interest. After supplying some starting values, the difference equation provides a sequence explaining how the variable changes over time. For example, consider a case in which the first value for a series is $z_1 = 1$ and then z_t follows a difference equation

$$z_t = z_{t-1} + 2 \tag{3.4}$$

This will give $z_2 = 3$, $z_3 = 5$, $z_4 = 7$ and so on. More relevant to our case is the multiplicative model

$$z_t = bz_{t-1} \tag{3.5}$$

For a starting value of $z_1 = x$, this difference equation delivers a sequence of values that looks like this: x, xb, xb^2, xb^3, xb^4.....

Note that if b is between zero and one, then this sequence converges to zero over time no matter what value x takes but if $b > 1$, the sequence will explode off towards either plus or minus infinity depending on whether the initial value was positive or negative. The same logic prevails if we add a constant term to the difference equation. Consider this equation:

$$z_t = a + bz_{t-1} \tag{3.6}$$

If b is between zero and one, then no matter what the starting value is, the sequence converges over time to $\frac{a}{1-b}$ but if $b > 1$, the sequence explodes towards infinity. Similarly, if we add random shocks to the model—making it what is known as a first-order autoregressive or AR(1) model—the key thing remains the value of b. If the model is

$$z_t = a + bz_{t-1} + \epsilon_t \tag{3.7}$$

where ϵ_t is a series of independently drawn zero-mean random shocks, then the presence of the shocks will mean the series won't simply converge to a constant or steadily explode. But as long as we have $0 < b < 1$ then the series will tend to oscillate above and below the average value of $\frac{a}{1-b}$ while if $b > 1$ the series will tend to explode to infinity over time.

The Taylor Principle and Macroeconomic Stability

These considerations explain why the Taylor principle is so important. If $\beta_\pi > 1$ then inflation dynamics in the IS-MP-PC model can be described by an AR(1) model with a coefficient on past inflation that is between zero and one (the θ in equation 3.3 plays the role of the coefficient b in the models just considered.) So a policy rule that satisfies the Taylor principle produces a stable time series for inflation under adaptive expectations. And because output depends on the gap between inflation expectations and the central bank's inflation target, stable inflation translates into stable output.

In contrast, once β_π falls below 1, the coefficient on past values of inflation in equation (3.3) becomes greater than one and the coefficient on the inflation target becomes negative. In this case, any change in inflation produces a greater change in the same direction next period and inflation ends up exploding off to either plus or minus infinity. Similarly output either collapses or explodes.

Why does β_π matter so much for macroeconomic stability? Obeying the Taylor principle means that shocks that boost inflation (whether they be supply or demand shocks) raise real interest rates (because nominal rates go up by more than inflation does) and thus reduce output, which contains the increase in inflation and keeps the economy stable. In contrast, when the β_π falls below 1, shocks that raise inflation result in lower real interest rates and higher output which further fuels the initial increase in inflation (similarly declines in inflation are further magnified). This produces an unstable explosive spiral.

You might be tempted to think that the arguments in favour of obeying the Taylor principle as explained here depends crucially on the assumption of adaptive expectations but this isn't the case. Even before assuming adaptive expectations, from equation (3.1) we can see that when $\theta > 1$, the coefficient on the central bank's inflation target is negative. So if you introduced a more sophisticated model of expectations formation, the public would realise that the central bank's inflation target doesn't have its intended influence on inflation and so there would no reason to expect this value of inflation to come about. But if people know that changes in expected inflation are translated more than one-for-one into changes in actual inflation then this could produce self-fulfilling inflationary spirals, even if the public had a more sophisticated method of forming expectations than the adaptive one employed here.

Graphical Representation

We can use graphs to illustrate the properties of the IS-MP-PC model when the Taylor principle is not obeyed. Recall that the IS-MP curve is given by this equation

$$y_t = y_t^* - \alpha\left(\beta_\pi - 1\right)\left(\pi_t - \pi^*\right) + \epsilon_t^y \tag{3.8}$$

The slope of the curve depends on whether or not $\beta_\pi > 1$. In the previous chapter, we assumed $\beta_\pi > 1$ and so the slope $-\alpha\left(\beta_\pi - 1\right) < 0$, meaning the IS-MP curve slopes down. With $\beta_\pi < 1$, the IS-MP curve slopes up. Figure 3.1 illustrates the IS-MP-PC model in this case under the assumption that $\pi_t^e = \pi^* = \pi_1$, i.e. that the public expects inflation to equal the central bank's target.

One technical point about this graph is worth noting. I have drawn the upward-sloping IS-MP curve as a steeper line than the upward-sloping Phillips curve. On the graph as we've drawn it in inflation-output space, the slope of this curve is $\frac{1}{\alpha(1-\beta_\pi)}$ while the slope of the Phillips curve is γ. One can show that the condition that $\frac{1}{\alpha(1-\beta_\pi)} > \gamma$ is the same as showing that $\theta > 0$, i.e. that we are ruling out values of β_π associated with the strange third case noted above.

Now consider what happens when there is an increase in inflation expectations when β_π falls below one. Figure 3.2 shows a shift in the Phillips curve due to inflation expectations increasing from π_1 to π_h (You can see that the value of inflation on the red Phillips curve when $y_t = y_t^*$ is $\pi_t = \pi_h$). Notice now that, because the IS-MP curve is steeper than the Phillips curve, inflation increases above π_h to take the higher value of π_2. Inflation overshoots the public's expected value.

Figure 3.3 shows what happens next if the public have adaptive expectations. In this next period, we have $\pi_t^e = \pi_2$ and inflation jumps all the way up to the even higher value of π_3. We won't show any more graphs but the process would continue with inflation increasing every period. These figures thus show graphically what we've already demonstrated with equations. The IS-MP-PC model generates explosive dynamics when the monetary policy rule fails to obey the Taylor principle.

Figure 3.1: The IS-MP-PC Model when $\left(1 - \frac{1}{\alpha\gamma}\right) < \beta_\pi < 1$

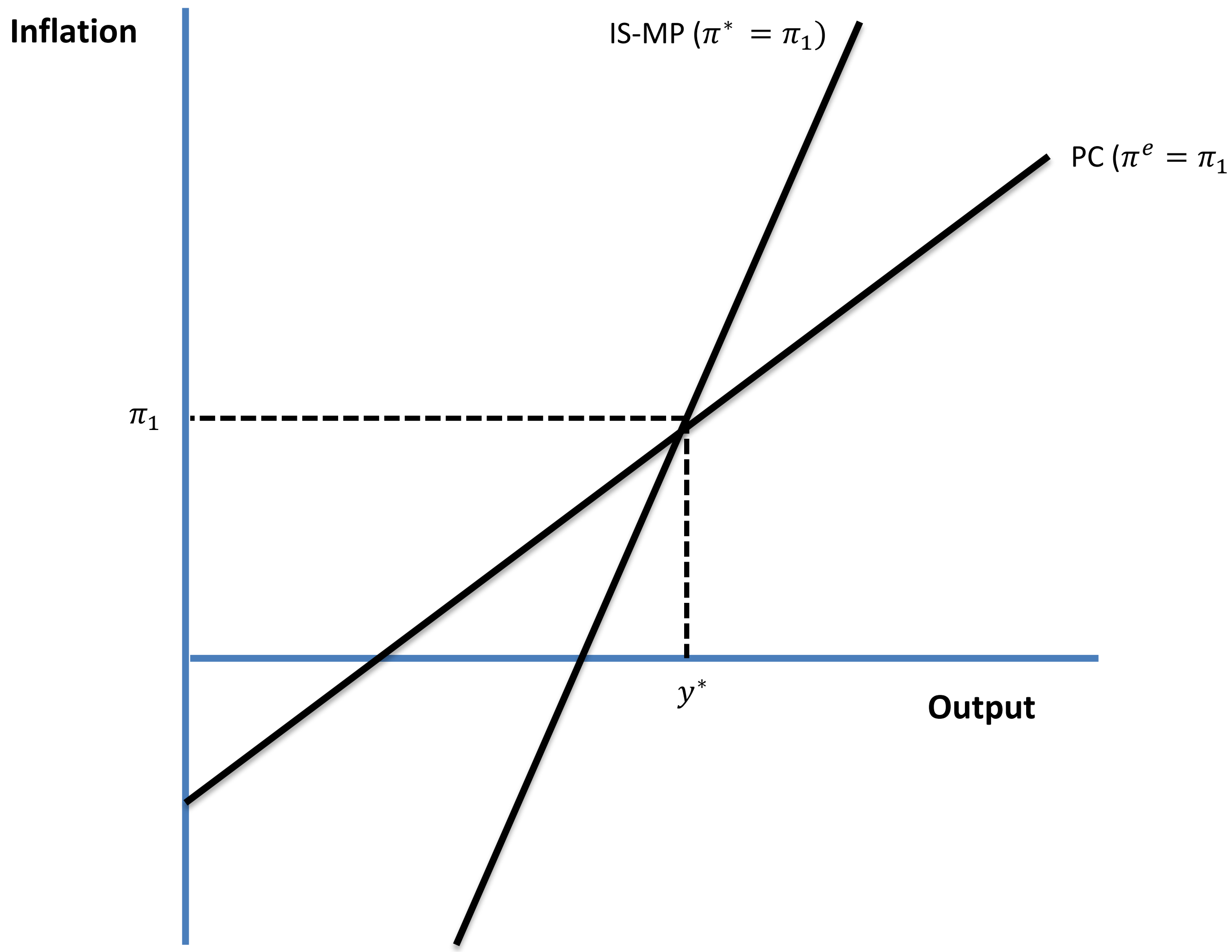


Figure 3.2: An Increase in π^e when $\left(1-\frac{1}{\alpha\gamma}\right) < \beta_\pi < 1$

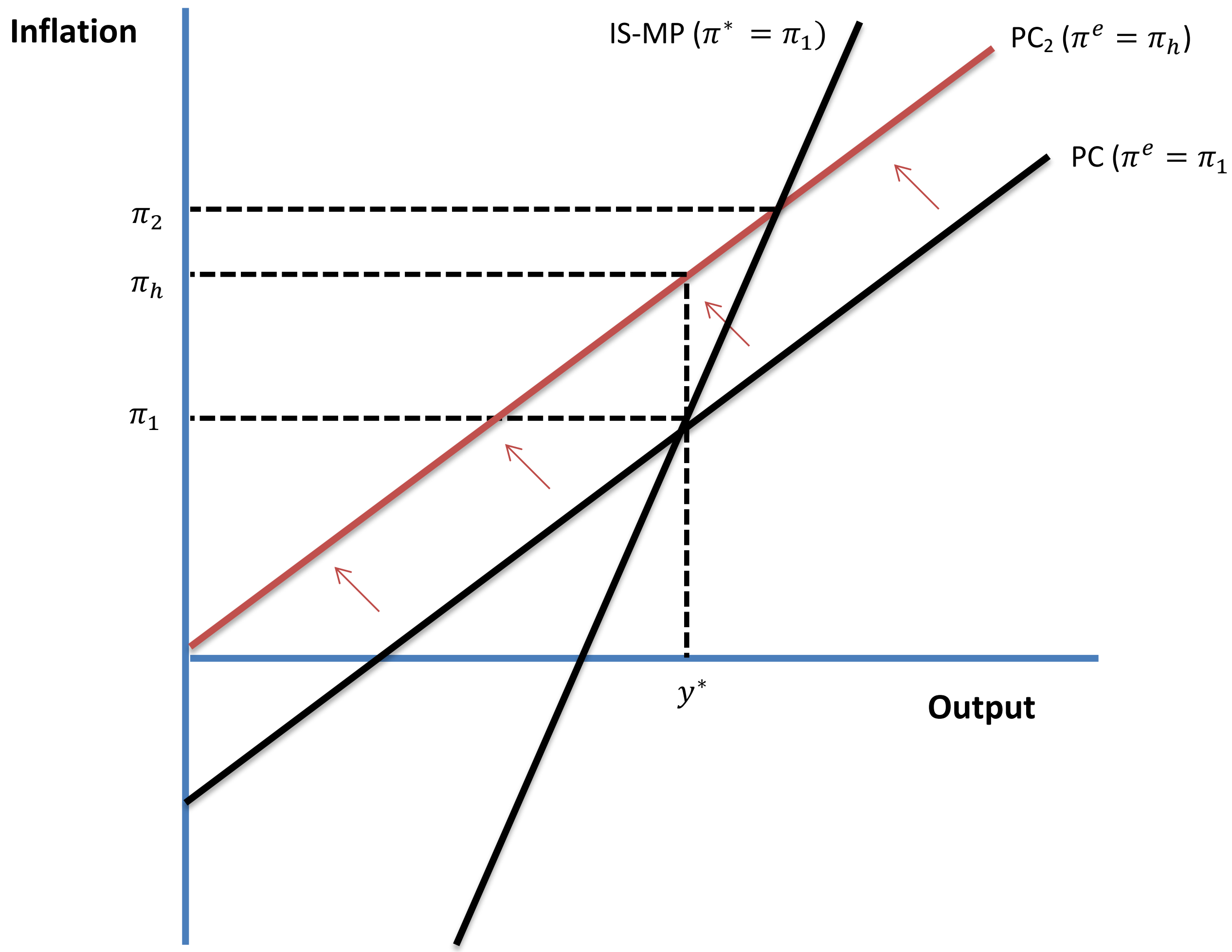
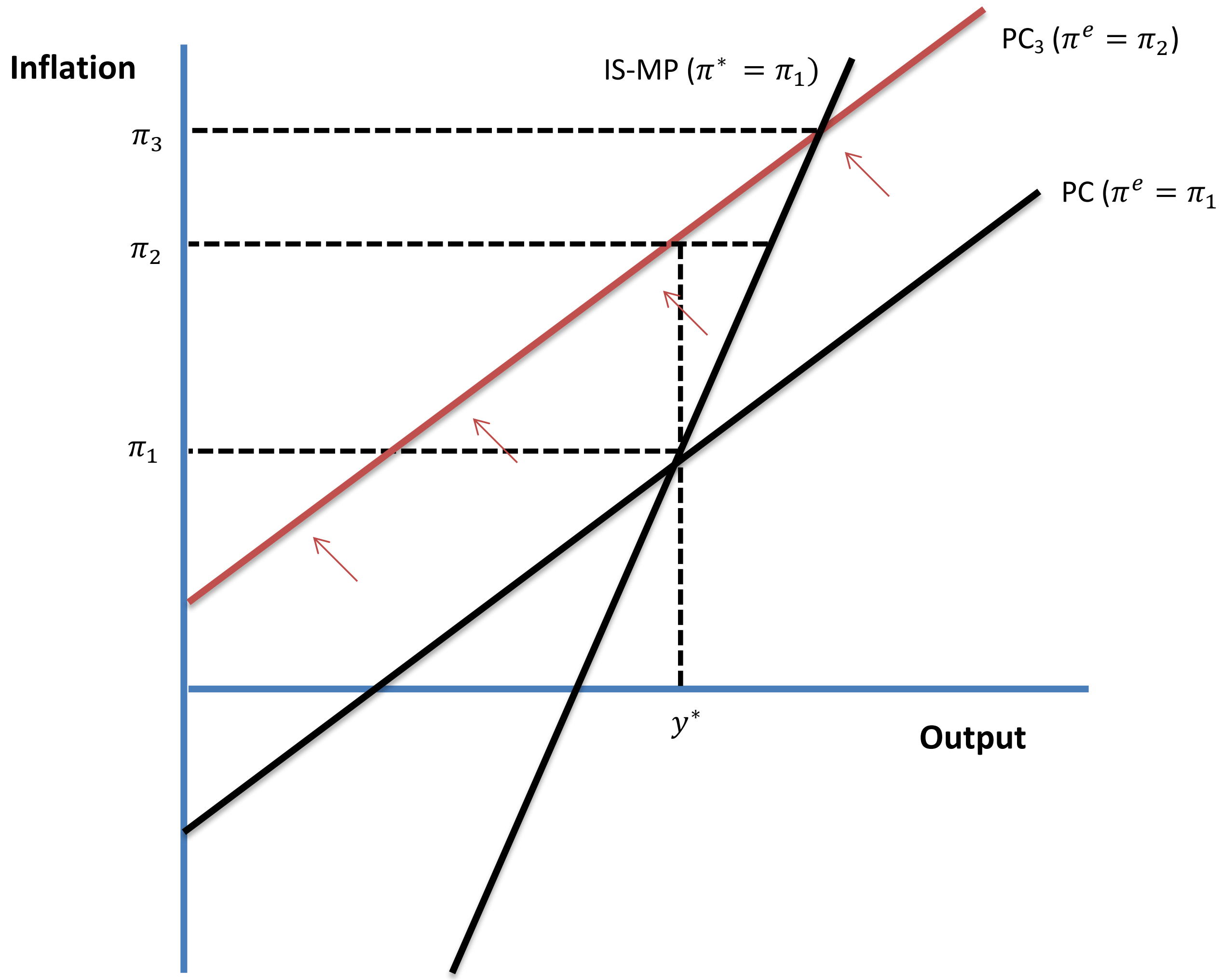


Figure 3.3: Explosive Dynamics when $\left(1 - \frac{1}{\alpha\gamma}\right) < \beta_\pi < 1$



An Increase in the Inflation Target

Figure 3.4 illustrates what happens in the IS-MP-PC model when the central bank changes its inflation target. The increase in the inflation target shifts the IS-MP curve upwards i.e. each level of output is associated with a higher level of inflation. However, because the IS-MP curve is steeper than the Phillips curve, the outcome is a reduction in inflation. Output also falls.

Even though this is exactly what our earlier equations predicted (the coefficient on the inflation target is $1 - \theta$ which is negative in this case) this seems like a very strange outcome. The central bank sets a higher inflation target and then inflation falls. Why is this?

The answer turns out to reflect the particular form of the monetary policy rule that we are using. This rule is as follows:

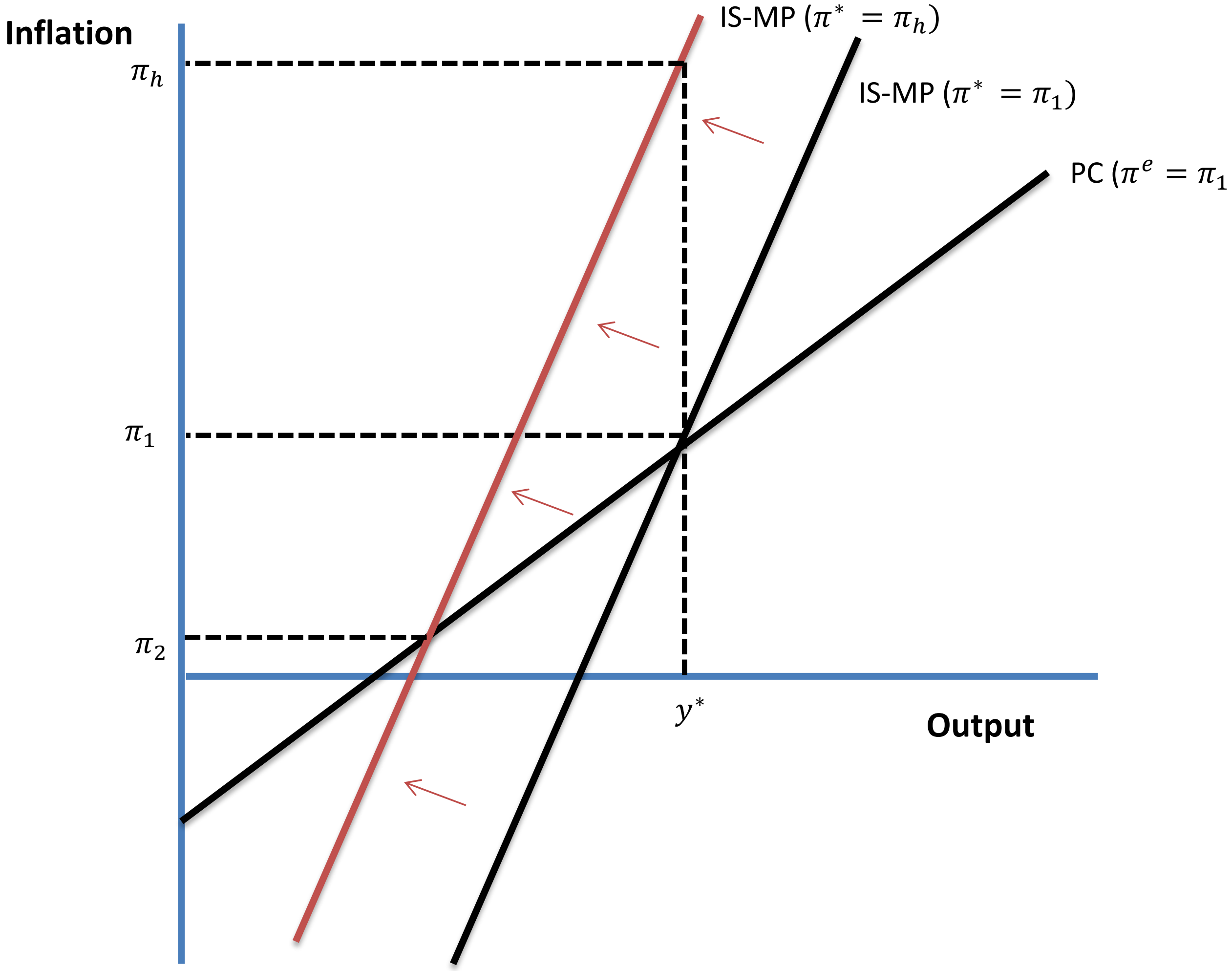
$$i_t = r^* + \pi^* + \beta_\pi \left(\pi_t - \pi^*\right) \tag{3.9}$$

You might expect that a higher inflation target would lead to the central bank setting a lower interest rate, i.e. they ease up to allow the economy to expand and let inflation move higher. However, if you look closely at this formula, you can see that an increase in the inflation target actually leads to a *higher* interest rates when $\beta_\pi < 1$.

This can be explained as follows. The inflation target appears twice in equation (3.9). It appears in brackets as part of the "'inflation gap" term $\pi_t - \pi^*$ which is muliplied by β_π. If this was the only place that it appeared, then indeed a higher inflation target would lead to lower interest rates. However, the first part of rule relates to setting the interest rate so that when inflation equalled its target, real interest rates would equal their "natural rate" r^*. The rule is set on the basis that if inflation is going to be higher on average, then the nominal interest rate also needs to be higher if real interest rates are to remain unchanged (this is commonly called the "Fisher effect" of inflation on interest rates).

Putting these two effects together, we see that an increases of x in the inflation target raises the nominal interest rate by x due to the real interest rate component and reduces it by $\beta_\pi x$ due to inflation now falling below target. If $\beta_\pi < 1$ then the higher inflation target results in higher interest rates and thus lower output. This is the pattern shown in Figure 3.4.

Figure 3.4: An Increase in π^* when $\left(1 - \frac{1}{\alpha\gamma}\right) < \beta_\pi < 1$



Evidence on Monetary Policy Rules and Macroeconomic Stability

Is there any evidence that obeying the Taylor principle provides greater macroeconomic stability? Some economists believe there is.

An important paper on this topic was "Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory" by Richard Clarida, Jordi Gali and Mark Gertler. These economists reported that estimated policy rules for the Federal Reserve appeared to show a change after Paul Volcker was appointed Chairman in 1979. They estimated that the post-1979 monetary policy appeared consistent with a rule in which the coefficient on inflation that was greater than one while the pre-1979 policy seemed consistent with a rule in which this coefficient was less than one. They also introduce a small model in which the public adopts rational expectations (more on what this means later) and show that failure to obey the Taylor principle can lead to the economy generating cycles based on self-fulfilling fluctuations. They argue that failure to obey the Taylor principle could have been responsible for the poor macroeconomic performance, featuring the stagflation combination of high inflation and recession, during the 1970s in the US.

There are a number of differences between the approach taken in Clarida, Gali, Gertler paper and our approach (in particular, their estimated policy rule is a "forward-looking" one in which policy reacts to expected future values of inflation and output) and the econometrics are perhaps more advanced than you have seen but it's still a pretty readable paper and a nice example of policy-relevant macroeconomic research.

That said, this being economics, there have been some dissenting voices on Clarida, Gali and Gertler's conclusions. In particular, there is the research of Athanasios Orphanides.[1]

[1] Athanasios Orphanides (2004) "Monetary Policy Rules, Macroeconomic Stability, and Inflation: A View from the Trenches" *Journal of Money, Credit and Banking*, vol. 36(2).

Orphanides is critical of Taylor rule regressions that use measures of the output gap that are based on detrending data from the full sample. This includes information that wasn't available to policy-makers when they were formulating policy in real time and so perhaps it is unfair to describe them as reacting to these estimates.

This point is particularly relevant for assessing monetary policy prior to 1979. During the 1970s, growth rates for major international economies slowed considerably. Policy-makers thought their economies were falling far short of its potential level. In retrospect it is clear that potential output growth rates were falling and true output gaps were small. Replacing the full-sample output gap estimates with real-time estimates that were available to the Fed at the time, Orphanides reports regressions which suggest that the 1970s Fed obeyed the Taylor rule with respect to reacting to inflation and that their mistake was over-reacting to inaccurate estimates of the output gap.

Chapter 4

The Zero Lower Bound and the Liquidity Trap

Up to now, we have assumed that the central bank in our model economy sets its interest rate according to a specific policy rule. Whatever the rule says the interest rate should be, the central banks sets that interest rate. But what if the rule predicts the central bank should set interest rates equal to a negative value? Will they?

In the past, the economics profession had a simple answer to this question. There should be a lower bound on interest rates of zero. If I loan you \$100 and only get \$101 back next period, I haven't earned much interest but at least I earned some. A negative interest rate would mean me loaning you \$100 and getting back less than that next year. Why would I do that? Since money maintains its nominal face value, I'd be better off just the keep the money in my bank account or under a mattress.

In practice, however, we have seen in recent years that negative interest rates can occur. For example, in the Euro Area, the ECB has charged banks for depositing money with it. This has essentially set the relevant marginal interest rate for these banks to a negative value and they are willing to make loans to other banks or purchase securities that have a negative interest rate, provided it is less negative than the deposit rate paid by the ECB.

There are, however, limits to how negative rates could get. At some point, banks would be better off to withdraw all of their money from their central bank deposit account and hold it in warehouses. This means there is effectively a lower bound on the interest rates set by monetary policy, though exactly what that lower limit might be is a bit unclear.

With these considerations in mind, we are going to adapt our model to take into account that there are times when the central bank would like to set i_t below zero but is not able to do so. We stick with zero rather than specifying a particular negative value for the lower bound: We could specify a specific non-zero value for the lower bound but this would just introduce an extra parameter into the model without gaining us much additional insight.

The Zero Lower Bound

When will the "zero lower bound" become a problem for a central bank? In our IS-MP-PC model, it depends on the form of the monetary policy rule. Up to now, we have been considering a monetary policy rule of the form

$$i_t = r^* + \pi^* + \beta_\pi (\pi_t - \pi^*) \tag{4.1}$$

This rule sees the nominal interest rate adjusted upwards and downwards as inflation changes. So the lower bound problem occurs when inflation goes below some critical value. This value might be negative, so it may occur when there is deflation, meaning prices are falling. Amending our model to remove the possibility that interest rates could become negative, our new monetary policy rule is

$$i_t = \text{Maximum} \left[r^* + \pi^* + \beta_\pi (\pi_t - \pi^*), 0 \right] \tag{4.2}$$

Because the intended interest rate of the central bank declines with inflation, this means that there is a particular inflation rate, π^{ZLB}, such that if $\pi_t < \pi^{ZLB}$ then the interest rate will equal zero. So what determines this specific value, π^{ZLB} that triggers the zero lower bound?

Algebraically, we can characterise π^{ZLB} as satisfying

$$r^* + \pi^* + \beta_\pi \left(\pi^{ZLB} - \pi^*\right) = 0 \tag{4.3}$$

This can be re-arranged as

$$\beta_\pi \pi^{ZLB} = \beta_\pi \pi^* - r^* - \pi^* \tag{4.4}$$

which can be solved to give

$$\pi^{ZLB} = \left(\frac{\beta_\pi - 1}{\beta_\pi}\right)\pi^* - \frac{r^*}{\beta_\pi} \tag{4.5}$$

Equation (4.5) tells us that three factors determine the value of inflation at which the central bank sets interest rates equal to zero.

1. **The inflation target**: The higher the inflation target π^* is, then the higher is the level of inflation at which a central bank will be willing to set interest rates equal to zero.

2. **The natural rate of interest**: A higher value of r^*, the "natural" real interest rate, lowers the level of inflation at which a central bank will be willing to set interest rates equal to zero. An increase in this rate makes central banks raise interest rates and so they will wait until inflation goes lower than previously to set interest rates to zero.

3. **The responsiveness of monetary policy to inflation**: Increases in β_π raise the coefficient on π^* in this formula, increasing the first term and it makes the second term (which has a negative sign) smaller. Both effects mean a higher β_π translates into a higher value for π^{ZLB}. Central banks that react more aggressively against inflation will wait for inflation to reach lower values before they are willing to set interest rates to zero.

The IS-MP Curve and the Zero Lower Bound

Given this characterisation of when the zero lower bound kicks in, we need to re-formulate the IS-MP curve. Once inflation falls below π^{ZLB}, the central bank cannot keep cutting interest rates in line with its monetary policy rule. Recalling that the IS curve is

$$y_t = y_t^* - \alpha \left(i_t - \pi_t - r^* \right) + \epsilon_t^y \tag{4.6}$$

We had previously derived the IS-MP curve by substituting in the monetary policy rule formula (4.1) for the i_t term. This gave us the IS-MP curve as:

$$y_t = y_t^* - \alpha \left(\beta_\pi - 1 \right) \left(\pi_t - \pi^* \right) + \epsilon_t^y \tag{4.7}$$

However, when $\pi_t \leq \pi^{ZLB}$ we need to substitute in zero instead of the negative value that the monetary policy rule would predict. So the IS-MP curve becomes

$$y_t = \begin{cases} y_t^* - \alpha \left(\beta_\pi - 1 \right) \left(\pi_t - \pi^* \right) + \epsilon_t^y & \text{when} \;\; \pi_t > \pi^{ZLB} \\ y_t^* + \alpha r^* + \alpha \pi_t + \epsilon_t^y & \text{when} \;\; \pi_t \leq \pi^{ZLB} \end{cases} \tag{4.8}$$

The effect of inflation on output in this revised IS-MP curve changes when inflation moves below π^{ZLB}. Above π^{ZLB}, higher values of inflation are associated with lower values of output. Below π^{ZLB}, higher values of inflation are associated with *higher* values of output. Graphically, this means the IS-MP curve shifts from being downward-sloping to being upward-sloping when inflation falls below π^{ZLB}. Figure 4.1 provides an example of how this looks.

Equation (4.8) also explains the conditions under which the zero lower bound is likely to be relevant. If there are no aggregate demand shocks, so $\epsilon_t^y = 0$, then the zero lower bound is likely to kick in at a point where output is above its natural rate; this is the case illustrated in Figure 4.1. But this combination of high output and low inflation is unlikely to be an equilibrium in the model unless the public expects very low inflation or deflation so the

Phillips curve intersects the IS-MP curve along the section that has output above its natural rate and inflation below π^{ZLB}.

However, if we have a large negative aggregate demand shock, so that $\epsilon_t^y < 0$, then it is possible to have output below its natural rate and inflation falling below π^{ZLB}. As illustrated in Figure 4.2, this situation is more likely to be an equilibrium (i.e. this position for the IS-MP curve is more likely to intersect with the Phillips curve) even if inflation expectations are close to the inflation target.

Figure 4.1: The IS-MP Curve with the Zero Lower Bound

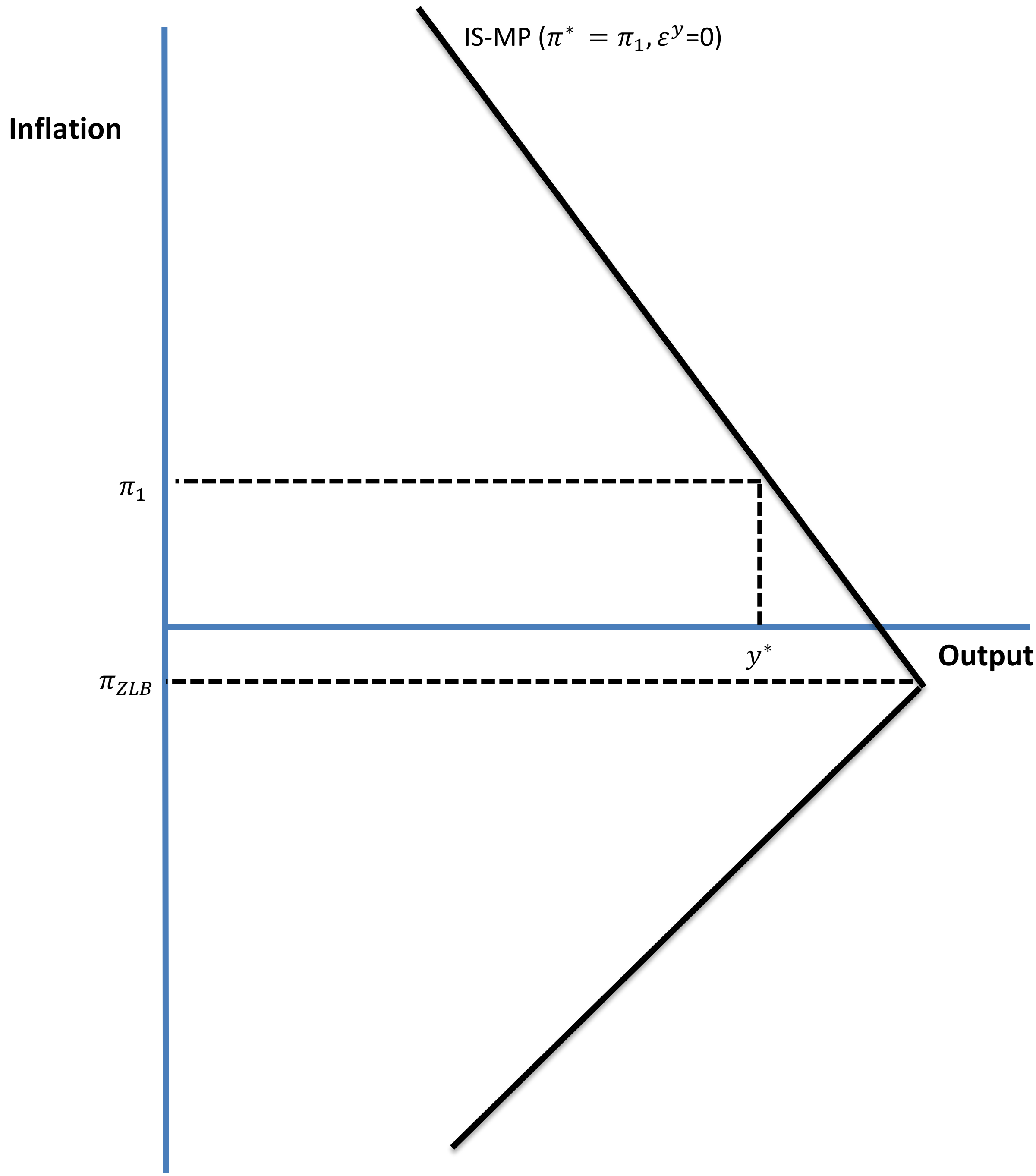
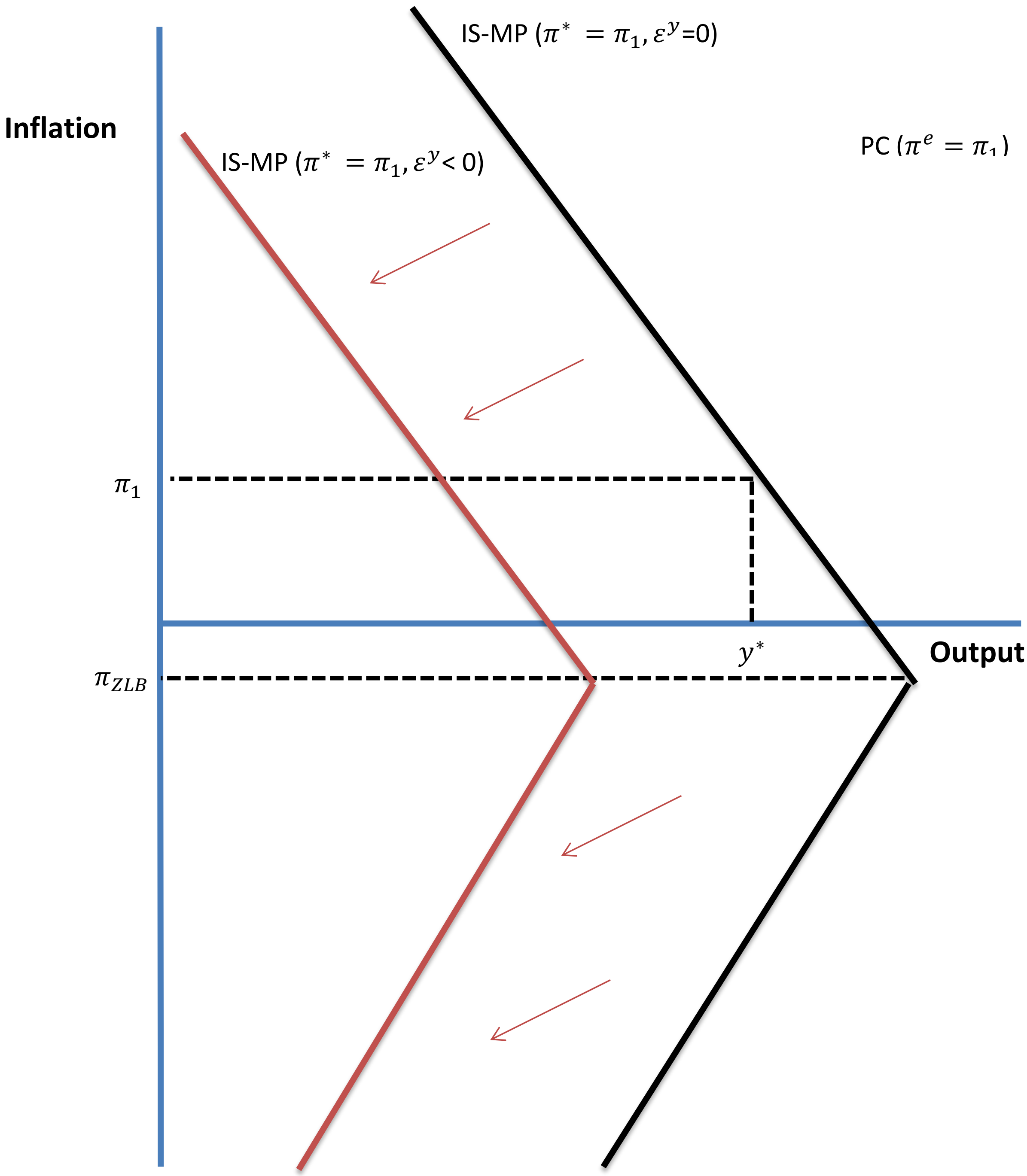


Figure 4.2: A Negative Aggregate Demand Shock



The Liquidity Trap

When inflation falls below the lower bound, output is determined by

$$y_t = y_t^* + \alpha r^* + \alpha \pi_t + \epsilon_t^y \tag{4.9}$$

Inflation is still determined by the Phillips curve

$$\pi_t = \pi_t^e + \gamma \left(y_t - y_t^* \right) + \epsilon_t^\pi \tag{4.10}$$

Using the expression for the output gap when the zero lower bound limit has been reached from equation (4.9) we get an expression for inflation under these conditions as follows

$$\pi_t = \pi_t^e + \gamma \left(\alpha r^* + \alpha \pi_t + \epsilon_t^y \right) + \epsilon_t^\pi \tag{4.11}$$

This can be re-arranged to give

$$\pi_t = \frac{1}{1-\alpha\gamma}\pi_t^e + \frac{\alpha\gamma}{1-\alpha\gamma} r^* + \frac{\gamma}{1-\alpha\gamma}\epsilon_t^y + \frac{1}{1-\alpha\gamma}\epsilon_t^\pi \tag{4.12}$$

The coefficient on expected inflation, $\frac{1}{1-\alpha\gamma}$ is greater than one. So, just as with the Taylor principle example from the previous chapter, changes in expected inflation translate into even bigger changes in actual inflation. As we discussed the last time, this leads to unstable dynamics. Because these dynamics take place only when inflation has fallen below the zero lower bound, the instability here relates to falling inflation expectations, leading to further declines in inflation and further declines in inflation expectations. Because output depends positively on inflation when the zero-bound constraint binds, these dynamics mean falling inflation (or increasing deflation) and falling output.

This position in which nominal interest rates are zero and the economy falls into a deflationary spiral is known as *the liquidity trap*. Figures 4.3 and 4.4 illustrate how the liquidity trap operates in our model. Figure 4.3 shows how a large negative aggregate demand shock can lead to interest rates hitting the zero bound even when expected inflation is positive.

Figure 4.4 illustrates how expected inflation has a completely different effect when the zero lower bound has been hit. It shows a fall in expected inflation after the negative demand shock (this example isn't adaptive expectations because I haven't drawn inflation expectations falling all the way to the deflationary outcome graphed in Figure 4.3). In our usual model set-up, a fall in expected inflation raises output. However, once at the zero bound, a fall in expected inflation reduces output, which further reduces inflation.

Figure 4.3: Equilibrium At the Lower Bound

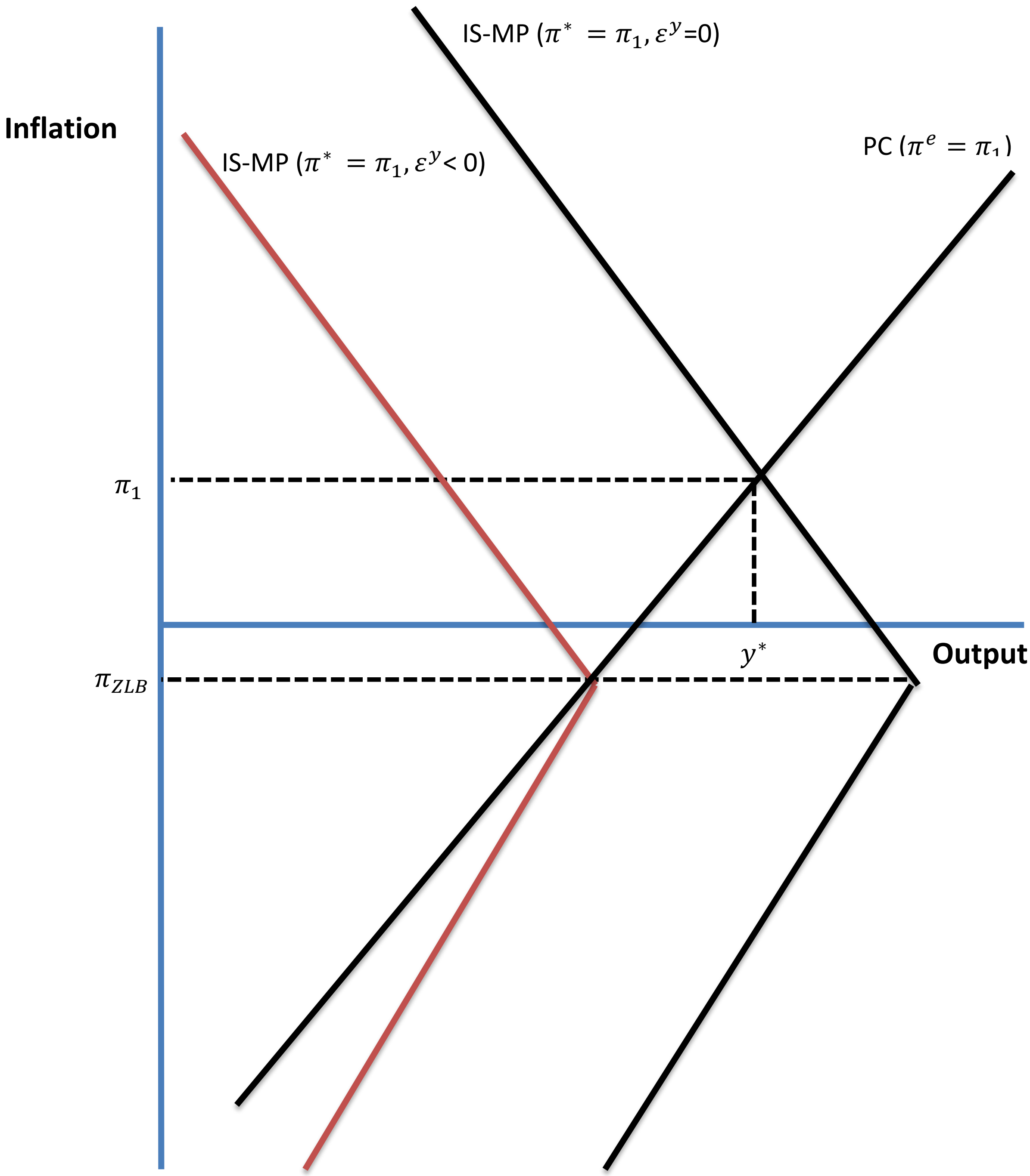
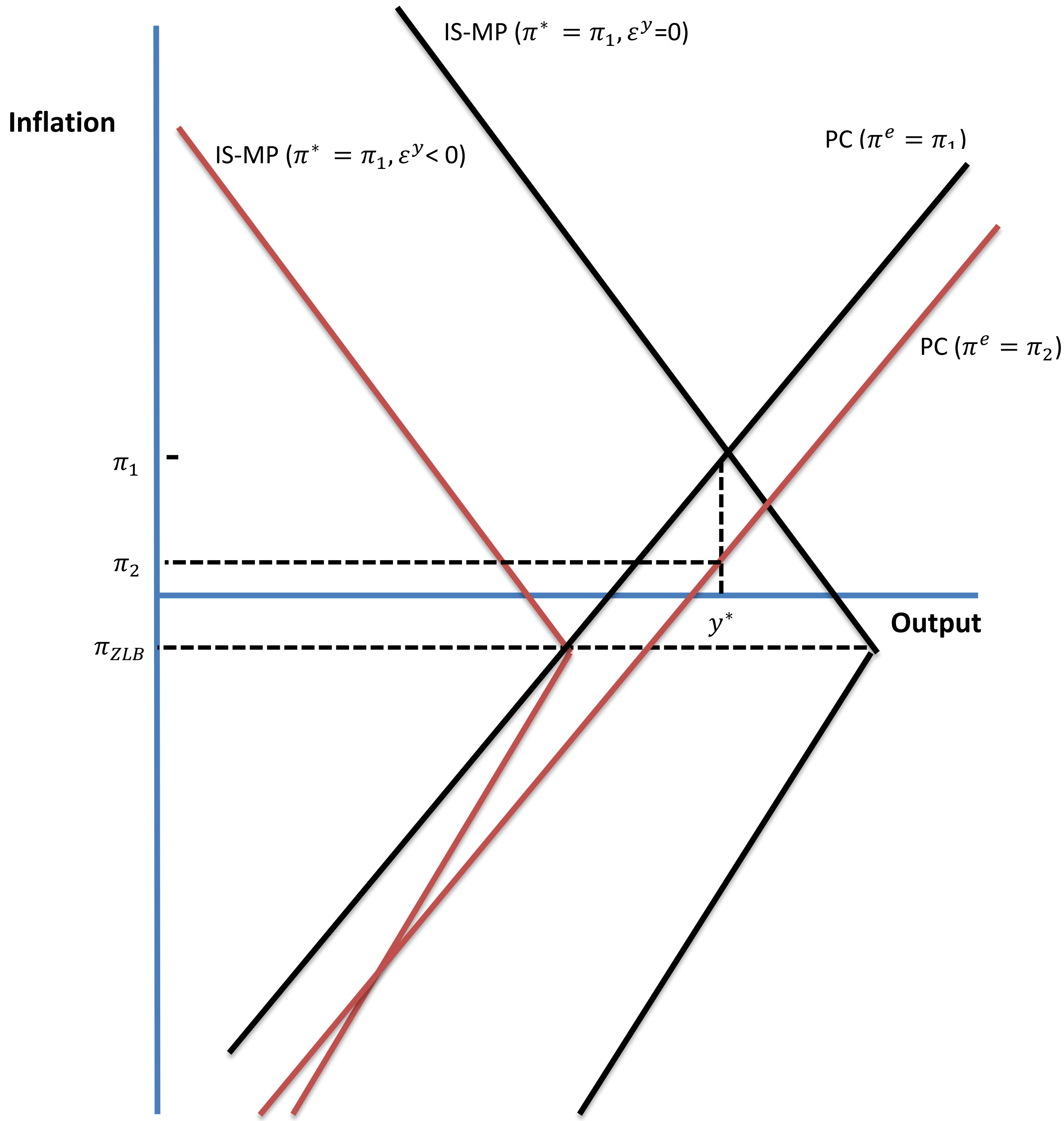


Figure 4.4: Falling Expected Inflation Worsens Slump



The Liquidity Trap with a Taylor Rule

For the simple monetary policy rule that we have been using, the zero lower bound is reached for a particular trigger level of inflation. Plugging in reasonable parameter values into equation (4.5) this trigger value will most likely be negative. In other words, with the monetary policy rule that we have been using, the zero lower bound will only be hit when there is deflation. However, if we have a different monetary policy rule, this result can be overturned. For example, remember the Taylor-type rule we considered earlier

$$i_t = r^* + \pi^* + \beta_\pi \left(\pi_t - \pi^*\right) + \beta_y \left(y_t - y_t^*\right) \tag{4.13}$$

Incorporating the zero lower bound, this would be adapted to be

$$i_t = \text{Maximum}\left[r^* + \pi^* + \beta_\pi \left(\pi_t - \pi^*\right) + \beta_y \left(y_t - y_t^*\right), 0\right] \tag{4.14}$$

For this rule, the zero lower bound is hit when

$$r^* + \pi^* + \beta_\pi \left(\pi_t - \pi^*\right) + \beta_y \left(y_t - y_t^*\right) = 0 \tag{4.15}$$

This condition can be re-written as

$$\beta_\pi \left(\pi_t - \pi^*\right) + \beta_y \left(y_t - y_t^*\right) = -r^* - \pi^* \tag{4.16}$$

In other words, there are a series of different combinations of inflation gaps and output gaps that can lead to monetary policy hitting the zero lower bound. For example, if $y_t = y_t^*$ the lower bound will be hit at the value of inflation given by equation (4.5), i.e. the level we have defined as π^{ZLB}. In contrast, inflation could equal its target level but policy would hit the zero bound if output fell as low as $y_t^* - \frac{r^*+\pi^*}{\beta_y}$.

Graphically, we can represent all the combinations of output and inflation that produce zero interest rates under the Taylor rule as the area under a downward-sloping line in Inflation-Output space. Figure 4.5 gives an illustration of what this area would look like. We showed in the first chapter that when we are above the zero bound, the IS-MP curve under the Taylor rule is of the same downward-sloping form as under our simple inflation targeting rule. At the zero bound, the arguments we've already presented here also apply so that the IS-MP curve becomes upward sloping.

Figure 4.6 illustrates two different cases of IS-MP curves when monetary policy follows a Taylor rule. The right-hand curve corresponds to the case $\epsilon_t^y = 0$ (no aggregate demand shocks) and this curve only interests with the zero bound area when there is a substantial deflation. In contrast, the left-hand curve corresponds to the case in which ϵ_t^y is highly negative (a large negative aggregate demand shocks) and this curve interests with the zero bound area even at levels of inflation that are positive and aren't much below the central bank's target.

Figure 4.5: Zero Bound is Binding in Blue Triangle Area

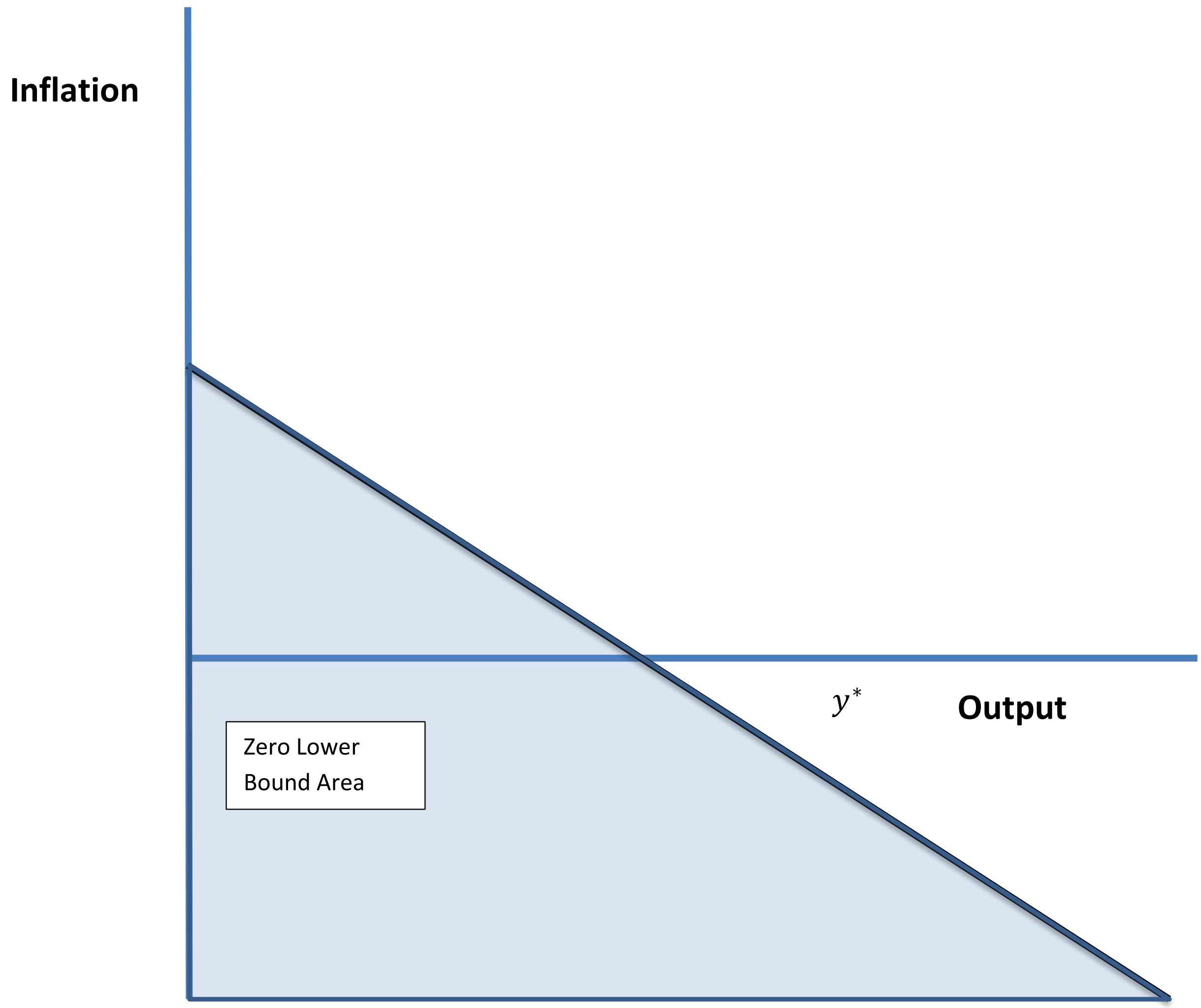
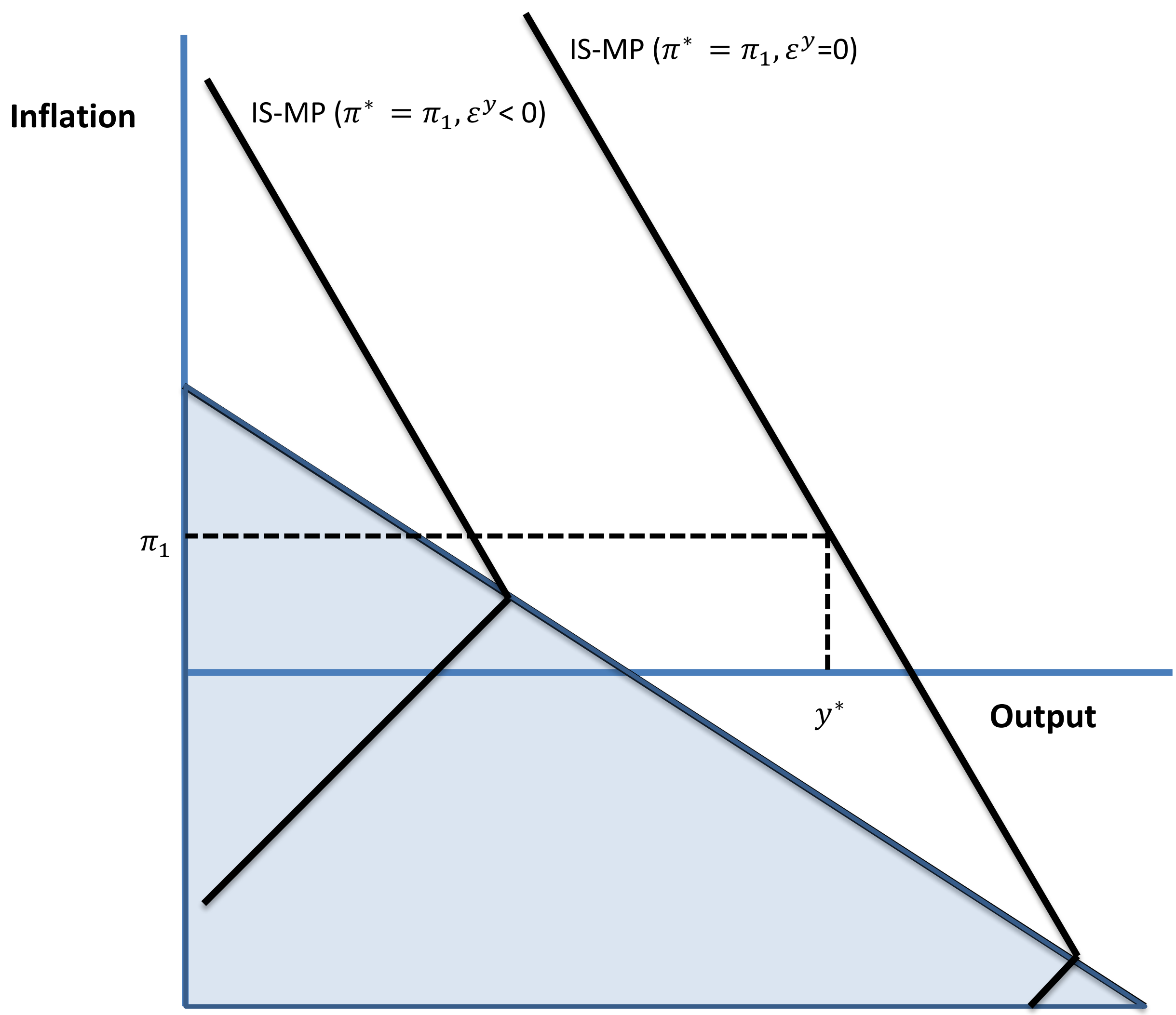


Figure 4.6: Zero Bound Can Be Hit With Positive Inflation



The Liquidity Trap: Reversing Conventional Wisdom

An important aspect of this model of the liquidity trap is it shows that some of the predictions that our model made (and which are now part of the conventional wisdom among monetary policy makers) do not hold when the economy is in a liquidity trap.

Up to now we have seen that as long as the central bank maintains its inflation target, then the model with adaptive expectations predicts that deviations of the public's inflation expectations from this target will be temporary and the economy will tend to converge back towards its natural level of output. However, once interest rates have hit the zero bound, this is no longer the case. Instead, the adaptive expectations model predicts the economy can spiral into an ever-declining slump.

Similarly, our earlier model predicted that a strong belief from the public that the central bank would keep inflation at target was helpful in stabilising the economy. However, once you reach the zero bound, convincing the public to raise its inflation expectations (perhaps by announcing a higher target for inflation) is helpful.

How to Get Out of the Liquidity Trap?

The most obvious way that a liquidity trap can end is if there is a positive aggregate demand shock that shifts the IS-MP curve back upwards so that the intersection with the Phillips curve occurs at levels of output and inflation that gets the economy out of the liquidity trap.

However, in reality, liquidity traps have often occurred during periods when there are ongoing and persistent slumps in aggregate demand. For example, after decades of strong growth, the Japanese economy went into a slump during the 1990s. Housing prices crashed and businesses and households were hit with serious negative equity problems. This type of "balance sheet" recession doesn't necessarily reverse itself quickly. The result in Japan was a long period in which prices were regularly falling and the Bank of Japan setting short-term interest rates close to zero throughout this period.

Given that economies in liquidity traps tend not to self correct with positive aggregate demand shocks from the private sector, governments can try to boost the economy by using fiscal policy to stimulate aggregate demand. Japan has used fiscal stimulus on various occasions with limited success.

What about monetary policy? With its policy interest rates at zero, can a central bank do any more to boost the economy? Debates on this topic have focused on several options.

The first option relates to the fact that while the short-term interest rates that are controlled by central banks may be zero, that doesn't mean the longer-term rates that many people borrow at will equal zero. By signalling that they intend to keep short-term rates low for a long period of time and perhaps by directly intervening in the bond market (i.e. quantitative easing) central banks can attempt to lower these longer-term rates.

The second option relates to inflation expectations. Our model tells us that output can be boosted when the economy is in a liquidity trap by raising inflation expectations. This acts to raise inflation (or reduce deflation) and this reduces real interest rates and boost output. Nobel prize-winner, Paul Krugman, recommends that central banks "'commit to being irresponsible" as a way out of these slumps. In other words, they should commit to a temporary period of inflation being higher than you would normally like. But central bankers are a conservative crowd and even temporary "irresponsibility" does not come easy to them.

A third option relates to exchange rates. To raise inflation, a central bank could announce targets for its exchange rate that would see it fall in value relative to the its major trading partners. Such a programme could be implemented by the central bank announcing that it is willing to buy and sell unlimited amounts of foreign exchange at an announced exchange rate. For example, suppose the euro is trading on the market so one euro equals \$1.20 and then the ECB announces that it is willing to swap euros for \$1. Anyone who wants to buy euros can now buy them at this lower rate, so there is a depreciation of the euro exchange rate. This currency depreciation would make imports more expensive, which would raise inflation. This latter approach has been labelled the "foolproof way to escape from a liquidity trap" by leading monetary policy expert Lars Svensson.[1] However, while this policy might work for a small country, the world's largest economies have agreed in recent years to refrain from "competitive devaluation" i.e. using depreciation of their exchange rate as a stimulus.

A fourth option is the level of inflation targeted by central banks. In our model, the stable equilibrium point sees a nominal interest rate of $(r^* + \pi^*)$ percent. If, for example, $\pi^* = 2$ (consistent with the preference of modern central bankers for a 2 percent inflation rate) and $r^* = 3$, then the nominal interest at equilibrium will be 5 percent. This gives quite a lot of room for cutting nominal interest rates before you reach zero. However, around the world, most central bankers think the equilibrium real interest rate (r^*) is not a fixed number and that it has declined a lot in recent years. To give an example, the members of the Federal Reserve's policy making committee, the Federal Open Market Committee (FOMC), are asked to provide forecasts of major macroeconomic variables over the next few years and also over the long-term. By subtracting their forecasts for long-run inflation (which basically always equal 2 percent) from their estimate of the long-run federal funds rate, you can calculate what

[1] Lars Svensson (2003). "Escaping from a Liquidity Trap and Deflation: The Foolproof Way and Others." *Journal of Economic Perspectives.*

the members think the long-run real interest rate will be. In January 2012, their median estimate of the long-run real federal funds rate was 2.25 percent. By June 2020, this estimate was 0.4 percent.

If the equilibrium real interest rate was only 0.4 percent, then the equilibrium nominal rate with a 2 percent inflation target will be only 2.4 percent, leaving much less room before you hit zero interest rates. One way to address this issue would be to target a higher inflation rate. With an inflation target of 4 percent, then the equilibrium nominal rate would be 4.4 percent and there would be an additional 200 basis points of monetary easing possible before the zero bound was hit. There has been some discussion of this idea in academic circles but most central bankers are generally very much against this idea.

Part II

Rational Expectations

Chapter 5

Rational Expectations and Asset Prices

One of the things we've focused on is how people formulate expectations about inflation. We put forward one model of how these expectations were formulated, an adaptive expectations model in which people formulated their expectations by looking at past values for a series. In this section of the book, we will look at an alternative approach that macroeconomists call "rational expectations". This approach is widely used in macroeconomics and we will cover its application to models of asset prices, consumption and other macroeconomic variables.

Rational Expectations and Macroeconomics

Almost all economic transactions rely crucially on the fact that the economy is not a "one-period game." In the language of macroeconomists, most economic decisions have an *intertemporal* element to them. Consider some obvious examples:

- We accept cash in return for goods and services because we know that, in the future, this cash can be turned into goods and services for ourselves.
- You don't empty out your bank account today and go on a big splurge because you're still going to be around tommorrow and will have consumption needs then.

- Conversely, sometimes you spend more than you're earning because you can get a bank loan in anticipation of earning more in the future, and paying the loan off then.

- Similarly, firms will spend money on capital goods like trucks or computers largely in anticipation of the benefits they will bring in the future.

Another key aspect of economic transactions is that they generally involve some level of *uncertainty*, so we don't always know what's going to happen in the future. Take two of the examples just given. While it is true that one can accept cash in anticipation of turning it into goods and services in the future, uncertainty about inflation means that we can't be sure of the exact quantity of these goods and services. Similarly, one can borrow in anticipation of higher income at a later stage, but few people can be completely certain of their future incomes.

For these reasons, people have to make most economic decisions based on their subjective *expectations* of important future variables. In valuing cash, we must formulate an expectation of future values of inflation; in taking out a bank loan, we must have some expectation of our future income. These expectations will almost certainly turn out to have been incorrect to some degree, but one still has to formulate them before making these decisions.

So, a key issue in macroeconomic theory is how people formulate expectations of economic variables in the presence of uncertainty. Prior to the 1970s, this aspect of macro theory was largely *ad hoc*. Different researchers took different approaches, but generally it was assumed that agents used some simple extrapolative rule whereby the expected future value of a variable was close to some weighted average of its recent past values. However, such models were widely criticised in the 1970s by economists such as Robert Lucas and Thomas Sargent. Lucas and Sargent instead promoted the use of an alternative approach which they called "rational expectations." This approach had been introduced in an important 1961 paper by John Muth.

The idea that agents expectations are somehow "rational" has various possible interpretations. However, when Muth's concept of rational expectations meant two very specific things:

- They use publicly available information in an efficient manner. Thus, they do not make systematic mistakes when formulating expectations.
- They understand the structure of the model economy and base their expectations of variables on this knowledge.

To many economists, this is a natural baseline assumption: We usually assume agents behave in an optimal fashion, so why would we assume that the agents don't understand the structure of the economy, and formulate expectations in some sub-optimal fashion? That said, rational expectations models generally produce quite strong predictions, and these can be tested. Ultimately, any assessment of a rational expectations model must be based upon its ability to fit the relevant macro data.

How We Will Describe Expectations

We will start with some terminology to explain how we will represent expectations. Suppose our model economy has a uncertainty so that people do not know what is going to happen in the future. Then we will write $E_t Z_{t+2}$ to mean the expected value the agents in the economy have at time t for what Z is going to be at time $t+2$. In other words, we assume people have a distribution of potential outcomes for Z_{t+2} and $E_t Z_{t+2}$ is mean of this distribution.

It is important to realise that E_t is not a number that is multiplying Z_{t+2}. Instead, it is a qualifier explaining that we are dealing with people's prior expectations of a Z_{t+2} rather than the actual realised value of Z_{t+2} itself.

We will be using some basic properties of the expected value of distributions. Specifically, we will use the fact that expected values of distributions is what is known as a *linear operator*. What is meant by that is that

$$E_t \left(\alpha X_{t+k} + \beta Y_{t+k} \right) = \alpha E_t X_{t+k} + \beta E_t Y_{t+k} \tag{5.1}$$

Some examples of this are the following. The expected value of five times a series equals five times the expected value of the series

$$E_t \left(5 X_{t+k} \right) = 5 E_t \left(X_{t+k} \right) \tag{5.2}$$

And the expected value of the sum of two series equals the sum of the two expected values.

$$E_t \left(X_{t+k} + Y_{t+k} \right) = E_t X_{t+k} + E_t Y_{t+k} \tag{5.3}$$

Asset Prices

The first class of rational expectations models that we will look at relate to the determination of asset prices. Asset prices are an increasingly important topic in macroeconomics. Movements in asset prices affect the wealth of consumers and businesses and have an important influence on spending decisions. In addition, while most of the global recessions that preceded the year 2000 were due to boom and bust cycles involving inflation getting too high and central banks slowing the economy to constrain it, more recent global recessions—the "dot com" recession of 2000/01 and the "great recession" of 2008/09—were triggered by big declines in asset prices following earlier large increases. A framework for discussing these movements is thus a necessary part of any training in macroeconomics.

In this chapter, we will start by considering the case of an asset that can be purchased today for price P_t and which yields a *dividend* of D_t. While this terminology obviously fits with the asset being a share of equity in a firm and D_t being the dividend payment, it could also be a house and D_t could be the net return from renting this house out, i.e. the rent minus any costs incurred in maintenance or managment fees. If this asset is sold tomorrow for price P_{t+1}, then it generates a rate of return on this investment of

$$r_{t+1} = \frac{D_t + \Delta P_{t+1}}{P_t} \tag{5.4}$$

This rate of return has two components, the first reflects the dividend received during the period the asset was held, and the second reflects the *capital gain* (or loss) due to the price of the asset changing from period t to period $t+1$. This can also be written in terms of the so-called *gross return* which is just one plus the rate of return.

$$1 + r_{t+1} = \frac{D_t + P_{t+1}}{P_t} \tag{5.5}$$

A useful re-arrangement of this equation that we will repeatedly work with is the following:

$$P_t = \frac{D_t}{1 + r_{t+1}} + \frac{P_{t+1}}{1 + r_{t+1}} \tag{5.6}$$

Asset Prices with Rational Expectations and Constant Expected Returns

We will now consider a *rational expectations* approach to the determination of asset prices. Rational expectations means investors understand equation (5.6) and that all expectations of future variables must be consistent with it. This implies that

$$E_t P_t = E_t \left[\frac{D_t}{1 + r_{t+1}} + \frac{P_{t+1}}{1 + r_{t+1}} \right] \tag{5.7}$$

where E_t means the expectation of a variable formulated at time t. The stock price at time t is observable to the agent so $E_t P_t = P_t$, implying

$$P_t = E_t \left[\frac{D_t}{1 + r_{t+1}} + \frac{P_{t+1}}{1 + r_{t+1}} \right] \tag{5.8}$$

A second assumption that we will make for the moment is that the expected return on assets equals some constant value for all future periods, unrelated to the dividend process.

$$E_t r_{t+k} = r \quad\quad k = 1, 2, 3, ... \tag{5.9}$$

One way to think of this is that there is a "required return", determined perhaps by the rate of return available on some other asset, which this asset must deliver. With this assumption in hand and assuming that everyone knows the value of D_t, equation (5.8) can be re-written as

$$P_t = \frac{D_t}{1 + r} + \frac{E_t P_{t+1}}{1 + r} \tag{5.10}$$

The Repeated Substitution Method

Equation (5.10) is a specific example of what is known as a *first-order stochastic difference equation.*[1] Because such equations occur commonly in macroeconomics, it will be useful to write down the general approach to solving these equations, rather than just focusing only on our current asset price example. In general, this type of equation can be written as

$$y_t = ax_t + bE_t y_{t+1} \tag{5.11}$$

Its solution is derived using a technique called *repeated substitution.* This works as follows. Equation (5.11) holds in all periods, so under the assumption of rational expectations, the agents in the economy understand the equation and formulate their expectation in a way that is consistent with it:

$$E_t y_{t+1} = aE_t x_{t+1} + bE_t E_{t+1} y_{t+2} \tag{5.12}$$

Note that this last term ($E_t E_{t+1} y_{t+2}$) should simplify to $E_t y_{t+2}$: It would not be rational if you expected that next period you would have a higher or lower expectation for y_{t+2} because it implies you already have some extra information and are not using it. This is known as the *Law of Iterated Expectations.* Using this, we get

$$E_t y_{t+1} = aE_t x_{t+1} + bE_t y_{t+2} \tag{5.13}$$

Substituting this into the previous equation, we get

$$y_t = ax_t + abE_t x_{t+1} + b^2 E_t y_{t+2} \tag{5.14}$$

Repeating this method by substituting in for $E_t y_{t+2}$, and then $E_t y_{t+3}$ and so on, we get a general solution of the form

$$y_t = ax_t + abE_t x_{t+1} + ab^2 E_t x_{t+2} + ... + ab^{N-1} E_t x_{t+N-1} + b^N E_t y_{t+N} \tag{5.15}$$

[1] Stochastic means random or incorporating uncertainty. It applies to this equation because agents do not actually know P_{t+1} but instead formulate expectations of it.

which can be written in more compact form as

$$y_t = a \sum_{k=0}^{N-1} b^k E_t x_{t+k} + b^N E_t y_{t+N} \tag{5.16}$$

For those of you unfamiliar with the summation sign terminology, summation signs work like this

$$\sum_{k=0}^{2} z_k = z_0 + z_1 + z_2 \tag{5.17}$$

$$\sum_{k=0}^{3} z_k = z_0 + z_1 + z_2 + z_3 \tag{5.18}$$

$$\sum_{k=0}^{4} z_k = z_0 + z_1 + z_2 + z_3 + z_4 \tag{5.19}$$

and so on.

The Dividend-Discount Model

Comparing equations (5.10) and (5.11), we can see that our asset price equation is a specific case of the first-order stochastic difference equation with

$$y_t = P_t \tag{5.20}$$

$$x_t = D_t \tag{5.21}$$

$$a = \frac{1}{1+r} \tag{5.22}$$

$$b = \frac{1}{1+r} \tag{5.23}$$

This implies that the asset price can be expressed as follows

$$P_t = \sum_{k=0}^{N-1} \left(\frac{1}{1+r}\right)^{k+1} E_t D_{t+k} + \left(\frac{1}{1+r}\right)^N E_t P_{t+N} \tag{5.24}$$

Another assumption usually made is that this final term tends to zero as N gets big:

$$\lim_{N\to\infty} \left(\frac{1}{1+r}\right)^N E_t P_{t+N} = 0 \tag{5.25}$$

What is the logic behind this assumption? One explanation is that if it did not hold then we could set all future values of D_t equal to zero, and the asset price would still be positive. But an asset that never pays out should be inherently worthless, so this condition rules this possibility out. With this imposed, our solution becomes

$$P_t = \sum_{k=0}^{\infty} \left(\frac{1}{1+r} \right)^{k+1} E_t D_{t+k} \tag{5.26}$$

This equation, which states that asset prices should equal a discounted present-value sum of expected future dividends, is known as the *dividend-discount model*.

Explaining the Solution Without Equations

The repeated substitution solution is really important to understand so let me try to explain it without equations. Suppose I told you that the right way to price a stock was as follows.

> Today's stock price should equal today's dividend plus half of tomorrow's expected stock price.

Now suppose it's Monday. Then that means the right formula should be

> Monday's stock price should equal Monday's dividend plus half of Tuesday's expected stock price.

It also means the following applies to Tuesday's stock price

> Tuesday's stock price should equal Tuesday's dividend plus half of Wednesday's expected stock price.

If people had rational expectations, then Monday's stock prices would equal

> Monday's dividend plus half of Tuesday's expected dividend plus one-quarter of Wednesday's expected stock price

And being consistent about it—factoring in what Wednesday's stock price should be—you'd get the price being equal to

> Monday's dividend plus half of Tuesday's expected dividend plus one-quarter of Wednesday's expected dividend plus one-eigth of Thursday's expected dividend and so on.

This is the idea being captured in equation (5.26).

Constant Expected Dividend Growth: The Gordon Growth Model

A useful special case that is often used as a benchmark for thinking about stock prices is the case in which dividend payments are expected to grow at a constant rate such that

$$E_t D_{t+k} = (1+g)^k D_t \tag{5.27}$$

In this case, the dividend-discount model predicts that the stock price should be given by

$$P_t = \frac{D_t}{1+r} \sum_{k=0}^{\infty} \left(\frac{1+g}{1+r} \right)^k \tag{5.28}$$

Now, remember the old multiplier formula, which states that as long as $0 < c < 1$, then

$$1 + c + c^2 + c^3 + ... = \sum_{k=0}^{\infty} c^k = \frac{1}{1-c} \tag{5.29}$$

This geometric series formula gets used *a lot* in modern macroeconomics, not just in examples involving the multiplier. Here we can use it as long as $\frac{1+g}{1+r} < 1$, i.e. as long as r (the expected return on the stock market) is greater than g (the growth rate of dividends). We will assume this holds. Thus, we have

$$P_t = \frac{D_t}{1+r}\frac{1}{1-\frac{1+g}{1+r}} \tag{5.30}$$

$$= \frac{D_t}{1+r}\frac{1+r}{1+r-(1+g)} \tag{5.31}$$

$$= \frac{D_t}{r-g} \tag{5.32}$$

When dividend growth is expected to be constant, prices are a multiple of current dividend payments, where that multiple depends positively on the expected future growth rate of dividends and negatively on the expected future rate of return on stocks. This formula is often called the *Gordon growth model*, after the economist that popularized it.[2] It is often used as a benchmark for assessing whether an asset is above or below the "fair" value implied by rational expectations. Valuations are often expressed in terms of dividend-price ratios, and the Gordon formula says this should be

$$\frac{D_t}{P_t} = r - g \tag{5.33}$$

Allowing for Variations in Dividend Growth

A more flexible way to formulate expectations about future dividends is to assume that dividends fluctuate around a steady-growth trend. An example of this is the following

$$D_t = A(1+g)^t + u_t \tag{5.34}$$

$$u_t = \rho u_{t-1} + \epsilon_t \tag{5.35}$$

These equations state that dividends are the sum of two processes: The first equals A at time $t = 0$ and then grows at rate g each period. The second, u_t, measures a cyclical component of

[2]The formula appeared in Myron Gordon's 1962 book *The Investment, Financing and Valuation of the Corporation.*

dividends, and this follows what is known as a first-order autoregressive process (AR(1) for short). Here ϵ_t is a zero-mean random "shock" term. Over large samples, we would expect u_t to have an average value of zero, but deviations from zero will be more persistent the higher is the value of the parameter ρ.

We will now derive the dividend-discount model's predictions for stock prices when dividends follow this process. The model predicts that

$$P_t = \sum_{k=0}^{\infty} \left(\frac{1}{1+r}\right)^{k+1} E_t \left(A(1+g)^{t+k} + u_{t+k}\right) \tag{5.36}$$

Let's split this sum into two. First the trend component,

$$\begin{aligned}
\sum_{k=0}^{\infty} \left(\frac{1}{1+r}\right)^{k+1} E_t \left(A(1+g)^{t+k}\right) &= \frac{A(1+g)^t}{1+r} \sum_{k=0}^{\infty} \left(\frac{1+g}{1+r}\right)^k && (5.37)\\
&= \frac{A(1+g)^t}{1+r} \frac{1}{1-\frac{1+g}{1+r}} && (5.38)\\
&= \frac{A(1+g)^t}{1+r} \frac{1+r}{1+r-(1+g)} && (5.39)\\
&= \frac{A(1+g)^t}{r-g} && (5.40)
\end{aligned}$$

Second, the cyclical component. Because $E(\epsilon_{t+k}) = 0$, we have

$$\begin{aligned}
E_t u_{t+1} &= E_t(\rho u_t + \epsilon_{t+1}) = \rho u_t && (5.41)\\
E_t u_{t+2} &= E_t(\rho u_{t+1} + \epsilon_{t+2}) = \rho^2 u_t && (5.42)\\
E_t u_{t+k} &= E_t(\rho u_{t+k-1} + \epsilon_{t+k}) = \rho^k u_t && (5.43)
\end{aligned}$$

So, this second sum can be written as

$$\begin{aligned}
\sum_{k=0}^{\infty} \left(\frac{1}{1+r}\right)^{k+1} E_t u_{t+k} &= \frac{u_t}{1+r} \sum_{k=0}^{\infty} \left(\frac{\rho}{1+r}\right)^k && (5.44)\\
&= \frac{u_t}{1+r} \frac{1}{1-\frac{\rho}{1+r}} && (5.45)
\end{aligned}$$

$$= \frac{u_t}{1+r}\frac{1+r}{1+r-\rho} \tag{5.46}$$
$$= \frac{u_t}{1+r-\rho} \tag{5.47}$$

Putting these two sums together, the stock price at time t is

$$P_t = \frac{A(1+g)^t}{r-g} + \frac{u_t}{1+r-\rho} \tag{5.48}$$

In this case, stock prices don't just grow at a constant rate. Instead they depend positively on the cyclical component of dividends, u_t, and the more persistent are these cyclical deviations (the higher ρ is), the larger is their effect on stock prices. To give a concrete example, suppose $r = 0.1$. When $\rho = 0.9$ the coefficient on u_t is

$$\frac{1}{1+r-\rho} = \frac{1}{1.1-0.9} = 5 \tag{5.49}$$

But if $\rho = 0.6$, then the coefficient falls to

$$\frac{1}{1+r-\rho} = \frac{1}{1.1-0.6} = 2 \tag{5.50}$$

Note also that when taking averages over long periods of time, the u components of dividends and prices will average to zero. Thus, over longer averages the Gordon growth model would be approximately correct, even though the dividend-price ratio isn't always constant. Instead, prices would tend to be temporarily high relative to dividends during periods when dividends are expected to grow at above-average rates for a while, and would be temporarily low when dividend growth is expected to be below average for a while. This is why the Gordon formula is normally seen as a guide to long-run average valuations rather than a prediction as to what the market should be right now.

Unpredictability of Stock Returns

The dividend-discount model has some very specific predictions for how stock prices should change over time. It implies that the change in prices from period t to period $t+1$ should be

$$P_{t+1} - P_t = \sum_{k=0}^{\infty} \left(\frac{1}{1+r} \right)^{k+1} E_{t+1} D_{t+k+1} - \sum_{k=0}^{\infty} \left(\frac{1}{1+r} \right)^{k+1} E_t D_{t+k} \tag{5.51}$$

Taking away the summation signs and writing this out in long form, it looks like this

$$\begin{aligned} P_{t+1} - P_t \quad = \quad & \left[\left(\frac{1}{1+r} \right) D_{t+1} + \left(\frac{1}{1+r} \right)^2 E_{t+1} D_{t+2} + \left(\frac{1}{1+r} \right)^3 E_{t+1} D_{t+3} + ... \right] \\ & - \left[\left(\frac{1}{1+r} \right) D_t + \left(\frac{1}{1+r} \right)^2 E_t D_{t+1} + \left(\frac{1}{1+r} \right)^3 E_t D_{t+2} + ... \right] \end{aligned} \tag{5.52}$$

We can re-arrange this equation in a useful way by grouping together each of the two terms that involve D_{t+1}, D_{t+2}, D_{t+3} and so on. (There is only one term involving D_t.) This can be written as follows

$$\begin{aligned} P_{t+1} - P_t \quad = \quad & - \left(\frac{1}{1+r} \right) D_t + \left[\left(\frac{1}{1+r} \right) D_{t+1} - \left(\frac{1}{1+r} \right)^2 E_t D_{t+1} \right] \\ & + \left[\left(\frac{1}{1+r} \right)^2 E_{t+1} D_{t+2} - \left(\frac{1}{1+r} \right)^3 E_t D_{t+2} \right] + \\ & + \left[\left(\frac{1}{1+r} \right)^3 E_{t+1} D_{t+3} - \left(\frac{1}{1+r} \right)^4 E_t D_{t+3} \right] + ... \end{aligned} \tag{5.53}$$

This equation explains three reasons why prices change from period P_t to period P_{t+1}.

- P_{t+1} differs from P_t because it does not take into account D_t – this dividend has been paid now and has no influence any longer on the price at time $t+1$. This is the first term on the right-hand side above.

- P_{t+1} applies a smaller discount rate to future dividends because have moved forward one period in time, e.g. it discounts D_{t+1} by $\left(\frac{1}{1+r}\right)$ instead of $\left(\frac{1}{1+r}\right)^2$.

- People formulate new expectations for the future path of dividends e.g. $E_t D_{t+2}$ is gone and has been replaced by $E_{t+1} D_{t+2}$

In general, the first few items above should not be too important. A single dividend payment being made shouldn't have too much impact on a stock's price and the discount rate shouldn't change too much over a single period (e.g. if r is relatively small, then $\left(\frac{1}{1+r}\right)$ and $\left(\frac{1}{1+r}\right)^2$ shouldn't be too different.) This means that changing expectations about future dividends should be the main factor driving changes in stock prices.

In fact, it turns out there is a very specific result linking the behaviour of stock prices with changing expectations. Ultimately, it is not stock prices, *per se*, that investors are interested in. Rather, they are interested in the combined return incorporating both price changes and dividend payments, as described by equation (5.4). It turns out that movements in stock returns are entirely driven by changes in dividend expectations.

With a number of lines of algebra (described in an appendix) equation (5.53) can be re-expressed as

$$P_{t+1} - P_t = -D_t + rP_t + \sum_{k=1}^{\infty} \left[\left(\frac{1}{1+r} \right)^k (E_{t+1} D_{t+k} - E_t D_{t+k}) \right] \tag{5.54}$$

Recalling the definition of the one-period return on a stock from equation (1), this return can be written as

$$r_{t+1} = \frac{D_t + \Delta P_{t+1}}{P_t} = r + \frac{\sum_{k=1}^{\infty} \left(\frac{1}{1+r}\right)^k (E_{t+1} D_{t+k} - E_t D_{t+k})}{P_t} \tag{5.55}$$

This is a very important result. It tells us that, if the dividend-discount model is correct, then the rate of return on stocks depends on how people change their minds about what they expect to happen to dividends in the future: The $E_{t+1} D_{t+k} - E_t D_{t+k}$ terms on the right-hand side of equation (5.55) describe the difference between what people expected at time $t+1$ for the dividend at time $t+k$ and what they expected for this same dividend payment at time t.

Importantly, if we assume that people formulate rational expectations, then the return on stocks should be unpredictable. This is because, if we could tell in advance how people were going to change their expectations of future events, then that would mean people have not been using information in an efficient manner. So, with rational expectations, the term in the summation sign in equation (5.55) must be zero on average and must reflect "news" that could not have been forecasted at time t. So the only thing determining changes in stock returns in the innately unforecastable process of people incorporating completely new information.

One small warning about this result. It is often mis-understood as a prediction that *stock prices* (rather than stock returns) should be unpredictable. This is not the case. The series that should be unpredictable is the total stock return including the dividend payment. Indeed, the model predicts that a high dividend payments at time t lowers stock prices at time $t+1$. Consider for example a firm that promises to make a huge dividend payment next month but says they won't make any payments after this for a long time. In that case, we would expect the price of the stock to fall after the dividend is payment. This shows that, even with rational expectations, stock prices movements can sometimes be predictable. Because dividend payments are only made on an occasional basis, this prediction can be tested and various studies have indeed found so-called "ex-dividend" effects whereby a stock price falls after a dividend is paid.

Evidence on Predictability and "Efficient Markets"

The theoretical result that stock returns should be unpredictable was tested in a series of empirical papers in the 1960s and 1970s, most notably by University of Chicago professor Eugene Fama and his co-authors. Fama's famous 1970 paper "Efficient Capital Markets: A Review of Theory and Empirical Work" reviewed much of this work. This literature came to a clear conclusion that stock returns did seem to be essentially unpredictable. The idea that you could not make easy money by "timing the market" entered public discussion with Burton Malkiel's famous 1973 book, *A Random Walk Down Wall Street* being particularly influential. A "random walk" is a series whose changes cannot be forecasted and rational expectations implies that changes in the cumulative return on a stock are unforecastable.

The work of Fama and his co-authors was very important in establishing key facts about how financial markets work. One downside to this research, though, was the introduction of a terminology that proved confusing. Fama's 1970 paper describes financial market as being "efficient" if they "fully reflect all available information." In general, the researchers contributing to this literature concluded financial markets were efficient because stock returns were difficult to forecast. However, this turned out to be a bit of a leap. It is certainly true that if stock prices incorporate all available information in the rational manner described above, then returns should be hard to forecast. But the converse doesn't necessarily apply: Showing that it was difficult to forecast stock returns turned out to not be the same thing as proving that stock markets were efficient.

Robert Shiller on Excess Volatility

The idea that financial markets were basically efficient was widely accepted in the economics profession by the late 1970s. Then, a Yale economist in his mid-thirties, Robert Shiller, dropped something of a bombshell on the finance profession. Shiller showed that the dividend-discount model beloved of finance academics completely failed to match the observed *volatility* of stock prices.[3] Specifically, stock prices were much more volatile than could be justified by the model.

To understand Shiller's basic point, we need to take a step back and think about some basic concepts relating to the formulation of expectations. First note that the *ex post* outcome for any variable can be expressed as the sum of its *ex ante* value expected by somebody and the unexpected component (i.e. the amount by which that person's expectation was wrong). This can be written in a formula as

$$X_t = E_{t-1}X_t + \epsilon_t \tag{5.56}$$

From statistics, we know that the variance of the sum of two variables equals the sum of their two variances plus twice their covariance. This means that the variance of X_t can be described by

$$\mathrm{Var}\left(X_t\right) = \mathrm{Var}\left(E_{t-1}X_t\right) + \mathrm{Var}\left(\epsilon_t\right) + 2\mathrm{Cov}\left(E_{t-1}X_t, \epsilon_t\right) \tag{5.57}$$

Now note that this last covariance term—between the "surprise" element ϵ_t and the ex-ante expectation $E_{t-1}X_t$—should equal zero if expectations are fully rational. If there was a correlation—for instance, so that a low value of the expectation tended to imply a high value for the error—then this would means that you could systematically construct a better forecast once you had seen the forecast that was provided. For example, if a low forecasted

[3] "Do Stock Prices Move Too Much to be Justified by Subsequent Changes in Dividends?" *American Economic Review*, June 1981

value tended to imply a positive error then you could construct a better forecast by going for a higher figure. But this contradicts the idea that investors have rational expectations and thus use all information efficiently.

So, if expectations are rational, then we have

$$\text{Var}\left(X_t\right) = \text{Var}\left(E_{t-1}X_t\right) + \text{Var}\left(\epsilon_t\right) \tag{5.58}$$

The variance of the observed series must equal the variance of the *ex ante* expectation plus the variance of the unexpected component. Provided there is uncertainty, so there is some unexpected component, then we must have

$$\text{Var}\left(X_t\right) > \text{Var}\left(E_{t-1}X_t\right) \tag{5.59}$$

In other words, the variance of the *ex post* outcome should be higher than the variance of *ex ante* rational expectation.

This reasoning has implications for the predicted volatility of stock prices. Equation (5.26) says that stock prices are an *ex ante* expectation of a discount sum of future dividends. Shiller's observation was that rational expectations should imply that the variance of stock prices be less than the variance of the present value of subsequent dividend movements:

$$\text{Var}(P_t) < \text{Var}\left[\sum_{k=0}^{\infty}\left(\frac{1}{1+r}\right)^{k+1} D_{t+k}\right] \tag{5.60}$$

A check on this calculation, using a wide range of possible values for r, reveals that this inequality does not hold: Stocks are actually much more volatile than suggested by realized movements in dividends.[4]

[4] While technically, the infinite sum of dividends can't be calculated because we don't have data going past the present, Shiller filled in all terms after the end of his sample based on plausible assumptions, and the results are not sensitive to these assumptions.

Figure 5.1 reproduces the famous graph from Shiller's 1981 paper showing actual stock prices (the solid line) moving around much more over time than his "discounted outcome of dividends" series (the dashed line).

Figure 5.1: Shiller's 1981 Chart Illustrating Excess Volatility

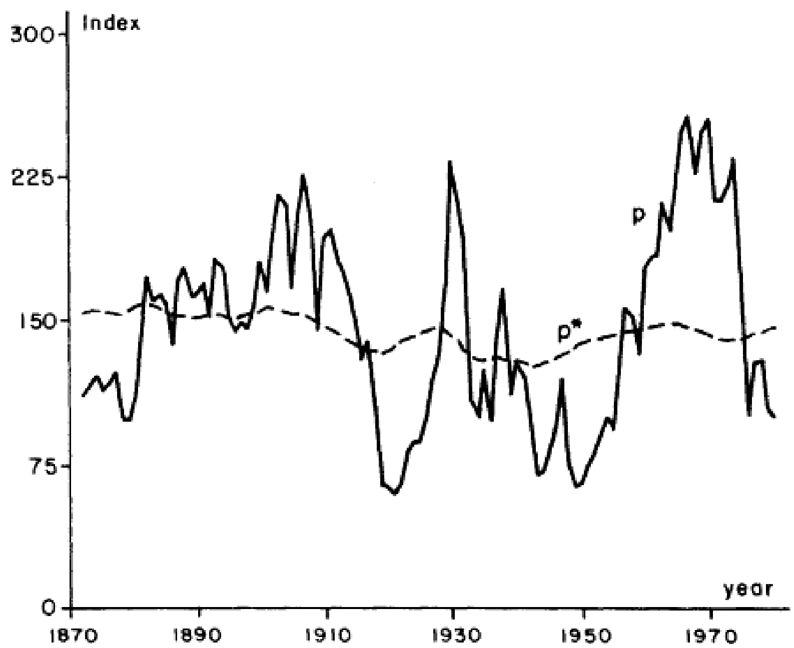


FIGURE 1

Note: Real Standard and Poor's Composite Stock Price Index (solid line p) and *ex post* rational price (dotted line p^*), 1871–1979, both detrended by dividing a long-run exponential growth factor. The variable p^* is the present value of actual subsequent real detrended dividends, subject to an assumption about the present value in 1979 of dividends thereafter. Data are from Data Set 1, Appendix.

Longer-Run Predictability

We saw earlier that the dividend-discount model predicts that when the ratio of dividends to prices is low, this suggests that investors are confident about future dividend growth. Thus, a low dividend-price ratio should help to predict higher future dividend growth. Shiller's volatility research pointed out, however, that there appears to be a lot of movements in stock prices that never turn out to be fully justified by later changes dividends. In fact, later research went a good bit further. For example, Campbell and Shiller (2001) show that over longer periods, dividend-price ratios are of essentially no use at all in forecasting future dividend growth.[5] In fact, a high ratio of prices to dividends, instead of forecasting high growth in dividends, tends to forecast lower future returns on the stock market albeit with a relatively low R-squared. See Figure 5.2.

This last finding seems to contradict Fama's earlier conclusions that it was difficult to forecast stock returns but these results turn out to be compatible with both those findings and the volatility results. Fama's classic results on predictability focused on explaining short-run stock returns e.g. can we use data from this year to forecast next month's stock returns? However, the form of predictability found by Campbell and Shiller (and indeed a number of earlier studies) related to predicting average returns over multiple years. It turns out an inability to find short-run predictability is not the same thing as an inability to find longer-run predictability.

To understand this, we need to develop some ideas about forecasting time series. Consider a series that follows the following $AR(1)$ time series process:

$$y_t = \rho y_{t-1} + \epsilon_t \tag{5.61}$$

[5] NBER Working Paper No. 8221.

where ϵ_t is a random and unpredictable "noise" process with a zero mean. If $\rho = 1$ then the change in the series is

$$y_t - y_{t-1} = \epsilon_t \tag{5.62}$$

so the series is what is what we described earlier as a random walk process whose changes cannot be predicted. Suppose, however, that ρ was close to but a bit less than one, say $\rho = 0.99$. The change in the series would now be given by

$$y_t - y_{t-1} = -0.01 y_{t-1} + \epsilon_t \tag{5.63}$$

Now suppose you wanted to assess whether you could forecast the change in the series based on last period's value of the series. You could run a regression of the change in y_t on last period's value of the series. The true coefficient in this relationship is -0.01 with the ϵ_t being the random error. This coefficient of -0.01 is so close to zero that you will probably be unable to reject that the true coefficient is zero unless you have far more data than economists usually have access to.

But what if you were looking at forecasting changes in the series over a longer time-horizon? To understand why this might be different, we can do another repeated substitution trick. The series y_t depends on its lagged value, y_{t-1} and a random shock. But y_{t-1} in turn depended on y_{t-2} and another random shock. And y_{t-2} in turn depended on y_{t-3} and another random shock. And so on. Plugging in all of these substitutions you get the following.

$$\begin{aligned}
y_t &= \rho y_{t-1} + \epsilon_t \\
&= \rho^2 y_{t-2} + \epsilon_t + \rho \epsilon_{t-1} \\
&= \rho^3 y_{t-3} + \epsilon_t + \rho \epsilon_{t-1} + \rho^2 \epsilon_{t-2} \\
&= \rho^N y_{t-N} + \sum_{k=0}^{N-1} \rho^k \epsilon_{t-k}
\end{aligned} \tag{5.64}$$

Now suppose you wanted to forecast the change in y_t over N periods with the value of the series from N periods ago. This change can be written as

$$y_t - y_{t-N} = \left(\rho^N - 1\right) y_{t-N} + \sum_{k=0}^{N-1} \rho^k \epsilon_{t-k} \tag{5.65}$$

Again the change in y_t over this period can be written as a function of a past value of the series and some random noise. The difference in this case is that the coefficient on the lagged value doesn't have to be small anymore even if we had a near-random-walk series. For example, suppose $\rho = 0.99$ and $N = 50$ so were looking at the change in the series over 50 periods. In this case, the coefficient is $(0.99^{50} - 1) = -0.4$. For this reason, regressions that seek to predict combined returns over longer periods have found statistically significant evidence of predictability even though this evidence cannot be found for predicting returns over shorter periods.

It is easy to demonstrate this result using any software that can generate random numbers. For example, in an appendix at the back of this chapter I provide a short programme written for the econometric package RATS. The code is pretty intuitive and could be repeated for lots of other packages. The programme generates random $AR(1)$ series with $\rho = 0.99$ by starting them off with a value of zero and then drawing random errors to generate full time series. Then regressions for sample sizes of 200 are run to see if changes in the series over one period, twenty periods and fifty periods can be forecasted by the relevant lagged values. This is done 10,000 times and the average t-statistics from these regressions are calculated.

Table 1 shows the results. The average t-statistic for the one-period forecasting regression is -1.24, not high enough to reject the null hypothesis that there is no forecasting power. In contrast, the average t-statistic for the 20-period forecasting regression is -5.69 and the average t-statistic for the 50-period forecasting regression is -8.95, so you can be very confident that there is statistically significant forecasting power over these horizons.

The intuition behind these results is fairly simple. Provided ρ is less than one in absolute value, $AR(1)$ series are what is known as mean-stationary. In other words, they tend to revert back to their average value. In the case of the y_t series here, this average value is zero. The speed at which you can expect them to return to this average value will be slow if ρ is high but they will eventually return. So if you see a high value of y_t, you can't really be that confident that it will fall next period but you can be very confident that it will eventually tend to fall back towards its average value of zero.

Pulling these ideas together to explain the various stock price results, suppose prices were given by

$$P_t = \sum_{k=0}^{\infty} \left[\left(\frac{1}{1+r} \right)^{k+1} E_t D_{t+k} \right] + u_t \tag{5.66}$$

where

$$u_t = \rho u_{t-1} + \epsilon_t \tag{5.67}$$

with ρ being close to one and ϵ_t being an unpredictable noise series. This model says that stock prices are determined by two elements. The first is the rational dividend-discount price and the second is a non-fundamental $AR(1)$ element reflecting non-rational market sentiment. The latter could swing up and down over time as various fads and manias affect the market.

In this case, statistical research would generate three results:

1. Short-term stock returns would be very hard to forecast. This is partly because of the rational dividend-discount element but also because changes in the non-fundamental element are hard to forecast over short-horizons.

2. Longer-term stock returns would have a statistically significant forecastable element,

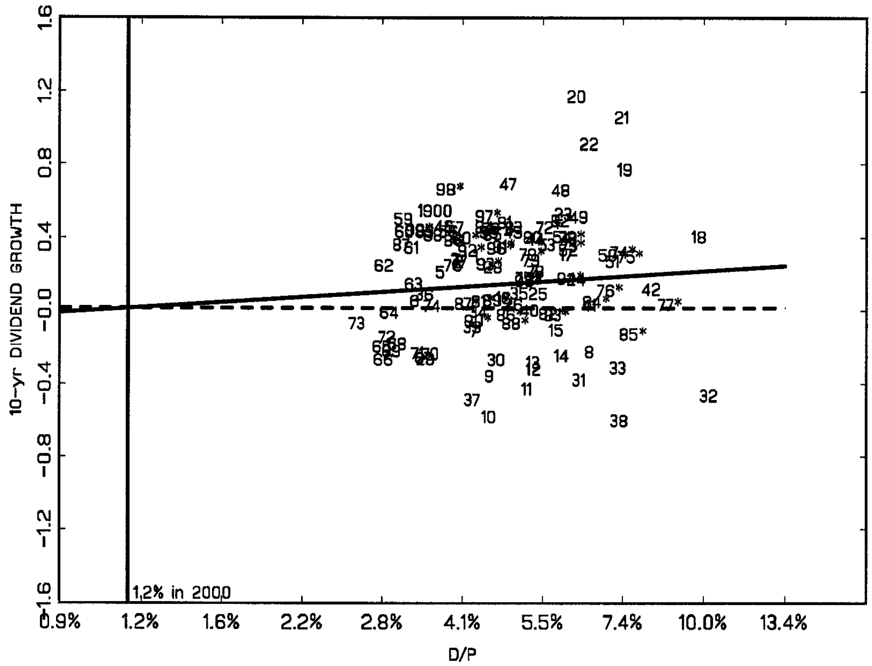
though with a relatively low R-squared. This is because the fundamental element that accounts for much of the variation cannot be forecasted while you can detect a statistically significant forecastable element for the non-fundamental component.

3. Stock prices would be more volatile than predicted by the dividend-discount model, perhaps significantly so. This is because non-fundamental series of the type described here can go through pretty long swings which adds a lot more volatility than the dividend-discount model would predict.

This suggests a possible explanation for the behaviour of stock prices. On average, they appear to be determined by something like the dividend-discount model but they also have a non-fundamental component that sees the market go through temporary (but potentially long) swings in which it moves away from the values predicted by this model.

Figure 5.2: Campbell and Shiller's 2001 Chart

Figure 3. 10-year DIVIDEND GROWTH vs D/P



10-year PRICE GROWTH vs D/P

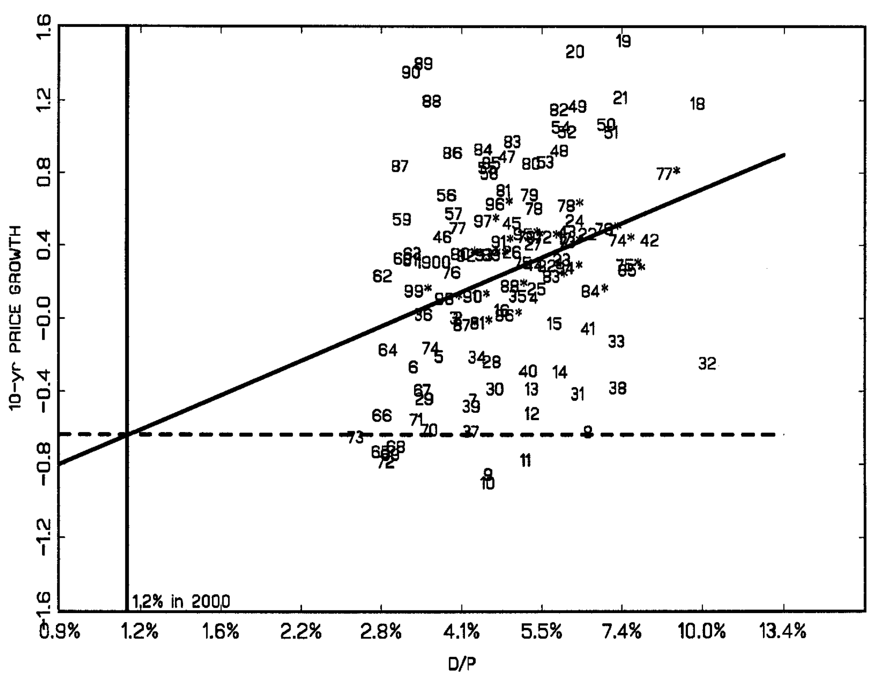


Table 1: Illustrating Long-Run Predictability

```
Statistics on Series TSTATS_1LAG
Observations                 10000
Sample Mean              -1.244998      Variance             0.463179
Standard Error            0.680573      of Sample Mean       0.006806
t-Statistic (Mean=0) -182.933896      Signif Level         0.000000
Skewness                 -0.123331      Signif Level (Sk=0) 0.000000
Kurtosis (excess)         0.512791      Signif Level (Ku=0) 0.000000
Jarque-Bera             134.915363      Signif Level (JB=0) 0.000000
```

```
Statistics on Series TSTATS_20LAG
Observations                 10000
Sample Mean              -5.693510      Variance             11.267177
Standard Error            3.356662      of Sample Mean       0.033567
t-Statistic (Mean=0) -169.618235      Signif Level         0.000000
Skewness                 -0.277797      Signif Level (Sk=0) 0.000000
Kurtosis (excess)         0.778124      Signif Level (Ku=0) 0.000000
Jarque-Bera             380.901249      Signif Level (JB=0) 0.000000
```

```
Statistics on Series TSTATS_50LAG
Observations                 10000
Sample Mean              -8.951905      Variance             31.554735
Standard Error            5.617360      of Sample Mean       0.056174
t-Statistic (Mean=0) -159.361419      Signif Level         0.000000
Skewness                 -0.338830      Signif Level (Sk=0) 0.000000
Kurtosis (excess)         0.926286      Signif Level (Ku=0) 0.000000
Jarque-Bera             548.845975      Signif Level (JB=0) 0.000000
```

Example: U.S. Stock Prices

These results would suggest that it may be possible to detect whether a stock market is over-valued or under-valued and thus to forecast its future path. Here we look at the US stock market as of July 2023.

Figure 5.3 shows the S&P 500 index which is a broad measure of large-capitalisation stocks. Figure 5.4 shows the ratio of dividends to prices for the S&P 500 index of US stocks over the period since the second world war. The measure of dividends used in the numerator is based on the average value of dividends over a twelve month period to smooth out volatile month-to-month movements. The chart shows that the dividend-price ratio for this index was 1.65% in July 2023. This is low by historical standards. The average value of this ratio over the period since 1945 is 3.2%. In other words, prices in July 2023 were twice as high relative to dividends as their historical average, which is normally a bad sign.

This ratio, however, has been indicating over-valuation for a very long time. The only point since the mid-1990s that the dividend-price ratio has exceeded that historical average was a brief period in early 2009 due to the plunge in stock prices after the Lehman Brothers bankruptcy and the emergence of the worst recession since the Great Depression. One reason for this change is that many firms have moved away from paying dividends towards using earnings to fund share repurchases. This is a way to return money to shareholders and increase the value of the remaining shares (each of them can get a higher share of future dividend payments) without the shareholders explicitly receiving dividend income at present, which would be taxable at a higher rate than capital gains.

Because of these factors, many analysts instead look at the ratio of total corporate earnings to prices. Figure 5.5 compares this to the dividend-price ratio. The historical post-War average value for the earnings-price ratio is 6.5%. The July 2023 value was 4.6%. This suggests prices were 40% higher relative to earnings than the historical average for this ratio. This is not as bad a signal as the dividend-price ratio but it is still relatively bearish.

A final consideration, however, is the discount rate being used to value stocks. We know from the Gordon growth model that one reason stock prices might be high relative to dividends (or earnings) is that the expected rate of return r may be low. Assuming the required rate of return on stocks reflects some premium over safe long-term investments such long-term government bonds, this could provide another explanation for high stock price valuations. Figure 5.6 shows the interest rate on 10-year US government bonds. In July 2023, these rates were higher than most of the previous decade but still relatively low by historical standards. What matters, however, is the relevant real interest rate and this depends on the level of inflation that is expected. It may be that the increase in interest rates experienced up to 2023 may simply reflect higher inflation expectations, so that real interest rates are still expected to be low. If this is the case, this could justify stock prices being high relative to historical averages for valuation metrics.

These considerations show that figuring out whether stocks are under- or over-valued is rarely as easy as examining one specific metric.

Figure 5.3: The S&P 500 Index of U.S. Stock Prices

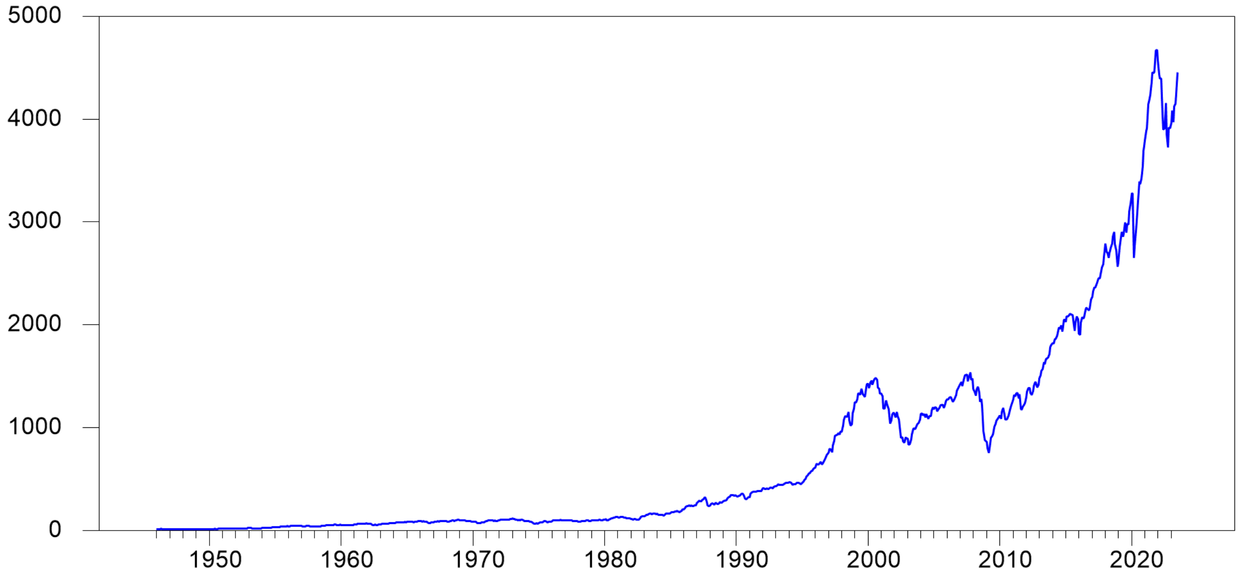


Figure 5.4: Dividend-Price Ratio for S&P 500

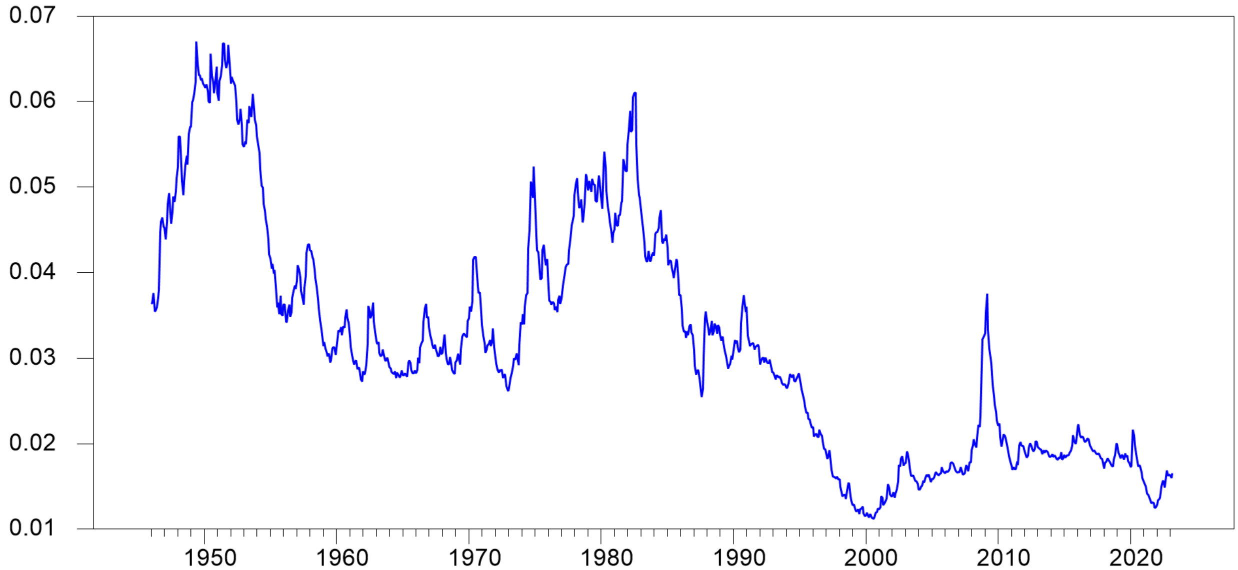


Figure 5.5: Dividend-Price and Earnings-Price Ratios for S&P 500

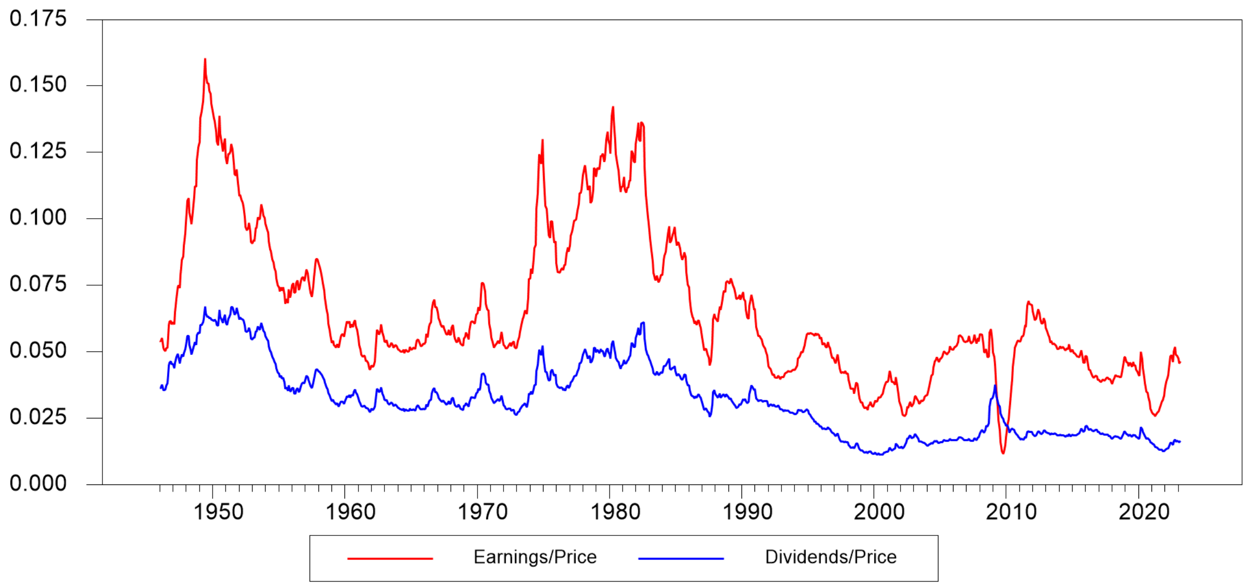
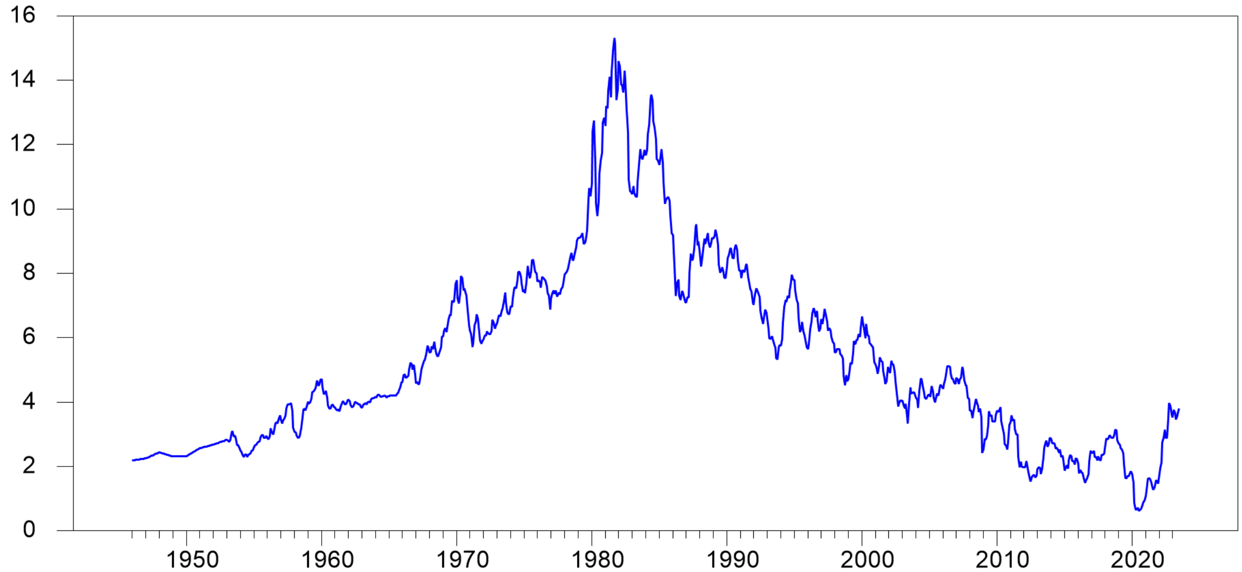


Figure 5.6: 10-Year Treasury Bond Rate



Time-Varying Expected Returns

I have suggested an alternative explanation of the facts to the dividend-discount model — one in which stock prices are also determined by a temporary but volatile non-fundamental component. However, the last observation in the previous section — about discount rates — suggests another way to "mend" the dividend-discount model and perhaps explain the extra volatility that affects stock prices: Change the model to allow for variations in expected returns. Consider the finding that a high value of the dividend-price ratio predicts poor future stock returns. Shiller suggests that this is due to temporary irrational factors gradually disappearing. But another possibility is that that the high value of this ratio is rationally anticipating low future returns.

We can reformulate the dividend-discount model with time-varying returns as follows. Let

$$R_t = 1 + r_t \tag{5.68}$$

Start again from the first-order difference equation for stock prices

$$P_t = \frac{D_t}{R_{t+1}} + \frac{P_{t+1}}{R_{t+1}} \tag{5.69}$$

where R_{t+1} is the return on stocks in period $t+1$. Moving the time-subscripts forward one period, this implies

$$P_{t+1} = \frac{D_{t+1}}{R_{t+2}} + \frac{P_{t+2}}{R_{t+2}} \tag{5.70}$$

Substitute this into the original price equation to get

$$\begin{aligned} P_t &= \frac{D_t}{R_{t+1}} + \frac{1}{R_{t+1}}\left(\frac{D_{t+1}}{R_{t+2}} + \frac{P_{t+2}}{R_{t+2}}\right) \\ &= \frac{D_t}{R_{t+1}} + \frac{D_{t+1}}{R_{t+1}R_{t+2}} + \frac{P_{t+2}}{R_{t+1}R_{t+2}} \end{aligned} \tag{5.71}$$

Applying the same trick to substitute for P_{t+2} we get

$$P_t = \frac{D_t}{R_{t+1}} + \frac{D_{t+1}}{R_{t+1}R_{t+2}} + \frac{D_{t+2}}{R_{t+1}R_{t+2}R_{t+3}} + \frac{P_{t+3}}{R_{t+1}R_{t+2}R_{t+3}} \tag{5.72}$$

The general formula is

$$P_t = \sum_{k=0}^{N-1} \left(\frac{D_{t+k}}{\prod_{m=1}^{k+1} R_{t+m}} \right) + \frac{P_{t+N}}{\prod_{m=1}^{N} R_{t+m}} \tag{5.73}$$

where $\prod_{n=1}^{h} x_i$ means the product of x_1, x_2 .... x_h. Again setting the limit of the $t+N$ term to zero and taking expectations, we get a version of the dividend-discount model augmented to account for variations in the expected rate of return.

$$P_t = \sum_{k=0}^{\infty} E_t \left(\frac{D_{t+k}}{\prod_{m=1}^{k+1} R_{t+m}} \right) \tag{5.74}$$

This equation gives a potential explanation for the failure of news about dividends to explain stock price fluctuations. Stock prices depend positively on expected future dividends. But they also depend negatively on the R_{t+k} values which measure the expected future return on stocks. So perhaps news about future stock returns explains movements in stock prices: When investors learn that future returns are going to be lower, this raises current stock prices.

In 1991, Eugene Fama provided his updated overview of the literature on the predictability of stock returns. By this point, Fama accepted the evidence on long-horizon predictability and had contributed to this literature. However, Fama and French (1988) put forward predictable time-variation in expected returns as the likely explanation for this result.[6] This explanation has also been put forward by some modern leading finance economists such as John Campbell and John Cochrane and is currently the leading hypothesis for reconciling the evidence on stock prices movements with rational expectations.[7]

[6] Eugene Fama and Kenneth French (1988). "Dividend Yields and Expected Stock Returns" *Journal of Financial Economics*, Volume 22, pages 3-25.

[7] Campbell's 1991 paper, "A Variance Decomposition for Stock Returns," *Economic Journal*, provides a nice framework for understanding this theory.

What About Interest Rates?

Changing interest rates on bonds are the most obvious source of changes in expected returns on stocks. Up to now, we only briefly discussed what determines the rate of return that investors require to invest in the stock market, but it is usually assumed that there is an arbitrage equation linking stock and bond returns, so that

$$E_t r_{t+1} = E_t i_{t+1} + \pi \tag{5.75}$$

In other words, next period's expected return on the stock market needs to equal next period's expected interest rate on bonds, i_{t+1}, plus a risk premium, π, which we will assume is constant.

Are interest rates the culprit accounting for the volatility of stock prices? They are certainly a plausible candidate. Stock market participants spend a lot of time monitoring the Fed and the ECB and news interpreted as implying higher interest rates in the future certainly tends to provoke declines in stock prices. Perhaps surprisingly, then, Campbell and Shiller (1988) showed that this type of equation still doesn't help that much in explaining stock market fluctuations.[8] Their methodology involved plugging in forecasts for future interest rates and dividend growth into the right-hand-side of (5.74) and checking how close the resulting series is to the actual dividend-price ratio. They concluded that expected fluctuations in interest rates contribute little to explaining the volatility in stock prices. A later study by Bernanke and Kuttner (2005) came to the same conclusions.[9]

[8] "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors," *Review of Financial Studies*, Autumn 1988.

[9] See "What Explains the Stock Market's Reaction to Federal Reserve Policy?", by Ben Bernanke and Kenneth Kuttner, *Journal of Finance*, June 2005.

Time-Varying Risk Premia or Behavioural Finance?

So, changes in interest rates do not appear to explain the volatility of stock market fluctuations. The final possible explanation for how the dividend-discount model may be consistent with the data is that changes in expected returns do account for the bulk of stock market movements, but that the principal source of these changes comes, not from interest rates, but from changes in the *risk premium* that determines the excess return that stocks must generate relative to bonds: The π in equation (5.75) must be changing over time. According to this explanation, asset price booms are often driven by investors being willing to take risks and receive a relatively low compensation for them (when investors are "risk-on" in the commonly-used market terminology) while busts often happen when investors start to demand higher risk premia (when they are "risk-off").

One problem with this conclusion is that it implies that, most of the time, when stocks are increasing it is because investors are anticipating lower stock returns at a later date. However, the evidence that we have on this seems to point in the other direction. For example, surveys have shown that even at the peak of "bull markets", average investors still anticipate high future returns on the market.

If one rejects the idea that, together, news about dividends and news about future returns explain all of the changes in stock prices, then one is forced to reject the rational expectations dividend-discount model as a complete model of the stock market. What is missing from this model? Many believe that the model fails to take into account of various human behavioural traits that lead people to act in a manner inconsistent with pure rational expectations. Economists like Shiller point to the various asset price "bubbles" of the past twenty years — such as the dot-com boom and bust and the rise and fall in house prices in countries like the U.S. and Ireland, as clear evidence that investors go through periods of "irrational exhuberance" which sees asset prices become completely detached from the fundamental values suggested by reasonable applications of the dividend-discount model.

Indeed, the inability to reconcile aggregate stock price movements with rational expectations is not the only well-known failure of modern financial economics. For instance, there are many studies documenting the failure of rational optimisation-based models to explain various cross-sectional patterns in asset returns, e.g. why the average return on stocks exceeds that on bonds by so much, or discrepancies in the long-run performance of small- and large-capitalisation stocks. Eugene Fama is the author of a number of famous papers with Kenneth French that have demonstrated these discrepancies though he interprets these results as most likely due to a rational pricing of the risk associated with certain kinds of assets.

For many, the answers to these questions lie in abandoning the pure rational expectations, optimising approach. Indeed, the field of *behavioural finance* is booming, with various researchers proposing all sorts of different non-optimising models of what determines asset prices. That said, at present, there is no clear front-runner "alternative" behavioural-finance model of the determination of aggregate stock prices.

Appendix 1: Proof of Equation (5.54)

We start by repeating equation (5.53):

$$
\begin{aligned}
P_{t+1} - P_t = \quad = \quad & -\left(\frac{1}{1+r}\right) D_t + \left[\left(\frac{1}{1+r}\right) D_{t+1} - \left(\frac{1}{1+r}\right)^2 E_t D_{t+1}\right] \\
& + \left[\left(\frac{1}{1+r}\right)^2 E_{t+1} D_{t+2} - \left(\frac{1}{1+r}\right)^3 E_t D_{t+2}\right] + \\
& + \left[\left(\frac{1}{1+r}\right)^3 E_{t+1} D_{t+3} - \left(\frac{1}{1+r}\right)^4 E_t D_{t+3}\right] + ...
\end{aligned}
$$

This shows that the change in stock prices is determined by a term relating to this period's dividend's "dropping out" and then a whole bunch of terms that involve period $t+1$ and period t expectations of future dividends. To be able to pull all the terms for each D_{t+k} together, we both add and subtract a set of terms of the form $\left(\frac{1}{1+r}\right)^k E_t D_{t+k}$. The equation then looks like this

$$
\begin{aligned}
P_{t+1} - P_t = \quad = \quad & \left(\frac{1}{1+r}\right) [D_{t+1} - E_t D_{t+1}] \\
& + \left(\frac{1}{1+r}\right)^2 [E_{t+1} D_{t+2} - E_t D_{t+2}] \\
& + \left(\frac{1}{1+r}\right)^3 [E_{t+1} D_{t+3} - E_t D_{t+3}] + ... \\
& - \left(\frac{1}{1+r}\right) D_t \\
& + \left(1 - \frac{1}{1+r}\right) \left(\frac{1}{1+r}\right) E_t D_{t+1} \\
& + \left(1 - \frac{1}{1+r}\right) \left(\frac{1}{1+r}\right)^2 E_t D_{t+2} \\
& + \left(1 - \frac{1}{1+r}\right) \left(\frac{1}{1+r}\right)^3 E_t D_{t+3} + ...
\end{aligned} \tag{5.76}
$$

The sequence summarised on the first three lines of equation (5.76) can be described using a summation sign as

$$
\sum_{k=1}^{\infty} \left[\left(\frac{1}{1+r}\right)^k (E_{t+1} D_{t+k} - E_t D_{t+k})\right] \tag{5.77}
$$

This is an infinite discounted sum of changes to people's expectations about future dividends.

The sequence summarised on the last three lines of equation (5.76) can be simplified to be

$$\left(\frac{r}{1+r}\right)(1+r)\left(P_t-\left(\frac{1}{1+r}\right)D_t\right)=rP_t-\left(\frac{r}{1+r}\right)D_t \tag{5.78}$$

Using these two simplifications, equation (5.76) can be re-written as

$$\begin{aligned}
P_{t+1}-P_t &= -\left(\frac{r}{1+r}\right)D_t-\left(\frac{1}{1+r}\right)D_t+rP_t+\sum_{k=1}^{\infty}\left[\left(\frac{1}{1+r}\right)^k\left(E_{t+1}D_{t+k}-E_tD_{t+k}\right)\right] \\
&= -D_t+rP_t+\sum_{k=1}^{\infty}\left[\left(\frac{1}{1+r}\right)^k\left(E_{t+1}D_{t+k}-E_tD_{t+k}\right)\right] \qquad (5.79)
\end{aligned}$$

which is the equation we were looking for: So the return on stocks can be written as

$$r_{t+1}=\frac{D_t+\Delta P_{t+1}}{P_t}=r+\frac{\sum_{k=1}^{\infty}\left(\frac{1}{1+r}\right)^k\left(E_{t+1}D_{t+k}-E_tD_{t+k}\right)}{P_t} \tag{5.80}$$

Appendix 2: Programme For Return Predictability Results

Below is the text of a programme to generate the return predictability results reported in Table 1. The programme is written for the econometric package RATS but a programme of this sort could be written for any package that has a random number generator.

```
allocate  10000

set y = 0
set tstats_1lag = 0
set tstats_20lag = 0
set tstats_50lag = 0

do k = 1,10000

set y 2 300 = 0.99*y{1} + %ran(1)
set dy   = y - y{1}
set dy20 = y - y{20}
set dy50 = y - y{50}

linreg(noprint) dy 101 300
# y{1}
comp tstats_1lag(k) = %tstats(1)

linreg(noprint) dy20 101 300
# y{20}
comp tstats_20lag(k) = %tstats(1)

linreg(noprint) dy50 101 300
# y{50}
comp tstats_50lag(k) = %tstats(1)

end do k

stats tstats_1lag
stats tstats_20lag
stats tstats_50lag
```

Chapter 6

Consumption and Asset Pricing

Elementary Keynesian macro theory assumes that households make consumption decisions based only on their current disposable income. In reality, of course, people have to base their spending decisions not just on today's income but also on the money they expect to earn in the future. During the 1950s, important research by Ando and Modigliani (the Life-Cycle Hypothesis) and Milton Friedman (the Permanent Income Hypothesis) presented significant evidence that people plan their expenditures in a systemic way, smoothing consumption over time even when their incomes fluctuated.

In this chapter, we will use the techniques developed in the previous chapter to derive a rational expectations version of the Permanent Income Hypothesis. We will use this model to illustrate some pitfalls in using econometrics to assess the effects of policy changes. We will discuss empirical tests of this model and present some more advanced topics. In particular, we will discuss the link between consumption spending and the return on various financial assets.

The Household Budget Constraint

We start with an identity describing the evolution of the stock of assets owned by households. Letting A_t be household assets, Y_t be labour income, and C_t stand for consumption spending, this identity is

$$A_{t+1} = (1 + r_{t+1})(A_t + Y_t - C_t) \tag{6.1}$$

where r_{t+1} is the return on household assets at time $t+1$. Note that Y_t is *labour* income (income earned from working) not total income because total income also includes the capital income earned on assets (i.e. total income is $Y_t + r_{t+1}A_t$.) Note, we are assuming that Y_t is take-home labour income, so it can considered net of taxes.

As with the equation for the return on stocks, this can be written as a first-order difference equation in our standard form

$$A_t = C_t - Y_t + \frac{A_{t+1}}{1 + r_{t+1}} \tag{6.2}$$

We will assume that agents have rational expectations. Also, in this case, we will assume that the return on assets equals a constant, r. This implies

$$A_t = C_t - Y_t + \frac{1}{1+r} E_t A_{t+1} \tag{6.3}$$

Using the same repeated substitution methods as before this can be solved to give

$$A_t = \sum_{k=0}^{\infty} \frac{E_t (C_{t+k} - Y_{t+k})}{(1+r)^k} \tag{6.4}$$

Note that we have again imposed the condition that the final term in our repeated substitution $\frac{E_t A_{t+k}}{(1+r)^k}$ goes to zero as k gets large. Effectively, this means that we are assuming that people consume some of their capital income (i.e. that assets are used to finance a level of consumption C_t that is generally larger than labour income Y_t). If this is the case, then this term tends to zero.

One way to understand this equation comes from re-writing it as

$$\sum_{k=0}^{\infty} \frac{E_t C_{t+k}}{(1+r)^k} = A_t + \sum_{k=0}^{\infty} \frac{E_t Y_{t+k}}{(1+r)^k} \tag{6.5}$$

This is usually called the *intertemporal budget constraint*. It states that the present value sum of current and future household consumption must equal the current stock of financial assets plus the present value sum of current and future labour income.

A consumption function relationship can be derived from this equation by positing some theoretical relationship between the expected future consumption values, $E_t C_{t+k}$, and the current value of consumption. This is done by appealing to the optimising behaviour of the consumer.

Piketty and $r > g$

Some of you may be aware of Thomas Piketty's famous book *Capital in the Twenty First Century*. If you're not, scroll down a few pages to check him out. Perhaps Piketty's most famous conjecture is there is a natural tendency in capitalist economies for wealth to accumulate faster than income. This conjecture can be understood on the basis of the simple budget identity we are working with here.

Consider the simple version of our budget constraint with a constant return on assets

$$A_{t+1} = (1+r)(A_t + Y_t - C_t) \tag{6.6}$$

If assets grew at rate r or faster, then this would likely mean they were growing faster than GDP, because r is generally higher than GDP growth. So what is the growth rate of the stock of assets? We can calculate the change in assets as

$$A_{t+1} - A_t = rA_t + (1+r)(Y_t - C_t) \tag{6.7}$$

So the growth rate of assets is given by

$$\frac{A_{t+1} - A_t}{A_t} = r + \frac{(1+r)(Y_t - C_t)}{A_t} \tag{6.8}$$

This means the growth rate of assets equals r plus an additional term that will be positive as long as $Y_t > C_t$ i.e. as long as labour income is greater than consumption. So this tells us that the growth rate of assets equals r plus a term that depends upon whether consumption is greater than or less than labour income. If consumption is less than labour income, assets grow at a rate that is greater than r while they will grow at a rate slower than r if consumption is greater than labour income.

Piketty bases his ideas about the tendency for wealth to rise faster than income on the fact that the rate of return on assets r has tended historically to be higher than the growth rate of GDP. If we observed $Y_t > C_t$, then assets would grow at a rate greater than r and so this would generally also be higher than the growth rate of GDP. However, we should expect to see $Y_t < C_t$, i.e. consumption to be greater than labour income: If the income people earn from their assets doesn't boost their consumption spending over what they could get just from their labour income, then what is the point of it? And indeed, the data generally show that consumption is greater than labour income, so that people consume some of their capital income (i.e. their income from assets) and total assets should generally grow at a rate that is less than r. Still, Piketty points out that it is possible for people to consume some of their capital income and still have assets growing at a rate smaller than r but greater than g.[1]

Under what conditions will assets grow at a faster rate than the growth rate of GDP, which

[1] For example, page 564: "If $r > g$, it suffices to reinvest a fraction of the return on capital equal to the growth rate g and consume the rest $(r - g)$.

Piketty terms g? Our previous equation tells us this happens when

$$g < r + \frac{(1+r)(Y_t - C_t)}{A_t} \tag{6.9}$$

This can be re-arranged to give

$$\frac{C_t - Y_t}{A_t} < \frac{r-g}{(1+r)} \tag{6.10}$$

So assets will grow faster than incomes if the amount of people's capital income that they consume (i.e. the amount they consume above their labour income) as a share of total assets is below the specific value on the right-hand-side.

Is there any result in economics that leads us to believe that this last inequality should generally hold? Not to my knowledge. In this sense, Piketty perhaps overstates the extent to which, on its own, the fact that $r > g$ is a "fundamental force for divergence." What is required for assets to steadily grow relative to income is not only this condition but also an additional, relatively arbitrary, restriction on how much people can consume and this latter condition may or may not hold at various times. However, what can be said is that during periods of high returns on capital, when the gap between r and g is particularly high, then the bigger the right-hand-side of equation (6.10) will be and it is perhaps more likely that the condition above will be held.

Most likely, however, the key empirical developments that Piketty's book focuses on—rising assets relative to income and growing inequality of wealth—are being driven by other forces that are making the income distribution more unequal and reducing the share of income going to workers rather than being related to some innate "law of capitalism" that drives wealth up at faster pace than incomes.

Figure 6.1: Thomas Piketty

Optimising Behaviour by the Consumer

We will assume that consumers wish to maximize a welfare function of the form

$$W = \sum_{k=0}^{\infty} \left(\frac{1}{1+\beta} \right)^k U\left(C_{t+k}\right) \tag{6.11}$$

where $U(C_t)$ is the instantaneous utility obtained at time t, and β is a positive number that describes the fact that households prefer a unit of consumption today to a unit tomorrow. Economists call this "discounting future utility." If the future path of labour income is known, consumers who want to maximize this welfare function subject to the constraints imposed by the intertemporal budget constraint must solve the following Lagrangian problem:

$$L\left(C_t, C_{t+1}, ...\right) = \sum_{k=0}^{\infty} \left(\frac{1}{1+\beta} \right)^k U\left(C_{t+k}\right) + \lambda \left[A_t + \sum_{k=0}^{\infty} \frac{Y_{t+k}}{(1+r)^k} - \sum_{k=0}^{\infty} \frac{C_{t+k}}{(1+r)^k} \right] \tag{6.12}$$

For every current and future value of consumption, C_{t+k}, this yields a first-order condition of the form

$$\left(\frac{1}{1+\beta} \right)^k U'\left(C_{t+k}\right) - \frac{\lambda}{(1+r)^k} = 0 \tag{6.13}$$

For $k = 0$, this implies

$$U'\left(C_t\right) = \lambda \tag{6.14}$$

For $k = 1$, it implies

$$U'\left(C_{t+1}\right) = \left(\frac{1+\beta}{1+r} \right) \lambda \tag{6.15}$$

Putting these two equations together, we get the following relationship between consumption today and consumption tomorrow:

$$U'\left(C_t\right) = \left(\frac{1+r}{1+\beta} \right) U'\left(C_{t+1}\right) \tag{6.16}$$

When there is uncertainty about future labour income, this optimality condition can just be re-written as

$$U'\left(C_t\right) = \left(\frac{1+r}{1+\beta} \right) E_t\left[U'\left(C_{t+1}\right)\right] \tag{6.17}$$

This implication of the first-order conditions for consumption is sometimes known as an *Euler equation*.

In an important 1978 paper, Robert Hall proposed a specific case of this equation.[2] Hall's special case assumed that

$$U(C_t) = aC_t + bC_t^2 \tag{6.18}$$

$$r = \beta \tag{6.19}$$

In other words, Hall assumed that the utility function was quadratic and that the real interest rate equalled the household discount rate. In this case, the Euler equation becomes

$$a + 2bC_t = E_t\left[a + 2bC_{t+1}\right] \tag{6.20}$$

which simplifies to

$$C_t = E_t C_{t+1} \tag{6.21}$$

This states that the optimal solution involves next period's expected value of consumption equalling the current value. Because, the Euler equation holds for all time periods, we have

$$E_t C_{t+k} = E_t C_{t+k+1} \qquad k = 1, 2, 3, ... \tag{6.22}$$

So, we can apply repeated iteration to get

$$C_t = E_t\left(C_{t+k}\right) \qquad k = 1, 2, 3, ... \tag{6.23}$$

In other words, all future expected values of consumption equal the current value. Because it implies that changes in consumption are unpredictable, this is sometimes called the *random walk* theory of consumption.

[2] "Stochastic Implications of the Life-Cycle Permanent Income Hypothesis: Theory and Evidence," *Journal of Political Economy*, December 1978.

The Rational Expectations Permanent Income Hypothesis

Hall's random walk hypothesis has attracted a lot of attention in its own right, but rather than focus on what should be unpredictable (changes in consumption), we are interested in deriving an explicit formula for what consumption should equal.

To do this, insert $E_t C_{t+k} = C_t$ into the intertemporal budget constraint, (6.5), to get

$$\sum_{k=0}^{\infty} \frac{C_t}{(1+r)^k} = A_t + \sum_{k=0}^{\infty} \frac{E_t Y_{t+k}}{(1+r)^k} \tag{6.24}$$

Now we can use the geometric sum formula to turn this into a more intuitive formulation:

$$\sum_{k=0}^{\infty} \frac{1}{(1+r)^k} = \frac{1}{1-\frac{1}{1+r}} = \frac{1+r}{r} \tag{6.25}$$

So, Hall's assumptions imply the following equation, which we will term the *Rational Expectations Permanent Income Hypothesis*:

$$C_t = \frac{r}{1+r} A_t + \frac{r}{1+r} \sum_{k=0}^{\infty} \frac{E_t Y_{t+k}}{(1+r)^k} \tag{6.26}$$

This equation is a rational expectations version of the well-known *permanent income hypothesis* (I will use the term RE-PIH below) which states that consumption today depends on a person's expected lifetime sequence of income.

Let's look at this equation closely. It states that the current value of consumption is driven by three factors:

- The expected present discounted sum of current and future labour income.
- The current value of household assets. This "wealth effect" is likely to be an important channel through which financial markets affect the macroeconomy.

- The expected return on assets: This determines the coefficient, $\frac{r}{1+r}$, that multiplies both assets and the expected present value of labour income. In this model, an increase in this expected return raises this coefficient, and thus boosts consumption.

A Concrete Example: Constant Expected Growth in Labour Income

This RE-PIH model can be made more concrete by making specific assumptions about expectations concerning future growth in labour income. Suppose, for instance, that households expect labour income to grow at a constant rate g in the future:

$$E_t Y_{t+k} = (1+g)^k Y_t \tag{6.27}$$

This implies

$$C_t = \frac{r}{1+r} A_t + \frac{rY_t}{1+r} \sum_{k=0}^{\infty} \left(\frac{1+g}{1+r}\right)^k \tag{6.28}$$

As long as $g < r$ (and we will assume it is) then we can use the geometric sum formula to simplify this expression

$$\sum_{k=0}^{\infty} \left(\frac{1+g}{1+r}\right)^k = \frac{1}{1-\frac{1+g}{1+r}} \tag{6.29}$$

$$= \frac{1+r}{r-g} \tag{6.30}$$

This implies a consumption function of the form

$$C_t = \frac{r}{1+r} A_t + \frac{r}{r-g} Y_t \tag{6.31}$$

Note that the higher is expected future growth in labour income g, the larger is the coefficient on today's labour income and thus the higher is consumption.

The Lucas Critique

The fact that the coefficients of so-called *reduced-form* relationships, such as the consumption function equation (6.31), depend on expectations about the future is an important theme in modern macroeconomics. In particular, in a famous 1976 paper, rational expectations pioneer Robert Lucas pointed out that the assumption of rational expectations implied that these coefficients would change if expectations about the future changed.[3] In our example, the marginal propensity to consume from current income will change if expectations about future growth in labour income change.

Lucas's paper focused on potential problems in using econometrically-estimated regressions to assess the impact of policy changes. He pointed out that changes in policy may change expectations about future values of important variables, and that these changes in expectations may change the coefficients of econometric relationships. This type of problem can limit the usefulness for policy analysis of reduced-form econometric models based on historical data. This problem is now known as the *Lucas critique* of econometric models.

To give a specific example, suppose the government is thinking of introducing a temporary tax cut on labour income. As noted above, we can consider Y_t to be after-tax labour income, so it would be temporarily boosted by the tax cut. Now suppose the policy-maker wants an estimate of the likely effect on consumption of the tax cut. They may get their economic advisers to run a regression of consumption on assets and after-tax labour income. If, in the past, consumers had generally expected income growth of g, then the econometric regressions will report a coefficient of approximately $\frac{r}{r-g}$ on labour income. So, the economic adviser might conclude that for each extra dollar of labour income produced by the tax cut, there will be an increase in consumption of $\frac{r}{r-g}$ dollars.

However, if households have rational expectations and operate according to equation (6.26)

[3] Robert Lucas, "Econometric Policy Evaluation: A Critique," *Carnegie-Rochester Series on Public Policy*, Vol. 1, pages 19-46, 1976.

then the true effect of the tax cut could be a lot smaller. For instance, if the tax cut is only expected to boost this period's income, and to disappear tomorrow, then each dollar of tax cut will produce only $\frac{r}{1+r}$ dollars of extra consumption. The difference between the true effect and the economic advisor's supposedly "scientific" regression-based forecast could be substantial. For instance, plugging in some numbers, suppose $r = 0.06$ and $g = 0.02$. In this case, the economic advisor concludes that the effect of a dollar of tax cuts is an extra 1.5 $(=\frac{.06}{.06-.02})$ dollars of consumption. In reality, the tax cut will produce only an extra 0.057 $(=\frac{.06}{1.06})$ dollars of extra consumption. This is a big difference.

The Lucas critique played an important role in the increased popularity of rational expectations economics. Examples like this one show the benefit in using a formulation such as equation (6.26) that explicitly takes expectations into account, instead of relying only on econometric regressions.

Implications for Fiscal Policy: Ricardian Equivalence

Like households, governments also have budget constraints. Here we consider the implications of these constraints for consumption spending in the Rational Expectations Permanent Income Hypothesis. First, let us re-formulate the household budget constraint to explicitly incorporate taxes. Specifically, let's write the period-by-period constraint as

$$A_{t+1} = (1+r)(A_t + Y_t - T_t - C_t) \tag{6.32}$$

where T_t is the total amount of taxes paid by households. Taking the same steps as before, we can re-write the intertemporal budget constraint as

$$\sum_{k=0}^{\infty} \frac{E_t C_{t+k}}{(1+r)^k} = A_t + \sum_{k=0}^{\infty} \frac{E_t (Y_{t+k} - T_{t+k})}{(1+r)^k} \tag{6.33}$$

Now let's think about the government's budget constraint. The stock of public debt, D_t evolves over time according to

$$D_{t+1} = (1+r)(D_t + G_t - T_t) \tag{6.34}$$

where G_t is government spending and T_t is tax revenue. Applying the repeated-substitution method we can obtain an intertemporal version of the government's budget constraint.

$$\sum_{k=0}^{\infty} \frac{E_t T_{t+k}}{(1+r)^k} = D_t + \sum_{k=0}^{\infty} \frac{E_t G_{t+k}}{(1+r)^k} \tag{6.35}$$

This states that the present discounted value of tax revenue must equal the current level of debt plus the present discounted value of government spending. In other words, in the long-run, the government must raise enough tax revenue to pay off its current debts as well as its current and future spending.

Consider the implications of this result for household decisions. If households have rational expectations, then they will understand that the government's intertemporal budget constraint, equation (6.35), pins down the present value of tax revenue. In this case, we can substitute the right-hand-side of (6.35) into the household budget constraint to replace the present value of tax revenue. Doing this, the household budget constraint becomes

$$\sum_{k=0}^{\infty} \frac{E_t C_{t+k}}{(1+r)^k} = A_t - D_t + \sum_{k=0}^{\infty} \frac{E_t (Y_{t+k} - G_{t+k})}{(1+r)^k} \tag{6.36}$$

Consider now the implications of this result for the impact of a temporary cut in taxes. Before, we had discussed how a temporary cut in taxes should have a small effect. This equation gives us an even more extreme result — unless governments plan to change the profile of government spending, then a cut to taxes today has no impact at all on consumption spending. This is because households anticipate that lower taxes today will just trigger higher taxes tomorrow.

This result – that rational expectations implied that a deficit-financed cut in taxes should have no impact on consumption – was first presented by Robert Barro in a famous 1974 paper.[4] It was later pointed out that some form of this result was alluded to in David Ricardo's writings in the nineteenth century. Economists love fancy names for things, so the result is now often referred to as *Ricardian equivalence*.

Evidence on the RE-PIH

There have been lots of macroeconomic studies on how well the RE-PIH fits the data. One problem worth noting is that there are some important measurement issues when attempting to test the theory. In particular, the model's assumption that consumption expenditures only yield a positive utility flow in the period in which the money is spent clearly does not apply to durable goods, such as cars or computers, which yield a steady flow of utility. For this reason, most empirical research has focused only on spending on nondurables (e.g. food) and services, with a separate literature focusing on spending on consumer durables.

There are various reasons why the RE-PIH may not hold. Firstly, it assumes that it is always feasible for households to "smooth" consumption in the manner predicted by the theory. For example, even if you anticipate earning lots of money in the future and would like to have a high level of consumption now, you may not be able to find a bank to fund a lavish lifestyle now based on your promises of future millions. These kinds of "liquidity constraints" may make consumption spending more sensitive to their current incomes than the RE-PIH predicts. Secondly, people may not have rational expectations and may not plan their spending decisions in the calculating optimising fashion assumed by the theory.

[4] Robert Barro (1974). "Are Government Bonds Net Wealth?" *Journal of Political Economy*, Volume 82(6).

Following Hall's 1978 paper, the 1980s saw a large amount of research on whether the RE-PIH fitted the data. The most common conclusion was that consumption was "excessively sensitive" to disposable income. In particular, changes in consumption appear to be more forecastable than they should be if Hall's random walk idea was correct. Campbell and Mankiw (1990) is a well-known paper that provides a pretty good summary of these conclusions.[5] They present a model in which a fraction of the households behave according to the RE-PIH while the rest simply consume all of their current income. They estimate the fraction of non-PIH consumers to be about a half. A common interpretation of this result is that liquidity constraints have an important impact on aggregate consumption. (A byproduct of this conclusion would be that financial sector reforms that boost access to credit could have an important impact on consumption spending.)

Evidence on Ricardian Equivalence

There is also a large literature devoted to testing the Ricardian equivalence hypothesis. In addition to the various reasons the RE-PIH itself may fail, there are various other reasons why Ricardian equivalence may not hold. Some are technical points. People don't actually live forever (as we had assumed in the model) and so they may not worry about future tax increases that could occur after they have passed away; taxes take a more complicated form than the simple lump-sum payments presented above; the interest rate in the government's budget constraint may not be the same as the interest rate in the household's constraint. (You can probably think of a few more.) More substantively, people may often be unable to tell whether tax changes are temporary or permanent. Most of the macro studies on this

[5] John Campbell and Gregory Mankiw (1990). "Permanent Income, Current Income, and Consumption," *Journal of Business and Economic Statistics*

topic (in particular those that use Vector Autoregressions) tend to find the effects of fiscal policy are quite different from the Ricardian equivalence predictions. Tax cuts and increases in government spending tend to boost the economy.

Perhaps the most interesting research on this area has been the use of micro data to examine the effect of changes in taxes that are explicitly predictable and temporary. One example is the paper by Parker, Souleles, Johnson and McClelland which examines the effect of tax rebates provided to U.S. taxpayers in 2008.[6] This programme saw the U.S. government send once-off payments to consumers in an attempt to stimulate the economy. Since these payments were being financed by expanding the government deficit, Ricardian equivalence predicts that consumers should not have responded. Parker et al, however, found the opposite using data from the Consumer Expenditure Survey. A quick summary:

> We find that, on average, households spent about 12-30% (depending on the specification) of their stimulus payments on nondurable expenditures during the three-month period in which the payments were received. Further, there was also a substantial and significant increase in spending on durable goods, in particular vehicles, bringing the average total spending response to about 50-90% of the payments.

You might suspect that these results are driven largely by liquidity constraints but the various microeconomic studies that have examined temporary fiscal policy changes have not always been consistent with this idea. For example, research by Parker (1999) showed the even relatively high-income consumers seemed to spend more in response to transitory changes in

[6] "Consumer Spending and the Economic Stimulus Payments of 2008." *American Economic Review*, 103(6), October 2013.

their social security taxes (which stop at a certain point in the year when workers reach a maximum threshold point) while Souleles (1999) found "excess sensitivity" results for consumer spending after people received tax rebate cheques.[7] These results show excess sensitivity even among groups of consumers that are unlikely to be liquidity constrained.

At the same time, this doesn't mean that households go on a splurge every time they get a large payment. For example, Hsieh (2003) examines how people in Alaska responded to large anticipated annual payments that they received from a state fund that depends largely on oil revenues.[8] Unlike the evidence on temporary tax cuts, Hsieh finds that Alaskan households respond to these payments in line with the predictions of the Permanent Income Hypothesis, smoothing out their consumption over the year. One possible explanation is that these large and predictable payments are easier for people to understand and plan around and the consequences of spending them too quickly are more serious than smaller once-off federal tax changes. There is clearly room for more research in this important area.

Precautionary Savings

I want to return to a subtle point that was skipped over earlier. If we keep the assumption $r = \beta$, then the consumption Euler equation is

$$U'(C_t) = E_t\left[U'(C_{t+1})\right] \tag{6.37}$$

You might think that this equation is enough to deliver the property of constant expected consumption. We generally assume declining marginal utility, so function U' is monotonically decreasing. In this case, surely the expectation of next period's marginal utility being the

[7] Jonathan Parker. "The Reaction of Household Consumption to Predictable Changes in Social Security Taxes," *American Economic Review*, Vol 89 No 4, September 1999. Nicholes Souleles. " The Response of Household Consumption to Income Tax Refunds," *American Economic Review*, Vol 89 No 4, September 1999

[8] Chang-Tai Hsieh. "Do Consumers React to Anticipated Income Changes? Evidence from the Alaska Permanent Fund" *American Economic Review*, March 2003.

same as this period's is the same as next period's expected consumption level being the same as this period's.

The problem with this thinking is the E_t here is a mathematical expectation, i.e. a weighted average over a set of possible outcomes. And for most functions F generally $E(F(X)) \neq F(E(X))$. In particular, if F is a concave function—meaning it has a negative second derivative—a famous result known as Jensen's inequality states that $E(F(X)) < F(E(X))$. This underlies the mathematical formulation of why people are averse to risk: The average utility expected from an uncertain level of consumption is less than from the "sure thing" associated with obtaining the average level of consumption. The sign of the Jensen's inequality result is reversed for convex functions, i.e. those with a positive second derivative.

In this example, we are looking at the properties of $E_t\left[U'\left(C_{t+1}\right)\right]$. Whether or not marginal utility is concave or convex depends on its second derivative, so it depends upon the third derivative of the utility function U'''. Most standard utility functions have positive third derivatives implying convex marginal utility and thus $E_t\left[U'\left(C_{t+1}\right)\right] > U'\left(E_t C_{t+1}\right)$. What we can see now is why the quadratic utility function was such a special case. Because this function has $U''' = 0$, its marginal utility is neither concave or convex and the Jensen relationship is an equality. So, in this very particular case, the utility function displays *certainty equivalence*: The uncertain outcome is treated the same way is if people were certain of achieving the average value of consumption.

Here's a specific example of when certainty equivalence doesn't hold.[9] Suppose consumers have a utility function of the form

$$U(C_t) = -\frac{1}{\alpha}\exp\left(-\alpha C_t\right) \tag{6.38}$$

[9] This particular example was first presented by Ricardo Caballero (1990), "Consumption Puzzles and Precautionary Savings" *Journal of Monetary Economics*, Volume 25, pages 113-136.

where exp is the exponential function. This implies marginal utility of the form

$$U'(C_t) = \exp(-\alpha C_t) \tag{6.39}$$

In this case, the Euler equation becomes

$$\exp(-\alpha C_t) = E_t(\exp(-\alpha C_{t+1})) \tag{6.40}$$

Now suppose the uncertainty about C_{t+1} is such that it is perceived to have a normal distribution with mean $E_t(C_{t+1})$ and variance σ^2. A useful result from statistics is that if a variable X is normally distributed has mean μ and variance σ^2:

$$X \sim N\left(\mu, \sigma^2\right) \tag{6.41}$$

then it can be shown that

$$E(\exp(X)) = \exp\left(\mu + \frac{\sigma^2}{2}\right) \tag{6.42}$$

In our case, this result implies that

$$\begin{aligned} E_t(\exp(-\alpha C_{t+1})) &= \exp\left(E_t(-\alpha C_{t+1}) + \frac{Var(-\alpha C_{t+1})}{2}\right) && (6.43) \\ &= \exp\left(-\alpha E_t(C_{t+1}) + \frac{\alpha^2\sigma^2}{2}\right) && (6.44) \end{aligned}$$

So, the Euler equation can be written as

$$\exp(-\alpha C_t) = \exp\left(-\alpha E_t(C_{t+1}) + \frac{\alpha^2\sigma^2}{2}\right) \tag{6.45}$$

Taking logs of both sides this becomes

$$-\alpha C_t = -\alpha E_t(C_{t+1}) + \frac{\alpha^2\sigma^2}{2} \tag{6.46}$$

which simplifies to

$$E_t(C_{t+1}) = C_t + \frac{\alpha\sigma^2}{2} \tag{6.47}$$

Even though expected marginal utility is flat, consumption tomorrow is expected to be higher than consumption today. Thus, uncertainty induces an "upward tilt" to the consumption profile. And this upward tilt has an affect on today's consumption: We cannot sustain higher consumption tomorrow without having lower consumption today.

Indeed, it turns out that this result allows us to calculate exactly what the effect of uncertainty is on consumption today. The Euler equation implies that

$$E_t\left(C_{t+k}\right) = C_t + \frac{k\alpha\sigma^2}{2} \tag{6.48}$$

Inserting this into the intertemporal budget constraint, we get

$$\sum_{k=0}^{\infty} \frac{C_t}{(1+r)^k} + \frac{\alpha\sigma^2}{2} \sum_{k=1}^{\infty} \frac{k}{(1+r)^k} = A_t + \sum_{k=0}^{\infty} \frac{E_t Y_{t+k}}{(1+r)^k} \tag{6.49}$$

It can be shown (mainly by repeatedly using the well-known geometric sum formula) that

$$\sum_{k=1}^{\infty} \frac{k}{(1+r)^k} = \frac{1+r}{r^2} \tag{6.50}$$

So, the intertemporal budget constraint simplifies to

$$\sum_{k=0}^{\infty} \frac{C_t}{(1+r)^k} + \frac{1+r}{r^2}\frac{\alpha\sigma^2}{2} = A_t + \sum_{k=0}^{\infty} \frac{E_t Y_{t+k}}{(1+r)^k} \tag{6.51}$$

and taking the same steps as before, consumption today is

$$C_t = \frac{r}{1+r} A_t + \frac{r}{1+r} \sum_{k=0}^{\infty} \frac{E_t Y_{t+k}}{(1+r)^k} - \frac{\alpha\sigma^2}{2r} \tag{6.52}$$

This is exactly as before apart from an additional "precautionary savings" term $-\frac{\alpha\sigma^2}{2r}$. The more uncertainty there is, the more lower the current level of consumption will be.

This particular result obviously relies on very specific assumptions about the form of the utility function and the distribution of uncertain outcomes. However, since almost all utility function feature positive third derivatives, the key property underlying the precautionary savings result—that higher uncertainty reduces consumption—will generally hold. It is an important result because some of the more important changes in the savings rate observed over time appear consistent with this type of precautionary savings behaviour. So, for example, during the global financial crisis of 2008, when there was so much uncertainty about how long the recession would last and what impact it would have, it is very likely that this greater uncertainty depressed consumption.

Incorporating Time-Varying Asset Returns

One simplification that we have made up to now is that consumers expect a constant return on assets. Here, we allow expected asset returns to vary. The first thing to note here is that one can still obtain an intertemporal budget constraint via the repeated substitution method. This now takes the form

$$\sum_{k=0}^{\infty} \frac{E_t C_{t+k}}{\left(\prod_{m=1}^{k+1} (1+r_{t+m})\right)} = A_t + \sum_{k=0}^{\infty} \frac{E_t Y_{t+k}}{\left(\prod_{m=1}^{k+1} (1+r_{t+m})\right)} \tag{6.53}$$

where $\prod_{n=1}^{h} x_i$ means the product of x_1, x_2 .... x_h. The steps to derive this are identical to the steps used to derive equation the dividend-discount model of asset prices with time-varying rates of return in the previous chapter.

The optimisation problem of the consumer does not change much. This problem now has the Lagrangian

$$L\left(C_t, C_{t+1}, ...\right) = \sum_{k=0}^{\infty} \left(\frac{1}{1+\beta}\right)^k U\left(C_{t+k}\right) + \lambda \left[A_t + \sum_{k=0}^{\infty} \frac{E_t Y_{t+k}}{\left(\prod_{m=1}^{k+1} (1+r_{t+m})\right)} - \sum_{k=0}^{\infty} \frac{E_t C_{t+k}}{\left(\prod_{m=1}^{k+1} (1+r_{t+m})\right)}\right] \tag{6.54}$$

And instead of the simple Euler equation (6.17), we get

$$U'(C_t) = E_t\left[\left(\frac{1+r_{t+1}}{1+\beta}\right)U'(C_{t+1})\right] \tag{6.55}$$

or, letting

$$R_t = 1 + r_t \tag{6.56}$$

we can re-write this as

$$U'(C_t) = E_t\left[\left(\frac{R_{t+1}}{1+\beta}\right)U'(C_{t+1})\right] \tag{6.57}$$

Consumption and Rates of Return on Assets

Previously, we had used an equation like this to derive the behaviour of consumption, given an assumption about the determination of asset returns. However, Euler equations have taken on a double role in modern economics because they are also used to consider the determination of asset returns, taking the path of consumption as given. The Euler equation also takes on greater importance than it might seem based on our relatively simple calculations because, once one extends the model to allow the consumer to allocate their wealth across multiple asset types, it turns out that equation (6.57) must hold for *all* of these assets. This means that for a set of different asset returns $R_{i,t}$, we must have

$$U'(C_t) = E_t\left[\left(\frac{R_{i,t+1}}{1+\beta}\right)U'(C_{t+1})\right] \tag{6.58}$$

for each of the assets.

So, for example, consider a pure risk-free asset that pays a guaranteed rate of return next period. The nearest example in the real-world is a short-term US treasury bill. Because there is no uncertainty about this rate of return, call it $R_{f,t}$, these terms can be taken outside the expectation term, and the Euler equation becomes

$$U'(C_t) = \frac{R_{f,t+1}}{1+\beta} E_t \left[U'(C_{t+1}) \right] \tag{6.59}$$

So, the risk-free interest rate should be determined as

$$R_{f,t+1} = \frac{(1+\beta) U'(C_t)}{E_t \left[U'(C_{t+1}) \right]} \tag{6.60}$$

To think about the relationship between risk-free rates and returns on other assets, it is useful to use a well-known result from statistical theory, namely

$$E(XY) = E(X)E(Y) + Cov(X, Y) \tag{6.61}$$

The expectation of a product of two variables equals the product of the expectations plus the covariance between the two variables. This allows one to re-write (6.58) as

$$U'(C_t) = \frac{1}{1+\beta} \left[E_t (R_{i,t+1}) E_t (U'(C_{t+1})) + Cov (R_{i,t+1}, U'(C_{t+1})) \right] \tag{6.62}$$

This can be re-arranged to give

$$\frac{(1+\beta) U'(C_t)}{E_t \left[U'(C_{t+1}) \right]} = E_t (R_{i,t+1}) + \frac{Cov (R_{i,t+1}, U'(C_{t+1}))}{E_t \left[U'(C_{t+1}) \right]} \tag{6.63}$$

Note now that, by equation (6.68), the left-hand-side of this equation equals the risk-free rate. So, we have

$$E_t (R_{i,t+1}) = R_{f,t+1} - \frac{Cov (R_{i,t+1}, U'(C_{t+1}))}{E_t \left[U'(C_{t+1}) \right]} \tag{6.64}$$

This equation tells us that expected rate of return on risky assets equals the risk-free rate *minus* a term that depends on the covariance of the risky return with the marginal utility of consumption. This equation is known as the *Consumption Capital Asset Pricing Model* or Consumption CAPM, and it plays an important role in modern finance. Most asset returns depend on payments generated by the real economy and so they are procyclical—they are better in expansions than during recessions. However, the usual assumption of diminishing marginal utility implies that U' depends negatively on consumption. This means that the covariance term is negative for assets whose returns are positively correlated with consumption and these assets will have a higher rate of return than the risk free rate. Indeed, the higher the correlation of the asset return with consumption, the higher will be the expected return.

Underlying this behaviour is the fact that consumers would like to use assets to hedge against consumption variations. Given two assets that have the same rate of return, a risk-averse consumer would prefer to have one that was negatively correlated with consumption than one that is positively correlated with consumption. For investors to be induced into holding both assets, the rate of return on the asset with a positive correlation with consumption needs to be higher.

Puzzles: Equity Premium and Risk-Free Rate

In theory, the consumption CAPM should be able to explain to us why some assets, such as stocks, tend to have such high returns while others, such as government bonds, have such low returns. However, it turns out that it has great difficulty in doing so. In the US, the average real return on stocks over the long run has been about six percent per year while the average return on Treasury bonds has been about one percent per year. In theory, this could be explained by the positive correlation between stock returns and consumption. In practice, this is not so easy. Most studies use simple utility functions such as the Constant Relative Risk Aversion (CRRA) preferences

$$U(C_t) = \frac{1}{1-\theta} C_t^{1-\theta} \tag{6.65}$$

so marginal utility is

$$U'(C_t) = C_t^{-\theta} \tag{6.66}$$

In this case, the consumption-CAPM equation becomes

$$E_t\left(R_{i,t+1}\right) = R_{f,t+1} - \frac{Cov\left(R_{i,t+1}, C_{t+1}^{-\theta}\right)}{E_t\left[C_{t+1}^{-\theta}\right]} \tag{6.67}$$

For values of θ considered consistent with standard estimates of risk aversion, this covariance on the right-hand side is not nearly big enough to justify the observed equity premium. It requires values such as $\theta = 25$, which turns out to imply people are incredibly risk averse: For instance, it implies they are indifferent between a certain 17 percent decline in consumption and 50-50 risk of either no decline or a 20 percent decline. One way to explain this finding is as follows. In practice, consumption tends to be quite smooth over the business cycle (our earlier model helps to explain why) so for standard values of θ, marginal utility doesn't change that much over the cycle and one doesn't need to worry too much equities being procyclical. However, if θ is very very high, then the gap between marginal utility in booms and recessions is much bigger: Marginal utility is really high in recessions and consumers really want an asset that pays off then. This leads to a high equity premium.

One route that doesn't seem to work is arguing that people really are that risk averse, i.e. that $\theta = 25$ somehow is a good value. The reason for this is that this value of θ would imply a much higher risk-free rate than we actually see. Plugging the CRRA utility function into the equation for the risk free rate

$$R_{f,t+1} = \frac{(1+\beta)\, C_t^{-\theta}}{E_t\left[C_{t+1}^{-\theta}\right]} \tag{6.68}$$

Neglecting uncertainty about consumption growth, this formula implies that on average, the risk-free rate should be

$$R_f = (1+\beta)\,(1+g_C)^{\theta} \tag{6.69}$$

where g_C is the growth rate of consumption. Plugging in the average growth rate of consumption, a value of $\theta = 25$ would imply a far higher risk-free rate than we actually see on government bonds.

There is now a very large literature dedicated to solving the equity premium and risk-free rate puzzles, but as of yet there is no agreed best solution.[10]

[10]The paper that started this whole literature is Rajnish Mehra and Edward Prescott, "The Equity Premium: A Puzzle" *Journal of Monetary Economics*, 15, 145-161. For a review, see Narayana Kocherlakota, "The Equity Premium: It's Still a Puzzle" *Journal of Economic Literature*, 34, 42-71.

Chapter 7

Exchange Rates, Interest Rates and Expectations

Our next example of the role of expectations in macroeconomics is an important one: The link between interest rates and exchange rates and the behaviour of flexible exchange rates.

Why Exchange Rates Matter

Why do exchange rates matter? Consider the Euro-Pound exchange rate, so that 1 = £X. Now suppose X goes up, so the Euro is worth more relative to the pound. What will happen to exports from the euro area to the UK and imports to the euro area from the UK?

1. *Exports*: For each pound in sterling revenues that an Irish firm earns, they now get less revenue in euros unless they increase their UK price. Because most of their costs (in particular wages) will be denominated in euros, this means that exporting will become less profitable at prevailing prices. Euro area firms may react to this by increasing the price they charge in the UK: This will reduce demand for their product, so exports will still decline. Alternatively, some firms that feel they cannot raises prices to restore profitability may simply exit from exporting. Between these two mechanisms, an increase in the value of the euro relative to the pound will reduce euro area exports to the UK.

2. *Imports*: Because the value of the euro has increased, UK firms will get more sterling revenues from exporting to the euro area at the same prices, so UK firms that hadn't previously been exporting to the euro area may start to do so. Alternatively, UK firms already exporting to the euro area may decide to lower their euro-denominated prices in Ireland and increase their market share while still getting the same sterling revenue per unit. Either way, imports to the euro area from the UK will increase.

So while an increase in the value of a country's currency may sound like a good thing, it tends to reduce exports, increase imports, and thus reduce the country's real GDP. In contrast, a depreciation of the currency boosts exports and has a positive effect on economic growth. For these reasons, a depreciation of the currency is often welcome in a recession and the absence of this tool when the exchange rate is fixed is often pointed to as a downside of such regimes.

That said, exchange rate depreciation has its downsides also:

1. *Inflation*: Depreciation tends to make imports more expensive and so add to inflation. This is one reason why central bankers tend to say they favour a strong currency—they are indicating their preference for low inflation. For small open economies that import a lot, the inflationary effects of depreciation are much bigger.

2. *Temporary Boost*: The boost to growth from a devaluation is often temporary. Over time, the increase in import prices may feed through to higher wages and this gradually erodes the competitive benefits from devaluation. The more open an economy is, the stronger the subsequent erosion of the competitive improvement.

Free Movement of Capital: Uncovered Interest Parity

Consider the case where there is free mobility of capital: In other words, people can move money from one country to another immediately and without incurring any fees or taxes. Specifically, consider the case where money can flow easily between the US and the Euro area.

Suppose now that investors can buy either US or European risk-free one-period bonds. European bonds have an interest rate of i_t^E and US bonds have an interest rate of i_t^{US}. Let e_t represent the amount of dollars that can be obtained in exchange for one Euro at time t.

Now let's think about about the return to a US investor who wants to invest \$1 in a Euro-denominated bond at time t and then convert the money back into dollars at time $t+1$. They do this as follows. First, they exchange their \$1 for for $\frac{1}{e_t}$ and use this money to buy a European bond worth. The bond pays an interest rate of i_t^E and then next period the US investor exchanges their $\frac{1+i_t^E}{e_t}$ back into dollars, so they end up with $\$\left(1+i_t^E\right)\left(\frac{e_{t+1}}{e_t}\right)$.

We are going to assume that the value of the exchange rate in period $t+1$ is not necessarily known at time t. In this case, investors will instead calculate an expected return on this investment strategy, replacing e_{t+1} with $E_t e_{t+1}$. In other words, they will use their expected value of the exchange rate next period instead of the unknown actual value. If we abstract from risk aversion, then the US investor will be indifferent between this buy-European-bond-and-swap-back-into-dollars strategy and purchasing a US bond as long as the expected return on the two strategies is the same. Mathematically, this means

$$\left(1+i_t^E\right)\left(\frac{E_t e_{t+1}}{e_t}\right) = 1 + i_t^{US} \tag{7.1}$$

An alternative expression for this is

$$\left(1+i_t^E\right)\left(1+\frac{E_t e_{t+1} - e_t}{e_t}\right) = 1 + i_t^{US} \tag{7.2}$$

which can be re-written as

$$1 + i_t^E + \frac{E_t e_{t+1} - e_t}{e_t} + i_t^E \left(\frac{E_t e_{t+1} - e_t}{e_t} \right) = 1 + i_t^{US} \tag{7.3}$$

Subtracting the 1 from each side, we get

$$i_t^E + \frac{E_t e_{t+1} - e_t}{e_t} + i_t^E \left(\frac{E_t e_{t+1} - e_t}{e_t} \right) = i_t^{US} \tag{7.4}$$

Since both i_t^E and $\frac{E_t e_{t+1} - e_t}{e_t}$ are going to be relatively small, the product of them will usually be close to zero, so the condition for the investor to be indifferent between the two investment strategies is

$$i_t^E + \frac{E_t e_{t+1} - e_t}{e_t} \approx i_t^{US} \tag{7.5}$$

This condition—which says that the foreign interest rate plus the expected percentage change in the value of the foreign currency should equal the domestic interest rate—is known as the *Uncovered Interest Parity* condition.

Why should we expect this condition to hold? Why would be expect investors to be indifferent between US and European bonds? Well, suppose it turned out that the European bonds offered a better deal than the US bonds: The combination of interest rate and expected exchange rate appreciation makes the rate of return on European bonds better than that on US bonds. Well, if there is perfect capital mobility, then this would mean that there would be a rush for investors to purchase European bonds rather than US bonds. European institutions who borrow via selling these bonds (governments, highly rated corporations) would figure out that they could borrow at a lower interest rate and still find investors willing to buy their bonds as well as US bonds. By this logic, deviations from Uncovered Interest Parity (UIP) should be temporary with borrowers adjusting the interest rates on their bonds to ensure that investors are indifferent between various international investments.

Note that it states that if European interest rates are lower than US rates, then the Euro must be expected to appreciate. This might seem counter-intuitive: Before reading this, you might expect the country that has higher interest rates to be the one with an appreciating currency. More on this below.

The Trilemma of International Finance

If the UIP relationship approximately holds, then this has important implications for the links between a country's choice of exchange rate regime and its choice of monetary policy. Specifically, if UIP holds, then it is not possible to have all three of the following:

1. Free capital mobility (money moving freely in and out of the country).
2. A fixed exchange rate.
3. Independent monetary policy.

You can have any two of these three things, but not the third:

1. You can have free capital mobility and a fixed exchange rate (so that $E_t e_{t+1} = e_t$) but then your interest rates must equal those of the area you have fixed exchange rates against ($i_t^{US} = i_t^E$). For example, Ireland had a fixed exchange rate with the UK for many years and interest rates in Ireland had to be adjusted upwards and downwards in line with what was happening with UK interest rates.
2. You can have free capital mobility and set your own monetary policy ($i_t^{US} \neq i_t^E$) but then your exchange rate cannot simply be fixed (so that $E_t e_{t+1} \neq e_t$). For example, in

the UK, the Bank of England sets short-term interest rates and the sterling exchange rate fluctuates freely in financial markets.

3. You can set your own monetary policy and fix your exchange rate against another country, but then you must intervene in capital markets to prevent people talking advantage of investment arbitrage opportunities. For example, China generally maintains a fixed exchange rate with the US dollar and also sets its own monetary policy but it does not allow free movement of capital.

This idea that you can only have two from three of free capital mobility, a fixed exchange rate and independent monetary policy is commonly known as the *trilemma* of international finance.

Flexible Exchange Rates Under Capital Mobility

Let's think about how exchange rates should behave under free capital mobility. Recall our example involving US and European bonds. The condition for the expected return on the two investments to be the same was

$$\left(1+i_t^E\right)\left(\frac{E_t e_{t+1}}{e_t}\right) = 1 + i_t^{US} \tag{7.6}$$

Equation (7.6) isn't a linear first-order stochastic difference equation of the type that we have studies up to know. However, if we take logs, it becomes

$$\log\left(1+i_t^E\right) + E_t \log e_{t+1} - \log e_t = \log\left(1+i_t^{US}\right) \tag{7.7}$$

This is a linear stochastic difference equation describing the properties of the log of the exchange rate. It can be re-arranged to be in our more familiar format as

$$\log e_t = \log\left(1+i_t^E\right) - \log\left(1+i_t^{US}\right) + E_t \log e_{t+1} \tag{7.8}$$

Going back to our description of first-order stochastic difference equations, this is an another example of one these equations of the form $y_t = ax_t + bE_t y_{t+1}$, this time with $y_t = \log e_t$, $x_t = \log\left(1 + i_t^E\right) - \log\left(1 + i_t^{US}\right)$, $a = b = 1$. If we apply the repeated substitution technique to this equation, we get

$$\log e_t = \sum_{k=0}^{\infty} E_t \left[\log\left(1 + i_{t+k}^E\right) - \log\left(1 + i_{t+k}^{US}\right)\right] \tag{7.9}$$

It turns out, however, that this is not the only possible solution. To see this, note that for any arbitrary number $\log \bar{e}$ we could re-arrange equation (7.8) as

$$\log e_t - \log \bar{e} = \log\left(1 + i_t^E\right) - \log\left(1 + i_t^{US}\right) + E_t \log e_{t+1} - \log \bar{e} \tag{7.10}$$

In other words, because the coefficient on the expected future exchange rate equals one (because the $b = 1$) then the repeated substitution method works not just for e_t but for any $e_t - \bar{e}$ where $\bar{e}$ is any arbitrary number. So, the general solution is

$$\log e_t = \log \bar{e} + \sum_{k=0}^{\infty} E_t \left[\log\left(1 + i_{t+k}^E\right) - \log\left(1 + i_{t+k}^{US}\right)\right] \tag{7.11}$$

where the theory does not predict what the value of $\bar{e}$ is. Because the natural log function has the property that $\log(1 + x) \approx x$, we can simplify this to read

$$\log e_t = \log \bar{e} + \sum_{k=0}^{\infty} E_t \left(i_{t+k}^E - i_{t+k}^{US}\right) \tag{7.12}$$

We can make a number of points about this equation.

- UIP tells us something about the *dynamics* of the exchange rate but it does not make definitive predictions about the level an exchange rate should be at, i.e. it does not pin down a unique value of $\bar{e}$. Other theories, such as Purchasing Power Parity (the idea that exchange rates should adjust so each type of currency has equivalent purchasing power) do make such predictions, though they don't work very well in practice.

- This unexplained $\bar{e}$ can be seen as a sort of long-run equilibrium exchange rate because this is the rate that holds when the average interest rate on European bonds in the future equals the average interest rate on US bonds.

- The model predicts that deviations from the long-run exchange rate $\bar{e}$ are determined by expectations that interest rates will differ across areas. In this example, the euro will be higher than $\bar{e}$ if people expect European interest rates to be higher in the future than US rates.

The model explains the slightly puzzling result we discussed earlier: That higher interest rates in Europe imply the euro is expected to depreciate. Suppose in period $t-1$, Euro and US interest rates were equal to each other and expected to stay that way. Equation (7.12) implies that under these circumstances we would have $\log e_{t-1} = \log \bar{e}$. Now suppose that, in period t, Euro interest rates unexpectedly went above US interest rates just for one period. What would happen? The Euro must end up back at $\bar{e}$ (because interest rates in the two areas are going to equal each other after period t) and the Euro must also be expected to depreciate (because of the higher current interest rate in Euro).

So, in response to the surprise temporary increase in European interest rates, the Euro immediately jumps upwards and then depreciates back to $\bar{e}$. This conforms with our intuition that a surprise announcement of higher European interest rates should make the Euro appreciate.

UIP and Exchange Rate Volatility

During the period after the Second World War up to the 1970s, most of the world's economies operated the so-called Bretton Woods system of quasi-fixed exchange rates. The 1970s saw the widespread introduction of market-determined flexible exchange rates. Prior to the introduction of this system, advocates of market-based flexible exchange rates had predicted that exchange rates would change very little over time under this system.

The truth turned out to be the opposite: Exchange rates change by large amounts on a daily, weekly and monthly basis. See Figure 7.1 which shows the Euro-dollar exchange rate. It has also gone through big swings: Reaching lows of 0.8 in 2000 and highs of 1.6 in 2008. In addition, there are often large day to day movements where the exchange rate will go up or down by one or two percent.

The model just developed—combining the UIP with rational expectations—helps to explain why exchange rates are so volatile. Using equation (7.12) for the level of exchange rates, we can derive the change in the exchange rate at time t as

$$\Delta \log e_t = \sum_{k=0}^{\infty} E_t \left(i^E_{t+k} - i^{US}_{t+k} \right) - \sum_{k=-1}^{\infty} E_{t-1} \left(i^E_{t+k} - i^{US}_{t+k} \right) \tag{7.13}$$

We will simplify this a bit via a slightly iffy bit of terminology, meaning that we will write $(E_t - E_{t-1})\, x_{t+k}$ to mean $E_t x_{t+k} - E_{t-1} x_{t+k}$, i.e. this means the change between time $t-1$ and time t in what people expect x_{t-k} to be. Given this, we can re-write the previous equation as

$$\Delta \log e_t = i^{US}_{t-1} - i^E_{t-1} + \sum_{k=0}^{\infty} (E_t - E_{t-1}) \left(i^E_{t+k} - i^{US}_{t+k} \right) \tag{7.14}$$

This equation tells us a lot about how exchange rates should behave if investors have rational expectations. Exchange rate changes reflect not only the *expected* change due to past interest rate differentials expiring (the $i^{US}_{t-1} - i^E_{t-1}$ term); they also reflect *unexpected* changes in the projected path of future interest rate differentials. This means that all information that affects expectations of future Euro-area and US interest rates feeds directly into today's exchange rate. Because interest rates are set by central banks in response to developments in the macroeconomy, this means that exchange rates should react to all types of macroeconomic news.

Figure 7.1: Daily Data on the Euro-Dollar Exchange Rate

Problems for the UIP-Rational Expectations Theory

The UIP theory helps to explain a number of important aspects of the behaviour of exchange rates. However, there have been many examples of where the theory just outlined does not seem to work well. Indeed, quite commonly, there have been examples where the theory predicts for an extended period of time that a currency depreciation or appreciation should be expected, when what actually happens is the opposite.

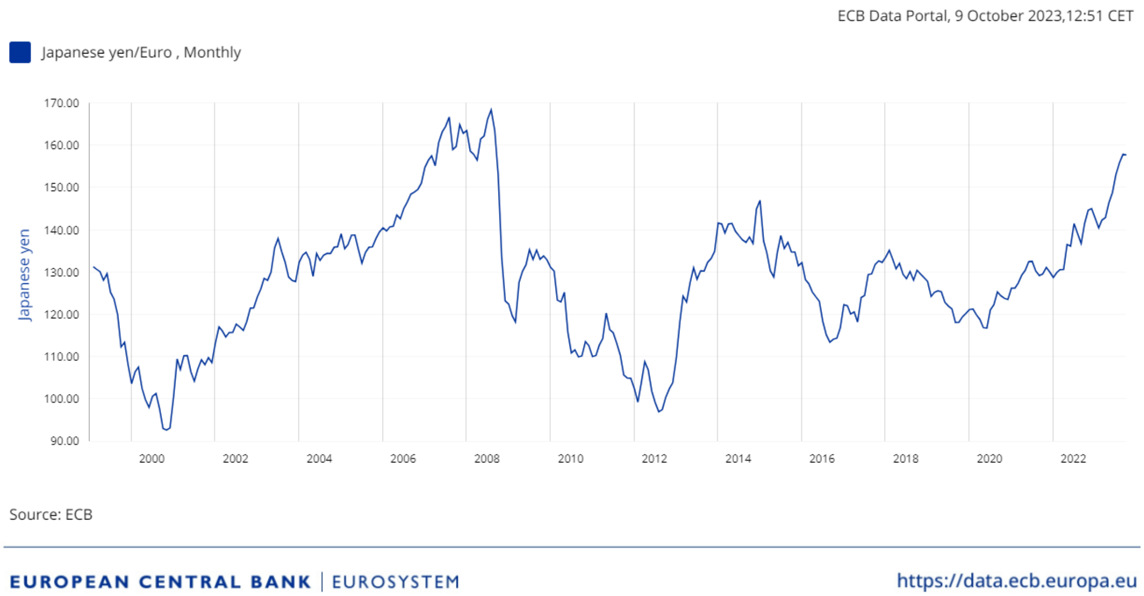
One potential explanation for this apparent failure that could still be consistent with the model is that $E_t e_{t+1} - e_t$ is not the same as $e_{t+1} - e_t$: The mathematical expectation of something and its actual outcome can sometimes differ from each other for quite a while. This is sometimes called *the Peso problem*. Sometimes interest rates in developing economies (such as Mexico, after which the term is named) are high because markets think there is a probability (perhaps a small probability) that a large depreciation may be coming. Just because the depreciation doesn't happen during a particular sample doesn't mean the expectation was unreasonable or that it won't be correct at some point.

But evidence also seems to exist of more systematic errors for the UIP theory. Take one example. For most of 2000s, Japanese interest rates were well below European levels for most of this decade. The UIP-Rational Expectations approach would have predicted that the Yen should have been appreciating against the Euro: In fact, the opposite happened systematically from 2001 to 2008. See Figure 7.2. Many traders systematically exploited this, borrowing at low interest rates in Yen, using the funds to buy Euro bonds that yielded higher interest rates and then repaying their debts in depreciated Yen—the so-called Yen carry trade. That said, as Figure 7.2 also shows, the "carry trade" unwound itself fairly spectacularly in 2008.

The leading explanations for the apparent failures of the UIP-RE theory involve introducing risk aversion (we have assumed investors are risk-neutral) and home-bias (the preference for assets denominated in your home currency). For instance, in relation to the theory's failure to explain the Yen carry trade period, it's worth noting that many Japanese investors have a strong preference for Yen-denominated assets and don't want to take on the extra currency-related risk of investing in dollar or euro-denominated assets.

These kinds of preferences may lead to short-term violations of the stronger predictions of the UIP-RE theory. However, they will not allow countries to escape from the restrictions of the Trilemma: A country that attempts to adopt a systematically different interest rate policy than another country simply will not be able to have a fixed exchange rate with that country unless it imposes capital controls.

Figure 7.2: Daily Data on the Euro-Yen Exchange Rate



Chapter 8

Sticky Prices and the Phillips Curve

One of the important themes of macroeconomics is that the behaviour of prices is crucial in determining how the macro-economy responded to shocks. The IS-LM model of intermediate macro assumed that prices were "sticky" in the short-run to obtain real effects for fiscal and monetary policy but we assumed that prices were flexible in the long-run so that the economy returned to its full employment level over time. In the IS-MP-PC theory, we formalised this idea a bit more: This model featured prices that adjusted over time in response to the real economy according to a Phillips curve. In this chapter, we will return to the topic of price setting and the relationship over time between inflation and the business cycle. We formally derive a kind of Phillips curve from "first principles" under the assumption that prices are sticky.

Evidence on Price Stickiness

When we discussed IS-LM, we assumed that the price level did not keep moving to constantly equate GDP with the level of output consistent with a natural rate of unemployment. Instead, we assumed that prices only changed gradually over time in response to the real economy. The idea that prices may be "sticky" has a long history in Keynesian macroeconomics but, until recent decades, there was comparatively little evidence on the extent to which prices changed over time.

This has changed since statistical agencies have made available the micro-data that underlie Consumer Price Indices (CPIs). To construct a CPI, agencies have to collect large numbers of quotes of prices on individual items (e.g. they can tell you the price in April of a bottle of Heinz ketchup at a particular store). These individual price quote data can be used to assess how often individual prices are changed.

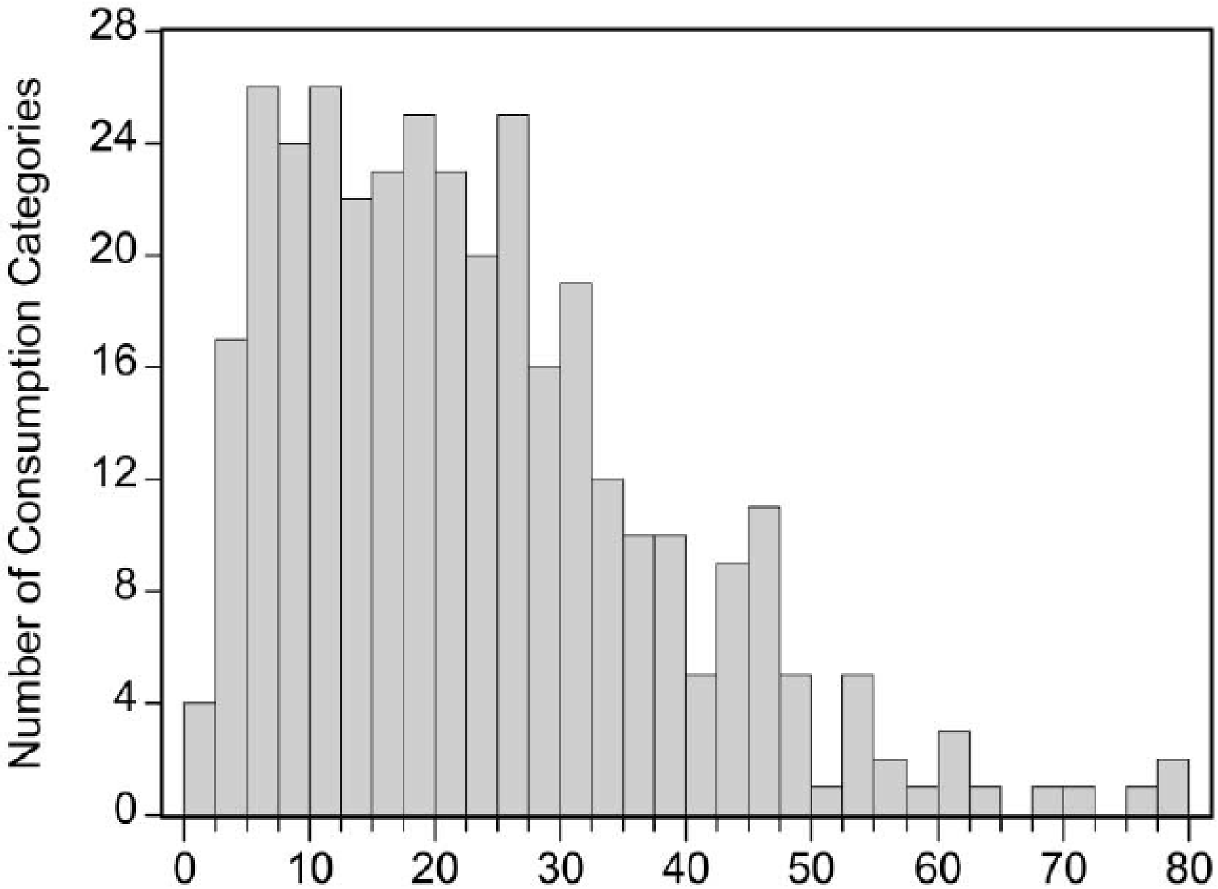
Studies of this type now exist for a large number of countries. For example, Bils and Klenow's 2004 paper provided evidence for consumer prices in the United States.[1] An important finding from this research is that the data show a very wide range of values across different goods and services for the frequency with which prices change. Figure 8.1 shows a histogram from Bils and Klenow's paper showing the distribution of the percentage probability that any price changes in a month. These vary from products whose prices have only a one percent probability of changing each month ("Coin-operated apparel laundry and dry cleaning") to those that have an 80 percent probability of changing each month (gasoline).

The table on the following page shows the median price duration is about four months. In other words, half of the prices quoted in the CPI change more than every four months, while the other half change less than every four months. Research for the euro area has shown that price durations are even longer in Europe. For example, Alvarez et al (2006) report a median price duration for the euro area of 10.6 months.[2]

[1]Mark Bils and Peter Klenow (2004). "Some Evidence on the Importance of Sticky Prices" *Journal of Political Economy*, Volume 112, Number 5.

[2]Luis Alvarez, Emmanuel Dhyne, Marco Hoeberichts, Claudia Kwapil, Herve Le Bihan. Patrick Lunnemann, Fernando Martins, Roberto Sabbatini, Harald Stahl, Philip Vermeulen and Jouko Vilmunen (2006). "Sticky Prices in the Euro Area: A Summary of New Micro-Evidence" *Journal of the European Economic Association*, Volume 4(2-3), pages 575-584.

Figure 8.1: The Distribution of Monthly Percent Probability of Price Changes



Bils and Klenow Evidence on Price Durations

TABLE 1
MONTHLY FREQUENCY OF PRICE CHANGES BY YEAR, 1995–2002

Year	Median Frequency (%)	Median Duration (Months)
1995	21.3	4.2
1996	20.8	4.3
1997	19.9	4.5
1998	21.2	4.2
1999	21.4	4.1
2000	21.7	4.1
2001–2	22.0	4.0

New Classical and New Keynesian Macroeconomics

After Milton Friedman's critique of the Phillips curve, macroeconomists began to pay more attention to the question of how expectations were formed. In particular, a number of papers by Robert Lucas and Thomas Sargent introduced rational expectations into macroeconomic modelling. These early papers tended to assume that prices were perfectly flexible, which limited the ability of fiscal and monetary policy to influence output. This school of thought became labelled *New Classical* economics.

In a number of famous New Classical papers, Robert Lucas argued that monetary policy could still have short-run effects even if prices were flexible and people had rational expectations. Lucas's model relied on the idea that firms had a difficulty in the short-run distinguishing between movements in their prices and movements in the overall price levels. For this reason, an increase in the money supply that provoked an increase in prices could, in the short-run, provoke higher output because firms may believe this is increasing their relative price and making production more profitable. Lucas emphasised, however, that once people had rational expectations, the impact of policy on output could only be short-lived. In particular, he stressed that only *unpredictable* fiscal and monetary policies would have an impact because people with rational expectations would anticipate the impact of predictable policy on the price level.

Once we allow prices to be sticky, however, these points no longer hold. Because some prices will not change even after the government changes fiscal or monetary policy, these policies will have the traditional short-run impacts described in simple Keynesian models even if people have rational expectations. There are lots of different ways of formulating the idea that prices may be sticky. Some of the best known formulations were those introduced in papers in the late seventies by John Taylor and Stanley Fischer.[3] These papers assumed that only a certain fraction of firms set prices each period but those who did change their prices would set them in an optimal manner using rational expectations. This work, which combined rational expectations with sticky prices, invented what is now known as *New Keynesian economics*.

Pricing à la Calvo

The New Keynesian literature contains a number of different formulations of sticky prices. For the rest of this chapter, we will use a formulation of sticky prices known as *Calvo pricing*, after the economist who first introduced it.[4] Though not the most realistic formulation of sticky prices, it turns out to provide analytically convenient expressions, and has implications that are very similar to those of more realistic (but more complicated) formulations.

The form of price rigidity faced by firms in the Calvo model is as follows. Each period, only a random fraction $(1-\theta)$ of firms are able to reset their price; all other firms keep their prices unchanged. When firms do get to reset their price, they must take into account that the price may be fixed for many periods. We assume they do this by choosing a log-price, z_t, that minimizes the "loss function"

$$L(z_t) = \sum_{k=0}^{\infty} (\theta\beta)^k E_t \left(z_t - p^*_{t+k}\right)^2 \tag{8.1}$$

where β is between zero and one, and p^*_{t+k} is the log of the optimal price that the firm would set in period $t+k$ if there were no price rigidity.

This expression probably looks a bit intimidating, so it's worth discussing it a bit to explain

[3]Stanley Fischer (1977), "Long-Term Contracts, Rational Expectations, and the Optimal Money Supply Rule," *Journal of Political Economy*, 85, 191-205, and John Taylor (1979), "Staggered Wage Setting in a Macro Model," *American Economic Review*, Papers and Proceedings, Vol. 69, 108-113.

[4]Guillermo Calvo, "Staggered Contracts in a Utility-Maximizing Framework" *Journal of Monetary Economics*, September 1983.

what it means. The loss function has a number of different elements:

- The term $E_t \left(z_t - p^*_{t+k} \right)^2$ describes the expected loss in profits for the firm at time $t+k$ due to the fact that it will not be able to set a frictionless optimal price that period. This quadratic function is intended just as an approximation to some more general profit function. What is important here is to note that because the firm may be stuck with the price z_t for some time, it will lose profits relative to what it would have been able to obtain if there were no price rigidities.

- The summation $\sum_{k=0}^{\infty}$ shows that the firm considers the implications of the price set today for all possible future periods.

- However, the fact that $\beta < 1$ implies that the firm places less weight on future losses than on today's losses. A dollar today is worth more than a dollar tomorrow because it can be re-invested. By the same argument, a dollar lost today is more important than a dollar lost tomorrow.

- Future losses are actually discounted at rate $(\theta\beta)^k$, not just β^k. This is because the firm only considers the *expected* future losses from the price being fixed at z_t. The chance that the price will be fixed until $t+k$ is θ^k. So the period $t+k$ loss is weighted by this probability. There is no point in the firm worrying too much about losses that might occur from having the wrong price far off in the future, when it is unlikely that the price will remained fixed for that long.

The Optimal Reset Price

After all that, the actual solution for the optimal value of z_t, (i.e. the price chosen by the firms who get to reset) is quite simple. Each of the terms featuring the choice variable z_t—that is, each of the $\left(z_t - p^*_{t+k}\right)^2$ terms—need to be differentiated with respect to z_t and then the sum of these derivatives is set equal to zero. This means

$$L'(z_t) = 2\sum_{k=0}^{\infty} (\theta\beta)^k E_t \left(z_t - p^*_{t+k}\right) = 0 \tag{8.2}$$

Separating out the z_t terms from the p^*_{t+k} terms, this implies

$$\left[\sum_{k=0}^{\infty} (\theta\beta)^k\right] z_t = \sum_{k=0}^{\infty} (\theta\beta)^k E_t p^*_{t+k} \tag{8.3}$$

Now, we can use our old pal the geometric sum formula to simplify the left side of this equation. In other words, we use the fact that

$$\sum_{k=0}^{\infty} (\theta\beta)^k = \frac{1}{1-\theta\beta} \tag{8.4}$$

to re-write the equation as

$$\frac{z_t}{1-\theta\beta} = \sum_{k=0}^{\infty} (\theta\beta)^k E_t p^*_{t+k} \tag{8.5}$$

implying a solution of the form

$$z_t = (1-\theta\beta)\sum_{k=0}^{\infty} (\theta\beta)^k E_t p^*_{t+k} \tag{8.6}$$

Stated in English, all this equation says is that the optimal solution is for the firm to set its price equal to a weighted average of the prices that it would have expected to set in the future if there weren't any price rigidities. Unable to change price each period, the firm chooses to try to keep close "on average" to the right price.

And what is this "frictionless optimal" price, p^*_t? We will assume that the firm's optimal pricing strategy without frictions would involve setting prices as a fixed markup over marginal cost:

$$p_t^* = \mu + mc_t \tag{8.7}$$

Thus, the optimal reset price can be written as

$$z_t = (1 - \theta\beta) \sum_{k=0}^{\infty} (\theta\beta)^k E_t (\mu + mc_{t+k}) \tag{8.8}$$

The New-Keynesian Phillips Curve

Now, we can show how to derive the behaviour of aggregate inflation in the Calvo economy. The aggregate price level in this economy is just a weighted average of last period's aggregate price level and the new reset price, where the weight is determined by θ:

$$p_t = \theta p_{t-1} + (1 - \theta) z_t, \tag{8.9}$$

This can be re-arranged to express the reset price as a function of the current and past aggregate price levels

$$z_t = \frac{1}{1-\theta} (p_t - \theta p_{t-1}) \tag{8.10}$$

Now, let's examine equation (8.8) for the optimal reset price again. We have shown that the first-order stochastic difference equation

$$y_t = a x_t + b E_t y_{t+1} \tag{8.11}$$

can be solved to give

$$y_t = a \sum_{k=0}^{\infty} b^k E_t x_{t+k} \tag{8.12}$$

Examining equation (8.8), we can see that z_t must obey a first-order stochastic difference equation with

$$y_t \;\; = \;\; z_t \tag{8.13}$$

$$
\begin{aligned}
x_t &= \mu + mc_t && (8.14)\\
a &= 1-\theta\beta && (8.15)\\
b &= \theta\beta && (8.16)
\end{aligned}
$$

In other words, we can write the reset price as

$$z_t = \theta\beta E_t z_{t+1} + (1-\theta\beta)(\mu + mc_t) \tag{8.17}$$

Substituting in the expression for z_t in equation (8.10) we get

$$\frac{1}{1-\theta}(p_t - \theta p_{t-1}) = \frac{\theta\beta}{1-\theta}(E_t p_{t+1} - \theta p_t) + (1-\theta\beta)(\mu + mc_t) \tag{8.18}$$

After a bunch of re-arrangements, this equation can be shown to imply

$$\pi_t = \beta E_t \pi_{t+1} + \frac{(1-\theta)(1-\theta\beta)}{\theta}(\mu + mc_t - p_t) \tag{8.19}$$

where $\pi_t = p_t - p_{t-1}$ is the inflation rate.

This equation is known as the *New-Keynesian Phillips Curve* (NKPC). It states that inflation is a function of two factors:

- Next period's expected inflation rate, $E_t \pi_{t+1}$.
- The gap between the frictionless optimal price level $\mu + mc_t$ and the current price level p_t. Another way to state this is that inflation depends positively on *real marginal cost*, $mc_t - p_t$.

Why is real marginal cost a driving variable for inflation? Firms in the Calvo model would like to keep their price as a fixed markup over marginal cost. If the ratio of marginal cost to price is getting high (i.e. if $mc_t - p_t$ is high) then this will spark inflationary pressures because those firms that are re-setting prices will, on average, be raising them.

Real Marginal Cost and Output

For simplicity, we will denote the deviation of real marginal cost from its frictionless level of $-\mu$ as

$$\hat{mc}_t^r = \mu + mc_t - p_t \tag{8.20}$$

so we can write the NKPC as

$$\pi_t = \beta E_t \pi_{t+1} + \frac{(1-\theta)(1-\theta\beta)}{\theta} \hat{mc}_t^r \tag{8.21}$$

One problem with attempting to implement this model empirically, is that we don't actually observe data on real marginal cost. National accounts data contain information on the factors that affect *average* costs such as wages, but do not tell us about the cost of producing an additional unit of output. That said, it seems very likely that marginal costs are procyclical, and more so than prices. When production levels are high relative to potential output, there is more competition for the available factors of production, and this leads to increases in real costs, i.e. increases in the costs of the factors over and above increases in prices. Some examples of the procyclicality of real marginal costs are fairly obvious. For example, the existence of overtime wage premia generally means a substantial jump in the marginal cost of labour once output levels are high enough to require some people to work more than the standard workweek.

For these reasons, many researchers have implemented the NKPC using a measure of the *output gap* (the deviation of output from its potential level) as a proxy for real marginal cost.

In other words, they assume a relationship such as

$$\hat{mc}_t^r = \lambda \tilde{y}_t \tag{8.22}$$

where $\tilde{y}_t$ is the output gap. This implies a New-Keynesian Phillips curve of the form

$$\pi_t = \beta E_t \pi_{t+1} + \gamma \tilde{y}_t \tag{8.23}$$

where

$$\gamma = \frac{\lambda (1-\theta)(1-\theta\beta)}{\theta} \tag{8.24}$$

And this approach can be implemented empirically using various measures for estimating potential output.[5]

The "Asset-Price-Like" Behaviour of NKPC Inflation

The New-Keynesian approach assumes that firms have rational expectations. Thus, we can apply the repeated substitution method to equation (8.23) to arrive at

$$\pi_t = \gamma \sum_{k=0}^{\infty} \beta^k E_t \tilde{y}_{t+k} \tag{8.25}$$

Inflation today depends on the whole sequence of expected future output gaps. Thus, the NKPC sees inflation as behaving according to the classic "asset-price" logic that we saw with the dividend-discount stock price model.

The NKPC and the Lucas Critique

Friedman's critique of the original Phillips curve, which means there is little or scope for

[5]Roberts (1995) shows that a number of other models of sticky prices also imply a formulation for inflation similar to the New Keynesian Phillips curve.

policy-makers to choose a tradeoff between inflation and output, is now widely accepted. However, as we discussed earlier, there is empirical evidence for a relationship of the form

$$\pi_t = \pi_{t-1} + \alpha - \beta u_t \tag{8.26}$$

So there is a relationship between the *change* in inflation and the level of unemployment. In this formulation, the prior inflation term reflects how last period's level of inflation changes people's expectations and so feeds into today's inflation. Based on this equation, economists often speak of the so-called NAIRU—the non-accelerating inflation rate of unemployment. This is the inflation rate consistent with constant inflation and it is defined implicitly by

$$\alpha - \beta u^* = 0 \Rightarrow u^* = \frac{\alpha}{\beta} \tag{8.27}$$

Empirical estimates of the NAIRU have often been used in real-world policy discussions, with the policy recommendations made on the basis of whether unemployment is above or below this NAIRU level.[6]

The NKPC model provides a different view of this empirical relationship. If this model is correct, then equation 8.26 can fit the data reasonably well without it being the true underlying structural relationship. In other words, if the true model is

$$\pi_t = \beta E_t \pi_{t+1} + \gamma \tilde{y}_t \tag{8.28}$$

then equation 8.26 might have a good statistical fit because π_{t-1} is likely to be correlated with $E_t \pi_{t+1}$. However, the Lucas critique means policy-makers should not rely on this relationship, because changes in policy may produce a break the correlation between $E_t \pi_{t+1}$ and π_{t-1} and at this point the econometric Phillips curve will break down.

[6]Note though the NAIRU terminology is actually a misnomer. If unemployment is below u^*, then inflation will be increasing, but not accelerating. The price level is what will be accelerating.

The NKPC and Disinflation

The NKPC also has important implications for how a government can approach reducing inflation. Consider again the Phillips curve based on adaptive expectations, equation (8.26). The fact that inflation depends on its own lagged values in this formulation means then it would be very difficult to reduce inflation quickly without a significant increase in unemployment. So, this Phillips curve suggests that gradualist policies are way to reduce inflation.

But the implications of the NKPC are completely different. There may be a statistical relationship between current and lagged inflation but the NKPC says that there is no structural relationship at all. Thus, there is no need for gradualist policies to reduce inflation. According to the NKPC, low inflation can be achieved immediately by the central bank announcing (and the public believing) that it is committing itself to eliminating positive output gaps in the future: This can be seen from equation 8.25.

Whether the empirical evidence fits with the NKPCs predictions is open for debate. For example, there has been plenty of evidence that reductions in inflation do tend to be costly in terms of lost output and high unemployment. Some, however, have put this down to the failure of governments and central banks to credibly convince the public of their commitment to lower inflation rates.

Chapter 9

Investment With Adjustment Costs

In the previous chapters, we have seen a number examples of forward-looking first-order stochastic difference equations of the form

$$y_t = ax_t + bE_t y_{t+1} \tag{9.1}$$

The solution that we have derived has been of the form

$$y_t = a \sum_{k=0}^{\infty} b^k E_t x_{t+k} \tag{9.2}$$

so that y_t is a completely forward-looking variable. Note that this means that y_t does not depend at all on its own past values. But this does not provide a good description of all macroeconomic variables. In some cases, variables depend not just on the present and expectations of the future but also depend on past values of the variable itself. We will discuss an example of this case.

Specifically, we will look at a theory of the determination of the capital stock (and thus investment). Empirical studies show that the capital stock does not change very much from period to period. Economists usually rationalise this by assuming that there are some form of "adjustment costs" that prevent firms from changing their capital stock too quickly. In this chapter, we will consider a model of investment with adjustment costs, show that it implies a *second-order* stochastic difference equation, and examine the methods used to solve these types of equations.

The Firm's Problem

Consider now the following model of firm investment. We will assume that, each period, there is a level of the log of the capital stock, k_t^*, that the firm would choose if there were no adjustment costs. We will call this the *frictionless optimal capital stock*. With adjustment costs the firm has to choose a planned sequence of capital stocks $E_t \{k_{t,} k_{t+1,} k_{t+2}, ...\}$ minimise the following "loss function"

$$L(k_t, k_{t+1}, t_{t+2}, ...) = \quad E_t \left[\sum_{m=0}^{\infty} \theta^m \left\{ \left(k_{t+m} - k_{t+m}^* \right)^2 + \alpha \left(k_{t+m} - k_{t+m-1} \right)^2 \right\} \right] \tag{9.3}$$

As with our previous loss function for the Calvo model, this looks intimidating but it's not too complicated once you take it bit by bit:

- Firstly, for each period, $t+m$, there is a term $\left(k_{t+m} - k_{t+m}^* \right)^2$ that describes the loss in profits suffered by the firm from not having its capital stock equated with the frictionless optimal level.

- Secondly, there is a term $\alpha \left(k_{t+m} - k_{t+m-1} \right)^2$ which describes the concept of adjustment costs formally: Ceteris paribus, changes in the capital stock have a negative effect on firm profits.

- The reason we are assuming that k_t is actually the log of the stock, as opposed to the stock itself, is that this way these losses can be viewed in percentage terms: It is the percentage gap between capital and its frictionless optimal that matters and also the percentage change in the stock. This makes more sense than levels of these gaps mattering because economic growth will make levels of these variables grow over time.

- Finally, the parameter θ is a discount rate less than one, which tells us that firms care more about profits today than profits tomorrow.

This loss function can be re-written as

$$\begin{aligned} L\left(k_t, k_{t+1}, t_{t+2}, ...\right) &= \left(k_t - k_t^*\right)^2 + \alpha\left(k_t - k_{t-1}\right)^2 + \theta E_t\left[\left(k_{t+1} - k_{t+1}^*\right)^2 + \alpha\left(k_{t+1} - k_t\right)^2\right] \\ &\quad + \theta^2 E_t\left[\left(k_{t+2} - k_{t+2}^*\right)^2 + \alpha\left(k_{t+2} - k_{t+1}\right)^2\right] + ... \end{aligned} \tag{9.4}$$

An optimal plan is arrived at by differentiating this with respect to each of the capital stock terms k_{t+m} and setting these derivatives equal to zero. Consider first differentiating with respect to k_t. This gives

$$2\left(k_t - k_t^*\right) + 2\alpha\left(k_t - k_{t-1}\right) - 2\alpha\theta E_t\left(k_{t+1} - k_t\right) = 0 \tag{9.5}$$

Again, try differentiating with respect to k_{t+1}. This gives

$$E_t\left[2\theta\left(k_{t+1} - k_{t+1}^*\right) + 2\alpha\theta\left(k_{t+1} - k_t\right) - 2\alpha\theta^2\left(k_{t+2} - k_{t+1}\right)\right] = 0 \tag{9.6}$$

This is the exact same as the previous first-order condition, only shifted forward one period. In fact, all of the FOCs describing the optimal dynamics of the capital are consistent with the same second-order stochastic difference equation

$$E_t\left[\left(k_t - k_t^*\right) + \alpha\left(k_t - k_{t-1}\right) - \alpha\theta\left(k_{t+1} - k_t\right)\right] = 0 \tag{9.7}$$

Drawing terms together, this gives

$$-\alpha\theta E_t k_{t+1} + \left(1 + \alpha + \alpha\theta\right) k_t - \alpha k_{t-1} = k_t^* \tag{9.8}$$

which can be re-written as

$$E_t k_{t+1} - \left(1 + \frac{1}{\theta} + \frac{1}{\alpha\theta}\right) k_t + \frac{1}{\theta} k_{t-1} = -\frac{1}{\alpha\theta} k_t^* \tag{9.9}$$

Because the maximum difference between time subscripts is two, this is a second-order stochastic difference equation. There are two different methods that are commonly used to solve equations of this form. I will discuss the so-called *factorization* method. For completeness, I have also attached the derivation of the solution using the other method known as the *method of undetermined coefficients*, but you can ignore this if you wish.

Lag Operators

The factorization method makes use what are known as *lag* and *forward* operators. These are commonly used in calculations relating to time series, and they work as follows. The lag operator turns a variable dated time t into a variable dated time $t-1$:

$$Ly_t = y_{t-1} \tag{9.10}$$

Lag operators can be multiplied and added just like normal variables. So, for instance, one can write

$$L^k y_t = y_{t-k} \tag{9.11}$$

The forward operator has the reverse effect of the lag operator

$$F^k y_t = y_{t+k} \tag{9.12}$$

Lag and forward operators also obey a form of the geometric sum formula. Recall that for $-1 < \beta < 1$, we have

$$\sum_{m=0}^{\infty} \beta^m = \frac{1}{1-\beta} \tag{9.13}$$

Recall also that if $-1 < \beta < 1$ and

$$y_t = \beta E_t y_{t+1} + x_t \tag{9.14}$$

then the solution is

$$y_t = \sum_{m=0}^{\infty} \beta^m E_t x_{t+k} \tag{9.15}$$

Equation (9.14) can be re-written as

$$y_t = E_t \left[\frac{1}{1 - \beta F} x_t \right] \tag{9.16}$$

So equation (9.15) means that

$$\frac{1}{1 - \beta F} = \sum_{m=0}^{\infty} \beta^m F^m \tag{9.17}$$

The same applies for lag operators

$$\frac{1}{1 - \beta L} = \sum_{m=0}^{\infty} \beta^m L^m \tag{9.18}$$

To verify that this is the case, note that if

$$y_t = \beta y_{t-1} + x_t \tag{9.19}$$

then one can apply repeated substitution to re-write this as

$$y_t = x_t + \beta x_{t-1} + \beta^2 x_{t-2} + \beta^3 x_{t-3} + ... \tag{9.20}$$

Armed with this knowledge of lag and forward operators we can solve the second-order stochastic difference equation using the factorization method.

Solution via Factorization

This method first re-writes equation (9.9) in terms of lag and forward operators. Written this way it is

$$E_t \left[\left(F - \left(1 + \frac{1}{\theta} + \frac{1}{\alpha\theta} \right) + \frac{1}{\theta} L \right) k_t \right] = -\frac{1}{\alpha\theta} k_t^* \tag{9.21}$$

Next, the method re-expresses the left-hand-side in terms of a quadratic equation in F multiplied by L:

$$E_t\left[\left(F^2-\left(1+\frac{1}{\theta}+\frac{1}{\alpha\theta}\right)F+\frac{1}{\theta}\right)Lk_t\right]=-\frac{1}{\alpha\theta}k_t^* \tag{9.22}$$

You may recall that polynominals of the form

$$g(x)=x^2+bx+c \tag{9.23}$$

can be re-written in terms of their roots as

$$g(x)=(x-\lambda_1)(x-\lambda_2) \tag{9.24}$$

where

$$\lambda_1+\lambda_2 = -b \tag{9.25}$$

$$\lambda_1\lambda_2 = c \tag{9.26}$$

In this case, one can show that the polynomial

$$x^2-\left(1+\frac{1}{\theta}+\frac{1}{\alpha\theta}\right)x+\frac{1}{\theta} \tag{9.27}$$

has two roots such that one root (λ) is between zero and one while the other equals $\frac{1}{\theta\lambda}$. This means that the optimality condition for the capital stock can be re-expressed as

$$E_t\left[(F-\lambda)\left(F-\frac{1}{\theta\lambda}\right)Lk_t\right]=-\frac{1}{\alpha\theta}k_t^* \tag{9.28}$$

Dividing across by $\left(F-\frac{1}{\theta\lambda}\right)$, this becomes

$$E_t\left[(F-\lambda)Lk_t\right]=-\frac{1}{\alpha\theta}E_t\left[\frac{1}{F-\frac{1}{\theta\lambda}}k_t^*\right] \tag{9.29}$$

Now we can use the properties of lag operators just derived to show that

$$\frac{1}{F-\frac{1}{\theta\lambda}}=\frac{-\theta\lambda}{1-\theta\lambda F}=-\theta\lambda\sum_{k=0}^{\infty}(\theta\lambda)^k F^k \tag{9.30}$$

So, the capital stock process has a solution of the form

$$k_t = \lambda k_{t-1} + \frac{\lambda}{\alpha} E_t \left[\sum_{n=0}^{\infty} (\theta\lambda)^n k^*_{t+n} \right] \tag{9.31}$$

Note now how adding adjustment costs changes the solution for a rational expectations model. This produces a second-order difference equation, and the solution is no longer completely forward-looking. Instead, the capital stock has a forward-looking component, which is a geometric discounted sum, but it also has a backward-looking component, whereby it depends on its own lagged value.

An Example: Investment, Output, and the Cost of Capital

The model can be fleshed out by stating what are the determinants of the frictionless optimal capital stock. For instance, if the production function was of the Cobb-Douglas form, then the optimal capital stock would take the form

$$K^*_t = \frac{Y_t}{C_t} \tag{9.32}$$

where Y_t is output and C_t is the cost of capital. Using lower-case letters to denote logs, this can be written as

$$k^*_t = y_t - c_t \tag{9.33}$$

So, the capital stock is determined by

$$k_t = \lambda k_{t-1} + \frac{\lambda}{\alpha} E_t \left[\sum_{n=0}^{\infty} (\theta\lambda)^n (y_{t+n} - c_{t+n}) \right] \tag{9.34}$$

Now assume that output and the cost of capital both follow AR(1) processes

$$y_t = \rho_y y_{t-1} + \epsilon^y_t \tag{9.35}$$

$$c_t = \rho_c c_{t-1} + \epsilon^c_t \tag{9.36}$$

The infinite sum component of the solution can now be written as

$$E_t \sum_{n=0}^{\infty} (\theta\lambda)^n y_{t+n} = \left[\sum_{n=0}^{\infty} (\theta\lambda\rho_y)^n\right] y_t = \frac{1}{1-\theta\lambda\rho_y} y_t \tag{9.37}$$

while

$$E_t \sum_{n=0}^{\infty} (\theta\lambda)^n c_{t+n} = \frac{1}{1-\theta\lambda\rho_c} c_t \tag{9.38}$$

So, the capital stock process is

$$k_t = \lambda k_{t-1} + \frac{\lambda}{\alpha}\frac{1}{1-\theta\lambda\rho_y} y_t - \frac{\lambda}{\alpha}\frac{1}{1-\theta\lambda\rho_c} c_t \tag{9.39}$$

This gives us a "reduced-form" relationship between the capital stock, the lagged capital stock, output and the cost of capital.

Note that the magnitudes of the coefficients on output and the cost of capital depend positively on the *persistence* of these variables. If ρ_y is close to one, then the coefficient on output will be high, with the same applying for ρ_c and the cost of capital. One example of an application of this type of reasoning is Tevlin and Whelan (2003).[1] This paper presents regressions of equipment investment on output and the cost of capital. It reports much larger coefficients on the cost of capital for investment in computers than for non-computing equipment, and uses a model of this sort to provide an explanation. The cost of capital for computing equipment is largely determined by very persistent shocks that tend to produce ever-decreasing computer prices. In contrast, for non-computing equipment, the cost of capital depends on a set of less persistent variables such as interest rates and tax incentives. This suggests that the cost of capital should have a smaller coefficient in a regression for the non-computing capital stock.

[1]Stacey Tevlin and Karl Whelan. "Explaining the Investment Boom of the 1990s," *Journal of Money, Credit, and Banking*, Volume 35, pages 1-22, 2003.

Appendix: The Undetermined Coefficients Method

The other method used to solve these models starts by assuming that one knows the form of the solution. So, one "guesses" that the solution is of the form

$$k_t = \lambda_1 k_{t-1} + \gamma E_t \left[\sum_{n=0}^{\infty} \lambda_2^n k_{t+n}^* \right]$$

From there, one goes on to figure out a unique set of values for λ_1, λ_2 andγ that are consistent with this equation, and with the optimality conditions for the capital stock. In this case

$$E_t k_{t+1} = \lambda_1 k_t + \gamma E_t \left[\sum_{n=0}^{\infty} \lambda_2^n k_{t+n+1}^* \right]$$

So, we have

$$\begin{aligned} -\alpha\theta \left[\lambda_1 k_t + \gamma E_t \left[\sum_{n=0}^{\infty} \lambda_2^n k_{t+n+1}^* \right] \right] + (1+\alpha+\alpha\theta) k_t - \alpha k_{t-1} &= k_t^* \\ (1+\alpha+\alpha\theta-\alpha\theta\lambda_1) k_t &= \alpha k_{t-1} + k_t^* + \alpha\theta\gamma E_t \left[\sum_{n=0}^{\infty} \lambda_2^n k_{t+n+1}^* \right] \end{aligned}$$

This can be re-written as

$$k_t = \frac{\alpha}{(1+\alpha+\alpha\theta-\alpha\theta\lambda_1)} k_{t-1} + \frac{k_t^*}{(1+\alpha+\alpha\theta-\alpha\theta\lambda_1)} + \frac{\alpha\theta\gamma}{(1+\alpha+\alpha\theta-\alpha\theta\lambda_1)} E_t \left[\sum_{n=0}^{\infty} \lambda_2^n k_{t+n+1}^* \right]$$

So, one can begin to make inferences about the coefficients:

$$\begin{aligned} \lambda_1 &= \frac{\alpha}{(1+\alpha+\alpha\theta-\alpha\theta\lambda_1)} \\ \gamma &= \frac{1}{(1+\alpha+\alpha\theta-\alpha\theta\lambda_1)} = \frac{\lambda_1}{\alpha} \\ \lambda_2 &= \alpha\theta\gamma = \theta\lambda_1 \end{aligned}$$

The solution is

$$k_t = \lambda k_{t-1} + \frac{\lambda}{\alpha} E_t \left[\sum_{n=0}^{\infty} (\theta\lambda)^n k_{t+n}^* \right] \tag{9.40}$$

where λ solves

$$\lambda (1+\alpha+\alpha\theta-\alpha\theta\lambda) = \alpha \tag{9.41}$$

This can be re-written as

$$\lambda^2 - \left(1 + \frac{1}{\theta} + \frac{1}{\alpha\theta}\right)\lambda + \frac{1}{\theta} = 0 \tag{9.42}$$

so the solution is the same as that derived from the factorization method above.

Personally, I am less fond of this method because it involves guessing the form of the solution, which is a bit of a cheat, because it is still quite algebra-intensive, and because it becomes impractical to apply once one moves to higher-order difference equations. In contrast, the factorization method can be used to characterize the solutions of difference equations of any order.

Part III

Long-Run Growth

Chapter 10

Growth Accounting

The chapters in this section will focus on what is known as "growth theory." Unlike most of macroeconomics, which concerns itself with what happens over the course of the business cycle (why unemployment or inflation go up or down during expansions and recessions), this branch of macroeconomics concerns itself with what happens over longer periods of time. In particular, it looks at the question "What determines the growth rate of the economy over the long run and what can policy measures do to affect it?" As we will also discuss, this is related to the even more fundamental question of what makes some countries rich and others poor.

In this chapter, we will cover what is known as "growth accounting" – a technique for explaining the factors that determine growth.

Production Functions

The usual starting point for growth accounting is the assumption that total real output in an economy is produced using an aggregate production function technology that depends on the total amount of labour and capital used in the economy. For illustration, assume that this takes the form of a Cobb-Douglas production function:

$$Y_t = A_t K_t^{\alpha} L_t^{\beta} \tag{10.1}$$

where K_t is capital input and L_t is labour input. Note that an increase in A_t results in higher output without having to raise inputs. Macroeconomists usually call increases in A_t "technological progress" and often refer to this as the "technology" term. As such, it is easy to imagine increases in A_t to be associated with people inventing new technologies that allow firms to be more productive. Ultimately, however, A_t is simply a measure of productive efficiency and it may go up or down for other reasons, e.g. with the imposition or elimination of government regulations. Because an increase in A_t increases the productiveness of the other factors, it is also sometimes known as *Total Factor Productivity* (TFP), and this is the term most commonly used in empirical papers that attempt to calculate this series.

Usually, we will be more interested in the determination of output per person in the economy, rather than total output. Output per person is often labelled *productivity* by economists with increases in output per worker called *productivity growth*. Productivity is obtained by dividing both sides of equation (10.1) by L_t to get

$$\frac{Y_t}{L_t} = A_t K_t^{\alpha} L_t^{\beta-1} \tag{10.2}$$

which can be re-arranged to give

$$\frac{Y_t}{L_t} = A_t \left(\frac{K_t}{L_t}\right)^{\alpha} L_t^{\alpha+\beta-1} \tag{10.3}$$

This equation shows that there are three potential ways to increase productivity:

- Technological progress: Improving the efficiency with which an economy uses its inputs, i.e. increases in A_t.
- Capital deepening (i.e. increases in capital per worker)
- Increases in the number of workers: Note that this only adds to growth if $\alpha + \beta > 1$, i.e. if there are increasing returns to scale. Most growth theories assumes constant returns

to scale: A doubling of inputs produces a doubling of outputs. If a doubling of inputs manages to more than double outputs, you could argue that the efficiency of production has improved and so perhaps this should be considered an increase in A rather than something that stems from higher inputs. If, there are constant returns to scale, then $\alpha + \beta - 1 = 0$ and this term disappears and production function can be written as

$$\frac{Y_t}{L_t} = A_t \left(\frac{K_t}{L_t}\right)^{\alpha} \tag{10.4}$$

The Determinants of Growth

Let's consider what determines growth with a constant returns to scale Cobb-Douglas production function (so $\beta = 1 - \alpha$)

$$Y_t = A_t K_t^{\alpha} L_t^{1-\alpha} \tag{10.5}$$

and let's assume that time is continuous. In other words, the time element t evolves smoothly instead of just taking integer values like $t = 1$ and $t = 2$.

How do we characterise how this economy grows over time? Let's denote the growth rate of Y_t by G_t^Y. This can be defined as

$$G_t^Y = \frac{1}{Y_t}\frac{dY_t}{dt} \tag{10.6}$$

In other words, the growth rate at any point in time is the change in output (the derivative of output with respect to time, $\frac{dY_t}{dt}$) divided by the level of output. We can characterise the growth rate of Y_t as a function of the growth rates of labour, capital and technology by differentiating the right-hand-side of equation 10.5 with respect to time. Before we do this, you should recall the product rule for differentiation, i.e. that

$$\frac{dAB}{dx} = B\frac{dA}{dx} + A\frac{dB}{dx} \tag{10.7}$$

For products of three variables (like we have in this case) this implies

$$\frac{dABC}{dx} = BC\frac{dA}{dx} + AC\frac{dB}{dx} + AB\frac{dC}{dx} \tag{10.8}$$

In our case, we have

$$\frac{dY_t}{dt} = \frac{dA_t K_t^{\alpha} L_t^{1-\alpha}}{dt} = K_t^{\alpha} L_t^{1-\alpha}\frac{dA_t}{dt} + A_t L_t^{1-\alpha}\frac{dK_t^{\alpha}}{dt} + A_t K_t^{\alpha}\frac{dL_t^{1-\alpha}}{dt} \tag{10.9}$$

We can use the chain rule to calculate the terms involving the impact of changes in capital and labour inputs:

$$\frac{dK_t^{\alpha}}{dt} = \frac{dK_t^{\alpha}}{dK_t}\frac{dK_t}{dt} = \alpha K_t^{\alpha-1}\frac{dK_t}{dt} \tag{10.10}$$

$$\frac{dL_t^{1-\alpha}}{dt} = \frac{dL_t^{1-\alpha}}{dL_t}\frac{dL_t}{dt} = (1-\alpha)\, L_t^{-\alpha}\frac{dL_t}{dt} \tag{10.11}$$

Plugging these formulae into the right places in equation (10.9) we get

$$\frac{dY_t}{dt} = K_t^{\alpha} L_t^{1-\alpha}\frac{dA_t}{dt} + \alpha A_t K_t^{\alpha-1} L_t^{1-\alpha}\frac{dK_t}{dt} + (1-\alpha)\, A_t K_t^{\alpha} L_t^{-\alpha}\frac{dL_t}{dt} \tag{10.12}$$

The growth rate of output is calculated by dividing both sides of this by Y_t which is the same as dividing by $A_t K_t^{\alpha} L_t^{1-\alpha}$.

$$\frac{1}{Y_t}\frac{dY_t}{dt} = \left(\frac{K_t^{\alpha} L_t^{1-\alpha}}{A_t K_t^{\alpha} L_t^{1-\alpha}}\right)\frac{dA_t}{dt} + \alpha\left(\frac{A_t K_t^{\alpha-1} L_t^{1-\alpha}}{A_t K_t^{\alpha} L_t^{1-\alpha}}\right)\frac{dK_t}{dt} + (1-\alpha)\left(\frac{A_t K_t^{\alpha} L_t^{-\alpha}}{A_t K_t^{\alpha} L_t^{1-\alpha}}\right)\frac{dL_t}{dt} \tag{10.13}$$

Cancelling the various terms that appear multiple times in the terms inside the brackets and we get

$$\frac{1}{Y_t}\frac{dY_t}{dt} = \frac{1}{A_t}\frac{dA_t}{dt} + \alpha\frac{1}{K_t}\frac{dK_t}{dt} + (1-\alpha)\,\frac{1}{L_t}\frac{dL_t}{dt} \tag{10.14}$$

This can written in more intuitive form as

$$G_t^Y = G_t^A + \alpha G_t^K + (1-\alpha)\, G_t^L \tag{10.15}$$

The growth rate of output equals the growth rate of the technology term plus a weighted average of capital growth and labour growth, where the weight is determined by the parameter

α. This is the key equation in growth accounting studies. These studies provide estimates of how much GDP growth over a certain period comes from growth in the number of workers, how much comes from growth in the stock of capital and how much comes from improvements in Total Factor Productivity.

One can also show that the growth rate of output per worker is the growth rate of output minus the growth in the number of workers, so this is determined by

$$G_t^Y - G_t^L = G_t^A + \alpha \left(G_t^K - G_t^L \right) \tag{10.16}$$

This is a re-statement in growth rate terms of our earlier decomposition of output growth into technological progress and capital deepening when the production function has constant returns to scale.

It is good to understand how equation 10.15 was derived but, more generally, it is useful to know how to derive growth rates of "Cobb-Douglas"-style variables.

- For example, remember the production function is $Y_t = A_t K_t^{\alpha} L_t^{1-\alpha}$. The reason an increase of x percent in A_t translates into an increase of x percent in output is because A_t multiplies the other terms.

- In contrast, K_t is taken to the power of α. An increase in K_t, say by replacing it with $(1+x) K_t$ is equivalent to multiplying the existing level of output by $(1+x)^{\alpha}$. Because α is assumed to be less than one, this is a smaller increase than comes from increasing A_t by a factor of $(1+x)$.

- To understand why a 1% increase in both K_t and L_t leads to a 1% increase in output, note that if we multiplied both the inputs in $K_t^{\alpha} L_t^{1-\alpha}$ by $(1+x)$, we would get

$$A_t \left((1+x) K_t \right)^{\alpha} \left((1+x) L_t \right)^{1-\alpha} = (1+x)^{\alpha} (1+x)^{1-\alpha} A_t K_t^{\alpha} L_t^{1-\alpha} = (1+x) A_t K_t^{\alpha} L_t^{1-\alpha}$$

How to Calculate the Sources of Growth: Solow (1957)

For most economies, we can calculate GDP, as well as the number of workers and also get some estimate of the stock of capital (this last is a bit trickier and usually relies on assumptions about how investment cumulates over time to add to the stock of capital.) We don't directly observe the value of the Total Factor Productivity term, A_t. However, if we knew the value of the parameter α, we could figure out the growth rate of TFP from the following equation based on re-arranging (10.15)

$$G_t^A = G_t^Y - \alpha G_t^K - (1-\alpha) G_t^L \tag{10.17}$$

But where would we get a value of α from? In a famous 1957 paper, MIT economist Robert Solow pointed out that we could arrive at an estimate of α by looking at the shares of GDP paid to workers and to capital.[1]

To see how this method works, consider the case of a perfectly competitive firm that is seeking to maximise profits. Suppose the firm sells its product for a price P_t (which it has no control over), pays wages to its workers of W_t and rents its capital for a rental rate of R_t (this last assumption—that the firm rents its capital—isn't important for the points that follow but it makes the calculations simpler.) This firm's profits are given by

$$\Pi_t = P_t Y_t - R_t K_t - W_t L_t \tag{10.18}$$

$$= P_t A_t K_t^{\alpha} L_t^{1-\alpha} - R_t K_t - W_t L_t \tag{10.19}$$

Now consider how the firm chooses how much capital and labour to use. It will maximise profits by differentiating the profit function with respect to capital and labour and setting the resulting derivatives equal to zero. This gives two conditions

$$\frac{\partial \Pi_t}{\partial K_t} = \alpha P_t A_t K_t^{\alpha-1} L_t^{1-\alpha} - R_t = 0 \tag{10.20}$$

[1] "Technical Change and the Aggregate Production Function", *Review of Economics and Statistics.*

$$\frac{\partial \Pi_t}{\partial L_t} = (1-\alpha) P_t A_t K_t^{\alpha} L_t^{-\alpha} - W_t = 0 \tag{10.21}$$

These can be simplified to read

$$\frac{\partial \Pi_t}{\partial K_t} = \alpha \frac{P_t Y_t}{K_t} - R_t = 0 \tag{10.22}$$

$$\frac{\partial \Pi_t}{\partial L_t} = (1-\alpha) \frac{P_t Y_t}{L_t} - W_t = 0 \tag{10.23}$$

Solving these we get

$$\alpha = \frac{R_t K_t}{P_t Y_t} \tag{10.24}$$

$$1-\alpha = \frac{W_t L_t}{P_t Y_t} \tag{10.25}$$

Take a close look these equations.

- $P_t Y_t$ is total nominal GDP (the price level times real output)
- $W_t L_t$ is the total amount of income paid out as wages (the wage rate times number of workers).
- $R_t K_t$ is the total amount of income paid to capital (the rental rate times the amount of capital).

These equations tell us that we can calculate $1-\alpha$ as the fraction of income paid to workers rather than to compensate capital. (In real-world economies, non-labour income mainly takes the form of interest, dividends, and retained corporate earnings). National income accounts come with various decompositions. One of them describes how different types of incomes add up to GDP. In most countries, these statistics show that wage income accounts for most of GDP, meaning $\alpha < 0.5$. A standard value that gets used in many studies, based on US estimates, is $\alpha = \frac{1}{3}$. I would note, however, that some studies do this calculation assuming firms are imperfectly competitive – if this is the case (as it is in the real world) then the shares of income earned by labour and capital depend on the degree of monopoly power. So one needs to be cautious about growth accounting calculations as they rely on theoretical assumptions that could potentially be misleading.

Solow's 1957 paper concluded that capital deepening had not been that important for U.S. growth for the period that he examined (1909-1949). In fact, he calculated that TFP growth accounted for 87.5% of growth in output per worker over that period. The calculation became very famous – it was one his papers that was cited by the Nobel committee when awarding Solow the prize for economics in 1987. TFP is sometimes called "the Solow residual" because it is a "backed out" calculation that makes things add up: You calculate it as the part of output growth not due to input growth in the same way as regression residuals in econometrics are the part of the dependent variable not explained by the explanatory variables included in the regression.

Example: The BLS Multifactor Productivity Figures

Most growth accounting calculations are done as part of academic studies. However, in some countries the official statistical agencies produce growth accounting calculations. In the U.S. the Bureau of Labor Statistics (BLS) produces them under the name "multifactor productivity" calculations, (i.e. they use the term MFP instead of the term TFP but conceptually they are the same thing.) Many of the studies add some "bells and whistles" to the basic calculations just described. For example, the BLS try to account for improvements in the "quality" of the labour force by accounting for improvements in the level of educational qualifications and work experience of employees. In other words, they view the production function as being of the form

$$Y_t = A_t K_t^{\alpha} \left(q_t L_t\right)^{1-\alpha} \tag{10.26}$$

where q_t is a measure of the "quality" of the labor input.

Figure 10.1 shows a summary of the BLS's calculations of the sources of growth in the US from 1987 to 2018. They conclude that average growth of 2.0 percent in the U.S. private nonfarm economy can be explained as follows: 0.8 percent comes from capital deepening, 0.4 percent comes from changes in labour composition and 0.8 percent comes from changes in what they call multifactor productivity. Looking at different samples, however, we can see large changes between different periods in the contribution of MFP.

- From 1987-1995, productivity growth averaged only 1.5 percent and MFP growth was weak, contributing only 0.5 percent per year to growth. During this period, there was a lot of discussion about the slowdown in growth relative to previous eras, with much of the focus on the poor performance of TFP growth. Paul Krugman's first popular economics book was called *The Age of Diminished Expectations* because people seemed to have accepted that the US economy was doomed to low productivity growth.

- From 1995-2007, productivity growth averaged a very respectable 2.8 percent, with MFP growth contributing 1.4 percent. During this period, there was a lot of discussion of the impact of new Internet-related technologies that improved efficiency. While the peak of this enthusiasm was around the dot-com bubble of the 2000s when there was lots of talk of a "New Economy", post-tech-bubble productivity performance was also pretty good.

- Since 2007, productivity growth has been weaker than in the previous decade, averaging only 1.3 percent. MFP growth has been particularly weak, averaging only 0.4 percent over this period.

In addition to the poor performance of U.S. productivity growth, another factor that is weighing on the potential for output growth is a slow growth rate of the labour force. After years of increasing numbers of people available for work due to normal population growth, immigration and increased female labour participation, the growth rate of the US labour force has been weaker over the past decade (see Figure 10.2). This is being driven by long-run demographic trends as the large “‘baby boom” generation starts to retire. This trend is set to continue over the next few decades. Figure 3 shows that the dependency ratio (the ratio of non-working to working people) is projected to increase significantly as the populations grows older on average.

In addition to the poor performance of U.S. productivity growth, another factor that is weighing on the potential for output growth is a slow growth rate of the labour force. After years of increasing numbers of people available for work due to normal population growth, immigration and increased female labour participation, the US labour force has flattened out (see Figure 10.2). This is being driven by long-run demographic trends as the large “‘baby boom” generation starts to retire. This trend is set to continue over the next few decades. Figure 10.3 shows that the dependency ratio (the ratio of non-working to working people) is projected to increase significantly as the population grows older on average.

Figure 10.1: Growth Accounting Calculations for the U.S.

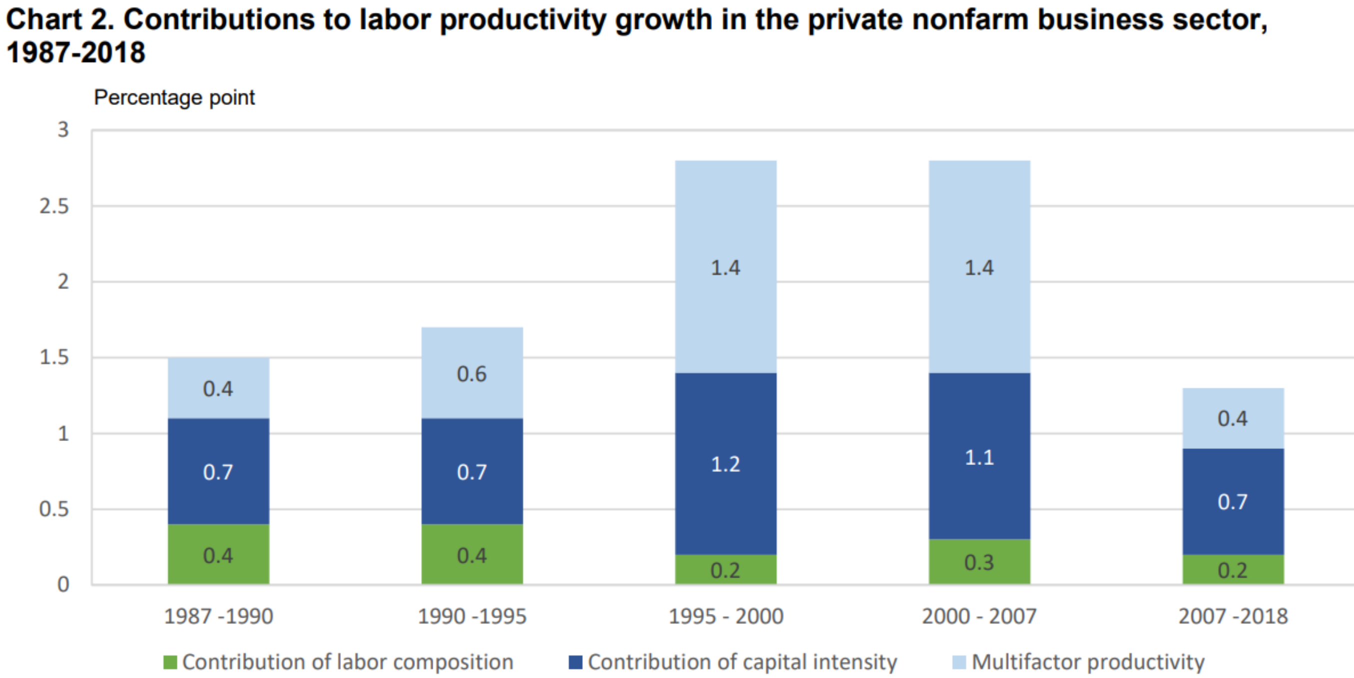


Figure 10.2: The U.S. Labour Force

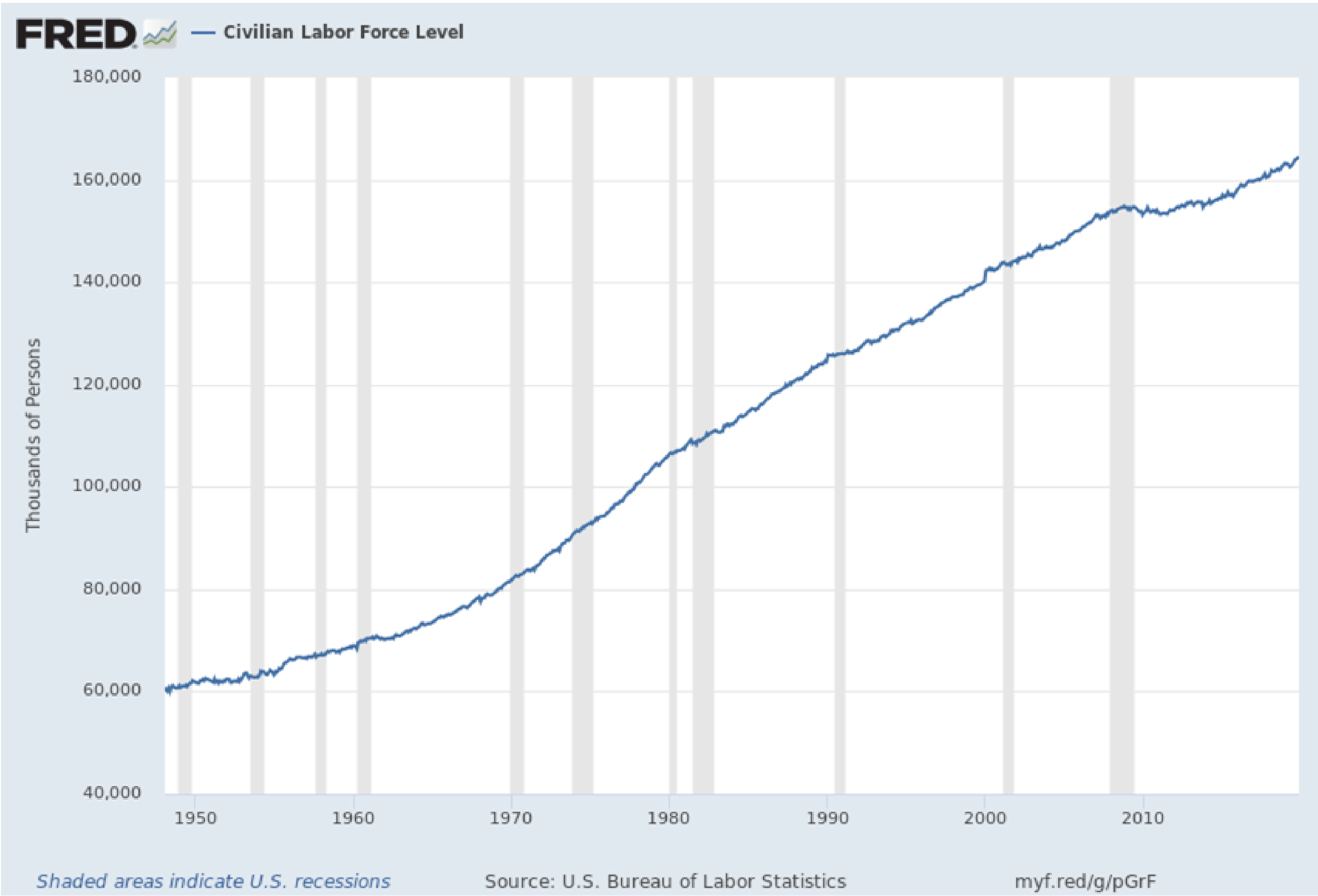
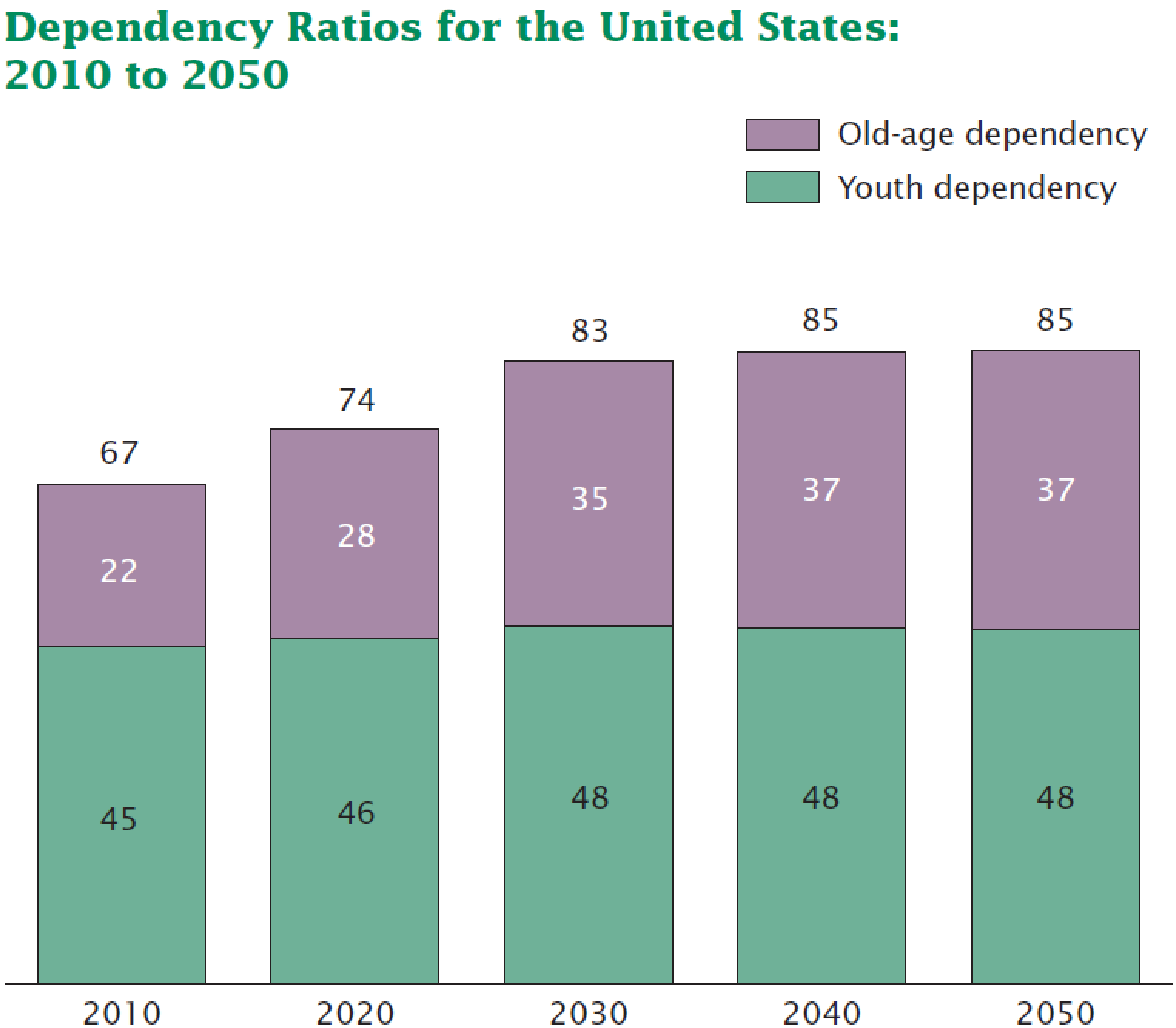


Figure 10.3: The Ratio of Non-Working to Working People in U.S.



Example: The Euro Area

Longer-term growth prospects in Europe appear to be worse than in the United States. My paper with Kieran McQuinn ("Europe's Long-Term Growth Prospects: With and Without Structural Reforms") reports a growth accounting analysis for the euro area and constructs longer-term growth projections. The following discussion is based on this work.

Figure 10.4 shows that growth in output per worker in the countries that make up the euro area has gradually declined over time. In particular, TFP growth has collapsed. From 2.7 percent per year over 1970-76, TFP growth has fallen to an average of 0.2 percent per year over the period 2000-2016. Figure 10.5 shows that weak performances for TFP growth was seen widely across different European countries.

Europe is also going through significant demographic change that will reduce the potential for GDP growth: See Figure 10.6. Population growth is slowing and total population is set to peak in before the middle of this century. The population is also ageing significantly. Indeed the total amount of people aged between 15 and 64 (i.e. the usual definition of work-age population) has peaked and is set to decline substantially over the next half century. Maintaining growth rates at close to those experienced historically will likely require policy changes aimed at increasing the size of the labour force (such as raising retirement ages and immigration) and boosting productivity.

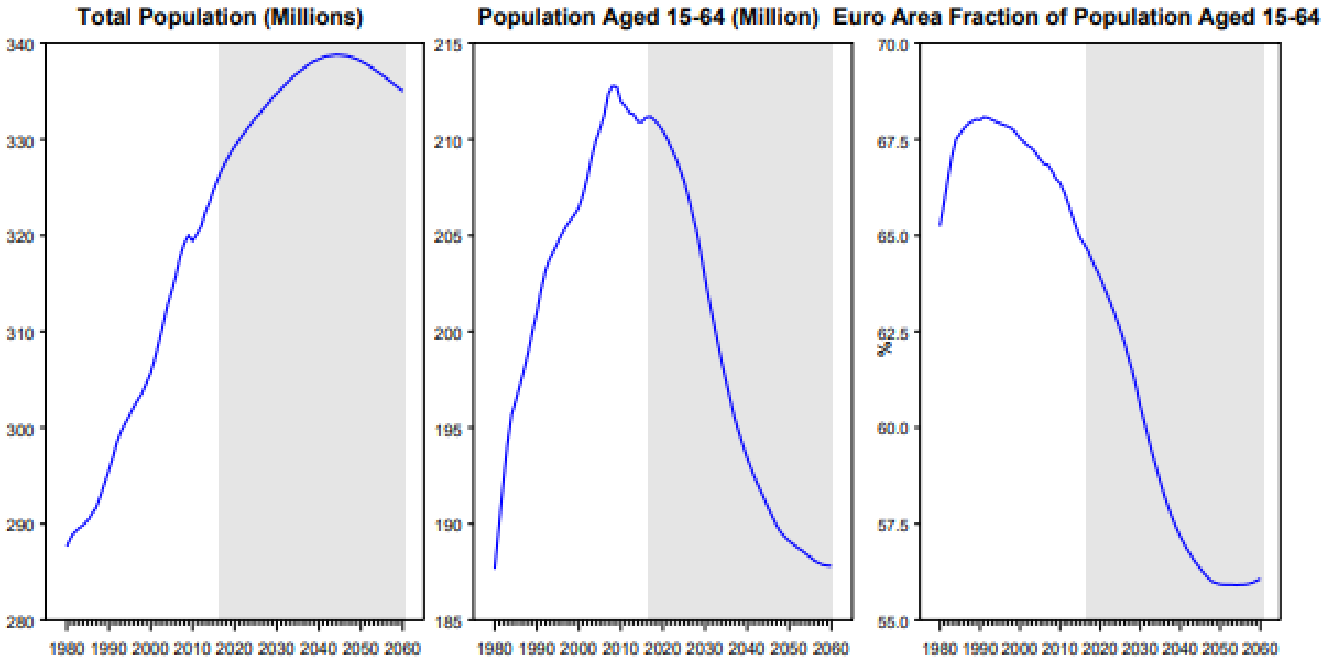
Figure 10.4: The Euro Area's Growth Performance

	Euro Area			
Period	$\triangle y$	$\triangle a$	$\triangle k$	$\triangle l$
1970-1976	3.6	2.7	1.5	-0.6
1977-1986	2.1	1.6	0.8	-0.4
1987-1996	2.3	1.5	0.8	0.0
1997-2006	2.2	0.7	0.8	0.7
2007-2016	0.3	0.1	0.5	-0.2

Figure 10.5: Country-by-Country Growth Performance 2000-2016

Country	2007-2016				2000-2016			
	$\triangle y$	$\triangle a$	$\triangle k$	$\triangle l$	$\triangle y$	$\triangle a$	$\triangle k$	$\triangle l$
Belgium	0.8	-0.1	0.7	0.1	1.4	0.3	0.7	0.3
Germany	1.0	0.2	0.4	0.4	1.2	0.4	0.5	0.3
France	0.6	0.1	0.6	-0.1	1.1	0.2	0.7	0.3
Greece	-3.4	-1.7	0.0	-1.8	-0.2	-0.1	0.5	-0.6
Ireland	3.3	2.9	1.0	-0.7	4.1	2.0	1.6	0.5
Italy	-0.8	-0.7	0.2	-0.3	0.1	-0.6	0.5	0.1
Spain	-0.1	0.4	0.6	-1.0	1.5	-0.1	1.0	0.6
United Kingdom	0.9	-0.2	0.6	0.5	1.7	0.4	0.9	0.5
Finland	-0.4	-0.5	0.5	-0.4	1.2	0.5	0.6	0.0
Lux	2.1	-0.9	1.3	1.6	2.9	0.2	1.3	1.4
Portugal	-0.5	0.2	0.1	-0.8	0.3	0.1	0.6	-0.5
Austria	0.7	-0.2	0.5	0.3	1.4	0.4	0.6	0.4
Netherlands	0.6	0.3	0.6	-0.3	1.2	0.4	0.7	0.1

Figure 10.6: Demographic Projections for the Euro Area from Eurostat



Example: A Tale of Two Cities

Alwyn Young's 1992 paper "A Tale of Two Cities: Factor Accumulation and Technical Change in Hong Kong and Singapore" is an interesting example of a growth accounting study.[2] He compared the growth experiences of these two small Asian economies from the early 1970s to 1990. Young explained his motivation for picking these two economies in terms of their similarities and their differences:

> In the prewar era, both economies were British colonies that served as entrepot trading ports, with little domestic manufacturing activity ... In the postwar era, however, both economies developed large export-dependent domestic manufacturing sectors. Both economies have passed through a similar set of industries, moving from textiles, to clothing, to plastics, to electronics, and then, in the 1980s, gradually moving from manufacturing into banking and financial services ... The postwar population of both was composed primarily of immigrant Chinese from Southern China ... While the Hong Kong government has emphasized a policy of laissez faire, the Singaporean government has, since the early 1960s, pursued the accumulation of physical capital via forced national saving."

Both economies were successful: Hong Kong had total growth of 147% between the early 1970s and 1990 and Singapore had growth of 154%. But Young was interested in exploring the extent to which TFP contributed to growth in these two economies. The results of his growth accounting calculations are shown on the next page. He found that Singapore's approach did not produce any TFP growth while Hong Kong's more free market approach lead to strong TFP growth with this element accounting for almost half of the growth in output per worker.

[2] Available at www.nber.org/chapters/c10990.pdf

One can argue this was a better outcome because Hong Kong achieved the growth without having to divert a huge part of national income towards investment rather than consumption. As we will see later, however, TFP-based growth has an advantage over growth based on capital accumulation because it is more sustainable.

Table from Alwyn Young's 1992 Paper

Table 5 CRUDE ESTIMATE OF TOTAL FACTOR PRODUCTIVITY GROWTH

Time period	Growth of: Output	Labor	Capital	Average capital share	Percentage contribution of: Labor	Capital	TFP Δ
Hong Kong							
71–76	0.406	0.165	0.447	0.330	0.27	0.36	0.36
76–81	0.512	0.253	0.527	0.386	0.30	0.40	0.30
81–86	0.294	0.095	0.388	0.421	0.19	0.55	0.26
86–90	0.260	0.036	0.237	0.414	0.08	0.38	0.54
71–90	1.472	0.549	1.599	0.384	0.23	0.42	0.35
Singapore							
70–75	0.454	0.247	1.005	0.553	0.24	1.22	−0.47
75–80	0.408	0.256	0.503	0.548	0.28	0.68	0.04
80–85	0.300	0.069	0.620	0.491	0.12	1.01	−0.13
85–90	0.383	0.252	0.273	0.468	0.35	0.33	0.31
70–90	1.545	0.825	2.402	0.533	0.25	0.83	−0.08

Chapter 11

The Solow Model

We have discussed how economic growth can come from either capital deepening (increased amounts of capital per worker) or from improvements in total factor productivity (sometimes termed technological progress). This suggests that economic growth can come about from saving and investment (so that the economy accumulates more capital) or from improvements in productive efficiency. In this chapter, we consider a model that explains the role these two elements play in generating sustained economic growth. The model is also due to Robert Solow, whose work on growth accounting we discussed before, and was first presented in his 1956 paper "A Contribution to the Theory of Economic Growth."

The Solow Model's Assumptions

The Solow model assumes that output is produced using a production function in which output depends upon capital and labour inputs as well as a technological efficiency parameter, A.

$$Y_t = AF(K_t, L_t) \tag{11.1}$$

It is assumed that adding capital and labour raises output

$$\frac{\partial Y_t}{\partial K_t} > 0 \tag{11.2}$$

$$\frac{\partial Y_t}{\partial L_t} > 0 \tag{11.3}$$

However, the model also assumes there are diminishing marginal returns to capital accumulation. In other words, adding extra amounts of capital gives progressively smaller and smaller increases in output. This means the second derivative of output with respect to capital is negative.

$$\frac{\partial^2 Y_t}{\partial K_t} < 0 \tag{11.4}$$

See Figure 11.1 for an example of how output can depend on capital with diminishing returns. Think about why diminishing marginal returns is probably sensible: If a firm acquires an extra unit of capital, it will probably be able to increase its output. But if the firm keeps piling on extra capital without raising the number of workers available to use this capital, the increases in output will probably taper off. A firm with ten workers would probably like to have at least ten computers. It might even be helpful to have a few more; perhaps a few laptops for work from home or some spare computers in case others break down. But at some point, just adding more computers doesn't help so much.

We will use a very stylized description of the other parts of this economy: This helps us to focus in on the important role played by diminishing marginal returns to capital. We assume a closed economy with no government sector or international trade. This means all output takes the form of either consumption or investment

$$Y_t = C_t + I_t \tag{11.5}$$

and that savings equals investment

$$S_t = Y_t - C_t = I_t \tag{11.6}$$

The economy's stock of capital is assumed to change over time according to

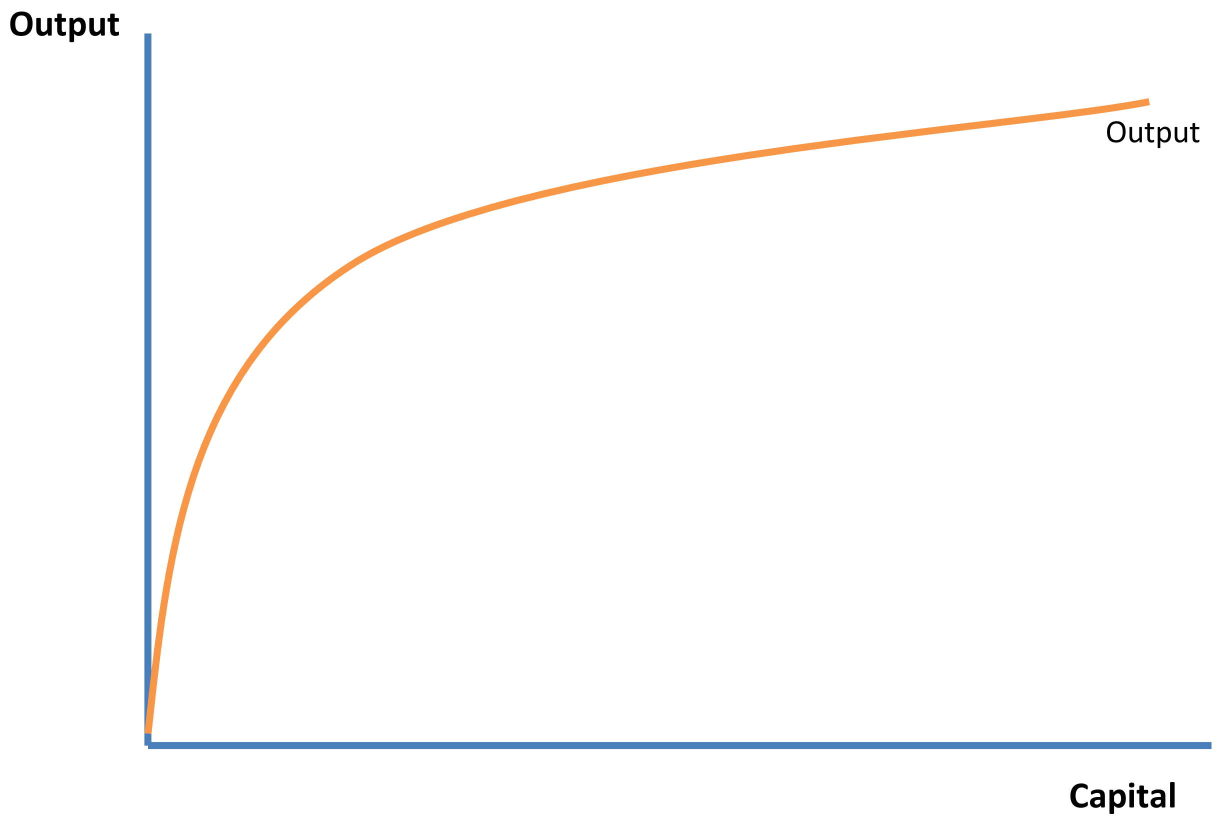
$$\frac{dK_t}{dt} = I_t - \delta K_t \tag{11.7}$$

In other words, the addition to the capital stock each period depends positively on investment and negatively on depreciation, which is assumed to take place at rate δ.

The Solow model does not attempt to model the consumption-savings decision. Instead it assumes that consumers save a constant fraction s of their income

$$S_t = sY_t \tag{11.8}$$

Figure 11.1: Diminishing Marginal Returns to Capital



Capital Dynamics in the Solow Model

Because savings equals investment in the Solow model, equation 11.8 means that investment is also a constant fraction of output

$$I_t = sY_t \tag{11.9}$$

which means we can re-state the equation for changes in the stock of capital

$$\frac{dK_t}{dt} = sY_t - \delta K_t \tag{11.10}$$

Whether the capital stock expands, contracts or stays the same depends on whether investment is greater than, equal to or less than depreciation.

$$\frac{dK_t}{dt} > 0 \quad \text{if} \quad \delta K_t < sY_t \tag{11.11}$$

$$\frac{dK_t}{dt} = 0 \quad \text{if} \quad \delta K_t = sY_t \tag{11.12}$$

$$\frac{dK_t}{dt} < 0 \quad \text{if} \quad \delta K_t > sY_t \tag{11.13}$$

In other words, if the ratio of capital to output is such that

$$\frac{K_t}{Y_t} = \frac{s}{\delta} \tag{11.14}$$

then the stock of capital will stay constant. If the capital-ouput ratio is lower than this level, then the capital stock will be increasing and if it is higher than this level, it will be decreasing.

Figure 11.2 provides a graphical illustration of this process. Depreciation is a simple straight-line function of the stock of capital while output is a curved function of capital, featuring diminishing marginal returns. When the level of capital is low, sY_t is greater than δK. As the capital stock increases, the additional investment due to the extra output tails off but the additional depreciation does not, so at some point sY_t equals δK and the stock of capital stops increasing. Figure 11.2 labels the particular point at which the capital stock remains unchanged as K^*. At this point, we have $\frac{K_t}{Y_t} = \frac{s}{\delta}$.

In the same way, if we start out with a high stock of capital, then depreciation, δK, will tend to be greater than investment, sY_t. This means the stock of capital will decline. When it reaches K^* it will stop declining. This an example of what economists call *convergent dynamics*. For any fixed set of the model parameters (s and δ) and other inputs into the production function (A_t and L_t) there will be a defined level of capital such that, no matter where the capital stock starts, it will converge over time towards this level.

Figure 11.3 provides an illustration of how the convergent dynamics determine the level of output in the Solow model. It shows output, investment and depreciation as a function of the capital stock. The gap between the green line (investment) and the orange line (output) shows the level of consumption. The economy converges towards the level of output associated with the capital stock K^*.

An Increase in the Savings Rate

Now consider what happens when the economy has settled down at an equilibrium unchanging level of capital K_1 and then there is an increase in the savings rate from s_1 to s_2.

Figure 11.4 shows what happens to the dynamics of the capital stock. The line for investment shifts upwards: For each level of capital, the level of output associated with it translates into more investment. So the investment curve shifts up from the green line to the red line. Starting at the initial level of capital, K_1, investment now exceeds depreciation. This means the capital stock starts to increase. This process continues until capital reaches its new equilibrium level of K_2 (where the red line for investment intersects with the black line for depreciation.) Figure 11.4 illustrates how output increases after this increase in the savings rate.

Figure 11.2: Capital Dynamics in The Solow Model

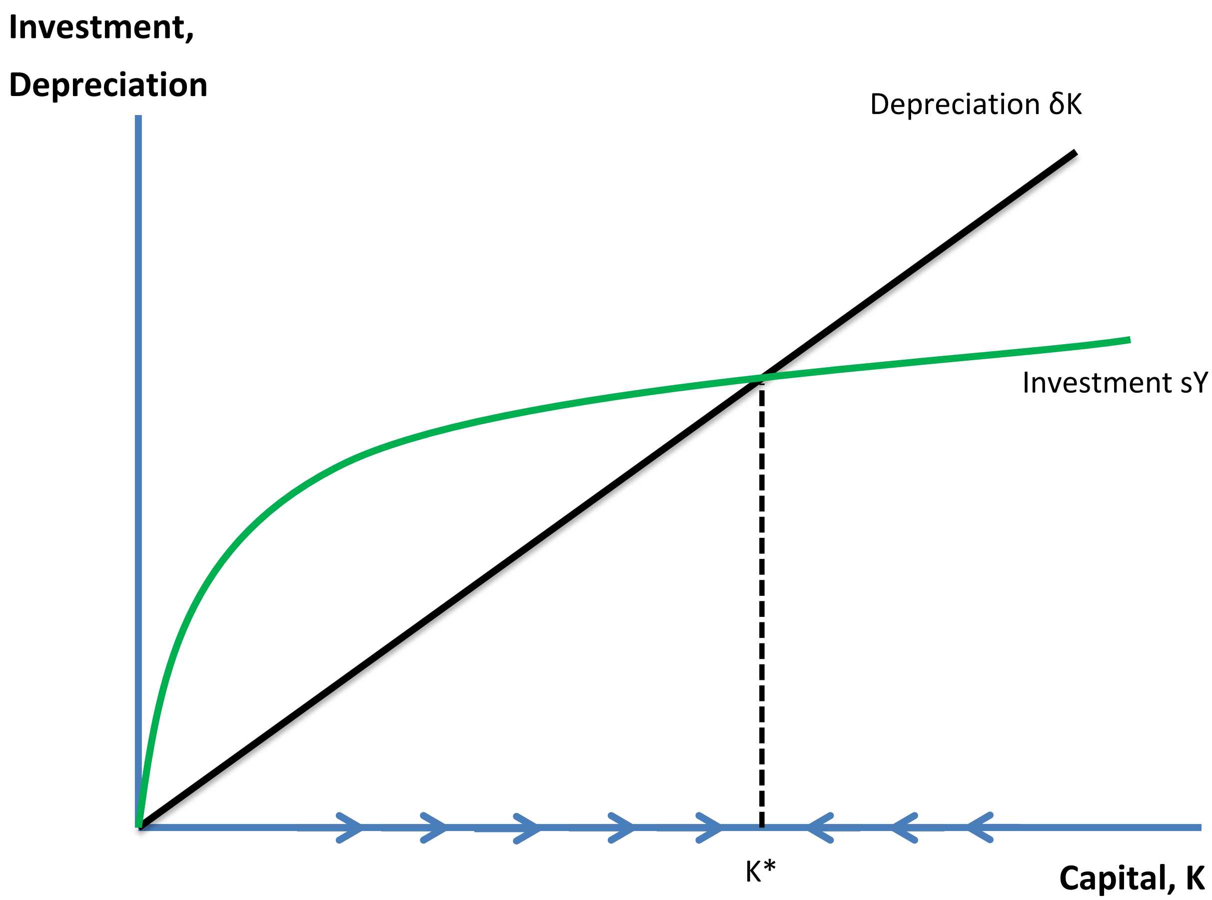


Figure 11.3: Capital and Output in the Solow Model

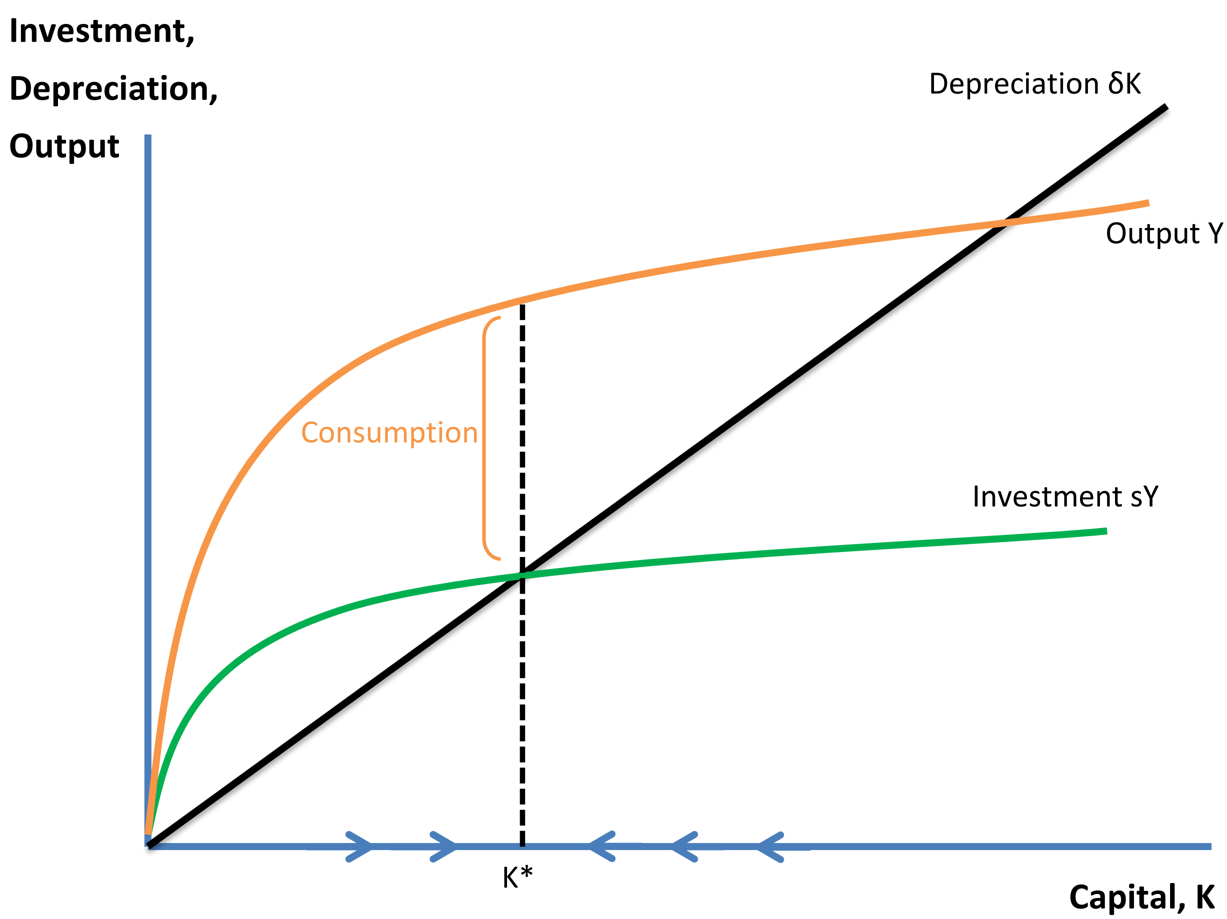


Figure 11.4: An Increase in the Saving Rate

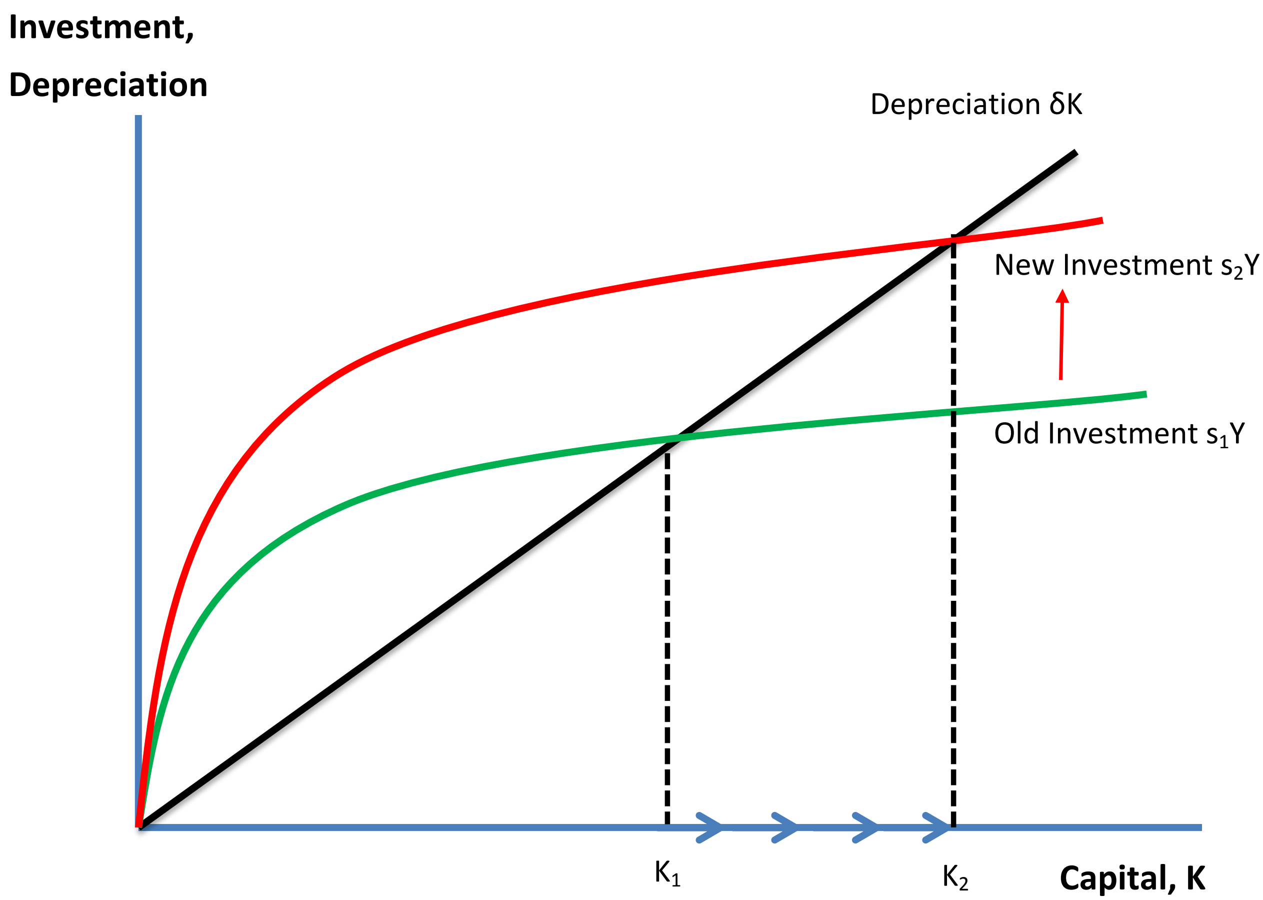
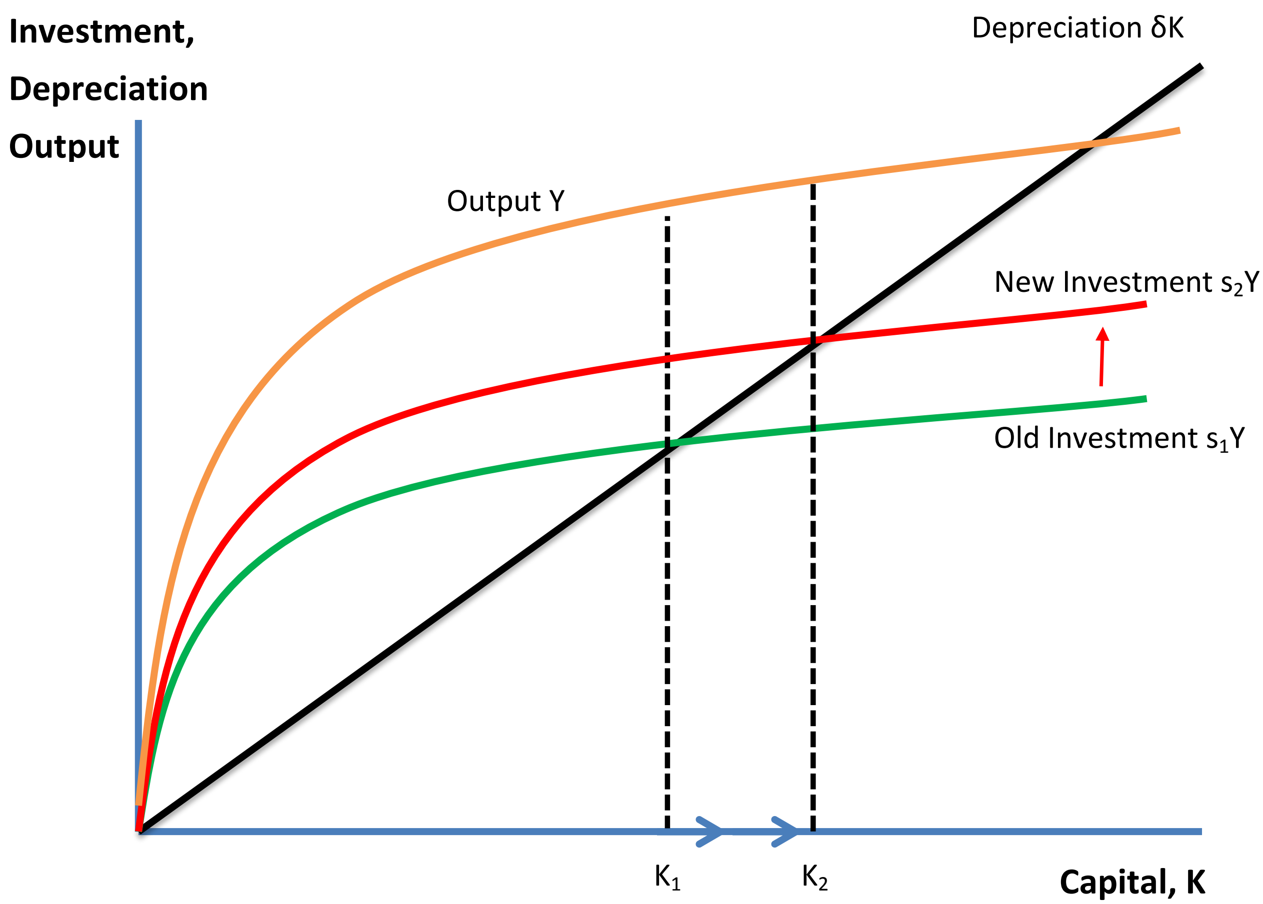


Figure 11.5: Effect on Output of Increased Saving



An Increase in the Depreciation Rate

Now consider what happens when the economy has settled down at an equilibrium unchanging level of capital K_1 and then there is an increase in the depreciation rate from δ_1 to δ_2.

Figure 11.6 shows what happens in this case. The depreciation schedule shifts up from the black line associated with the original depreciation rate, δ_1, to the new red line associated with the new depreciation rate, δ_2. Starting at the initial level of capital, K_1, depreciation now exceeds investment. This means the capital stock starts to decline. This process continues until capital falls to its new equilibrium level of K_2 (where the red line for depreciation intersects with the green line for investment.) So the increase in the depreciation rate leads to a decline in the capital stock and in the level of output.

An Increase in Technological Efficiency

Now consider what happens when technological efficiency A_t increases. Because investment is given by

$$I_t = sY_t = sAF(K_t, L_t) \tag{11.15}$$

a one-off increase in A thus has the same effect as a one-off increase in s. Capital and output gradually rise to a new higher level. Figure 11.7 shows the increase in capital due to an increase in technological efficiency.

Figure 11.6: An Increase in Depreciation

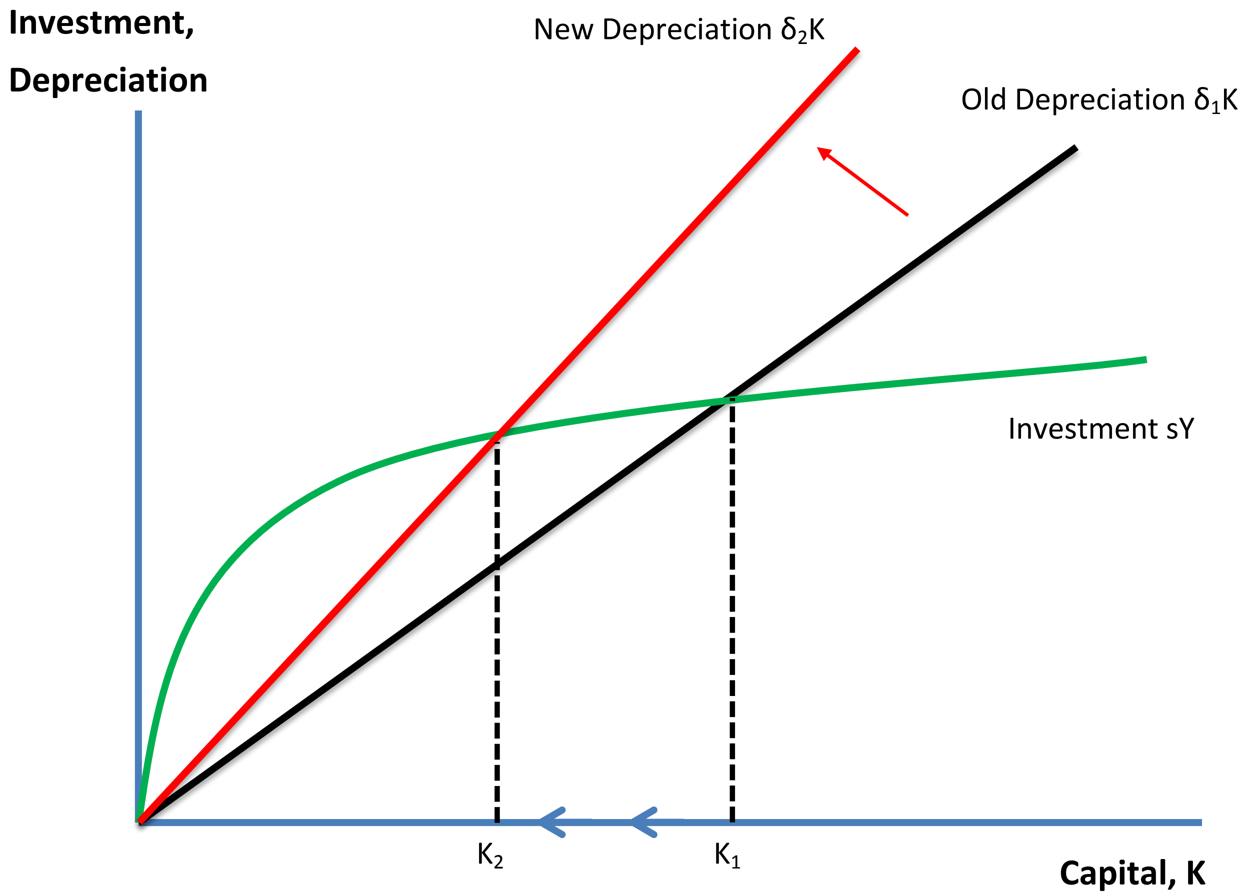
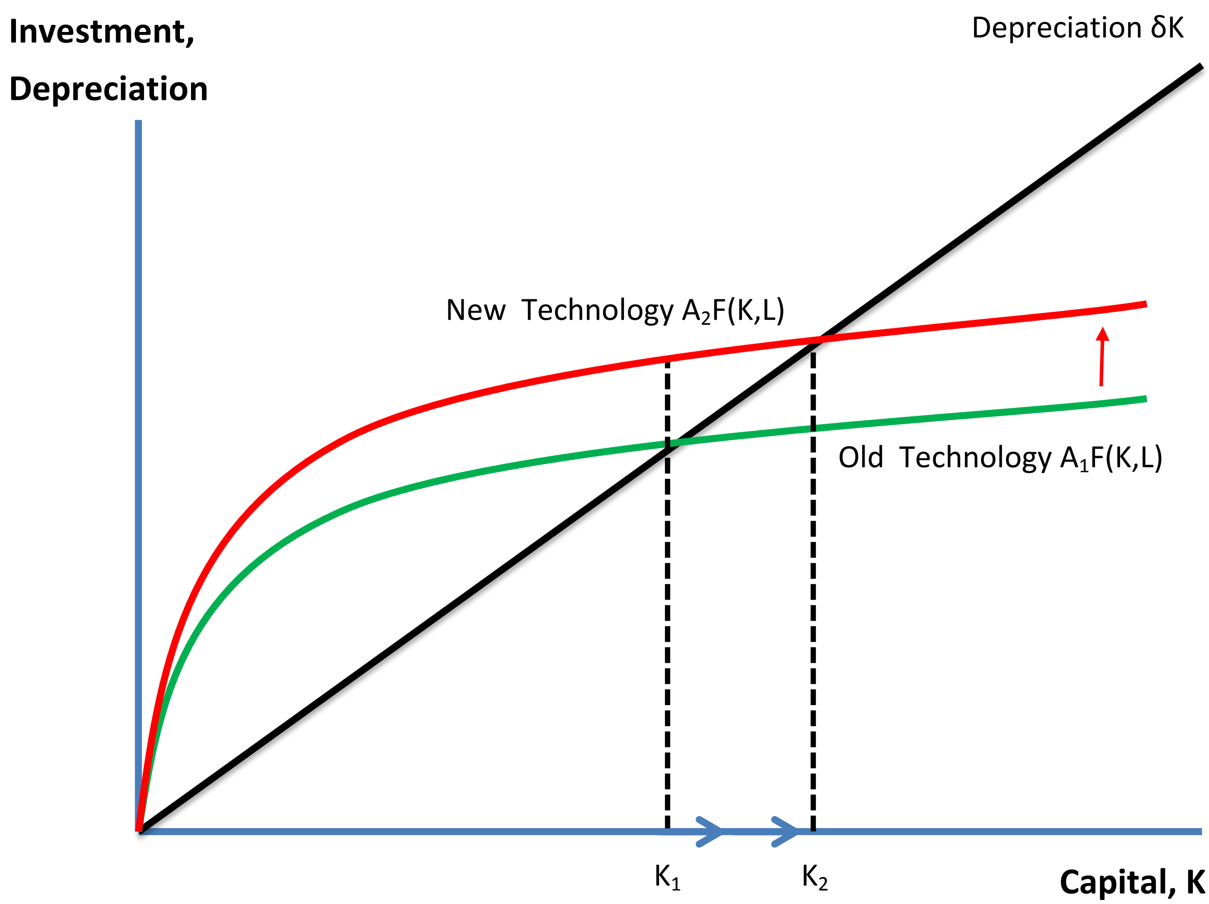


Figure 11.7: An Increase in Technological Efficiency



Solow and the Sources of Growth

In the previous chapter, we described how capital deepening and technological progress were the two sources of growth in output per worker. Specifically, we derived an equation in which output growth was a function of growth in the capital stock, growth in the number of workers and growth in technological efficiency.

Our previous discussion had pointed out that a one-off increase in technological efficiency, A_t, had the same effects as a one-off increase in the savings rate, s. However, there are important differences between these two types of improvements. The Solow model predicts that economies can only achieve a temporary boost to economic growth due to a once-off increase in the savings rate. If they want to sustain economic growth through this approach, then they will need to keep raising the savings rate. However, there are likely to be limits in any economy to the fraction of output that can be allocated towards saving and investment, particularly if it is a capitalist economy in which savings decisions are made by private citizens.

Unlike the savings rate, which will tend to have an upward limit, there is no particular reason to believe that technological efficiency A_t has to have an upper limit. Indeed, growth accounting studies tend to show steady improvements over time in A_t in most countries. Going back to Young's paper on Hong Kong and Singapore discussed in the previous chapter, you can see now why it matters whether an economy has grown due to capital deepening or TFP growth. The Solow model predicts that a policy of encouraging growth through more capital accumulation will tend to tail off over time producing a once-off increase in output per worker. In contrast, a policy that promotes the growth rate of TFP can lead to a sustained higher growth rate of output per worker.

The Capital-Output Ratio with Steady Growth

Up to now, we have only considered once-off changes in output. Here, however, we consider how the capital stock behaves when the economy grows at steady constant rate G^Y. Specifically, we can show in this case that the ratio of capital to output will tend to converge to a specific value. Recall from the previous chapter that if we have something of the form

$$Z_t = U_t^\alpha W_t^\beta \tag{11.16}$$

then we have the following relationship between the various growth rates

$$G_t^Z = \alpha G_t^U + \beta G_t^W \tag{11.17}$$

The capital output ratio $\frac{K_t}{Y_t}$ can be written as $K_t Y_t^{-1}$. So the growth rate of the capital-output ratio can be written as

$$G_t^{\frac{K}{Y}} = G_t^K - G_t^Y \tag{11.18}$$

Starting from equation 11.10, the growth rate of the capital stock can be written as

$$G_t^K = \frac{1}{K_t}\frac{dK_t}{dt} = s\frac{Y_t}{K_t} - \delta \tag{11.19}$$

so the growth rate of the capital-output ratio is

$$G_t^{\frac{K}{Y}} = s\frac{Y_t}{K_t} - \delta - G^Y \tag{11.20}$$

This gives a slightly different form of convergence dynamics from those we saw earlier. This equation shows that the growth rate of the capital-output ratio depends negatively on the level of this ratio. This means the capital-output ratio displays convergent dynamics. When it is above a specific equilibrium value it tends to fall and when it is below this equilibrium value it tends to increase. Thus, the ratio is constantly moving towards this equilibrium value.

We can express this formally as follows:

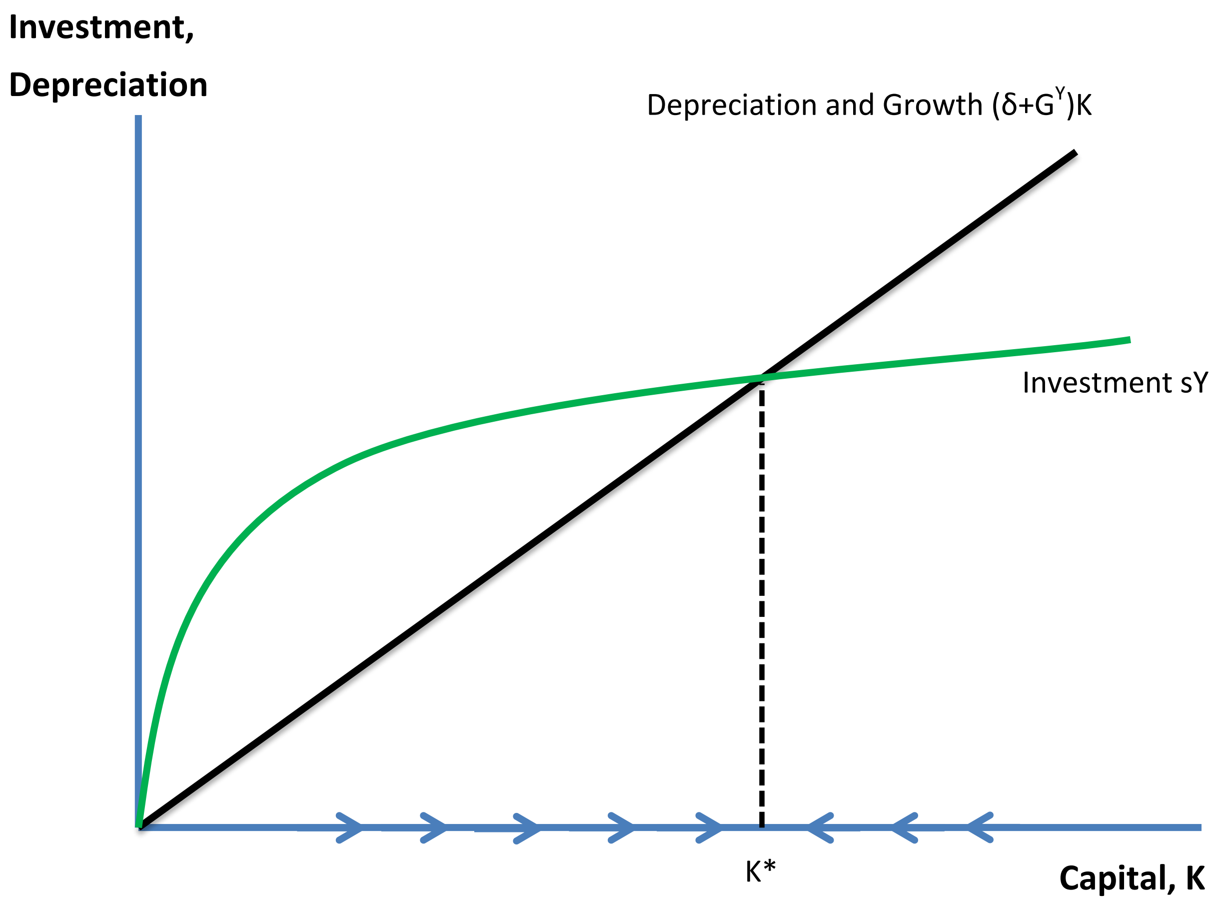
$$G_t^{\frac{K}{Y}} > 0 \quad \text{if} \quad \frac{K_t}{Y_t} < \frac{s}{\delta + G^Y} \tag{11.21}$$

$$G_t^{\frac{K}{Y}} = 0 \quad \text{if} \quad \frac{K_t}{Y_t} = \frac{s}{\delta + G^Y} \tag{11.22}$$

$$G_t^{\frac{K}{Y}} < 0 \quad \text{if} \quad \frac{K_t}{Y_t} > \frac{s}{\delta + G^Y} \tag{11.23}$$

We can illustrate these dynamics using a slightly altered version of our earlier graph. Figure 11.8 changes the depreciation line to be the amount of capital necessary not just to replace depreciation but also to have a percentage increase in the capital stock that matches the increase in output. The diagram shows that the economy will tend to move towards a capital stock such that $sY_t = \left(\delta + G^Y\right) K_t$ meaning the capital-output ratio is $\frac{K_t}{Y_t} = \frac{s}{\delta + G^Y}$.

Figure 11.8: The Equilibrium Capital Stock in a Growing Economy



Briefly, Back to Piketty

In Chapter 6, we discussed one of Thomas Piketty's explanations for why capital may tend to growth faster than income (the $r > g$ argument). Piketty has a different argument for why capital may grow faster than income that relates to the result we have just derived.

In his book, Piketty describes a different assumption about savings in the economy from the one we have just derived. Specifically, he works with a net savings rate, $\tilde{s}$, which is defined as follows

$$I_t - \delta K_t = \tilde{s} Y_t \tag{11.24}$$

In other words, defined like this, $\tilde{s}$ is a savings rate that subtracts off the share of GDP taken up by capital depreciation. In the same way, net national product is defined as GDP minus depreciation. Given this definition, we can write the change in the capital stock as

$$\delta K_t = \tilde{s} Y_t \tag{11.25}$$

Repeating the calculations from above with this model, the growth rate of capital

$$G_t^K = \frac{1}{K_t} \frac{dK_t}{dt} = \frac{I_t - \delta K_t}{K_t} \tag{11.26}$$

becomes

$$G_t^K = \frac{1}{K_t} \frac{dK_t}{dt} = \tilde{s} \frac{Y_t}{K_t} \tag{11.27}$$

So the growth rate of the capital-output ratio is

$$G_t^{\frac{K}{Y}} = \tilde{s} \frac{Y_t}{K_t} - G^Y \tag{11.28}$$

This gives a convergence dynamics in terms of this net savings rate.

$$G_t^{\frac{K}{Y}} > 0 \quad \text{if} \quad \frac{K_t}{Y_t} < \frac{\tilde{s}}{G^Y} \tag{11.29}$$

$$G_t^{\frac{K}{Y}} = 0 \quad \text{if} \quad \frac{K_t}{Y_t} = \frac{\tilde{s}}{G^Y} \tag{11.30}$$

$$G_t^{\frac{K}{Y}} < 0 \quad \text{if} \quad \frac{K_t}{Y_t} > \frac{\tilde{s}}{G^Y} \tag{11.31}$$

So the capital output ratio converges to $\frac{K_t}{Y_t} = \frac{\tilde{s}}{G_Y}$. Again showing his gift for grand terminology, Piketty calls this result *the second fundamental law of capitalism.* His research has argued that growth appears to be slowing around the world and thus, with G^Y in the denominator heading towards zero, the capital-output ratio is likely to be ever-rising.

This prediction will, of course, also hold for the standard Solow model formulation in which the capital-output ratio converges to $\frac{s}{\delta+G^Y}$. The most obvious difference, however, is that Piketty's formulation suggests that when G^Y tends towards zero that we could see the capital-output ratio head towards infinity because his steady-state ratio does not have the δ in the denominator. However, this is somewhat misleading. In the standard formulation of the model, you can show that the net savings rate along a steady growth path will be

$$\tilde{s} = \frac{I_t}{Y_t} - \delta\frac{K_t}{Y_t} \tag{11.32}$$

$$= s - \frac{s\delta}{G_Y + \delta} \tag{11.33}$$

$$= \frac{sG_Y}{G_Y + \delta} \tag{11.34}$$

So when output growth goes to zero, the net savings rate also goes to zero. This means we shouldn't just look at Piketty's formula of $\frac{\tilde{s}}{G_Y}$ for the steady-state capital-output ratio and imagine the denominator (G_Y) heading to zero while the numerator $\tilde{s}$ is fixed. From this discussion, you can take that slower output growth is likely to raise the ratio of capital to income, but it is not likely to head towards infinity!

Why Growth Accounting Can Be Misleading

Of the cases just considered in which output and capital both increase—an increase in the savings rate and an increase in the level of TFP—the evidence points to increases in TFP being more important as a generator of long-term growth. Rates of savings and investment for most countries tend to stay within certain ranges while large increases in TFP over time have been recorded for many countries. It's worth noting then that growth accounting studies can perhaps be a bit misleading when considering the ultimate sources of growth.

Consider a country that has a constant share of GDP allocated to investment but is experiencing steady growth in TFP. The Solow model predicts that this economy should experience steady increases in output per worker and increases in the capital stock. A growth accounting exercise may conclude that a certain percentage of growth stems from capital accumulation but ultimately, in this case, all growth (including the growth in the capital stock) actually stems from growth in TFP. The moral here is that pure accounting exercises may miss the ultimate cause of growth.

Krugman on "The Myth of Asia's Miracle"

I encourage you to read Paul Krugman's 1994 article "The Myth of Asia's Miracle."[1] It discusses a number of examples of cases where economies where growth was based on largely on capital accumulation. In addition to the various Asian countries covered in Alwyn Young's research, Krugman (correctly) predicted a slowdown in growth in Japan, even though at the time many US commentators were focused on the idea that Japan was going to overtake US levels of GDP per capita.

[1] www.foreignaffairs.com/articles/50550/paul-krugman/the-myth-of-asias-miracle

Perhaps most interesting is his discussion of growth in the Soviet Union. Krugman notes that the Soviet economy grew strongly after World War 2 and many in the West believed they would become more prosperous than capitalist economies. The Soviet Union's achievement in placing the first man in space provoked Kennedy's acceleration in the space programme, mainly to show the U.S. was not falling behind communist systems. However, some economists that had examined the Soviet economy were less impressed. Here's an extended quote from Krugman's article:

> When economists began to study the growth of the Soviet economy, they did so using the tools of growth accounting. Of course, Soviet data posed some problems. Not only was it hard to piece together usable estimates of output and input (Raymond Powell, a Yale professor, wrote that the job "in many ways resembled an archaeological dig"), but there were philosophical difficulties as well. In a socialist economy one could hardly measure capital input using market returns, so researchers were forced to impute returns based on those in market economies at similar levels of development. Still, when the efforts began, researchers were pretty sure about what: they would find. Just as capitalist growth had been based on growth in both inputs and efficiency, with efficiency the main source of rising per capita income, they expected to find that rapid Soviet growth reflected both rapid input growth and rapid growth in efficiency.
>
> But what they actually found was that Soviet growth was based on rapid–growth in inputs–end of story. The rate of efficiency growth was not only unspectacular, it was well below the rates achieved in Western economies. Indeed, by some estimates, it was virtually nonexistent.

> The immense Soviet efforts to mobilize economic resources were hardly news. Stalinist planners had moved millions of workers from farms to cities, pushed millions of women into the labor force and millions of men into longer hours, pursued massive programs of education, and above all plowed an ever-growing proportion of the country's industrial output back into the construction of new factories.
>
> Still, the big surprise was that once one had taken the effects of these more or less measurable inputs into account, there was nothing left to explain. The most shocking thing about Soviet growth was its comprehensibility.
>
> This comprehensibility implied two crucial conclusions. First, claims about the superiority of planned over market economies turned out to be based on a misapprehension. If the Soviet economy had a special strength, it was its ability to mobilize resources, not its ability to use them efficiently. It was obvious to everyone that the Soviet Union in 1960 was much less efficient than the United States. The surprise was that it showed no signs of closing the gap.
>
> Second, because input-driven growth is an inherently limited process, Soviet growth was virtually certain to slow down. Long before the slowing of Soviet growth became obvious, it was predicted on the basis of growth accounting.

The Soviet leadership did a good job for a long time of hiding from the world that their economy had stopped growing but ultimately the economic failures of the centrally planning model (combined with its many political and ethnic tensions) ended in a dramatic implosion of the communist system in Russia and the rest of Eastern Europe.

A Formula for Steady Growth

All of the results so far apply for any production function with diminishing marginal returns to capital. However, we can also derive some useful results by making specific assumptions about the form of the production function. Specifically, we will consider the constant returns to scale Cobb-Douglas production function

$$Y_t = A_t K_t^{\alpha} L_t^{1-\alpha} \tag{11.35}$$

This means output growth is determined by

$$G_t^Y = G_t^A + \alpha G_t^K + (1-\alpha) G_t^L \tag{11.36}$$

Now consider the case in which the growth rate of labour input is fixed at n

$$G_t^L = n \tag{11.37}$$

and the growth rate of total factor productivity is fixed at g.

$$G_t^A = g \tag{11.38}$$

The formula for output growth becomes

$$G_t^Y = g + \alpha G_t^K + (1-\alpha) n \tag{11.39}$$

This means all variations in the growth rate of output are due to variations in the growth rate for capital. If output is growing at a constant rate, then capital must also be growing at a constant rate. And we know that the capital-output ratio tends to move towards a specific equilibrium value. So along a steady growth path, the growth rate of output equals the growth rate of capital. Thus, the previous equation can be re-written

$$G_t^Y = g + \alpha G_t^Y + (1-\alpha) n \tag{11.40}$$

which can be simplified to

$$G_t^Y = \frac{g}{1-\alpha} + n \tag{11.41}$$

The growth rate of output per worker is

$$G_t^Y - n = \frac{g}{1-\alpha} \tag{11.42}$$

So the economy tends to converge towards a steady growth path and the growth rate of output per worker along this path is $\frac{g}{1-\alpha}$. Without growth in technological efficiency, there can be no steady growth in output per worker.

A Useful Formula for Output Per Worker

In this case of the Cobb-Douglas production function, output per worker can be written as

$$\frac{Y_t}{L_t} = A_t \left(\frac{K_t}{L_t} \right)^{\alpha} \tag{11.43}$$

In other words, output per worker is a function of technology and of capital per worker. A drawback of this representation is that we know that increases in A_t also increase capital per worker, so this has the misleading implications about the role of capital accumulation discussed above. It is useful, then, to derive an alternative characterisation of output per worker, one that we will use again. First, we'll define the capital-output ratio as

$$x_t = \frac{K_t}{Y_t} \tag{11.44}$$

So, the production function can be expressed as

$$Y_t = A_t \left(x_t Y_t \right)^{\alpha} L_t^{1-\alpha} \tag{11.45}$$

Here, we are using the fact that

$$K_t = x_t Y_t \tag{11.46}$$

Dividing both sides of this expression by Y_t^{α}, we get

$$Y_t^{1-\alpha} = A_t x_t^{\alpha} L_t^{1-\alpha} \tag{11.47}$$

Taking both sides of the equation to the power of $\frac{1}{1-\alpha}$ we arrive at

$$Y_t = A_t^{\frac{1}{1-\alpha}} x_t^{\frac{\alpha}{1-\alpha}} L_t \tag{11.48}$$

So, output per worker is

$$\frac{Y_t}{L_t} = A_t^{\frac{1}{1-\alpha}} x_t^{\frac{\alpha}{1-\alpha}} \tag{11.49}$$

This equation states that all fluctuations in output per worker are due to either changes in technological progress or changes in the capital-output ratio. When considering the relative role of technological progress or policies to encourage accumulation, we will see that this decomposition is more useful than equation 11.43 because the level of technology does not affect x_t in the long run while it does affect $\frac{K_t}{L_t}$. So, this decomposition offers a cleaner picture of the part of growth due to technology and the part that is not.

A Formal Model of Convergence Dynamics

Because A_t is assumed to grow at a constant rate each period, this means that all of the interesting dynamics for output per worker in this model stem from the behaviour of the capital-output ratio. We will now describe in more detail how this ratio behaves. Before doing so, I want to introduce a new piece of terminology.

A useful mathematical shorthand that saves us from having to write down derivatives with respect to time everywhere is to write

$$\dot{Y}_t = \frac{dY_t}{dt} \tag{11.50}$$

What we are really interested in, though, is *growth rates* of series, so we need to scale this by the level of output itself. Thus, $\frac{\dot{Y}_t}{Y_t}$, and this is our mathematical expression for the growth rate of a series. For our Cobb-Douglas production function, we can use the result we derived earlier to express the growth rate of output as

$$\frac{\dot{Y}_t}{Y_t} = \frac{\dot{A}_t}{A_t} + \alpha \frac{\dot{K}_t}{K_t} + (1-\alpha)\frac{\dot{L}_t}{L_t} \tag{11.51}$$

The Solow model assumes

$$\frac{\dot{A}_t}{A_t} = g \tag{11.52}$$

$$\frac{\dot{N}_t}{N_t} = n \tag{11.53}$$

So this can be re-written as

$$\frac{\dot{Y}_t}{Y_t} = g + \alpha \frac{\dot{K}_t}{K_t} + (1-\alpha)n \tag{11.54}$$

Similarly, because

$$x_t = K_t Y_t^{-1} \tag{11.55}$$

its growth rate can be written as

$$\frac{\dot{x}_t}{x_t} = \frac{\dot{K}_t}{K_t} - \frac{\dot{Y}_t}{Y_t} \tag{11.56}$$

To get an expression for the growth rate of the capital stock, w re-write the capital accumulation equation as

$$\dot{K}_t = sY_t - \delta K_t \tag{11.57}$$

and divide across by K_t on both sides

$$\frac{\dot{K}_t}{K_t} = s\frac{Y_t}{K_t} - \delta \tag{11.58}$$

This means we write, the growth rate of the capital stock as

$$\frac{\dot{K}_t}{K_t} = \frac{s}{x_t} - \delta \tag{11.59}$$

Now using equation (11.54) for output growth and equation (11.59) for capital growth, we can derive a useful equation for the dynamics of the capital-output ratio:

$$\begin{aligned}
\frac{\dot{x}_t}{x_t} &= (1-\alpha)\frac{\dot{K}_t}{K_t} - g - (1-\alpha)n && (11.60)\\
&= (1-\alpha)(\frac{s}{x_t} - \frac{g}{1-\alpha} - n - \delta) && (11.61)
\end{aligned}$$

This dynamic equation has a very important property: The growth rate of x_t depends negatively on the value of x_t. In particular, when x_t is over a certain value, it will tend to decline, and when it is under that value it will tend to increase. This provides a specific illustration of the convergent dynamics of the capital-output ratio.

What is the long-run steady-state value of x_t, which we will label x^*? It is the value consistent with $\frac{\dot{x}}{x} = 0$. This implies that

$$\frac{s}{x^*} - \frac{g}{1-\alpha} - n - \delta = 0 \qquad (11.62)$$

This solves to give

$$x^* = \frac{s}{\frac{g}{1-\alpha} + n + \delta} \qquad (11.63)$$

Given this equation, we can derive a more intuitive-looking expression to describe the convergence properties of the capital-output ratio. The dynamics of x_t are given by

$$\frac{\dot{x}_t}{x_t} = (1-\alpha)(\frac{s}{x_t} - \frac{g}{1-\alpha} - n - \delta) \qquad (11.64)$$

Multiplying and dividing the right-hand-side of this equation by $(\frac{g}{1-\alpha} + n + \delta)$:

$$\frac{\dot{x}_t}{x_t} = (1-\alpha)(\frac{g}{1-\alpha} + n + \delta)\left(\frac{s/x_t - \frac{g}{1-\alpha} - n - \delta}{\frac{g}{1-\alpha} + n + \delta}\right) \qquad (11.65)$$

The last term inside the brackets can be simplified to give

$$\frac{\dot{x}_t}{x_t} = (1-\alpha)(\frac{g}{1-\alpha} + n + \delta)\left(\frac{1}{x_t}\frac{s}{\frac{g}{1-\alpha} + n + \delta} - 1\right) \qquad (11.66)$$

$$= (1-\alpha)(\frac{g}{1-\alpha}+n+\delta)\left(\frac{x^*}{x_t}-1\right) \quad (11.67)$$

$$= (1-\alpha)(\frac{g}{1-\alpha}+n+\delta)\left(\frac{x^*-x_t}{x_t}\right) \quad (11.68)$$

This equation states that each period the capital-output ratio closes a fraction equal to $\lambda = (1-\alpha)(\frac{g}{1-\alpha}+n+\delta)$ of the gap between the current value of the ratio and its steady-state value.

Illustrating Convergence Dynamics

Often, the best way to understand dynamic models is to load them onto the computer and see them run. This is easily done using spreadsheet software such as Excel or econometrics-oriented packages such as RATS. Figures 11.9 to 11.11 provide examples of the behaviour over time of two economies, one that starts with a capital-output ratio that is half the steady-state level, and other that starts with a capital output ratio that is 1.5 times the steady-state level.

The parameters chosen were $s = 0.2, \alpha = \frac{1}{3}, g = 0.02, n = 0.01, \delta = 0.06$. Together these parameters are consistent with a steady-state capital-output ratio of 2. To see, this plug these values into (11.63):

$$x^* = \left(\frac{K}{Y}\right)^* = \frac{s}{\frac{g}{1-\alpha}+n+\delta} = \frac{0.2}{1.5*0.02+0.01+0.06} = 2 \quad (11.69)$$

Figure 11.9 shows how the two capital-output ratios converge, somewhat slowly, over time to their steady-state level. This slow convergence is dictated by our choice of parameters: Our "convergence speed" is:

$$\lambda = (1-\alpha)(\frac{g}{1-\alpha}+n+\delta) = \frac{2}{3}(1.5*0.02+0.01+0.06) = 0.067 \quad (11.70)$$

So, the capital-output ratio converges to its steady-state level at a rate of about 7 percent per period. These are fairly standard parameter values for annual data, so this should be understood to mean 7 percent per year.

Figure 11.10 shows how output per worker evolves over time in these two economies. Both economies exhibit growth, but the capital-poor economy grows faster during the convergence period than the capital-rich economy. These output per worker differentials may seem a little small on this chart, but the Figure 11.11 shows the behaviour of the growth rates, and this chart makes it clear that the convergence dynamics can produce substantially different growth rates depending on whether an economy is above or below its steady-state capital-output ratio. During the initial transition periods, the capital-poor economy grows at rates over 6 percent, while the capital-rich economy grows at under 2 percent. Over time, both economies converge towards the steady-state growth rate of 3 percent.

Figure 11.9: Convergence Dynamics for the Capital-Output Ratio

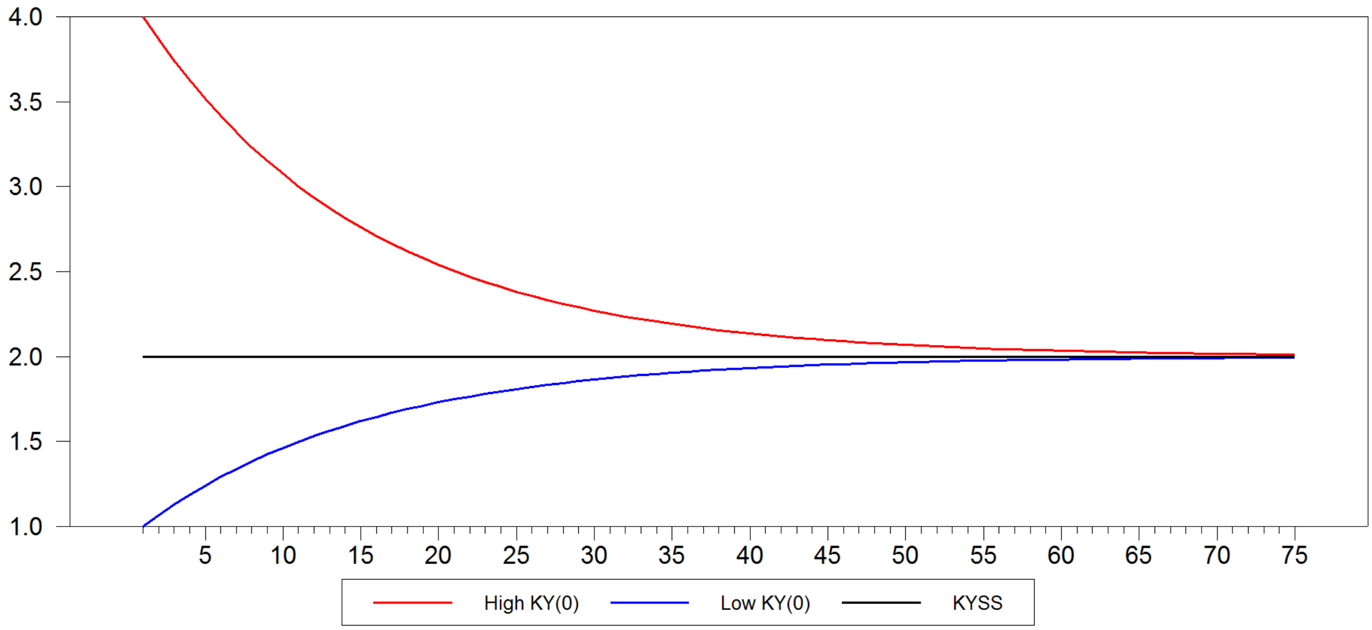


Figure 11.10: Convergence Dynamics for Output Per Worker

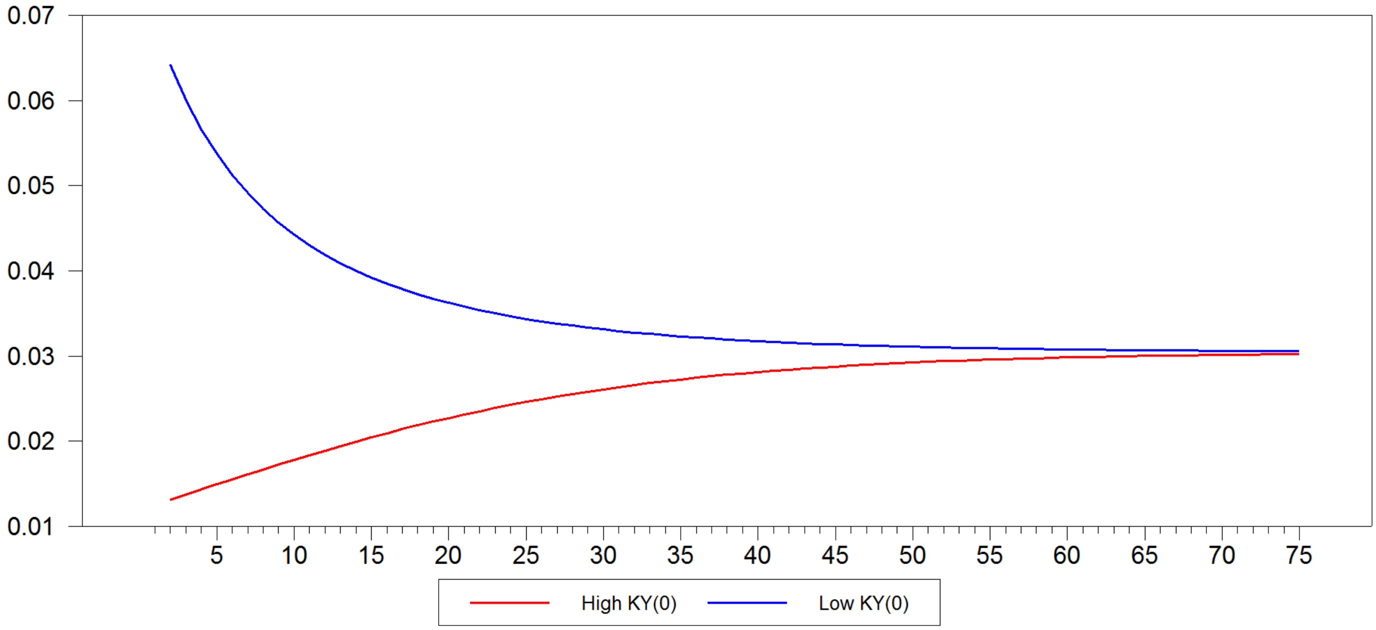
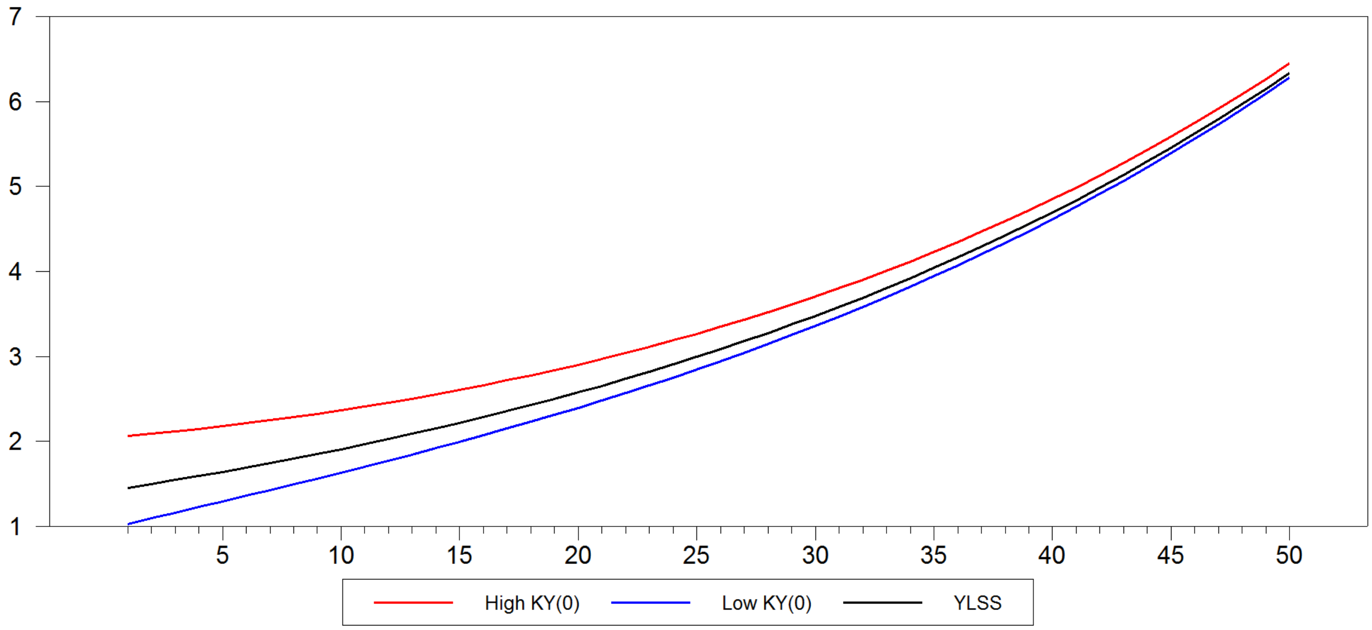


Figure 11.11: Convergence Dynamics for the Growth Rate of Output Per Worker



Illustrating Changes in Key Parameters

Figures 11.12 to 11.14 examine what happens when the economy is moving along the steady-state path consistent with the parameters just given, and then one of the parameters is changed. Specifically, they examine the effects of changes in s, δ and g.

Consider first an increase in the savings rate to $s = 0.25$. This has no effect on the steady-state growth rate. But it does change the steady-state capital-output ratio from 2 to 2.5. So the economy now finds itself with too little capital relative to its new steady-state capital-output ratio. The growth rate jumps immediately and only slowly returns to the long-run 3 percent value. The faster pace of investment during this period gradually brings the capital-output ratio into line with its new steady-state level.

The increase in the savings rate permanently raises the level of output per worker relative to the path that would have occurred without the change. However, for our parameter values, this effect is not that big. This is because the long-run effect of the savings rate on output per worker is determined by $s^{\frac{\alpha}{1-\alpha}}$, which in this case is $s^{0.5}$. So in our case, 25 percent increase in the savings rate produces an 11.8 percent increase in output per worker ($1.25^{0.5} = 1.118$). More generally, a doubling of the savings rate raises output per worker by 41 percent ($2^{0.5} = 1.41$).

The charts also show the effect of an increase in the depreciation rate to $\delta = 0.11$. This reduces the steady-state capital-output ratio to 4/3 and the effects of this change are basically the opposite of the effects of the increase in the savings rate.

Finally, there is the increase in the rate of technological progress. I've shown the effects of a change from $g = 0.02$ to $g = 0.03$. This increases the steady-state growth rate of output per worker to 0.045. However, as the charts show there is another effect: A faster steady-state growth rate for output reduces the steady-state capital-output ratio. Why? The increase in

g raises the long-run growth rate of output; this means that each period the economy needs to accumulate more capital than before just to keep the capital-output ratio constant. Again, without a change in the savings rate that causes this to happen, the capital-output ratio will decline. So, the increase in g means that—as in the depreciation rate example—the economy starts out in period 25 with too much capital relative to its new steady-state capital-output ratio. For this reason, the economy doesn't jump straight to its new 4.5 percent growth rate of output per worker. Instead, after an initial jump in the growth rate, there is a very gradual transition the rest of the way to the 4.5 percent growth rate.

Figure 11.12: Capital-Output Ratios: Effect of Increases In ...

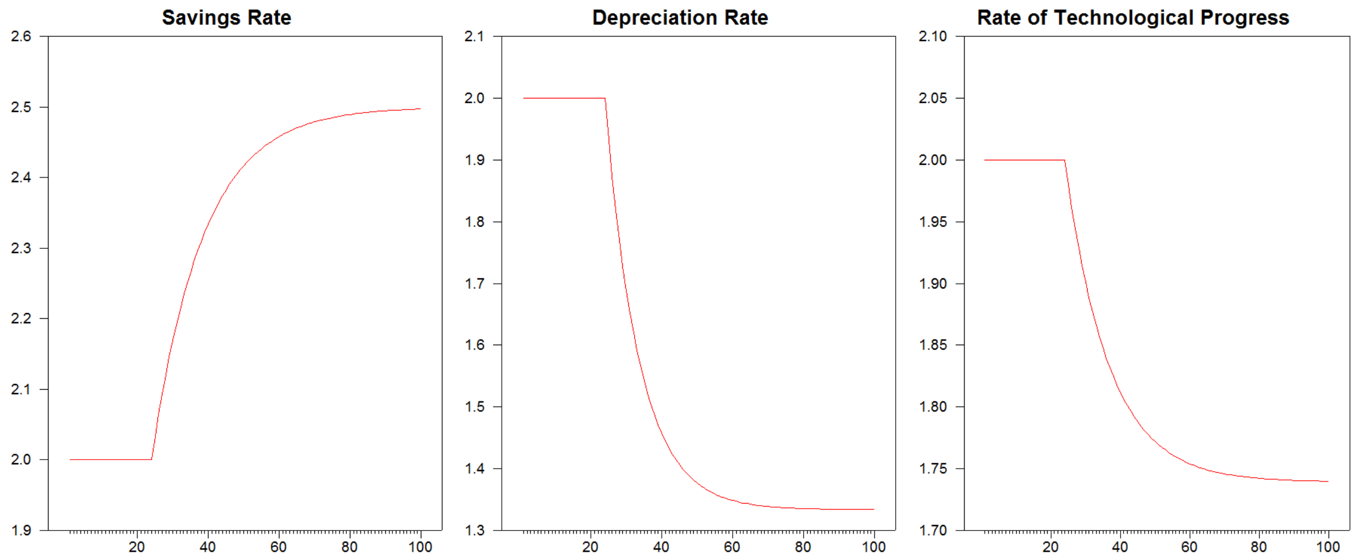


Figure 11.13: Growth Rates of Output Per Hour: Effect of Increases In ...

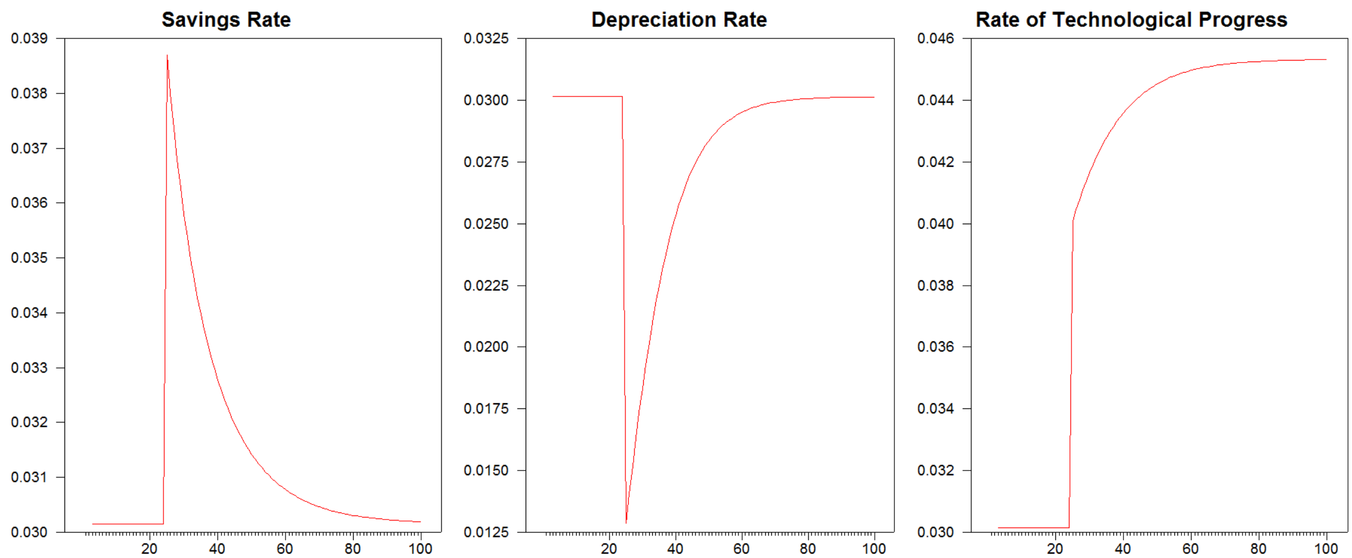
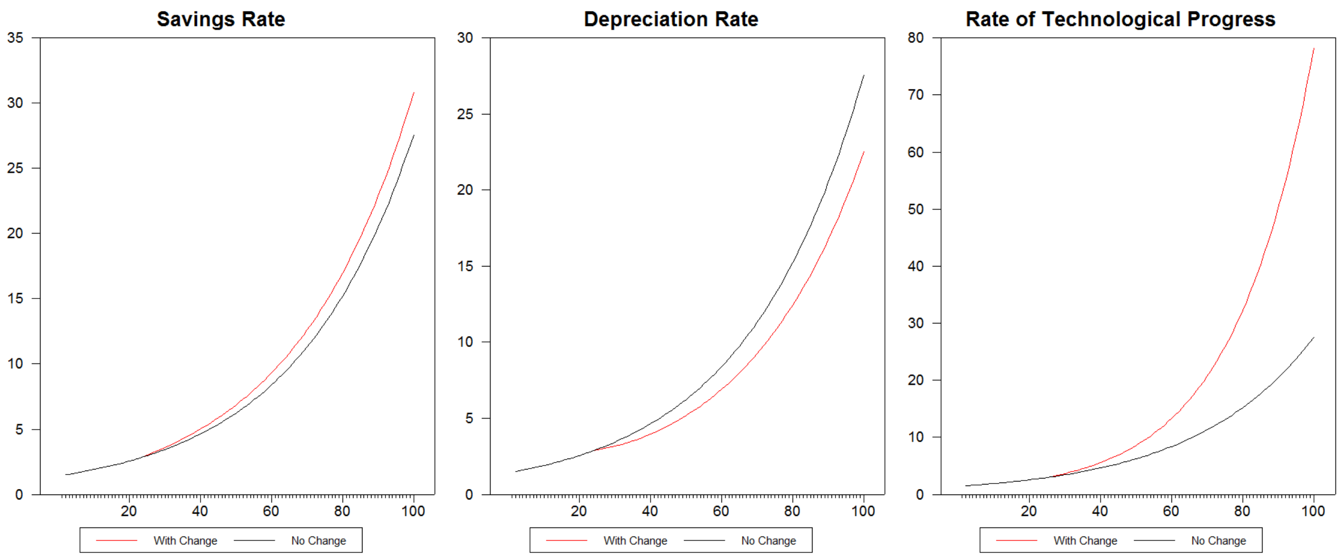


Figure 11.14: Output Per Hour: Effect of Increases In ...



Convergence Dynamics in Practice

The Solow model predicts that no matter what the original level of capital an economy starts out with, it will tend to revert to the equilibrium levels of output and capital indicated by the economy's underlying features. Does the evidence support this idea?

Unfortunately, history has provided a number of extreme examples of economies having far less capital than is consistent with their fundamental features. Wars have provided the "natural experiments" in which various countries have had huge amounts of their capital destroyed. The evidence has generally supported Solow's prediction that economies that experience negative shocks should tend to recover from these setbacks and return to their pre-shock levels of capital and output. For example, both Germany and Japan grew very strongly after the war, recovering prosperity despite the massive damage done to their stocks of capital by war bombing.

A more extreme example, perhaps, is study by Edward Miguel and Gerard Roland of the long-run impact of U.S. bombing of Vietnam in the 1960s and 1970s.[2] Miguel and Roland found large variations in the extent of bombing across the various regions of Vietnam. Despite large differences in the extent of damage inflicted on different regions, Miguel and Roland found little evidence for lasting relative damage on the most-bombed regions by 2002. (Note this is not the same as saying there was no damage to the economy as a whole — the study is focusing on whether those areas that lost more capital than average ended up being poorer than average).

[2] "The Long-Run Impact of Bombing Vietnam, *Journal of Development Economics* Volume 96, pages 1-15, 2011

Chapter 12

Endogenous Technological Change

The Solow model identified technological progress or improvements in total factor productivity (TFP) as the key determinant of growth in the long run, but did not provide any explanation of what determines it. In the technical language used by macroeconomists, long-run growth in the Solow framework is determined by something that is *exogenous* to the model.

In this chapter, we consider a specific model that makes technological progress *endogeous*, meaning determined by the actions of the economic agents described in the model. The model is due to Paul Romer ("Endogenous Technological Change," *Journal of Political Economy*, 1990) who was awarded the Nobel Prize in Economics in 2018 for his work in this area. The model starts by accepting the Solow model's result that technological progress is what determines long-run growth in output per worker. But, unlike the Solow model, Romer provides an explanation of what determines technological progress.

TFP Growth as Invention of New Inputs

So what is this technology term A anyway? The Romer model takes a specific concrete view on this issue. Romer describes the aggregate production function as

$$Y = L_Y^{1-\alpha}\left(x_1^\alpha + x_2^\alpha + ... + x_A^\alpha\right) = L_Y^{1-\alpha}\sum_{i=1}^{A} x_i^\alpha \tag{12.1}$$

where L_Y is the number of workers producing output and the x_i's are different types of capital goods. The crucial feature of this production function is that diminishing marginal returns applies, not to capital as a whole, but separately to each of the individual capital goods (because $0 < \alpha < 1$).

If A was fixed, the pattern of diminishing returns to each of the separate capital goods would mean that growth would eventually taper off to zero. However, in the Romer model, A is not fixed. Instead, there are L_A workers engaged in R&D and this leads to the invention of new capital goods. This is described using a "production function" for the change in the number of capital goods:

$$\dot{A} = \gamma L_A^{\lambda} A^{\phi} \tag{12.2}$$

The change in the number of capital goods depends positively on the number of researchers (λ is an index of how slowly diminishing marginal productivity sets in for researchers) and also on the prevailing value of A itself. This latter effect stems from the "giants shoulders" effect.[1] For instance, the invention of a new piece of software will have relied on the previous invention of the relevant computer hardware, which itself relied on the previous invention of semiconductor chips, and so on.

Romer's model contains a full description of the factors that determines the fraction of workers that work in the research section. The research sector gets rewarded with patents that allow it to maintain a monopoly in the product it invents; wages are equated across sectors, so the research sector hire workers up to point where their value to it is as high as it is to producers of final output. In keeping with the spirit of the Solow model, I'm going to just treat the share of workers in the research sector as an exogenous parameter (but will discuss

[1] Stemming from Isaac Newton's observation "If I have seen farther than others, it is because I was standing on the shoulders of giants."

later some of the factors that should determine this share). So, we have

$$L = L_A + L_Y \tag{12.3}$$

$$L_A = s_A L \tag{12.4}$$

And again we assume that the total number of workers grows at an exogenous rate n:

$$\frac{\dot{L}}{L} = n \tag{12.5}$$

Simplifying the Aggregate Production Function

We can define the aggregate capital stock as

$$K = \sum_{i=1}^{A} x_i \tag{12.6}$$

Again, we'll treat the savings rate as exogenous and assume

$$\dot{K} = s_K Y - \delta K \tag{12.7}$$

One observation that simplifies the analysis of the model is the fact that all of the capital goods play an identical role in the production process. For this reason, we can assume that the demand from producers for each of these capital goods is the same, implying that

$$x_i = \bar{x} \qquad i = 1, 2, ...A \tag{12.8}$$

This means that the production function can be written as

$$Y = A L_Y^{1-\alpha} \bar{x}^{\alpha} \tag{12.9}$$

Note now that

$$K = A\bar{x} \Rightarrow \bar{x} = \frac{K}{A} \tag{12.10}$$

so output can be re-expressed as

$$Y = AL_Y^{1-\alpha}\left(\frac{K}{A}\right)^{\alpha} = (AL_Y)^{1-\alpha}K^{\alpha} \tag{12.11}$$

This looks just like the Solow model's production function. The TFP term is written as $A^{1-\alpha}$ as opposed to just A as it was in our first handout, but this makes no difference to the substance of the model.

Steady-State Growth in The Romer Model

You can use the same arguments as before to show that this economy converges to a steady-state growth path in which capital and output grow at the same rate. So, we can derive the steady-state growth rate as follows. Re-write the production function as

$$Y = (As_Y L)^{1-\alpha}K^{\alpha} \tag{12.12}$$

where

$$s_Y = 1 - s_A \tag{12.13}$$

Our usual procedure for taking growth rates give us

$$\frac{\dot{Y}}{Y} = (1-\alpha)\left(\frac{\dot{A}}{A} + \frac{\dot{s_Y}}{s_Y} + \frac{\dot{L}}{L}\right) + \alpha\frac{\dot{K}}{K} \tag{12.14}$$

Now use the fact that the steady-state growth rates of capital and output are the same to derive that this steady-state growth rate is given by

$$\left(\frac{\dot{Y}}{Y}\right)^{*} = (1-\alpha)\left(\frac{\dot{A}}{A} + \frac{\dot{s_Y}}{s_Y} + \frac{\dot{L}}{L}\right) + \alpha\left(\frac{\dot{Y}}{Y}\right)^{*} \tag{12.15}$$

Finally, because the share of labour allocated to the non-research sector cannot be changing along the steady-state path (otherwise the fraction of researchers would eventually go to zero or become greater than one, which would not be feasible) we have

$$\left(\frac{\dot{Y}}{Y} - \frac{\dot{L}}{L}\right)^* = \frac{\dot{A}}{A} \tag{12.16}$$

The steady-state growth rate of output per worker equals the steady-state growth rate of A. The only difference from the Solow model is that writing the TFP term as $A^{1-\alpha}$ makes this growth rate $\frac{\dot{A}}{A}$ as opposed to $\frac{1}{1-\alpha}\frac{\dot{A}}{A}$.

Deriving the Steady-State Growth Rate

The big difference relative to the Solow model is that the A term is determined within the model as opposed to evolving at some fixed rate unrelated to the actions of the agents in the model economy. To derive the steady-state growth rate in this model, note that the growth rate of the number of capital goods is

$$\frac{\dot{A}}{A} = \gamma \left(s_A L\right)^{\lambda} A^{\phi-1} \tag{12.17}$$

The steady-state of this economy features A growing at a constant rate. This can only be the case if the growth rate of the right-hand-side of (12.17) is zero. Using our usual procedure for calculating growth rates of Cobb-Douglas-style items, we get

$$\lambda\left(\frac{\dot{s}_A}{s_A} + \frac{\dot{L}}{L}\right) - (1-\phi)\frac{\dot{A}}{A} = 0 \tag{12.18}$$

Again, in steady-state, the growth rate of the fraction of researchers $(\frac{\dot{s}_A}{s_A})$ must be zero. So, along the model's steady-state growth path, the growth rate of the number of capital goods (and hence output per worker) is

$$\left(\frac{\dot{A}}{A}\right)^* = \frac{\lambda n}{1-\phi} \tag{12.19}$$

The long-run growth rate of output per worker in this model depends on positively on three factors:

- The parameter λ, which describes the extent to which diminishing marginal productivity sets in as we add researchers.

- The strength of the "standing on shoulders" effect, ϕ. The more past inventions help to boost the rate of current inventions, the faster the growth rate will be.

- The growth rate of the number of workers n. The higher this, the faster the economy adds researchers. This may seem like a somewhat unusual prediction, but it holds well if one takes a very long view of world economic history. Prior to the industrial revolution, growth rates of population and GDP per capita were very low. The past 200 years have seen both population growth and economic growth rates increases. See the figures on the next two pages (the first comes from Greg Clark's book *A Farewell to Alms* which provides a very interesting discussion of pre-Industrial-Revolution economies.)

Figure 12.1: World Economic History

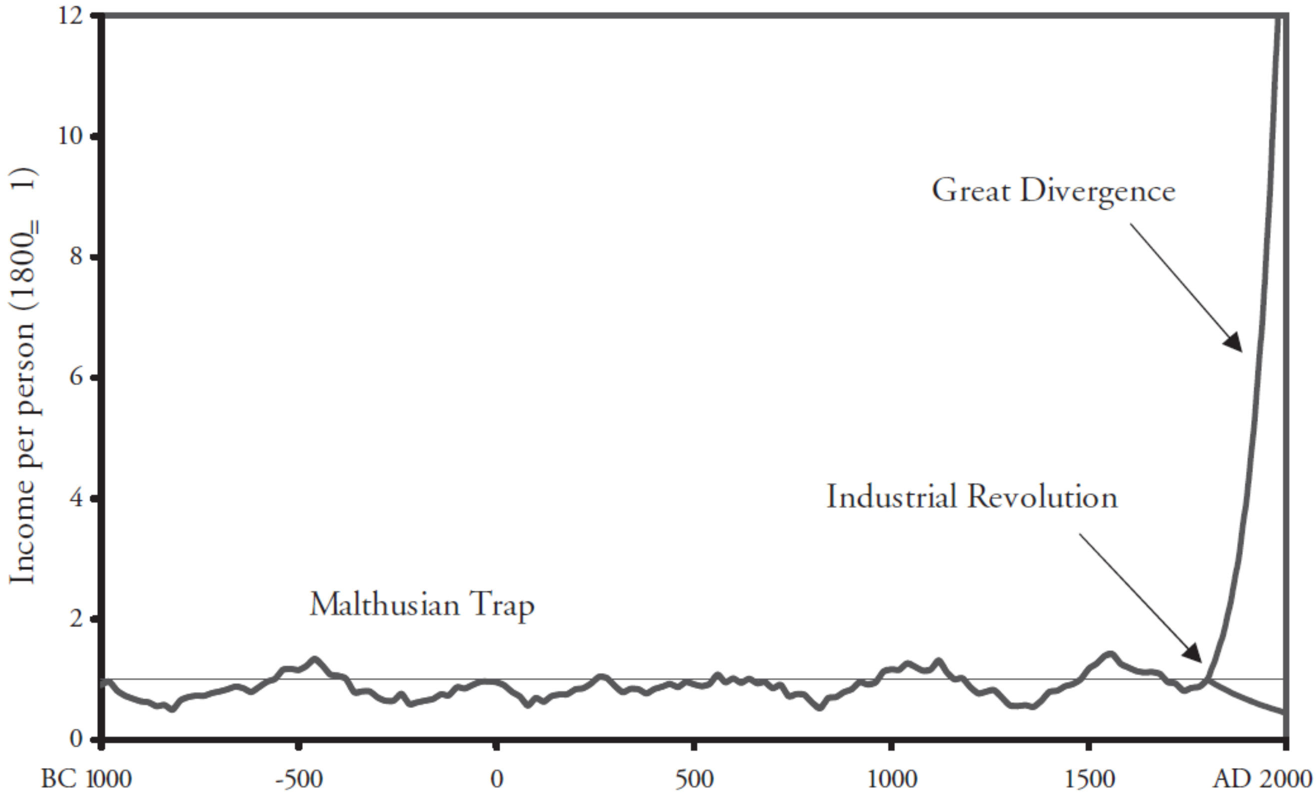
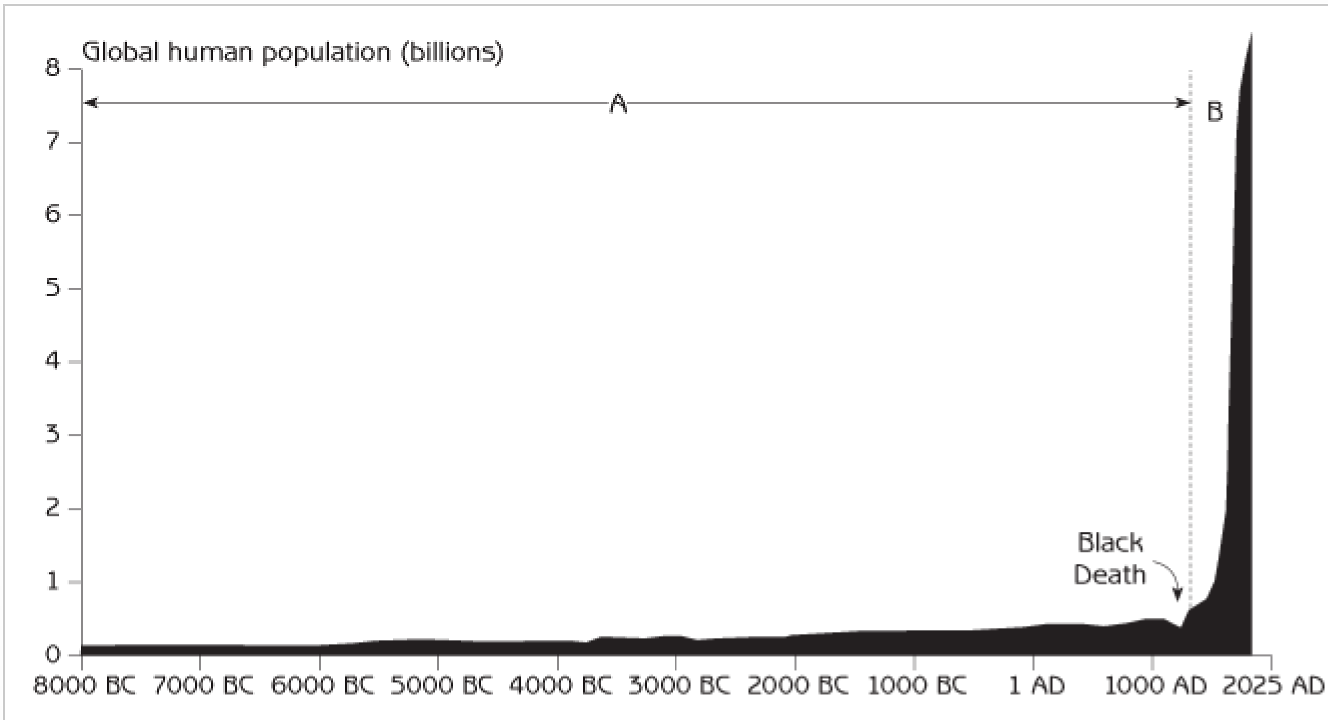


Figure 1.1 World economic history in one picture. Incomes rose sharply in many countries after 1800 but declined in others.

Figure 12.2: Global Population



The Steady-State Level of Output Per Worker

Just as with our discussion of the Solow model, we can decompose output per worker into a capital-output ratio component and a TFP component. In other words, one can re-arrange equation (12.11) to get

$$\frac{Y}{L_Y} = \left(\frac{K}{Y}\right)^{\frac{\alpha}{1-\alpha}} A \tag{12.20}$$

and use the fact that $L_Y = (1 - s_A)L$ to get

$$\frac{Y}{L} = (1 - s_A)\left(\frac{K}{Y}\right)^{\frac{\alpha}{1-\alpha}} A \tag{12.21}$$

Note that the s_A term reflects the reduction in the production of goods and services due to a fraction of the labour force being employed as researchers. One can also use the same arguments to show that, along the steady-state growth path the capital-output ratio is

$$\left(\frac{K}{Y}\right)^* = \frac{s_K}{n + \frac{\lambda n}{1-\phi} + \delta} \tag{12.22}$$

(The $\frac{\lambda n}{1-\phi}$ here takes the place of the $\frac{g}{1-\alpha}$ in the first handout's expression for the steady-state capital-output ratio because this is the new formula for the growth rate of output per worker). Finally, we can also figure out the level of A along the steady-state growth path as follows. Along the steady-state path, we have

$$\frac{\dot{A}}{A} = \gamma \left(s_A L\right)^{\lambda} A^{\phi-1} = \frac{\lambda n}{1-\phi} \tag{12.23}$$

This latter equality can be re-arranged as

$$A^* = \left(\frac{\gamma\left(1-\phi\right)}{\lambda n}\right)^{\frac{1}{1-\phi}} \left(s_A L\right)^{\frac{\lambda}{1-\phi}} \tag{12.24}$$

So, along the steady-state growth path, output per worker is

$$\left(\frac{Y}{L}\right)^* = (1 - s_A)\left(\frac{s_K}{n + \frac{\lambda n}{1-\phi} + \delta}\right)^{\frac{\alpha}{1-\alpha}} \left(\frac{\gamma\left(1-\phi\right)}{\lambda n}\right)^{\frac{1}{1-\phi}} \left(s_A L\right)^{\frac{\lambda}{1-\phi}} \tag{12.25}$$

Convergence Dynamics for A

We noted already that the arguments showing that the capital-output ratio tends to converge towards its steady-state are the same here as in the Solow model. What about the A term? How do we know, for instance, that A always reverts back eventually to the path given by equation (12.24)? To see that this is the case, let

$$g_A = \frac{\dot{A}}{A} = \gamma \left(s_A L\right)^{\lambda} A^{\phi-1} \tag{12.26}$$

The growth rate of the right-hand-side of this equation is

$$\frac{\dot{g}_A}{g_A} = \lambda \left(\frac{\dot{s}_A}{s_A} + n\right) - \left(1 - \phi\right) g_A \tag{12.27}$$

One can use this equation to show that g_A will be falling whenever

$$g_A > \frac{\lambda n}{1 - \phi} + \frac{\lambda}{1 - \phi} \frac{\dot{s}_A}{s_A} \tag{12.28}$$

So, apart from periods when the share of researchers is changing, the growth rate of A will be declining whenever it is greater than its steady-state value of $\frac{\lambda n}{1-\phi}$. The same argument works in reverse when g_A is below its steady-state value. Thus, the growth rate of A displays convergent dynamics, always tending back towards its steady-state value. And equation (12.24) tells us exactly what the *level* of A has to be if the growth rate of A is at its steady-state value.

Optimal R&D?

We haven't discussed the various factors that may determine the share of the labour force allocated to the research sectors, s_A. However, in equation (12.25) we have diagnosed two separate offsetting effects that s_A has on output: A negative one caused by the fact the researchers don't actually produce output, and a positive one due to the positive effect of the share of researchers on the level of technology.

Equation (12.25) looks very complicated but it looks simpler if we just take all the terms that don't involve s_A and bundle them together calling them X and also write $Z = \frac{\lambda}{1-\phi}$. In this case, the equation becomes

$$\left(\frac{Y}{L}\right)^* = X\left(1 - s_A\right)\left(s_A\right)^Z \tag{12.29}$$

Written like this, it is a relatively simple calculus problem to figure out the level of s_A that maximises the level of output per worker along the steady-state growth path. In other words, one can can differentiate equation (12.25) with respect to s_A, set equal to zero, and solve to obtain that this optimizing share of researchers is

$$s_A^{**} = \frac{Z}{1+Z} = \frac{\frac{\lambda}{1-\phi}}{1+\frac{\lambda}{1-\phi}} = \frac{\lambda}{1-\phi+\lambda} \tag{12.30}$$

When one fills in the model to determine s_A endogenously, does the economy generally arrive at this optimal level? No. The reason for this is that research activity generates *externalities* that affect the level of output per worker, but which are not taken into account by private individuals or firms when they make the choice of whether or not to conduct research. Looking at the "ideas" production function, equation (12.2), one can see both positive and negative externalities:

- A positive externality due to the "giants shoulders" effect. Researchers don't take into account the effect their inventions have in boosting the future productivity of other researchers. The higher is θ, the more likely it is that the R&D share will be too low.

- A negative externality due to the fact that $\lambda < 1$, so diminishing marginal productivity applies to the number of researchers.

Whether there is too little or too much research in the economy relative to the optimal level depends on the strength of these various externalities. However, using empirical estimates of the parameters of equation (12.2), Charles Jones and John Williams have calculated that it is far more likely that the private sector will do too little research relative to the social optimum.[2]

To give some insight into this result, note that the steady-state growth rate in this model is $\frac{\lambda n}{1-\phi}$, so $\frac{\lambda}{1-\phi}$ is the ratio of the growth rate of output per worker to the growth rate of population. Suppose this equals one, so growth in output per worker equals growth in population—perhaps a reasonable ballpark assumption. In this case $\frac{\lambda}{1-\phi} = 1$ and the optimal share of researchers is one-half. Indeed, for any reasonable steady-state growth rate, the optimal share of researchers is very high, so it is hardly surprising that the economy does not automatically generate this share.

This result points to the potential for policy interventions to boost the rate of economic growth by raising the number of researchers. For instance, laws to strengthen patent protection may raise the incentives to conduct R&D. This points to a potential conflict between macroeconomic policies aimed at raising growth and microeconomic policies aimed at reducing the inefficiencies due to monopoly power: Some amount of monopoly power for patent-holders

[2] Charles I. Jones and John C. Williams, "Too Much of a Good Thing? The Economics of Investment in R&D", *Journal of Economic Growth*, March 2000, Vol. 5, No. 1, pp. 65-85.

may be necessary if we want to induce a high level of R&D and thus a high level of output.

Robert Gordon on The Past and Future of New Technologies

Many of the facts about economic history back up Romer's vision of economic growth. Robert Gordon's paper "Is US economic growth over? Faltering innovation confronts the six headwinds" provides an excellent description of the various phases of technological invention and also provides an interesting perspective on the potential for future technological progress.[3] Gordon highlights how economic history can be broken into different periods based on how the invention of technologies have impacted the economy.

The First Industrial Revolution: "centered in 1750-1830 from the inventions of the steam engine and cotton gin through the early railroads and steamships, but much of the impact of railroads on the American economy came later between 1850 and 1900. At a minimum it took 150 years for IR1 to have its full range of effects."

The Second Industrial Revolution: "within the years 1870-1900 created within just a few years the inventions that made the biggest difference to date in the standard of living. Electric light and a workable internal combustion engine were invented in a three-month period in late 1879. The number of municipal waterworks providing fresh running water to urban homes multiplied tenfold between 1870 and 1900. The telephone, phonograph, and motion pictures were all invented in the 1880s. The benefits of IR2 included subsidiary and complementary inventions, from elevators, electric machinery and consumer appliances; to the motorcar, truck, and airplane; to highways, suburbs, and supermarkets; to sewers to carry the wastewater away. All this had been accomplished by 1929, at least in urban America, although it took longer

[3]CEPR Policy Insight, Number 63

to bring the modern household conveniences to small towns and farms. Additional follow-up inventions continued and had their main effects by 1970, including television, air conditioning, and the interstate highway system. The inventions of IR2 were so important and far-reaching that they took a full 100 years to have their main effect."

The Third Industrial Revolution: "is often associated with the invention of the web and internet around 1995. But in fact electronic mainframe computers began to replace routine and repetitive clerical work as early as 1960."

Gordon's paper is very worth reading for understanding how the innovations associated with the "second industrial revolution" completely altered people's lives. He describes life in 1870 as follows

> most aspects of life in 1870 (except for the rich) were dark, dangerous, and involved backbreaking work. There was no electricity in 1870. The insides of dwelling units were not only dark but also smoky, due to residue and air pollution from candles and oil lamps. The enclosed iron stove had only recently been invented and much cooking was still done on the open hearth. Only the proximity of the hearth or stove was warm; bedrooms were unheated and family members carried warm bricks with them to bed.
>
> But the biggest inconvenience was the lack of running water. Every drop of water for laundry, cooking, and indoor chamber pots had to be hauled in by the housewife, and wastewater hauled out. The average North Carolina housewife in 1885 had to walk 148 miles per year while carrying 35 tonnes of water.

Gordon believes that the technological innovations associated with computer technologies are far less important than those associated with the “second industrial revolution” and that growth may sputter out over time. Figure 1 repeats a chart from Gordon’s paper showing the growth rate of per capita GDP for the world’s leading economies (first the UK, then the US). It shows growth accelerating until 1950 and declining thereafter. Figure 2 shows a hypothetical chart in which Gordon projects a continuing fall-off in growth.

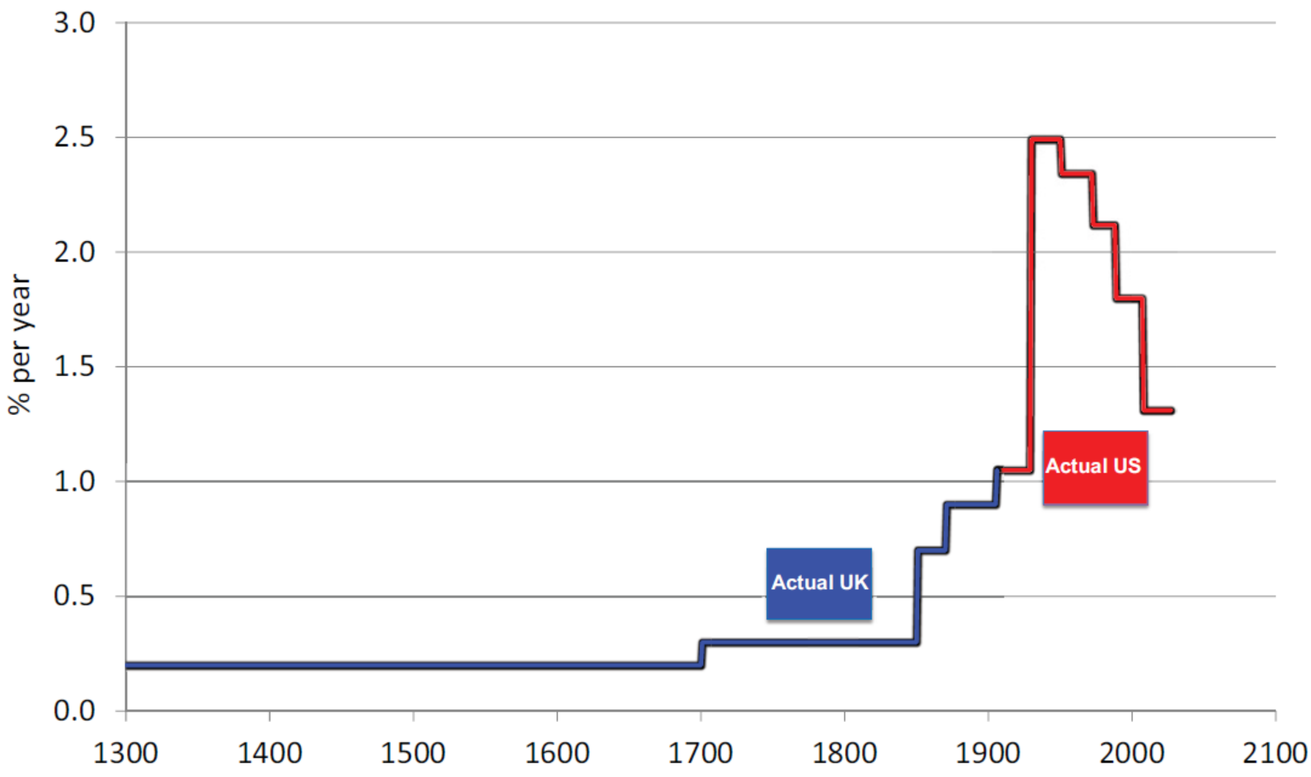
To illustrate why he believes modern inventions don’t match up with past improvements, Gordon offers the following thought experiment.

> You are required to make a choice between option A and option B. With option A you are allowed to keep 2002 electronic technology, including your Windows 98 laptop accessing Amazon, and you can keep running water and indoor toilets; but you can’t use anything invented since 2002.
>
> Option B is that you get everything invented in the past decade right up to Facebook, Twitter, and the iPad, but you have to give up running water and indoor toilets. You have to haul the water into your dwelling and carry out the waste. Even at 3am on a rainy night, your only toilet option is a wet and perhaps muddy walk to the outhouse. Which option do you choose?

You probably won’t be surprised to find out that most people pick option A.

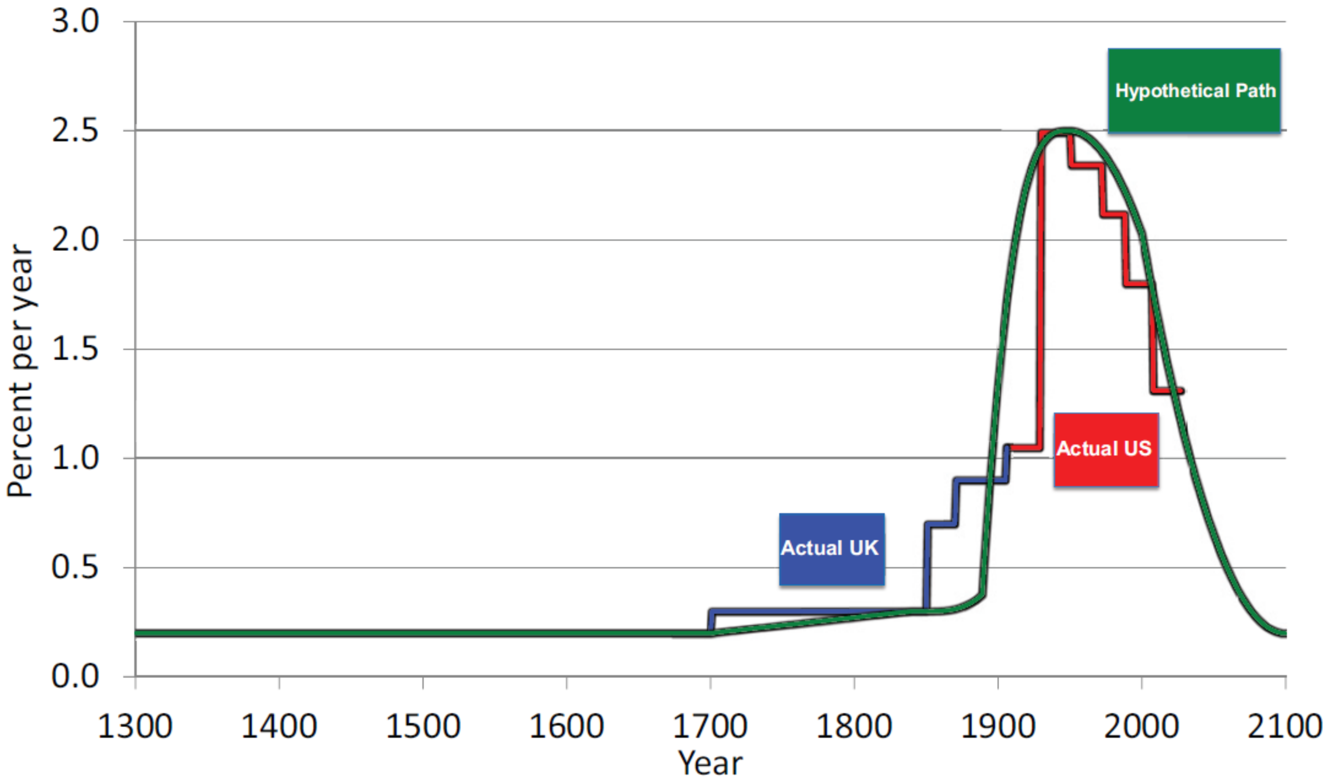
Gordon also discusses other factors likely to hold back growth in leading countries such as the leveling off of a long-run pattern of educational achievement, an aging population and energy-related constraints. It’s worth noting, though, that while Gordon’s paper is very well researched and well argued, economists are not very good at forecasting the invention of new technologies or their impact on the economy. For all we know, the next “industrial revolution" could be around the corner to spark a new era of rapid growth. Joel Mokyr's article "Is technological progress a thing of the past?" is a good counterpart to Gordon's scepticism.[4]

Figure 12.3: Gordon on the Growth Rate of Leading Economies



[4]Available at www.voxeu.org/article/technological-progress-thing-past

Figure 12.4: Gordon's Hypothetical Path for Growth



Chapter 13

Cross-Country Technology Diffusion

So far, we've been discussing how the invention of new technologies promotes economic growth by pushing out the "technological frontier" and allowing capital to be allocated across new and old technologies with diminishing returns setting in. This is clearly an important aspect of economic growth. However, we should remember that only a very few countries in the world are "on the technological frontier"—most places are not relying on Apple to invent a new gadget to promote efficiency. One way to illustrate this point is to estimate the level of total factor productivity for different countries in the world.

An important paper that did these calculations and used them to shed light on cross-country income differences is Hall and Jones (1999).[1] The basis of the study is a "levels accounting" exercise that starts from the following production function

$$Y_i = K_i^{\alpha} \left(h_i A_i L_i \right)^{1-\alpha} \tag{13.1}$$

Like the BLS multifactor productivity calculations that we discussed previously, Hall and Jones account for the effect of education on the productivity of the labour force. Specifically, they construct measures of *human capital* based on estimates of the return to education—this is the h_i in the above equation.

[1] Robert E. Hall and Charles I. Jones. "Why Do Some Countries Produce So Much More Output per Worker than Others?," *Quarterly Journal of Economics*, February 1999.

Hall and Jones show that their production function can be re-formulated as

$$\frac{Y_i}{L_i} = \left(\frac{K_i}{Y_i}\right)^{\frac{\alpha}{1-\alpha}} h_i A_i \tag{13.2}$$

Hall and Jones then constructed a measure h_i using evidence on levels of educational attainment and they also set $\alpha = 1/3$. This allowed them to use (13.2) to express all cross-country differences in output per worker in terms of three multiplicative terms: capital intensity, human capital per worker, and technology or total factor productivity. They found that output per worker in the richest five countries was 31.7 times that in the poorest five countries. This was explained as follows:

- Differences in capital intensity contributed a factor of 1.8.
- Differences in human capital contributed a factor of 2.2
- The remaining difference—a factor of 8.3—was due to differences in TFP.

The results from this paper show that differences in total factor productivity, rather than differences in factor accumulation, are the key explanation of cross-country variations in income levels. A more detailed table of Hall and Jones's calculations is reproduced on the next page. These calculations show that most countries are very far from the technological frontier, so their growth is not likely to be reliant on the invention of new technologies.

Figure 13.1: Table from Hall-Jones Paper

TABLE I
PRODUCTIVITY CALCULATIONS: RATIOS TO U. S. VALUES

Country	Y/L	Contribution from $(K/Y)^{\alpha/(1-\alpha)}$	H/L	A
United States	1.000	1.000	1.000	1.000
Canada	0.941	1.002	0.908	1.034
Italy	0.834	1.063	0.650	1.207
West Germany	0.818	1.118	0.802	0.912
France	0.818	1.091	0.666	1.126
United Kingdom	0.727	0.891	0.808	1.011
Hong Kong	0.608	0.741	0.735	1.115
Singapore	0.606	1.031	0.545	1.078
Japan	0.587	1.119	0.797	0.658
Mexico	0.433	0.868	0.538	0.926
Argentina	0.418	0.953	0.676	0.648
U.S.S.R.	0.417	1.231	0.724	0.468
India	0.086	0.709	0.454	0.267
China	0.060	0.891	0.632	0.106
Kenya	0.056	0.747	0.457	0.165
Zaire	0.033	0.499	0.408	0.160
Average, 127 countries:	0.296	0.853	0.565	0.516
Standard deviation:	0.268	0.234	0.168	0.325
Correlation with Y/L (logs)	1.000	0.624	0.798	0.889
Correlation with A (logs)	0.889	0.248	0.522	1.000

The elements of this table are the empirical counterparts to the components of equation (3), all measured as ratios to the U. S. values. That is, the first column of data is the product of the other three columns.

Leaders and Followers

The Romer model probably should not be thought of as a model of growth in any one particular country. No country uses only technologies that were invented in that country; rather, products invented in one country end up being used all around the world. Thus, the model is best thought of as a model of the leading countries in the world economy. How then should long-run growth rates be determined for individual countries? By itself, the Romer model has no clear answer, but it suggests a model in which ability to learn about the usage of new technologies should plays a key role in determining output per worker.

We will now describe such a model. The mathematics of the model are also formally equivalent to a well-known model of Nelson and Phelps (*AER*, 1966), though the application there is different, their subject being the diffusion of technological knowledge over time within an individual country.

The Model

We will assume that there is a "lead" country in the world economy that has technology level, A_t at time t which grows at rate g every period, so that

$$\frac{\dot{A}_t}{A_t} = g \tag{13.3}$$

All other countries in the world, indexed by j, have technology levels given by $A_{jt} < A_t$. The growth rate of technology in country j is determined by

$$\frac{\dot{A}_{jt}}{A_{jt}} = \lambda_j + \sigma_j \frac{(A_t - A_{jt})}{A_{jt}} \tag{13.4}$$

where $\lambda_j < g$ and $\sigma_j > 0$. This tells us that technology growth in all countries apart from the lead country is determined by two factors

- Learning: The second term says that their technology level will grow faster the bigger is the percentage gap between its level of technology, A_{jt} and the level of the leader, A_t. The larger is the parameter σ_j, the better the country is at learning about the technologies being applied in the lead country.

- The first term, λ_j indicates the country's capacity for increasing its level of technology without learning from the leader. We impose the condition $\lambda_j < g$. This means that country j can't grow faster than the lead country without the learning that comes from having lower technology than the frontier.

Exponential Growth

You've probably heard about exponential functions before but, even if you have, it's worth a quick reminder. The number $e \approx 2.71828$ is a very special number such that the function

$$\frac{de^x}{dx} = e^x \tag{13.5}$$

One way to see why the number is 2.718 is to use something called the Taylor series approximation for a function, which states that you can approximate a function $f(x)$ as

$$f(x) = f(a)+f'(x)(x-a)+\frac{1}{2}f''(x))(x-a)^2+\frac{1}{3!}f'''(x))(x-a)^3+...\frac{1}{n!}f^n(x))(x-a)^n+... \tag{13.6}$$

where $n! = (1)(2)(3)...(n-1)(n)$. If there is a number, e that has the property that $e^x = f(x) = f'(x)$, then that means that all derivatives also equal e^x. In this case, we have

$$e^x = e^a + e^a(x-a) + \frac{1}{2}e^a(x-a)^2 + \frac{1}{3!}e^a(x-a)^3 + ... \tag{13.7}$$

Setting $x = 1, a = 0$, this becomes

$$e = 1 + \frac{1}{2} + \frac{1}{3!} + \frac{1}{4!} + ... \tag{13.8}$$

This converges to 2.71828. Ok, that's not on the test but worth knowing. Now note that

$$\frac{de^{gt}}{dt} = \frac{de^{gt}}{d(gt)}\frac{dg}{dt} = ge^{gt} \tag{13.9}$$

Now let's relate this back to our model. The fact that the lead country has growth such that

$$\frac{dA_t}{dt} = \dot{A}_t = gA_t \tag{13.10}$$

means that this country is characterised by what is known as exponential growth, i.e.

$$A_t = A_0 e^{gt} \tag{13.11}$$

We write the first term as A_0 because $e^{(g)(0)} = 1$ so whatever term multiplies e^{gt} that is the value that A_t takes in the first period.

Dynamics of Technology

Now we are going to try to figure out how the technology variable behaves in the follower country. First, lets take equation (13.4) and multiply across by A_{jt} to get

$$\dot{A}_{jt} = \lambda_j A_{jt} + \sigma_j \left(A_t - A_{jt}\right) \tag{13.12}$$

This is what is known as a first-order linear differential equation (differential equation because it involves a derivative; first-order because it only involves a first derivative; linear because it doesn't involve any terms taken to powers than are not one.) These equations can be solved to illustrate how A_j changes over time. To do this, we first draw some terms together to re-write it as

$$\dot{A}_{jt} + \left(\sigma_j - \lambda_j\right) A_{jt} = \sigma_j A_t \tag{13.13}$$

Recalling equation (13.11) for the technology level of the leader country, this differential equation can be re-written as

$$\dot{A}_{jt} + \left(\sigma_j - \lambda_j\right) A_{jt} = \sigma_j A_0 e^{gt} \tag{13.14}$$

Now we'll move on to illustrating how people figure out how an A_{jt} that satisfies this equation needs to behave.

One Possible Solution

Let's think about what we learned about exponential functions to help us see what form a potential solution might take. The derivative of A_{jt} with respect to time plus $(\sigma_j - \lambda_j)$ times A_{jt} can be written as a multiple of the exponential function.

Looked at this way, we might guess that one possible solution for an A_{jt} process that will satisfy this equation is something of the form $B_j e^{gt}$ where B_j is some unknown coefficient. Indeed, it turns out that this is the case. Let's figure out what B_j must be. It must satisfy

$$gB_j e^{gt} + (\sigma_j - \lambda_j) B_j e^{gt} = \sigma_j A_0 e^{gt} \tag{13.15}$$

Canceling the e^{gt} terms, we see that

$$B_j = \frac{\sigma_j A_0}{\sigma_j + g - \lambda_j} \tag{13.16}$$

So, this solution takes the form

$$A_{jt}^p = B_j e^{gt} = \left(\frac{\sigma_j}{\sigma_j + g - \lambda_j}\right) A_0 e^{gt} = \left(\frac{\sigma_j}{\sigma_j + g - \lambda_j}\right) A_t \tag{13.17}$$

A General Solution

Is that it or could we add on an additional term and still get a solution? Suppose we look for a solution of the form

$$A_{jt} = B_j e^{gt} + D_{jt} \tag{13.18}$$

Then the solution would have to obey

$$gBe^{gt} + \dot{D}_{jt} + (\sigma_j - \lambda_j)\left(Be^{gt} + D_{jt}\right) = \sigma_j A_0 e^{gt} \tag{13.19}$$

All the terms in e^{gt} cancel out because, by construction of B_j, they satisfy equation (13.15). This means the additional term D_{jt} must satisfy

$$\dot{D}_{jt} + (\sigma_j - \lambda_j)\, D_{jt} = 0 \tag{13.20}$$

Again using the properties of the exponential function, this equation is satisfied by anything of the form

$$D_{jt} = D_{j0} e^{-(\sigma_j - \lambda_j)t} \tag{13.21}$$

where D_{j0} is a parameter that can take on any value. So, given the differential equation (13.12), all possible solutions for technology in country j must take the form

$$A_{jt} = \left(\frac{\sigma_j}{\sigma_j + g - \lambda_j}\right) A_t + D_{j0} e^{-(\sigma_j - \lambda_j)t} \tag{13.22}$$

where D_{j0} is an arbitrary parameter than can take any value.

Properties of the Solution

Now we would like to examine the properties of this solution. Does technology in the follower country catch up and, if not, where does it end up and why? To answer these questions, it is useful to express A_{jt} as a ratio of the frontier level of technology. This can be written as

$$\frac{A_{jt}}{A_t} = \frac{\sigma_j}{\sigma_j + g - \lambda_j} + \frac{D_{j0}}{A_t} e^{-(\sigma_j - \lambda_j)t} \tag{13.23}$$

Now using the fact that $A_t = A_0 e^{gt}$, this becomes

$$\frac{A_{jt}}{A_t} = \frac{\sigma_j}{\sigma_j + g - \lambda_j} + \frac{D_{j0}}{A_0} e^{-(\sigma_j + g - \lambda_j)t} \tag{13.24}$$

To understand the properties of this solution, recall that we assumed $\lambda_j < g$, which means that on its own (without catch-up growth) the follower country's level of technology grows slower than the leader country and also that $\sigma_j > 0$ (some learning takes place). Putting these two assumptions together, we can say

$$\sigma_j + g - \lambda_j > 0 \tag{13.25}$$

That means that

$$e^{-(\sigma_j + g - \lambda_j)t} \to 0 \quad \text{as} \quad t \to \infty \tag{13.26}$$

This means that the second term in (13.24) tends towards zero. So, over time, as this term disappears, the country converges towards a level of technology that is a constant ratio, $\frac{\sigma_j}{\sigma_j + g - \lambda_j}$ of the frontier level, and its growth rate tends towards g.

Note that $g - \lambda_j > 0$ also means that

$$0 < \frac{\sigma_j}{\sigma_j + g - \lambda_j} < 1 \tag{13.27}$$

so each country never actually catches up to the leader but instead converges to some fraction of the lead country's technology level. This makes sense if you think about it. Because of their inferiority at developing their own technologies ($\lambda_j < g$) the follower countries will always be falling further behind the leader unless there is a gap between their level of technology and the leader. So, to have a steady-state in which everyone's technology is growing at the same rate, the followers must all have technology levels below that of the leader.

In addition, $g - \lambda_j > 0$ means that

$$\frac{d}{d\sigma_j}\left(\frac{\sigma_j}{\sigma_j + g - \lambda_j}\right) > 0 \tag{13.28}$$

The equilibrium ratio of the country's technology to the leader's depends positively on the "learning parameter" σ_j. The higher this parameter—the more fo the gap to the leader that it closes each period—then the close the ratio gets to one and the higher up the "pecking order" the country gets. It's also true that

$$\frac{d}{d\lambda_j}\left(\frac{\sigma_j}{\sigma_j + g - \lambda_j}\right) > 0 \tag{13.29}$$

In other words, the more growth the country can generate each period independent of learning from the leader, the higher will be its equilibrium ratio of technology relative to the leader.

Illustrating the Model

Going back to the equation for the ratio of technology in country j to the leader, equation (13.24), we noted already that the second term tends to disappear to zero over time. That doesn't mean it's unimportant. How a country behaves along its "transition path" depends on the value of the initial parameter D_{j0}.

- If $D_{j0} < 0$, then the term that is disappearing over time is a negative term that is a drag on the level of technology. This means that the country starts out below its equilbrium technology ratio, grows faster than the leader for some period of time with growth eventually tailing off to the growth rate of the leader.

- If $D_{j0} > 0$, then the term that is disappearing over time is a positive term that is boosting the level of technology. This means that the country starts out above its equilibrium technology ratio, grows slower than the leader for some period of time with growth eventually moves up towards the growth rate of the leader.

We have illustrated how these dynamics would work with the Figures 13.2 to 13.4. These charts show model simulations for a leader economy with $g = 0.02$ and a follower economy with $\lambda_j = 0.01$ and $\sigma_j = 0.04$. These values mean

$$\frac{\sigma_j}{\sigma_j + g - \lambda_j} = \frac{0.04}{0.04 + 0.02 - 0.01} = 0.8 \tag{13.30}$$

so the follower economy converges to a level of technology that is 20 percent below that of the leader. The first collection of charts show what happens when this economy has a value of $D_{j0} = -0.5$, so that it starts out with a technology level only 30 percent that of the leader. They grow faster than the leader country for a number of years before they approach the 0.8 equilibrium ratio and then their growth rate settles down to the same rate as that of the leader.

The second collection of charts show what happens when this economy has a value of $D_{j0} = 0.5$, so that it starts out with a technology level 30 percent above that of the leader, even though the equilibrium value is 20 percent below. Technology levels in this follower country never actually decline but they do go through a long-period of slow growth rates before eventually heading towards the same growth rate as the leader as they approach the 0.8 equilibrium ratio.

Finally, we show how the model may also be able to account for the sort of "growth miracles" that are occasionally observed when countries suddenly start experiencing rapid growth: If a country can increase its value of σ_j via education or science-related policies, its position in the steady-state distribution of income may move upwards substantially, with the economy then going through a phase of rapid growth. The third collection of charts show what happens when, in period 21, an economy changes from having $\sigma_j = 0.005$ to $\sigma_j = 0.04$. The equilibrium technology ratio changes from one-third to 0.8 and the economy experiences a long transitional period of rapid growth.

An important message from this model is that for most countries, it is not their ability to invent new capital goods that is key to high living standards, but rather their ability to learn from those countries that are more technologically advanced.

Figure 13.2: Catching Up

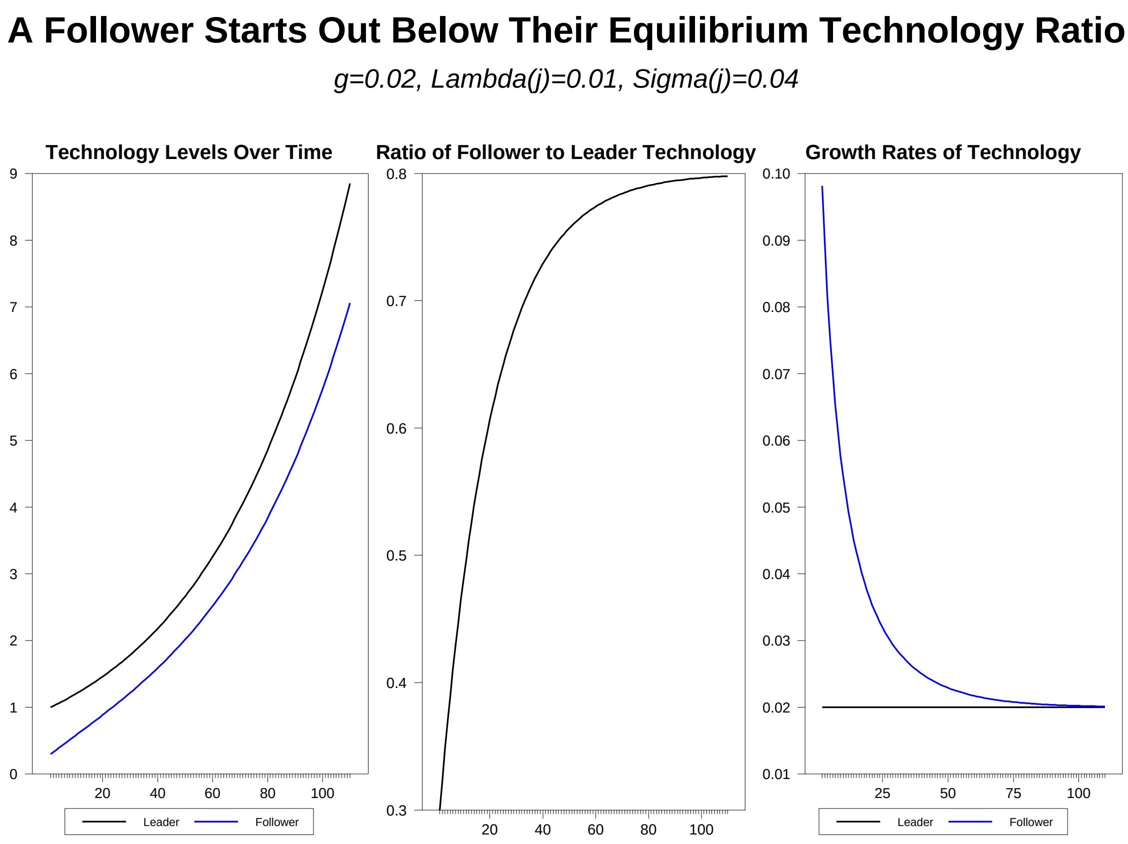


Figure 13.3: Falling Back

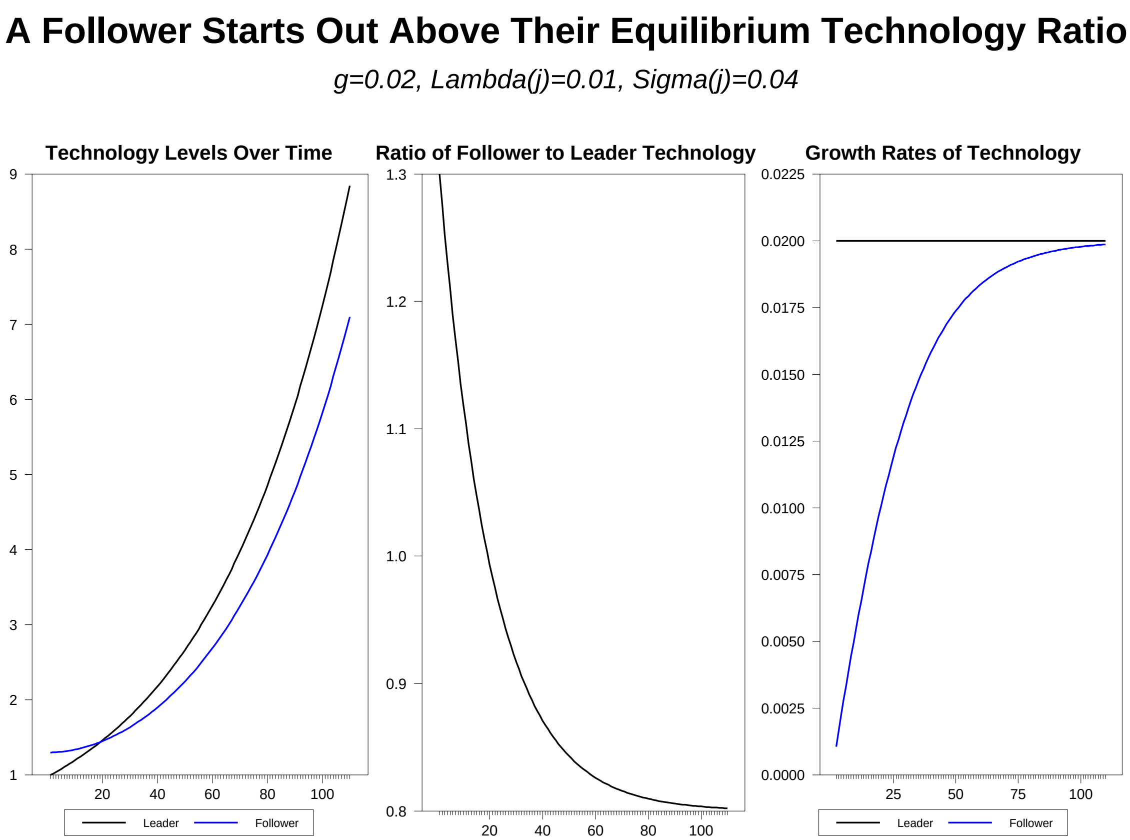
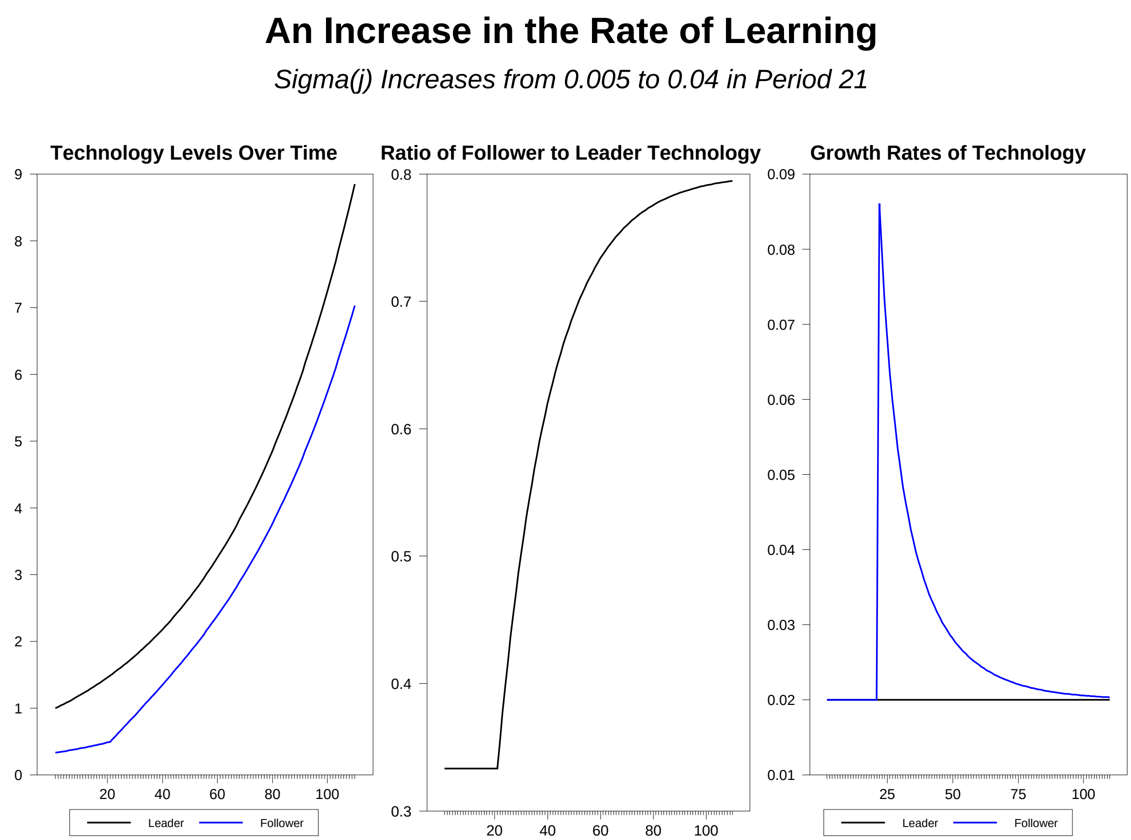


Figure 13.4: A Growth Miracle



Chapter 14

Institutions and Efficiency

We have documented huge differences in total factor productivity across countries. What determines these differences? One answer is provided by the combination of the Romer model and the leader-follower model. According to these models, large differences in TFP reflect variations in the extent to which countries have adopted the latest technologies.

However, this is perhaps too mechanistic a view of what generates cross-country differences in efficiency. TFP doesn't just reflect the technologies a country's people use. It is a measure of the efficiency with which an economy makes use of its resources and there are a whole range of other factors that can affect this. For example:

- *Bureaucratic Inefficiency and Corruption*: Satisfaction of bureaucratic requirements and bribing of officials can be important diversions of resources in poor economies.

- *Crime*: Time spent on crime does not produce output. Neither do resources devoted to protecting inviduals and firms from crime.

- *Restrictions on Market Mechanisms*: Protectionism, price controls, and central planning can all lead to resources being allocated in an inefficient manner.

In addition, while technology adoption certainly has an impact on differences in TFP, this still leaves open the question of what drives the pace of technology adoption in poorer countries. Ultimately, the models so far don't answer the question of the *deeper determinants* of economic success. We will now discuss on the idea that the ultimate explanation for patterns of economic efficiency relates to differences in institutions.

Douglass North and Institutions

There is now a large literature that focuses on the idea that differences in *institutions* provides the key to understanding TFP differences across countries. Economic activity does not take place in a vacuum. Firms need to take account of the legal and regulatory environment, the tax system, and the services provided by government as well as the political setting that determines these institutions.

The work of economic historian Douglass North, winner of the 1993 Nobel prize for economics, was particularly influential in stressing the key importance of good institutions for economic growth. The introduction to North's paper "Institutional Change: A Framework of Analysis" gives a flavour of his arguments:

> A theory of institutional change is essential for further progress in the social sciences in general and economics in particular. Essential because neo-classical theory (and other theories in the social scientist's toolbag) at present cannot satisfactorily account for the very diverse performance of societies and economies both at a moment of time and over time. The explanations derived from neo-classical theory are not satisfactory because, while the models may account for most of the differences in performance between economies on the basis of differential investment in education, savings rates, etc., they do not account for why economies would fail to undertake the appropriate activities if they had a high payoff. Institutions deter-

> mine the payoffs. While the fundamental neo-classical assumption of scarcity and hence competition has been robust (and is basic to this analysis), the assumption of a frictionless exchange process has led economic theory astray. Institutions are the structure that humans impose on human interaction and therefore define the incentives that (together with the other constraints (budget, technology, etc.) determine the choices that individuals make that shape the performance of societies and economies over time.

He goes to discuss the link between institutions and the profit-maximising decisions that people will take:

> Institutions consist of formal rules, informal constraints (norms of behavior, conventions, and self imposed codes of conduct) and the enforcement characteristics of both ... If institutions are the rules of the game, organizations are the players. They are groups of individuals engaged in purposive activity. The constraints imposed by the institutional framework (together with the other constraints) define the opportunity set and therefore the kind of organizations that will come into existence ... If the highest rates of return in a society are to be made from piracy, then organizations will invest in knowledge and skills that will make them better pirates; if organizations realize the highest payoffs by increasing productivity then they will invest in skills and knowledge to achieve that objective.

This paper contains a discussion of some aspects of the US's institutional history that have been positive for economic growth. Much of North's other work focuses on the development of institutions that made some countries such as the UK successful early developers through the industrial revolution while others lagged.

An Example of the Importance of Institutions

Korea provides an extreme example of the importance of institutions in determining the success of an economy. After World War II, Korea was split into a northern zone that became the Democratic People's Republic of Korea, a Soviet-style socialist republic, while South Korea became a capitalist economy.

North Korea received external support from the USSR for many years but no longer receives external aid. It remains a centrally planned economy with only one political party. The economy has failed to prosper and there are reliable reports of large amounts of death from famine in the 1990s. In contrast, South Korea has been a huge economic success and is home to many globally successful corporations such as Samsung and Hyundai.

The figure on the next page illustrates the gap between North and South Korea. While the two areas began with few substantive differences, sharing a common culture and identity, their different economic institutions mean that they are now completely different. Viewed from the sky, you can see development all over South Korea while North Korea is almost fully dark because of a lack of electricity.

Figure 14.1: The Korean Peninsula at Night

An Econometric Approach

The historical approach adopted by North and isolated examples of extreme events (such as the Korean split) has been very valuable in highlighting cases where good institutions have facilitated economic growth and where bad institutions have prevented it. In recent decades, there has been an attempt to also assess the role of institutions in economic development using more formal econometric techniques. An early paper in this literature was the 1999 *Quarterly Journal of Economics* paper by Robert Hall and Charles I. Jones that we discussed earlier. Hall and Jones used the term *social infrastructure* to describe the institutions that affect incentives to produce and invest. Their approach was to collect data on a large number of countries and then estimate regressions of the form

$$\frac{Y_i}{L_i} = \alpha + \beta S_i + \epsilon_i \tag{14.1}$$

where $\frac{Y}{L}$ is output per worker in country i and S_i is a variable that aims to measure the extent to which institutions in country i facilitate economic activity. Hall and Jones constructed their S_i variable as an average of two different variables:

1. An "index of government antidiversion policies". This is an average of five different variables relating to (i) law and order (ii) bureaucratic quality (iii) corruption (iv) risk of expropriation, and (v) government repudiation of contracts.

2. An index that focuses on the openness of a country to trade with other countries

There are two potentially serious econometric problems when assessing the linkage between productivity and institutions. The first is *endogeneity*. Do countries get rich because they have good institutions or do countries have good institutions because they are rich? The latter linkage certainly exists. Citizens in richer countries have substantial incentives to keep good institutions that promote productive efficiency because they would have lot to lose if their markets ceased to work well; these incentives may be weaker in the world's poorer countries. Hall and Jones thus describe their "social infrastructure" variable as being determined by

$$S_i = \gamma + \delta \frac{Y_i}{L_i} + \theta X_i + \eta_i \tag{14.2}$$

In this case, a simple OLS regression of $\frac{Y_i}{L_i}$ on S_i will produce a positive estimate of β—the effect of institutions on output per worker—even if the true value of β was zero.

The second econometric problem is *measurement error*. The variables used as measures of institutional quality can only ever be proxies, and possibly poor proxies, for the true measure of institutional quality that actually affects economic output. The use of proxies like this is the same as using variables that are affected by measurement error. One of the standard results from econometrics is that measurement error can result in downward bias in coefficients. In other words, the OLS coefficient might be less than the true coefficient.

So the presence of these econometric problems means OLS estimation will produce biased estimates, though whether the bias is upwards or downwards depends on the source of the bias. The usual solution to these econometric problems is estimation via instrumental variables. This means estimating β from

$$\frac{Y_i}{L_i} = \alpha + \beta \hat{S}_i + \epsilon_i \tag{14.3}$$

where $\hat{S}_i$ is the fitted value from a regression of S on a set of instruments (exogenous variables that that may be correlated with the institutions variable but that are not affected by the country's level of output per worker). By focusing on variations in institutions related to exogenous factors that are not determined by output per worker, the researcher can try to identify the true causal effect of institutions.

Hall and Jones's Findings

Finding good instruments for this problem can be tricky. Many of the papers in this literature have focused on either *geography* or *history* as their inspiration for truly exogenous sources of variations in institutions.

- A country's geography is certainly exogenous—it is not influenced by a country's level of prosperity. But certain types of geographical features may be correlated with whether a country has good institutions or not. Hall and Jones used the country's distance from the equator as an instrument. Other papers have also used coastal access, average temperature, rainfall and soil quality.

- In relation to history, many countries around the world were colonised by various European countries and their current institutions (e.g. whether a country uses a French or English legal systems) are often determined, in a somewhat random fashion, by which countries colonised them. Hall and Jones used instruments measuring the fraction of people speaking English as a native language and a variable measuring the fraction of people speaking other Western European languages.

Using their selected instrument set, Hall and Jones found a positive and significant effect of their "social infrastructure" variable when estimating using IV methods, with the coefficient being higher than the OLS estimate. They concluded from this that there is a large causal effect from institutions to productivity and that the measurement error is a more important source of bias in their OLS regressions than is endogeneity.

Some Other Papers

There is now a large empirical literature on this topic. Some examples:

- Acemoglu, Johnson, and Robinson (*AER*, 2001) assess the effect on GDP per capita of institutions, proxied by a measure of "protection against expropriation risk." They use a new instrument measuring settler mortality in different European colonies. They argue that countries where mortality for initial settlers was low were places where Europeans were more likely to settle and set up good institutions, with the reverse working when settler mortality was high. With this variable as an instrument, they find a very strong effect of certain measures of institutions on output per capita.

- Rodrik, Subramanian and Trebbi (*Journal of Economic Growth*, 2004). These authors assess the role of institutions (as proxied by a variable measuring the strength of the rule of law), openness to trade and geography (as measured by distance from the equator). To be able to assess whether geography has a direct effect on income per capita, they use other variables such as the Acemoglu-Johnson-Robinson settler mortality variable and language-related variables as instruments. They conclude that institutions, in the form of their rule of law variable, are the key determinant of economic success and do not find a significant role for trade or geography.

- Gillanders and Whelan (*Kyklos*, 2014) compare the effect of the Rule of Law variable preferred by Rodrik, Subramanian and Trebbi with a new variable that measures the "ease of doing business". Both are institutional variables but they measure different types of institutions. This paper also applies IV methods using geographical variables as instruments and concludes that it is the ease of doing business that is the key determinant of output per capita rather than Rule of Law variable.

Part IV

Growth and Resources

Chapter 15

The Malthusian Model

The previous chapters studied economies that grow steadily over time. For many countries around the world, that has been a reasonable description of their behaviour since the start of the Industrial Revolution. However, prior to around the year 1800, there is very little evidence of steady growth in income levels. The chart on the next page repeats the chart shown earlier from the book *A Farewell to Alms* by economic historian Greg Clark. It summarises world economic history as a long period in which living standards fluctuated over time showing no growth trend before the Industrial Revolution lead to steady growth over time (though Clark notes that this take-off did not occur in all countries and some remain exceptionally poor).

Measurement of living standards is an imprecise business even in modern economies with well-resourced statistical agencies. So it's hardly surprising that there is a lot of controversy over Clark's particular interpretation of the evidence as implying no trend growth at all in living standards prior to 1800. Other studies show slow but gradual increases in living standards prior to the Industrial Revolution but all agree that the average rate of economic growth was very low before 1800.[1] In addition, Figure 15.2 shows that global population growth was extremely slow until 1800 and then increased to much higher rates.

[1] "When Did Growth Begin? New Estimates of Productivity Growth in England from 1250 to 1870" by Paul Bouscasse, Emi Nakamura and Jon Steinsson suggests productivity growth in England began in 1600. https://eml.berkeley.edu/ enakamura/papers/malthus.pdf

What explains these patterns? Our previous models would suggest the pace of technological progress must have been slower before the Industrial Revolution and this is true. But cumulatively, there was a lot of technological progress in the years prior to 1800 with many advances made in science and in the organisation of economic life. One might have expected this to translate into growth in average living standards over time but the evidence suggests such progress was limited. In this chapter, we will present the Malthusian model, which explains how the world works very differently when rates of technological progress are slow.

Figure 15.1: World Economic History (from Greg Clark's book)

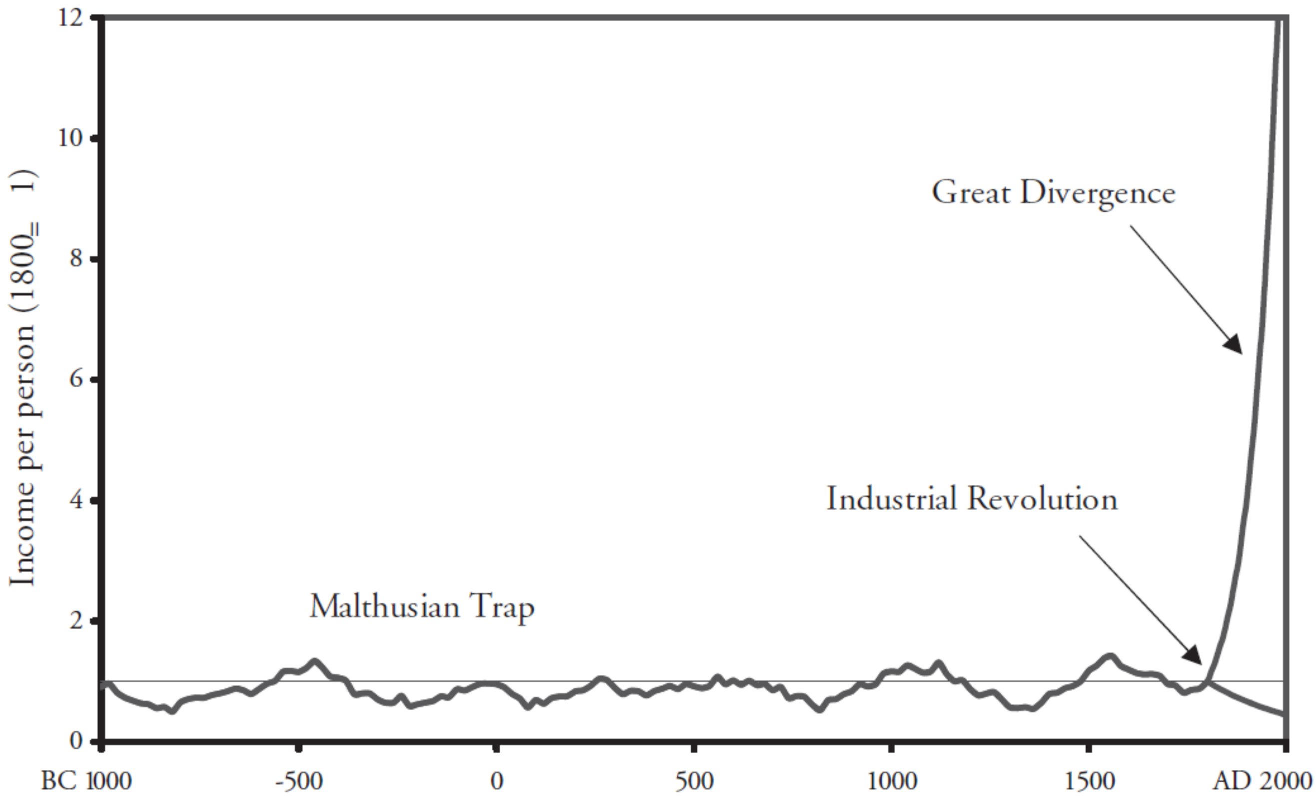
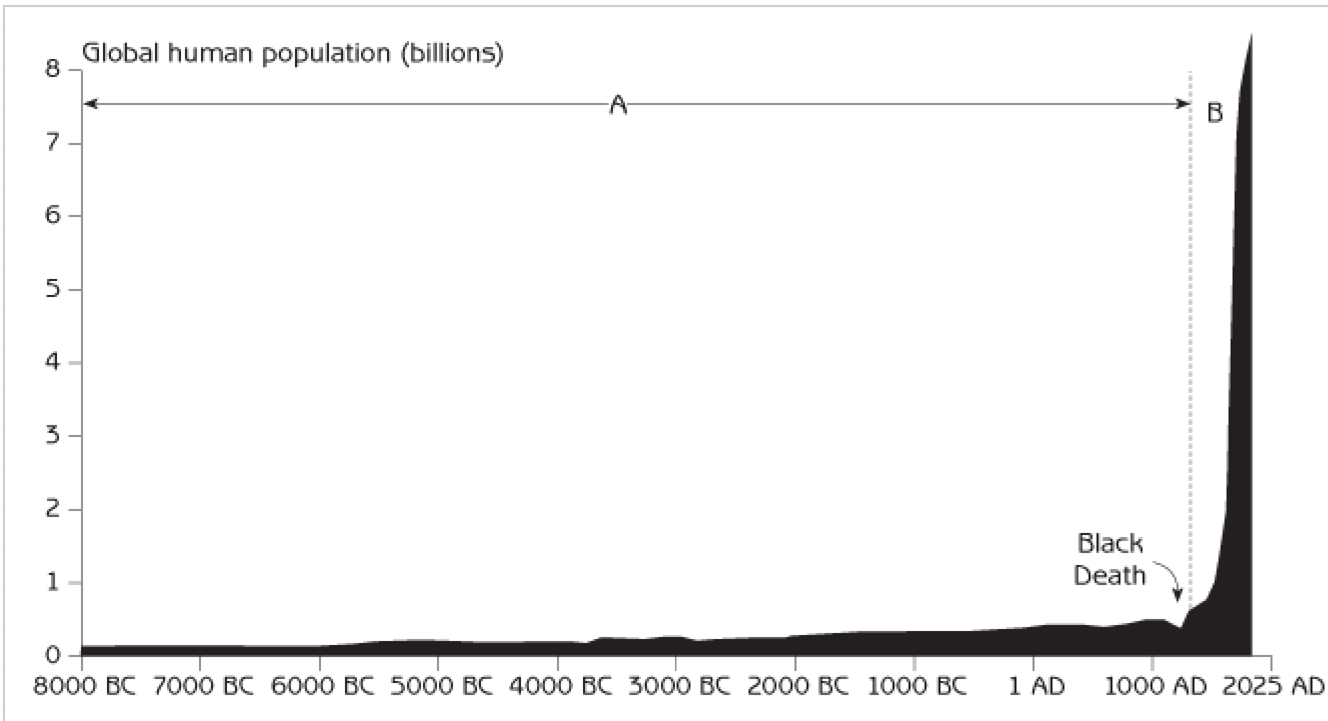


Figure 1.1 World economic history in one picture. Incomes rose sharply in many countries after 1800 but declined in others.

Figure 15.2: Global Population



Life Expectancy and Income Levels

The Malthusian model has two key elements: A negative relationship between income levels and the size of population and a positive relationship between income levels and population *growth*. Let's start with the second relationship

By definition, population growth increases with birth rates and falls with death rates. Death rates, in turn, are the key determinant of life expectancy. Throughout history, there has been a strong relationship between a country's average level of income per capita and its average life expectancy. This relationship still holds strongly today. Figure 3 shows a chart taken from a wonderful website called Gapminder which allows you to make animated charts showing developments over time and around the world in income levels, health outcomes and lots of other areas.

Figure 15.3 shows the relationship between average life expectancy and real income per person. Each dot corresponds to a country, with the size of the dot corresponding to its population. The chart shows that in a strong positive relationship between the average life expectancy and real income per person. Figure 15.4 shows a relationship of this kind holding inside a large country: It shows data for U.S. counties with higher income per capita have longer life expectancy.

Internationally, this pattern is partly related to the availability of medicines in advanced countries that allow people to live much longer. But it is more influenced by very high rates of child mortality. Figure 15.5 shows another Gapminder chart. This one shows that mortality among children under 5 is still very common in the world's poorest countries due to malnutrition and poor public health systems.

This relationship between income levels and the rate of death among the population will be a key element of the version of the Malthusian model that we will cover.

Figure 15.3: Life Expectancy and Real GDP Per Capita Around the World in 2022

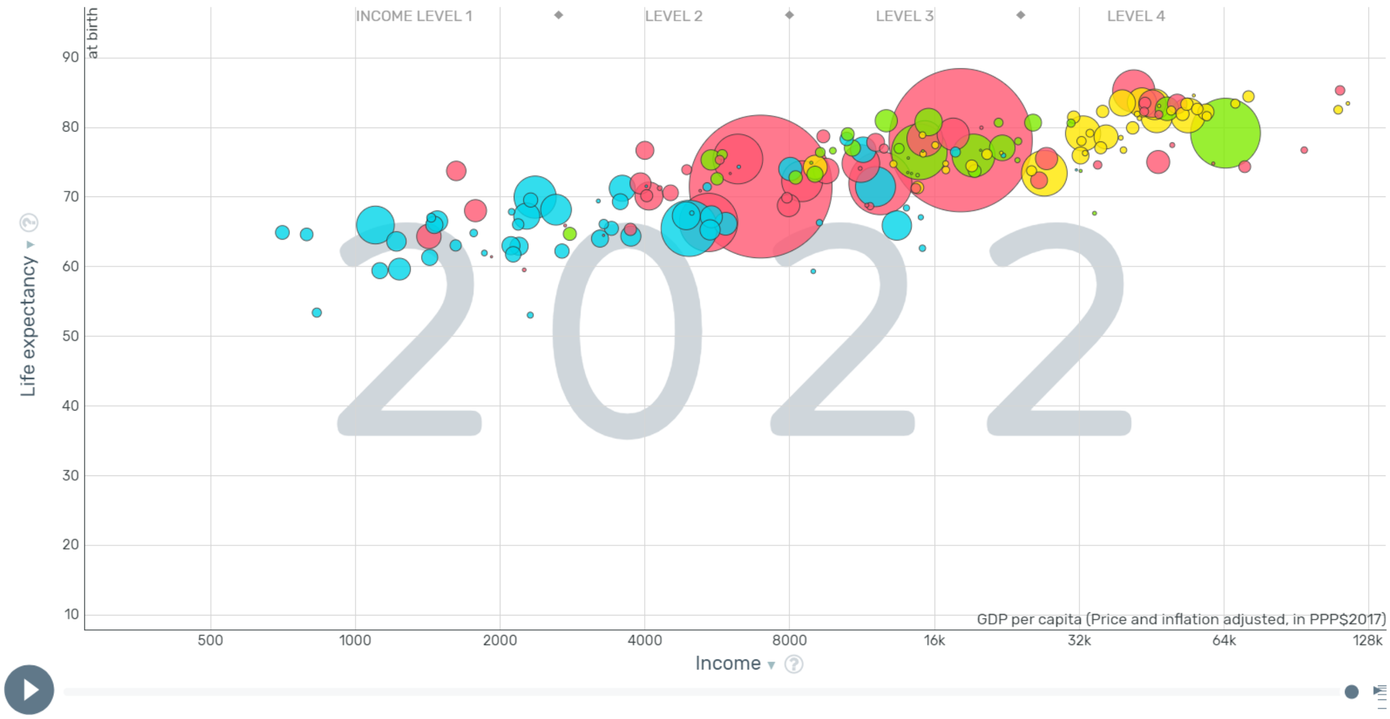


Figure 15.4: Life Expectancy and Income Levels: U.S. Counties

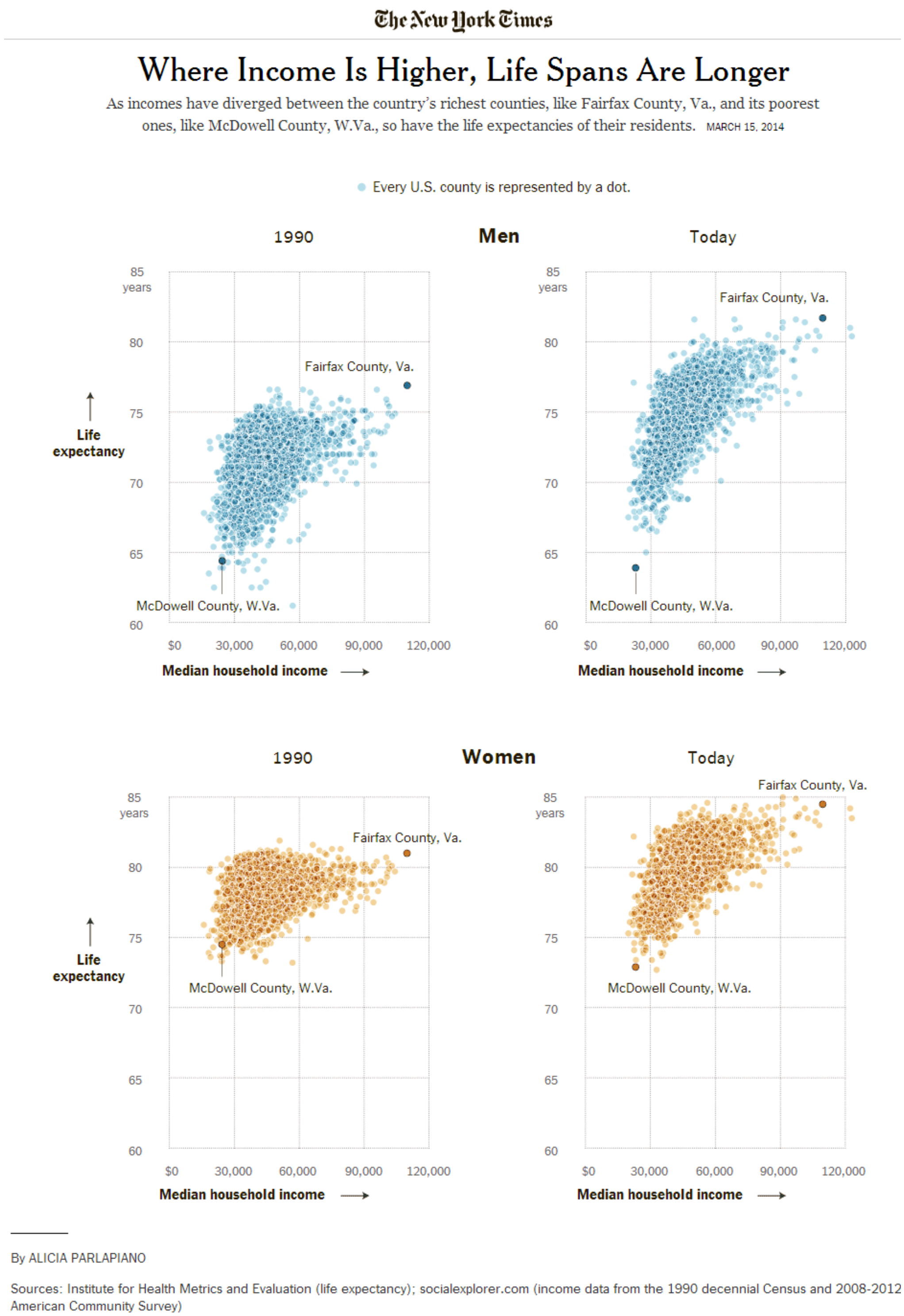
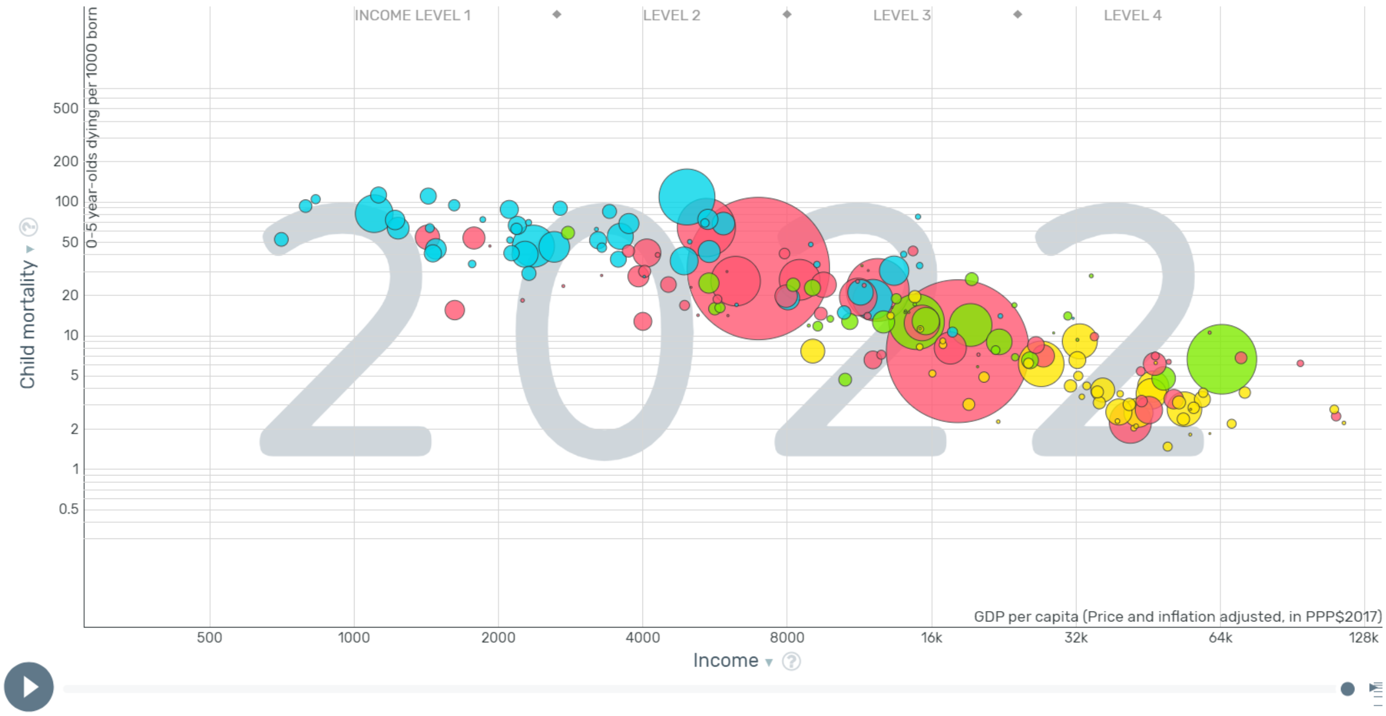


Figure 15.5: Child Mortality and Real GDP Per Capita Around the World in 2022



Population and Income Levels

The second element of the Malthusian model is a negative relationship between income levels and the level of population. Before discussing Malthus's thoughts on this issue, it's worth using the language of modern economics to describe this relationship.

Consider an economy in which aggregate output is determined by a Cobb-Douglas production function

$$Y_t = AK^{\alpha}L_t^{1-\alpha} \tag{15.1}$$

Here, I've assumed that both capital and technology are fixed (and so have no time subscript), so that labour input is the only factor that produces changes in output. We can figure out the demand for labour by assuming that the firms in the economy maximise profits in a competitive manner. Thus, firms are maximising

$$\pi = pAK^{\alpha}L_t^{1-\alpha} - wL - rK \tag{15.2}$$

where p is the price of output, w is the wage rate and r is an implicit rental rate for capital. The first-order condition for labour is

$$(1-\alpha)\,pAK^{\alpha}L^{-\alpha} - w = 0 \tag{15.3}$$

This can be re-arranged as

$$\frac{w}{p} = (1-\alpha)\,A\left(\frac{K}{L}\right)^{\alpha} \tag{15.4}$$

Assuming that a constant fraction θ of the population is working

$$L = \theta N \tag{15.5}$$

we get

$$\frac{w}{p} = (1-\alpha)\,A\left(\frac{K}{\theta N}\right)^{\alpha} \tag{15.6}$$

The higher the population, the lower will be the real wage. This is because of diminishing marginal returns to labour and the fact that workers are being paid their marginal wage product.

Now note that the direct link between higher population and lower wage rates (and thus lower living standards) works here because technology and capital are held constant. In the Solow growth model, there is both rising population and increasing wages because technology improvements and capital accumulation offset the negative effects on wages of rising population. In this example, we have assumed something quite different, i.e. no technological progress. We will return, however, to the question of what happens when there is a slow but steady rate of technological improvement.

Malthus (1798)

Thomas Malthus's 1798 book *An Essay on the Principle of Population* put together the two ideas that we have just discussed. He noted that rising living standards can lead to higher population growth but the famously-gloomy Malthus believed that this increase in population would ultimately undo the original increase in living standards.

Malthus placed a somewhat different emphasis on the various links than in our discussion. In relation to the link between demographics and living standards, Malthus focused on two mechanisms ("checks on living standards") that would cause population growth to increase as living standards rose and thus ultimately see the increase in living standards reversed.

The first mechanism, which Malthus labelled "the preventative check" was the tendency to see more births when real wages are high. In pre-Industrial Revolution Britain, the tradition was for people to marry relatively late as they waited to accumulate the wealth to be able to support a family. This tended to keep fertility rates relatively low. In practice, as discussed in Greg Clark's book on the Malthusian model, the evidence for a link between living standards and birth rates prior to the Industrial Revolution is fairly weak and I will assume a constant birth rate in the model treatment below (though the logic of the model is unchanged if you assume a positive relationship between birth rates and living standards.)

The second mechanism, which Malthus labelled "the positive check", was the negative effect of living standards on death rates. Evidence for this mechanism is stronger and still exists today. Malthus describes it as follows:

> *the actual distresses of some of the lower classes, by which they are disabled from giving the proper food and attention to their children, act as a positive check to the natural increase of population.*

This is the mechanism that we will focus on in our description of the model.

In relation to the negative effect of population on living standards, I've used a production function approach and emphasised the role played by the assumption of technology increases failing to offset the effect of increased population. Malthus focused more the idea of increased numbers of people putting a strain on food resources:

> *An increase of population without a proportional increase of food will evidently have the same effect in lowering the value of each man's patent. The food must necessarily be distributed in smaller quantities, and consequently a day's labour will purchase a smaller quantity of provisions. An increase in the price of provisions would arise either from an increase of population faster than the means of subsistence, or from a different distribution of the money of the society.*

The Model and its Convergent Dynamics

We will now describe a Malthusian model in somewhat more formal terms than Malthus did. Basically, I'm following Greg Clark's version of the model as described in Chapter 2 of his book, though I'm using a constant birth rate rather than one that depends on income levels.

The model has four equations. First, there is the definition of the change in the population, which just states that population equals last period's population plus last period's level of births minus deaths. (There are lots of different possible timing conventions here. I have in mind that the population level is measured at the start of each period, while births and deaths occur over the course of the period, but the particular timing convention adopted isn't important):

$$N_t = N_{t-1} + B_{t-1} - D_{t-1} \tag{15.7}$$

Births are a constant fraction of the population

$$\frac{B_t}{N_t} = b \tag{15.8}$$

While deaths are a decreasing function of real income per person

$$\frac{D_t}{N_t} = d_0 - d_1 Y_t \tag{15.9}$$

Finally, real income per person is a negative function of the population size:

$$Y_t = a_0 - a_1 N_t \tag{15.10}$$

Figure 15.6 shows how the death and birth rate equations combine together to make population dynamics a function of income per person. The death rate depends negatively on income per person, so at sufficiently high income levels—in this case, levels above Y^*—births are greater than deaths and population is growing, while population is falling at income levels below Y^*.

Figure 15.7 then shows that the economy tends to return to this equilibrium level of income. When income is above Y^*, population is growing. But the figure shows that growing population means income levels are falling. So income levels tend to fall when income is above Y^* and increase when it is below Y^*. Similarly population tends to fall when it is above the level of population associated with Y^*, call this N^*, and rise when it is below this level. This means that both income and population display what we have called *convergent dynamics* in our discussion of the Solow model: Wherever the economy starts out, it tends to converge towards these specific levels of income and population. Because the economy tends to revert back to the same levels of income and population, this phenomenon is often called *The Malthusian Trap*.

Figure 15.8 shows how the birth and death schedules, on the one hand, and the income-population schedule on the other, combine to determine the model's properties. Perhaps surprisingly, it is the birth and death schedules and not the income-population schedule that determines the long-run level of real income per person in the model. The income-population schedule then determines how many people are alive, given that level of income.

Figure 15.6: Birth and Death Rate Schedules

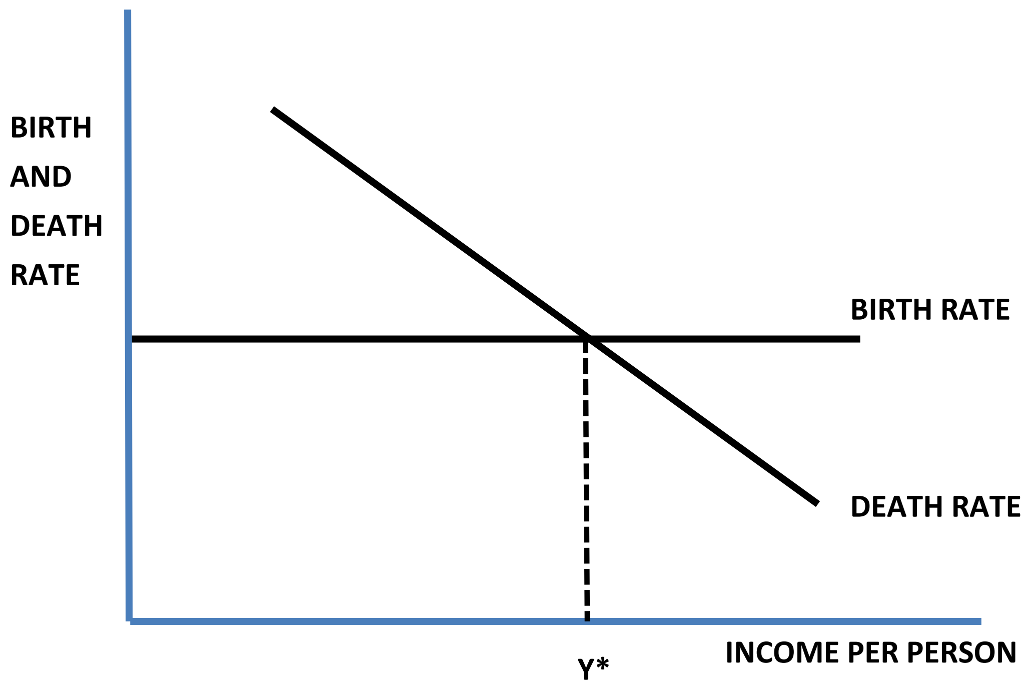


Figure 15.7: The Income-Population Schedule

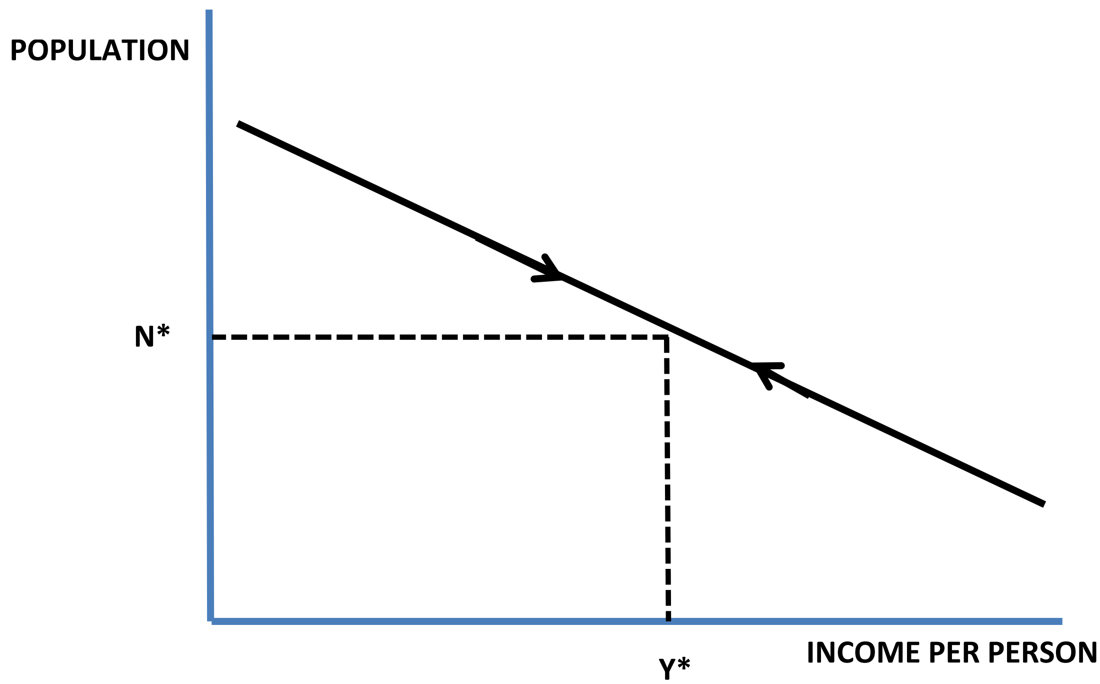
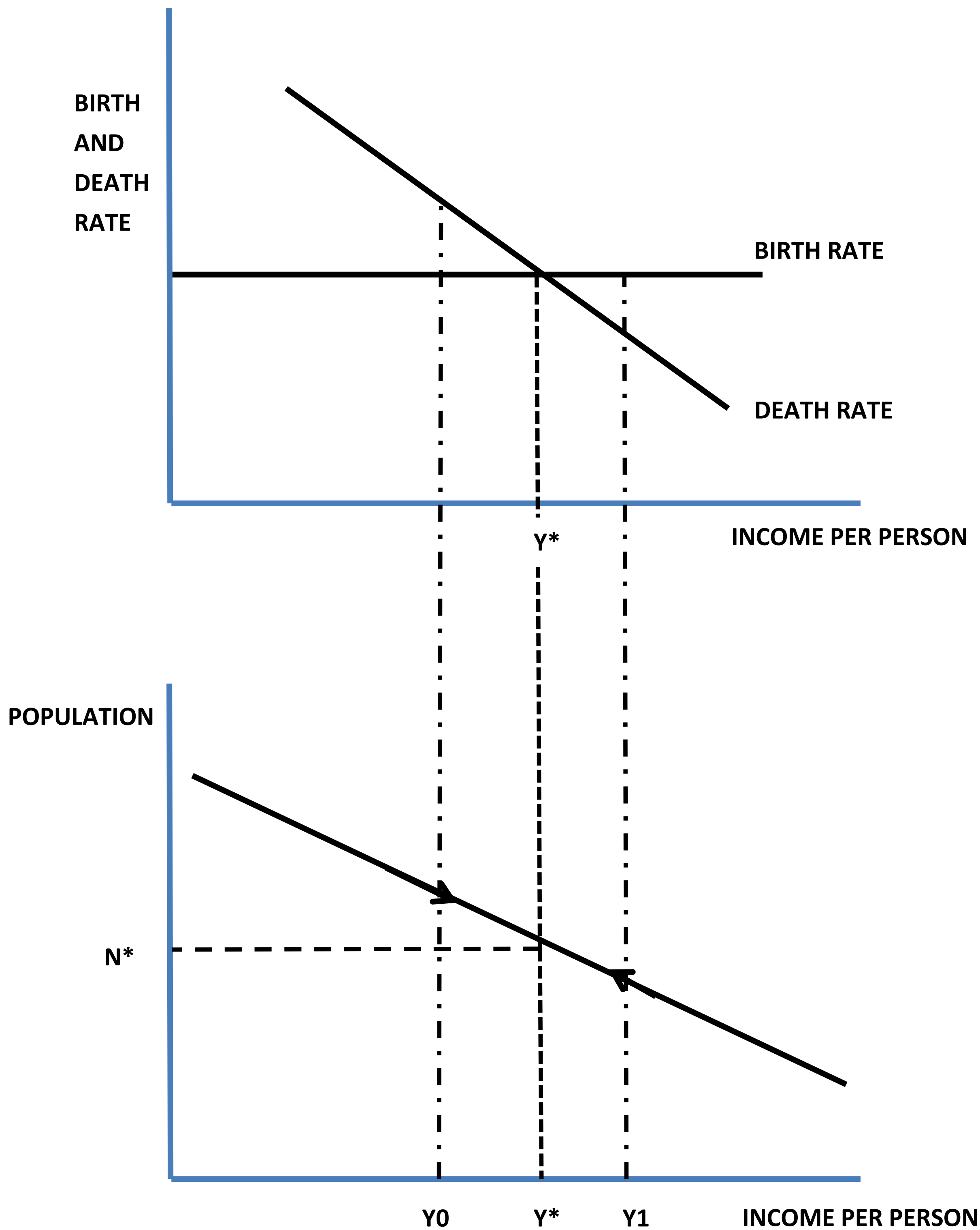


Figure 15.8: The Full Model



Calculating the Long-Run Equilibrium

We can figure N^* and Y^* out algebraically as follows. Combining the birth and death schedules with the equation for population change, we get

$$\frac{N_t - N_{t-1}}{N_{t-1}} = b - d_0 + d_1 Y_{t-1} \tag{15.11}$$

Inserting the dependence of income levels on wages, we get

$$\frac{N_t - N_{t-1}}{N_{t-1}} = b - d_0 + d_1 a_0 - d_1 a_1 N_{t-1} \tag{15.12}$$

This shows that the growth rate of population depends negatively on last period's level of population: This is what determines the convergent dynamics. The level of N such that population stays unchanged, shown in the Figure 15.7 as N^*, is given by

$$b - d_0 + d_1 a_0 - d_1 a_1 N^* = 0 \tag{15.13}$$

which solves to give

$$N^* = \frac{b - d_0 + d_1 a_0}{d_1 a_1} \tag{15.14}$$

The long-run equilibrium level of population depends positively on the birth rate, b, and on a_0, which effectively measures the level of technology in the model (if this increases it can offset the negative effect of higher population on income levels). The equilibrium level of population depends negatively on the exogenous element of the death rate (d_0), on the sensitivity of the death rate to income levels (d_1), and on the sensitivity of income levels to population (a_1).

The long-run equilibrium level of real income per person can be derived as the income level that gives a growth rate of population of zero

$$\frac{N_t - N_{t-1}}{N_{t-1}} = b - d_0 + d_1 Y^* = 0 \Rightarrow Y^* = \frac{d_0 - b}{d_1} \tag{15.15}$$

This level of income, as we noted above from the graphical illustration of the model, depends only on the parameters of the birth and death schedule and not at all on the parameters of the income-population schedule. So, for example, even if there was an increase in a_0 so that people could be paid more wages for each level of population, this would result, over time, only in higher population rather than higher income levels. Income levels depend negatively on birth rates, positively on death rates and negatively on the sensitivity of death rates to income levels.

A final way of illustrating the convergent dynamics of the model is to note that equation 15.12 for population growth can be re-written as

$$\frac{N_t - N_{t-1}}{N_{t-1}} = (d_1 a_1) \left(\frac{b - d_0 + d_1 a_0}{d_1 a_1} - N_{t-1} \right) \tag{15.16}$$

Using the formula for N^* in equation (15.14), this becomes

$$\frac{N_t - N_{t-1}}{N_{t-1}} = (d_1 a_1) \left(N^* - N_{t-1} \right) \tag{15.17}$$

In other words, the growth rate of population is determined by how far population is from its equilibrium level, with the speed of adjustment to this equilibrium, $d_1 a_1$, determined by the sensitivity of income levels to population and the sensitivity of the death rate to income levels.

How the Malthusian Economy Responds to Shocks

Finally, we consider three kinds of shocks to the Malthusian economy. In each case, we assume the economy starts at an equilbrium with population of N_0 and income levels Y_0. First, consider an increase in d_0 which shifts the death rate schedule up. Figure 15.9 illustrates what happens: At the starting level of income, Y_0, death rates now start to exceed birth rates.

Population falls and income rises until we reach the new higher equilibrium level of income Y_1 with its corresponding lower level of population N_1.

Figure 15.10 illustrates the consequences of an increase in the birth rate, b. This shock works in the opposite fashion to the death rate shock.

Finally, Figure 15.11 illustrates the consequences of a once-off increase in technology, i.e. an increase in a_0 so that people are able to earn more money at each level of population. The initial response to this shock is higher income levels. However, these higher income levels reduce the death rate and, over time, income levels return to their original equilibrium level. While income levels return to their original level, population is permanently higher because the new level of productivity permits a higher level of population than the old level.

There is an interesting contrast here between what happens when there is technological progress in the Solow model and when technology improves in the Malthusian model. The difference relates to the assumption in the Solow model that there is a consistent and non-trivial pace of technology increase. In the Malthusian model, the instantaneous effect of an increase in efficiency is an improvement of living standards. But this is offset over time by population increases if there aren't any further increases in technology.

In the Solow model, technology keeps increasing and keeps pushing up incomes every period, so the population can steadily increase without pushing income levels down. Greg Clark argues that while, cumulatively, there was a large increase in technology from ancient times to 1800, the pace of this increase was never fast enough to prevent population growth eroding its effects on living standards, so that prior to the Industrial Revolution, improvements in productive efficiency only translated into higher population.

Figure 15.9: A Shift in the Death Rate Schedule

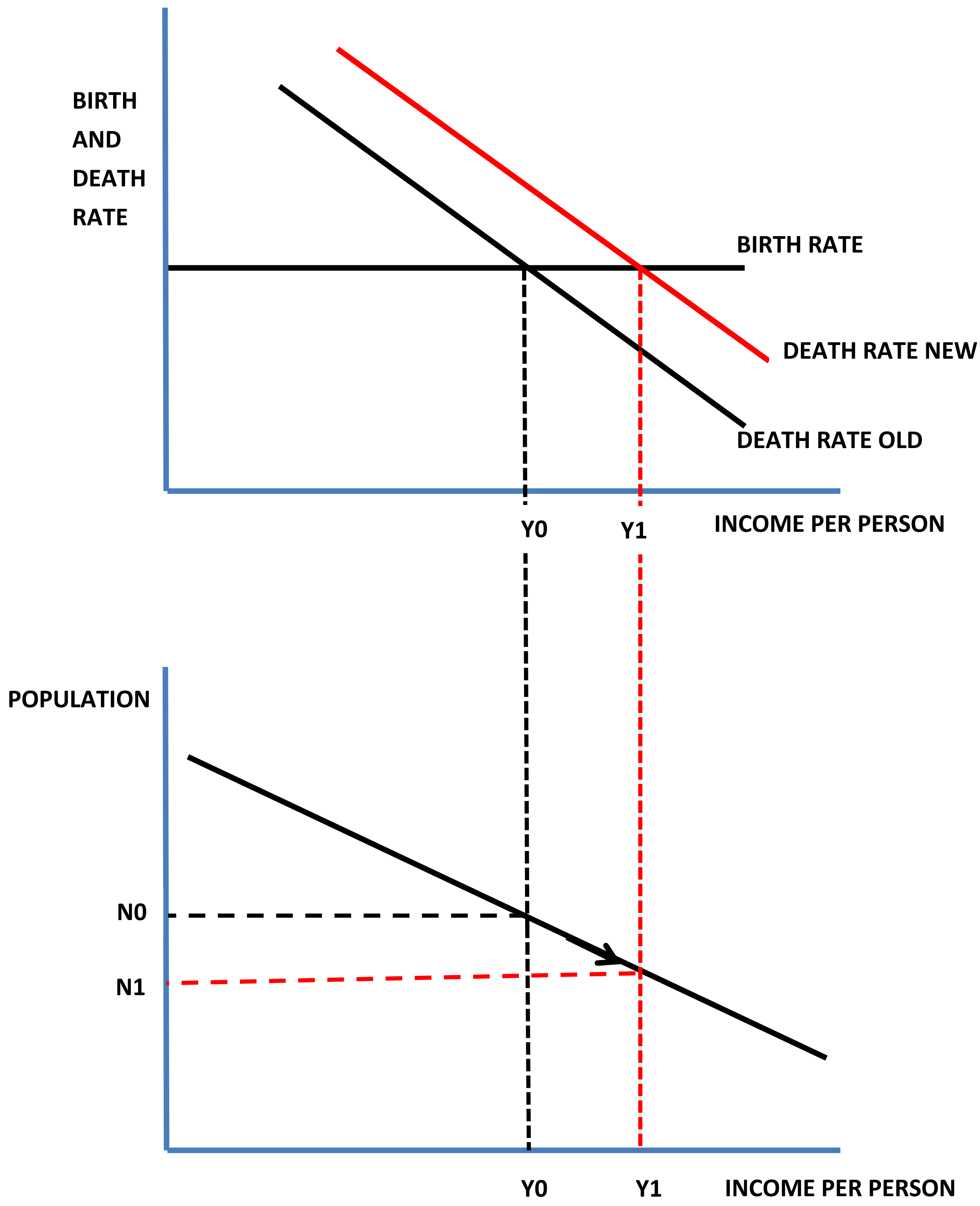


Figure 15.10: A Shift in the Birth Rate Schedule

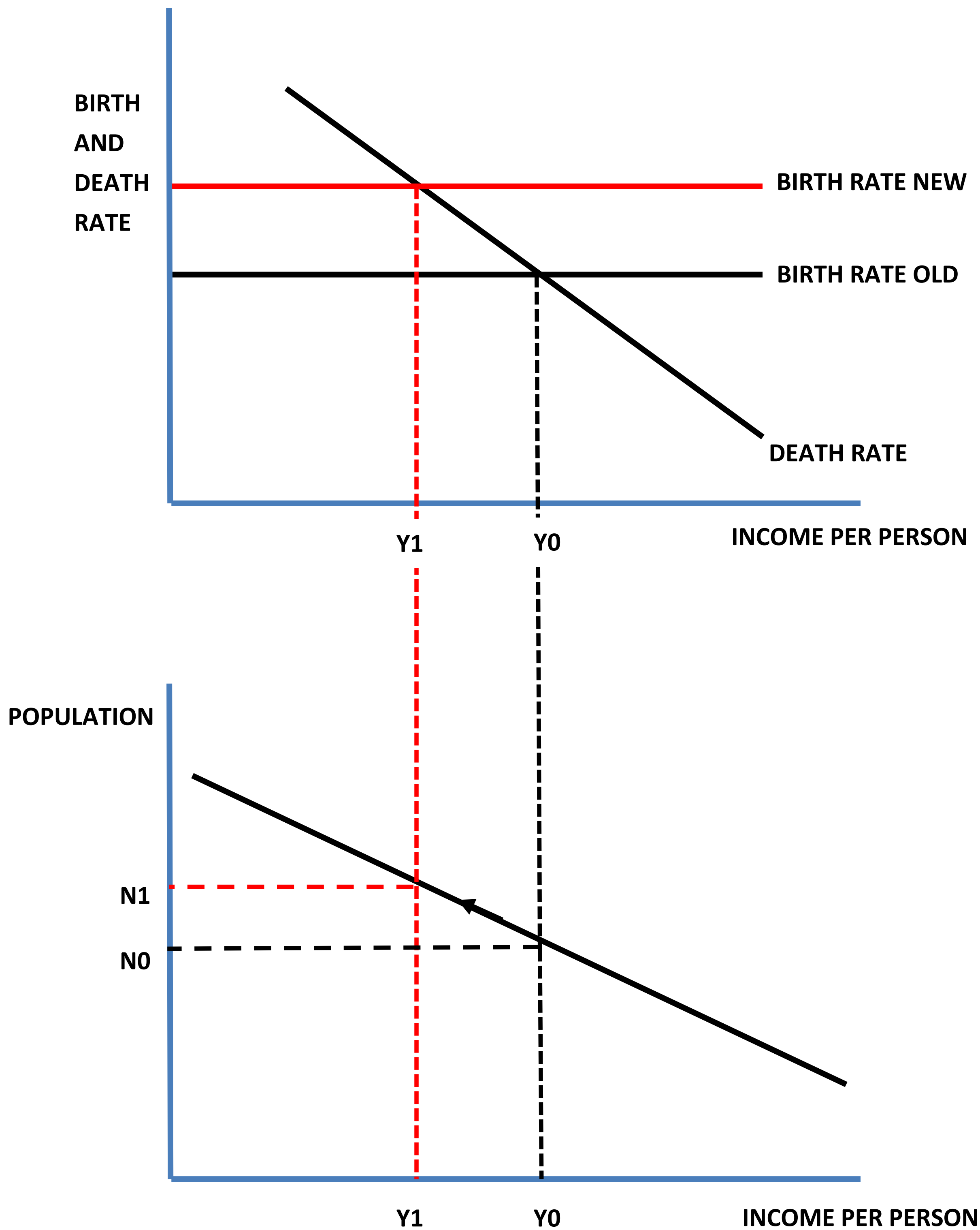
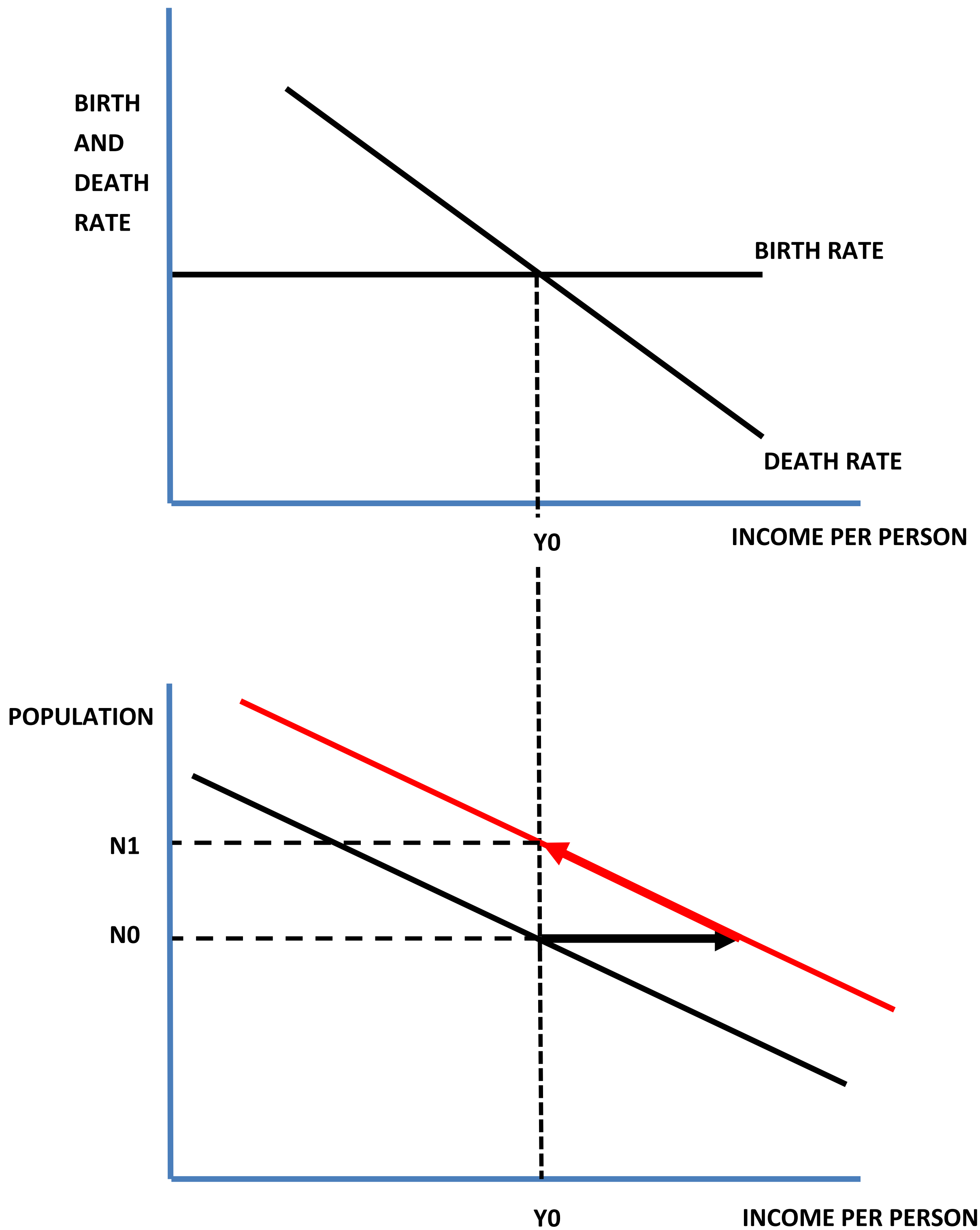


Figure 15.11: An Increase in Technological Efficiency



Malthus on the Poor Laws

The Malthusian model is one in which our usual understanding of what is good and what is bad is turned on its head. Things that we think are good, such as people living longer, turn out to be bad for average living standards, and things that we think are bad, like plagues and diseases, have a positive effect on those who survive. This non-intuitive worldview translated into Malthus's own policy recommendations. For example, he argued strongly against "poor laws" that provided assistance to the poor:

> *The poor laws of England tend to depress the general condition of the poor in these two ways. Their first obvious tendency is to increase population without increasing the food for its support. A poor man may marry with little or no prospect of being able to support a family in independence. They may be said therefore in some measure to create the poor which they maintain, and as the provisions of the country must, in consequence of the increased population, be distributed to every man in smaller proportions, it is evident that the labour of those who are not supported by parish assistance will purchase a smaller quantity of provisions than before and consequently more of them must be driven to ask for support.*
>
> *Secondly, the quantity of provisions consumed in workhouses upon a part of the society that cannot in general be considered as the most valuable part diminishes the shares that would otherwise belong to more industrious and more worthy members, and thus in the same manner forces more to become dependent. If the poor in the workhouses were to live better than they now do, this new distribution of the money of the society would tend more conspicuously to depress the condition of those out of the workhouses by occasioning a rise in the price of provisions.*

Over the years, Malthus has often been criticised for being overly-pessimistic about the fate of mankind and for opposing socially-progressive policies. However, it is worth noting the date that he wrote his famous essay—1798. Up until the time that he wrote his essay, his version of how the world worked actually described the economy remarkably well. It was only after his book was written that technological progress became fast enough to render his analysis less relevant.

Chapter 16

Malthus and the Environment

The Malthusian model may seem of interest today only for the light that it sheds on how the world worked before the Industrial Revolution ushered in an era of growth and increasing prosperity. Recall, however, that Malthus's views on how rising population reduced living standards focused on how increasing numbers of people placed pressures on the allocation of scarce resources, particularly food. In a world in which global population has just passed 8 billion, up from 4 billion in 1960 and 2 billion in 1927, it is reasonable to ask whether important global resources, such as energy sources, agricultural land and the global resource of a stable climate, can continue to withstand the strain of increasing population.

In this chapter, we will study a model that combines a Malthusian approach to population dynamics with an approach to modelling changes in a renewable resource base, which can expand or contract. The model was first presented by James A. Brander and M. Scott Taylor in their 1998 *American Economic Review* paper "The Simple Economics of Easter Island: A Ricardo-Malthus Model of Renewable Resource Use."

Easter Island

On Easter Sunday 1722, a Dutch explorer called Jacob Roggeveen came across a Pacific island that is believed to be the most remote inhabitable place in the world. Situated over two thousand miles west of Chile (see Figure 16.1) it is about 1300 miles east of its nearest inhabited neighbour, Pitcairn Island. Known as Easter Island since Roggeveen's brief visit, the island its inhabitants called Rapa Nui has had a long and fascinating history.

There is no written history of events at Easter Island prior to Roggeveen's visit so we are relying on the interpretation of archeological evidence to reconstruct what happened a long time ago. The interpretation I'm passing on in these brief notes comes from my reading of a chapter in Jared Diamond's book, *Collapse*, but there are archeologists and scientists who disagree with some aspects of this story.

Easter Island was probably first populated sometime around 900 AD. That it was ever populated, given its remoteness, is somewhat extraordinary. It seems likely that, once populated, it had little (and possibly no) contact with the outside world. The most remarkable feature of the island is its collection of hundreds of carved ceremonial statues featuring torsos and heads (see Figures 16.2 and 16.3) which were mainly built between 1100 and 1500. The natives most likely erected the statues as a form of religious worship. Evidence suggests that the island was divided into twelve tribes and they competed with each other (perhaps for local pride, perhaps for favour with the gods) by building larger and larger statues over time.

The statues were enormous. On average, they were 12 feet high and weighed 14 tons, while the largest weighs 82 tons. There is plenty of evidence to show that the statues required huge resources and that at least some of these resources were organised on a shared basis by a centralised leadership. Large teams of carvers were needed to create the statues and as many as 250 people were required to spend days transporting the statues around the island. When first populated, the island had large amounts of palm trees which supplied the resources for canoes, for tools for hunting and for materials used to transport the statues (sleds, rope, levers etc.) Estimates of peak population vary but it appears that the population peaked at about 15,000 in the early 1600s.

By the time Europeans began to visit the island one hundred years later, however, the island was largely deforested and population seemed to be as low as 3,000. Without palm trees, the islanders no longer had materials with which to build good canoes and this limited their abilities to catch fish. Without forests, the island lost most of its land birds, which had been an important source of meat. By the 1700s, the population survived mainly on farming, with chickens the main source of protein, but deforestation had also reduced water retention in the soil and led to soil erosion (the island is quite windy) so agricultural yields also declined.

Statue building had ceased by the early 1600s: Many of the statues remain today in various states of completion at the quarry at Rano Raraku where they were carved. Archeological evidence shows increasing numbers of spears and daggers appearing around this time, as well as evidence of people starting to live in caves and fortified dwellings. By the time Europeans arrived in the following century, tensions over food shortages had spilled over into intra-tribal rivalries with tribes knocking over the statues of their rivals. By the mid-1800s, all the statues had been toppled, so today's standing statues have been put in place in modern times.

There are many gaps in our understanding of what happened at Easter Island but the basic story appears to be that the population expanded to the point where the island's resources began to diminish and once population started to decline, the island went into a downward spiral. By the time Europeans visited in the seventeenth century, both population and resources had been greatly diminished from their peak levels.

The model laid out over the next few pages provides a description of how this can happen. We conclude with some thoughts about why it was allowed to happen and the potential implications for current global environmental problems.

Figure 16.1: The World's Most Remote Place

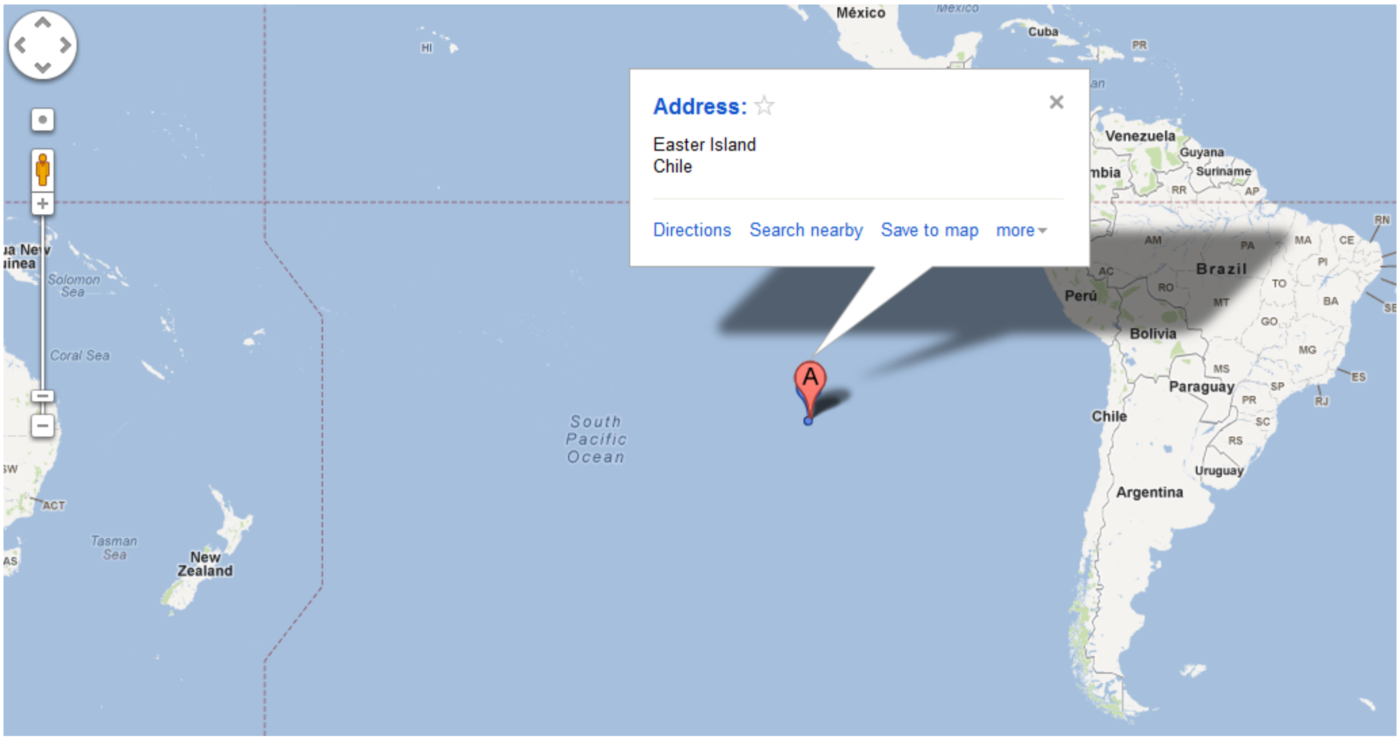


Figure 16.2: Easter Island Statues

Figure 16.3: Some Standing, Some Toppled

The Model

The model economy consists of population of N_t people at time t, who sustain themselves by collecting a harvest, H_t from a renewable resource stock denoted by S_t. Think of S_t as equivalent to a forest, or a herd of animals, or a stock of fishes; more realistically, think of it as the combination of a set of different resources of this type.

The model consists of three elements. The first element describes the change in population: This depends positively on the size of the harvest (a bigger harvest means less deaths and perhaps more births) and on an exogenous factor $d > 0$ such that without a harvest, there is a certain percentage reduction in population.

$$\frac{dN_t}{dt} = -dN_t + \theta H_t \tag{16.1}$$

The next element describes the harvest. We assume that the harvest reaped per person is a positive function of the size of the resource stock.

$$\frac{H_t}{N_t} = \gamma S_t \tag{16.2}$$

The final element, describing the change in the resource stock, is perhaps the most important. We are describing a resource stock that is renewable. It doesn't simply decline when harvested until it is all gone. Instead, it has its own capacity to increase. For example, stocks of fish can be depleted but will increase naturally again if fishing is cut back. So, our equation for the change in resources is

$$\frac{dS_t}{dt} = G\left(S_t\right) - H_t \tag{16.3}$$

The second term on the right-hand-side captures that the resource stock is reduced by the amount that is harvested. The first element is more interesting. It describes the ability of the resource to grow. Brander and Taylor use a logistic function to describe how the resource stock renews itself

$$G\left(S_t\right) = rS_t\left(1 - S_t\right) \tag{16.4}$$

This equation can be interpreted as follows. The maximum level of resources is $S_t = 1$: At this level, there can no further increase in S_t. Also, if $S_t = 0$ so the resource base has disappeared, then it cannot be regenerated. For all levels in between zero and one, we can note that

$$\frac{G\left(S_t\right)}{S_t} = r\left(1 - S_t\right) \tag{16.5}$$

So the amount of natural renewal as a fraction of the stock decreases steadily as the stock reaches its maximum value of one. This means that if the stock gets very low, it can grow at a fast rate if there is limited harvesting. However, if the stock is starting from a low base, the absolute size of this increase may still be small.

Dynamics of Population

We are going to describe the dynamics of this model using what is known as a *phase diagram*, which is a diagram that shows the direction in which variables are moving depending upon the values that they take. In our case, we are going to describe the joint dynamics of N_t and S_t.

Inserting the equation for the harvest, equation 16.2, into equation 16.1 for population growth, we get

$$\frac{dN_t}{dt} = -dN_t + \theta\gamma S_t N_t \tag{16.6}$$

This equations shows us that the change in population is a positive function of the resource stock. This means there is a particular value of the resource stock, S^*, for which population growth is zero. When resources are higher than S^* population increases and when it is lower than S^* population declines. The value of S^* can be calculated as the value for which the change in population is zero meaning

$$-dN_t + \theta\gamma S^* N_t = 0 \Rightarrow S^* = \frac{d}{\theta\gamma} \tag{16.7}$$

The resource stock consistent with an unchanged population depends positively on the exogenous death rate of the population, d, and negatively on the sensitivity of the population to the size of the harvest, θ, and on γ which describes the productivity of the harvesting technology.

Figure 16.4 shows how we illustrate the dynamics with a phase diagram. We put population on the x-axis and the stock of resources on the y-axis. Unchanged population corresponds to a straight line at S^*. For all values of resources above S^* population is increasing: Thus in the area above the line, we show an arrow pointing right, meaning population is increasing. In the area below this line, there is an arrow pointing left, meaning population is falling.

Dynamics of Resources

The dynamics of resources are derived by substituting in the logistic resource renewal function, equation (16.4), and the equation for the harvest, equation (16.2), into equation (16.3) to get

$$\frac{dS_t}{dt} = rS_t\left(1 - S_t\right) - \gamma N_t S_t \tag{16.8}$$

The stock of resources will be unchanged for all combinations of S_t and N_t that satisfy

$$rS_t\left(1 - S_t\right) - \gamma N_t S_t = 0 \Rightarrow N_t = \frac{r\left(1 - S_t\right)}{\gamma} \tag{16.9}$$

This means that there is downward sloping line in $N - S$ space along which each point is a point such that the change in resources is zero. This line is shown on Figure 16.5. The upper point crossing the S axis corresponds to no change because $S = 1$ and there are no people; as we move down the line we get points that correspond to no change in the stock of resources because while there are progressively larger numbers of people, the stock gets smaller and so can renew itself at a faster pace.

Remembering that equation 16.8 tells us that the change in the stock resources depends negatively on the size of the population, note now that every point that lies to the right of the downward-sloping $\dot{S}_t = \frac{dS}{dt} = 0$ line has a higher level of population than the points on line. That means that the stock of resources is declining for every point to the right of the line and increasing for every point to the left of it. Thus, in the area above the downward-sloping line on Figure 16.5, we show an arrow pointing down, meaning the stock of resources is falling. In the area below this line, there is an arrow pointing up, meaning the stock of resources is increasing.

The Joint Dynamics of Population and Resources

In Figure 16.6, we put together the four arrows drawn in Figures 16.4 and 16.5. This phase diagram shows that the joint dynamics of population and resources can be divided up into four different quadrants.

We can also see that there is one point at which both population and resources are unchanged, and thus the model stays at this point if it is reached. We know already from equation (16.7) that the level of the resource stock at this point is $S^* = \frac{d}{\theta\gamma}$. We can calculate the level of population associated with this point by inserting this formula into equation (16.9):

$$N^* = \frac{r\left(1 - \frac{d}{\theta\gamma}\right)}{\gamma} = \frac{r\left(\theta\gamma - d\right)}{\theta\gamma^2} \tag{16.10}$$

This level of population depends positively on r (so faster resource renewal raises population) and on θ (the sensitivity of population growth to the harvest) and negatively on d (the exogenous death rate coefficient).

This point is clearly some kind of "equilibrium" in the sense that once the economy reaches this point, it tends to stay there. But is the economy actually likely to end up at this point? The answer is yes: From any interior point (i.e. a point in which there is a non-zero population and resource stock) the economy eventually ends up at (N^*, S^*). It's beyond the scope of this class to prove formally that this is the case (the Brander-Taylor paper goes through all the gory details) but I can note that, after messing around with the equations, one can show that

$$\begin{aligned}
\frac{1}{N_t}\frac{dN_t}{dt} &= \theta\gamma\left(S_t - S^*\right) \\
\frac{1}{S_t}\frac{dS_t}{dt} &= \gamma\left(N^* - N_t\right) + r\left(S^* - S_t\right)
\end{aligned}$$

so the dynamics of both population and the resource stock are both driven by how far the economy is from this equilibrium point.

Harvesting and Long-Run Population

What does changing the parameter γ (which determines the fraction of the resources that is harvested) do to the equilibrium level of population? There are two different effects. On the one hand, a higher γ means people get to eat more of the harvest, which tends to increase population. On the other hand, the smaller stock of resources associated with the higher value of γ will tend to sustain fewer people. We can calculate the derivative of the equilibrium level of population with respect to γ as follows

$$\begin{aligned}
\frac{dN^*}{d\gamma} &= -\frac{r}{\gamma^2} + \frac{2rd}{\theta\gamma^3} && (16.11) \\
&= \frac{1}{\gamma^2}\left(\frac{2d}{\theta\gamma} - 1\right) && (16.12)
\end{aligned}$$

$$= \frac{r}{\gamma^2}\left(2S^* - 1\right) \qquad (16.13)$$

This shows that whether an increase in γ raises or reduces the equilibrium population depends on the size of the equilibrium level of resources. If the equilibrium level of resources is over half the maximum amount (which we have set equal to one) then we have $2S^* - 1 > 0$ and a more intensive rate of harvesting raises the population even though it reduces the total amount of resources. On the other hand, if the equilibrium level of resources is less than half the original maximum amount then we have $2S^* - 1 < 0$ and a more intensive rate of harvesting reduces the population.

For a place like Easter Island, where the economy ended up with a hugely diminished amount of resources, likely corresponds to the latter case, so it was an example of an economy that would have had a higher long-run population if they had harvested less.

Back to Easter Island

Let's go back to Easter Island and imagine the island in its early days with a full stock of resources and very few residents. What happens next? Figure 16.7 provides an illustration.

The economy starts out in what we can call "the happy quadrant" with resources above the long-run equilibrium and an expanding population. How do we know the dynamics take the "curved" form displayed in Figure 16.7? Well, when the economy crosses into the bottom right quadrant, in which population is now falling, the economy doesn't suddenly jump off in a different direction; the model's equations don't allow for any sudden jumps. Thus, the turnaround from increasing population to falling population must occur gradually over time.

So what happens to our theoretical Easter Island?

- For many years, the population expands and resources decline.
- Then, when it moves into the bottom right quadrant, population falls and resources keep declining.
- Then the economy moves into the bottom left quadrant where population keeps falling but resources finally start to recover.
- Then the economy moves into the quadrant in the triangle under the two curves and population starts to recover and resources increase.
- Finally, the economy moves back into the quadrant where it started but with less population and lower resources. The process is repeated with smaller fluctuations until it ends up at equilibrium with $S = S^*$ and $N = N^*$

Our theoretical Easter Island sees its population far overshoot its long-run equilibrium level before collapsing below this level and then oscillating around the long-run level and then finally settling down.

Figure 16.4: Population Dynamics

Figure 16.5: Resource Stock Dynamics

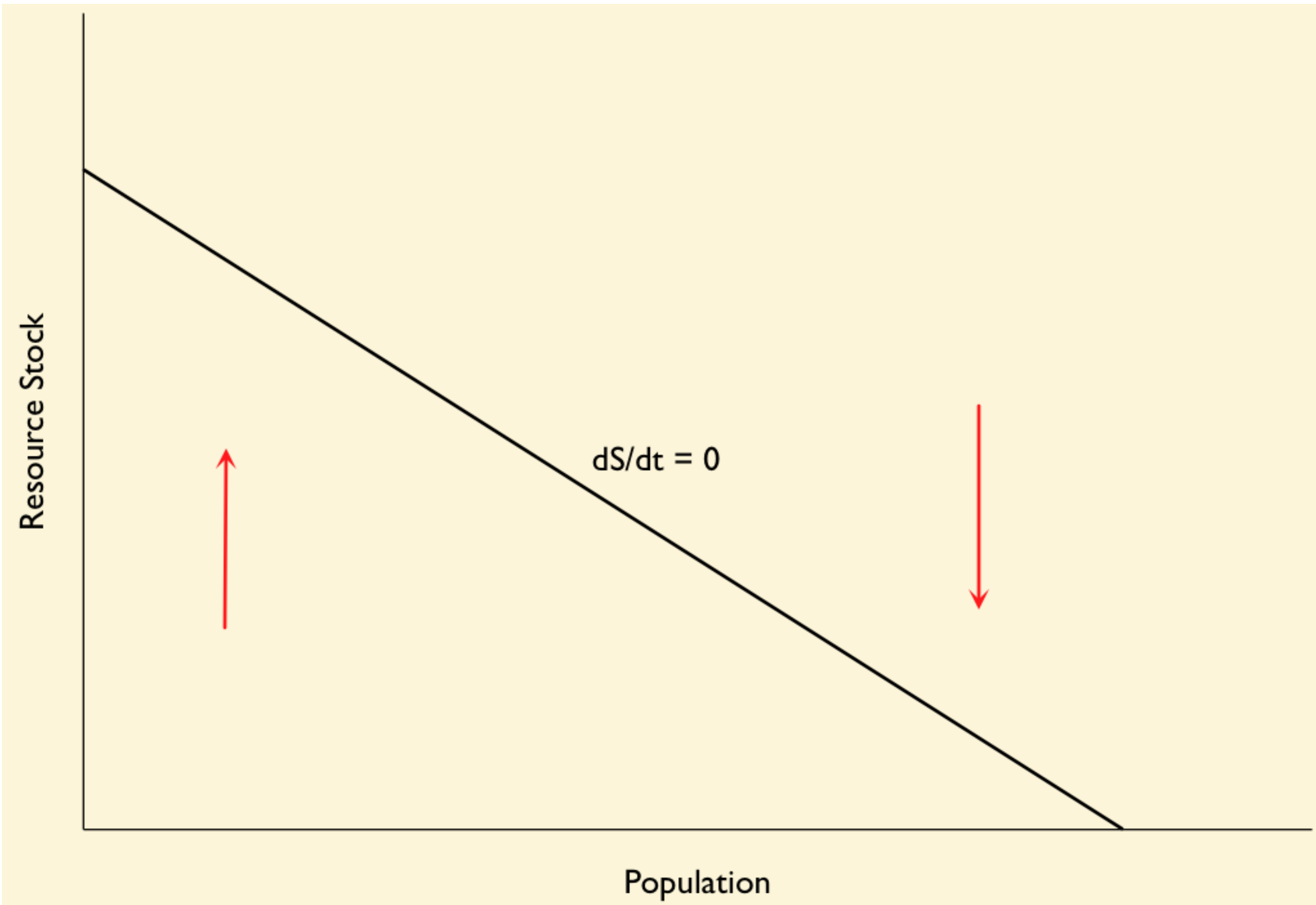


Figure 16.6: Dynamics Differ In Four Quadrants

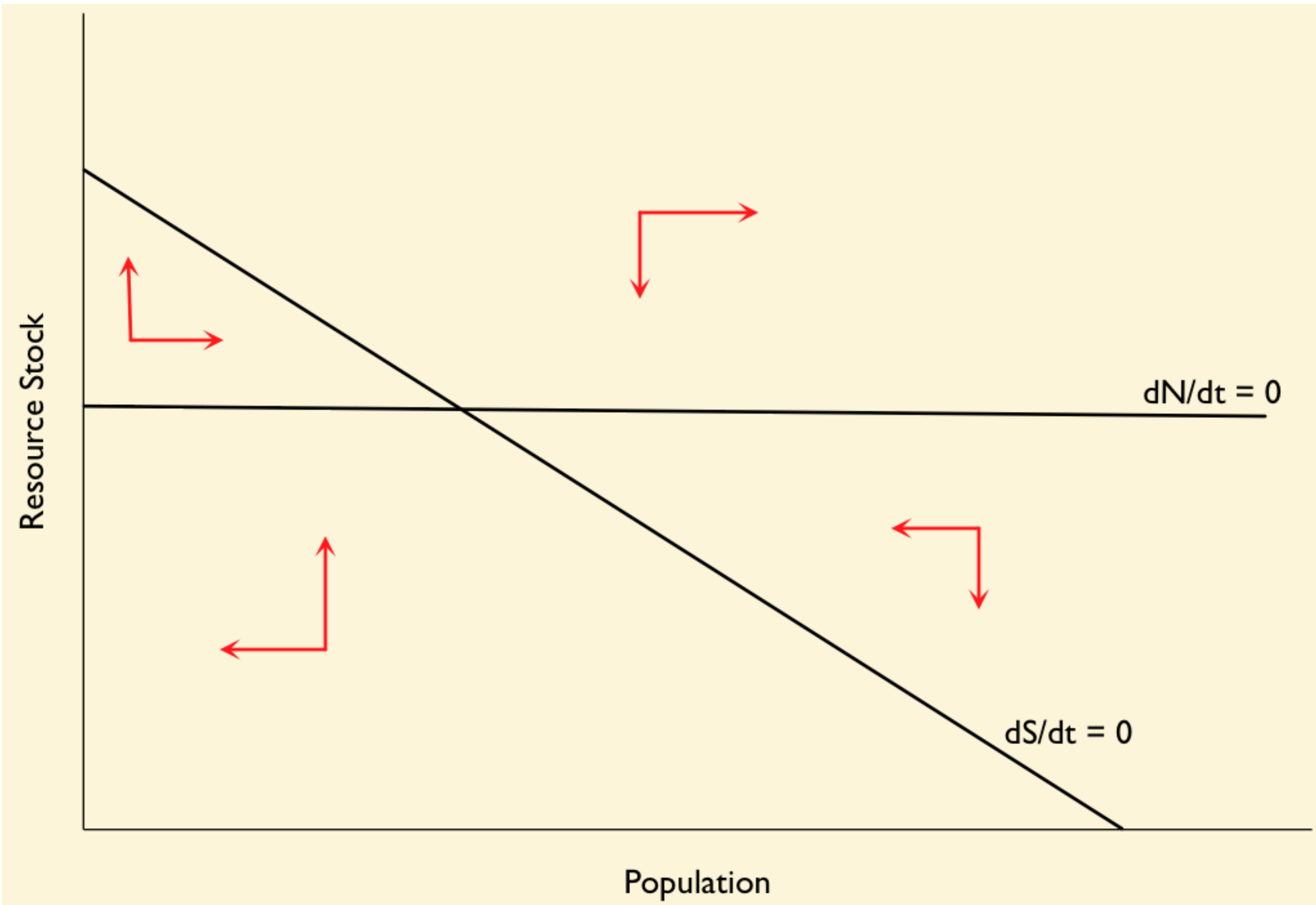
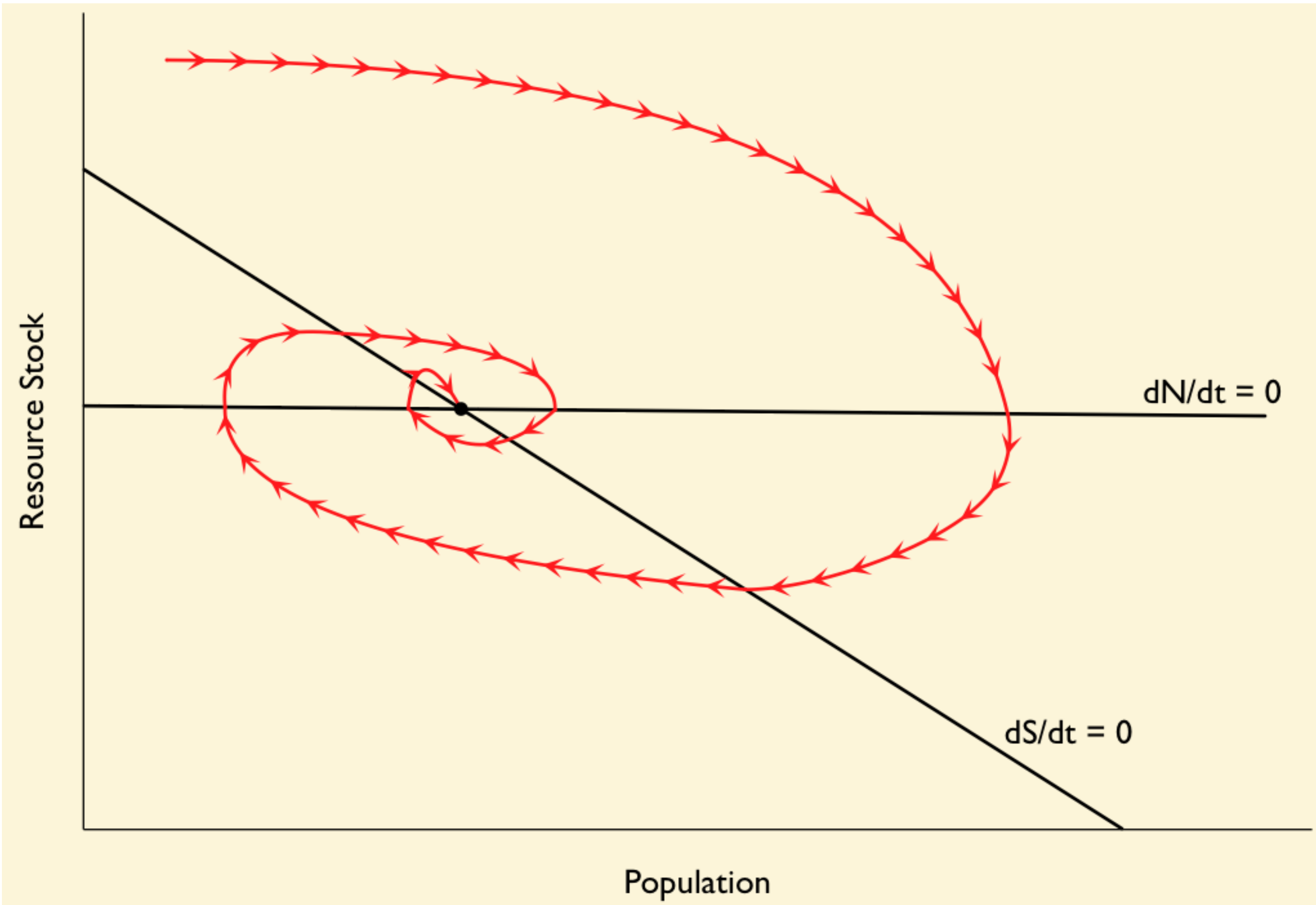


Figure 16.7: Illustrative Dynamics Starting from Low Population and High Resources



Numerical Example: A Lower Harvesting Rate

One of the ways to explore the properties of models like this one is to use software like Excel or more "programming-oriented" econometric software like RATS to simulate discrete-time versions of the model. Figures 16.8 and 16.9 show time series for resources and population generated from a RATS programme that simulates a discrete-time adaptation of the model. The programme is shown at the back in an appendix. It implements a version of the model with parameter values $d = 0.075, r = 0.075, \gamma = 2, \theta = 0.1$ and an initial population of $N_1 = 0.0001$. The parameter values are set so that the equilibrium level of resources is $S^* = \frac{d}{\theta\gamma} = \frac{0.075}{0.2} = 0.375$ while the equilibrium level of population is $N^* = \frac{r(1-S^*)}{\gamma} = \frac{(0.075)(1-0.375)}{2} = 0.023475$.

Figure 16.8 shows that, for these parameter values, the stock of resources falls to about half of its long-run equilibrium value, then rises and overshoots this value and then oscillates before settling down at this equilibrium level. Figure 16.9 shows the associated movements in population. We see population surge to levels that are over twice the long-run sustainable level, then dramatically drop to undershoot this level before eventually settling down.

Because we have chosen a base case in which $S^* < 0.5$, this is a case where there would be higher resources and population in the long-run if we had a somewhat lower rate of harvesting. Indeed, you can pick a rate of harvesting that avoids a collapse scenario altogether. Figures 16.10 and 16.11 compare the base case we have just looked at with a case in which the rate of harvesting was 40 percent lower, so $\gamma = 1.2$. In this case, the resource stock only slightly undershoots its long-run level and the population only slightly-overshoots. The economy ends up a similar level of population but arrives there in a less dramatic fashion.

Figure 16.8: Resources in a Simulated Easter Island Economy

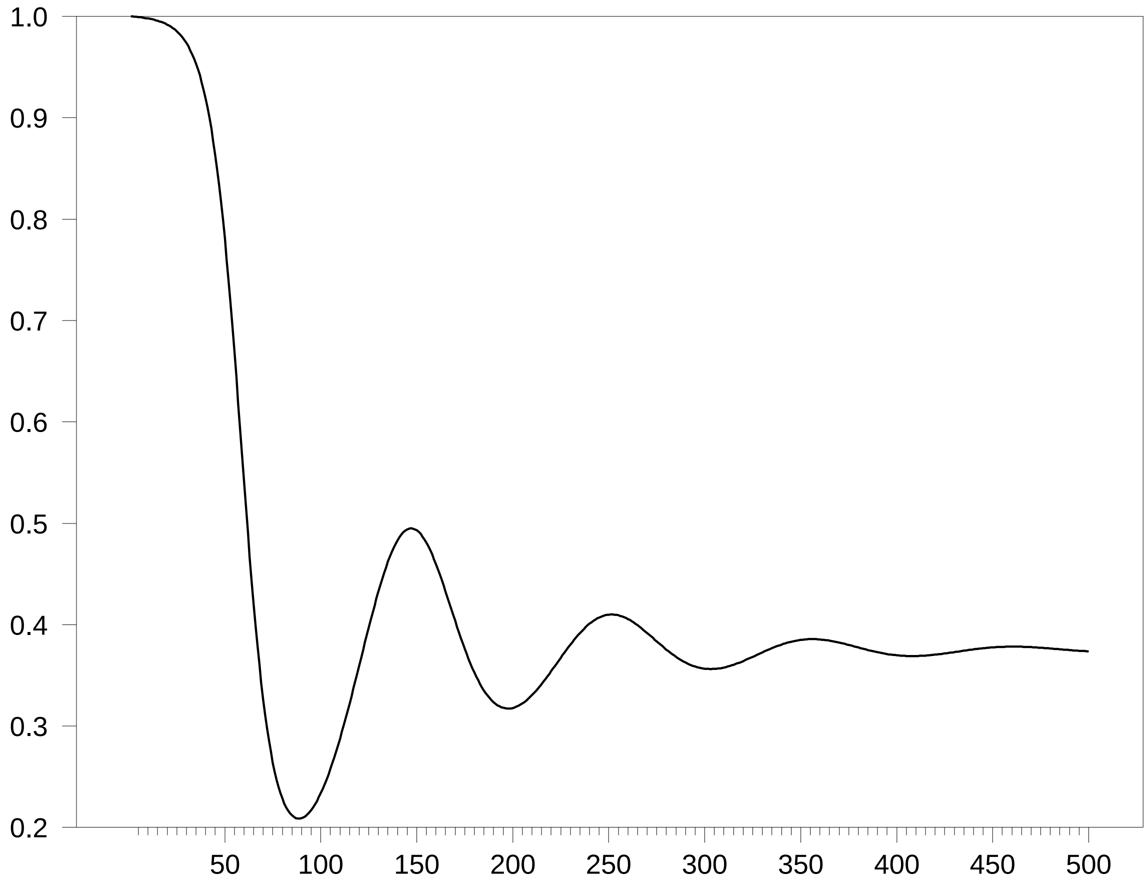


Figure 16.9: Population in a Simulated Easter Island Economy

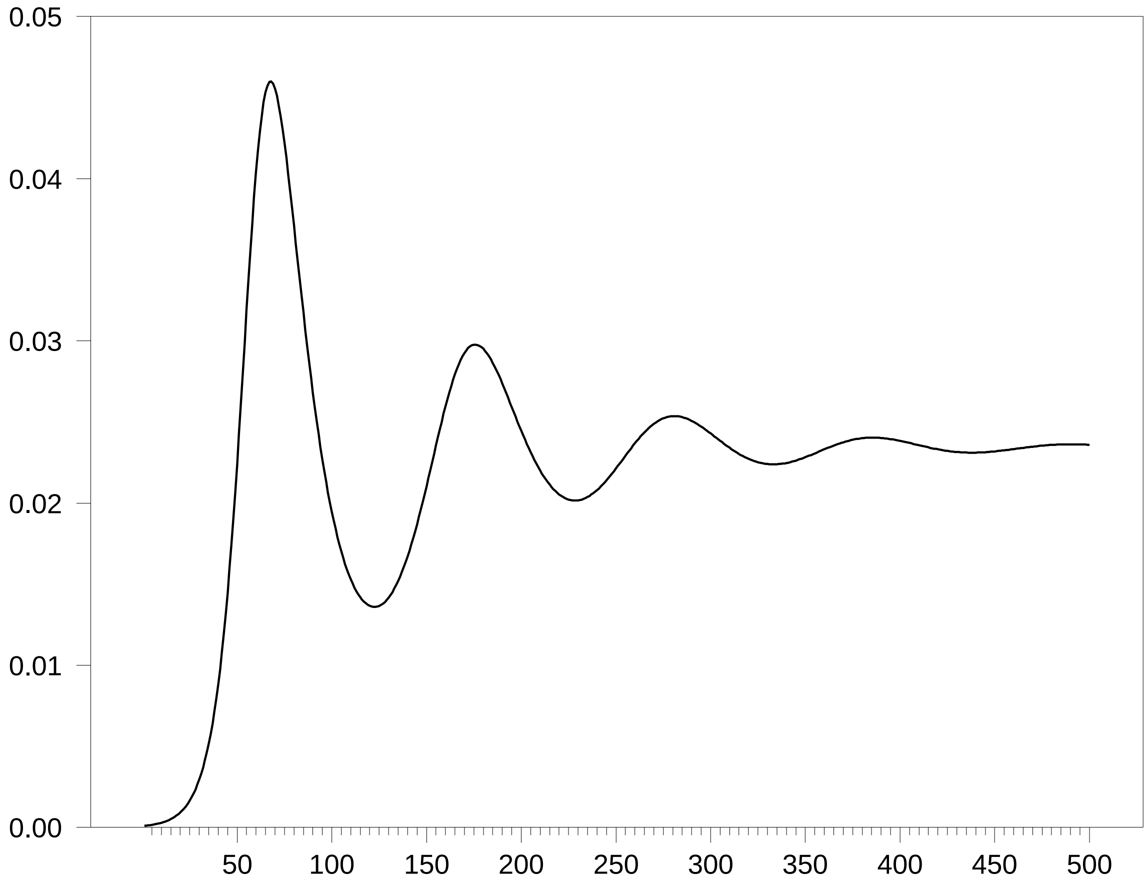


Figure 16.10: Resource Stock with Less Harvesting

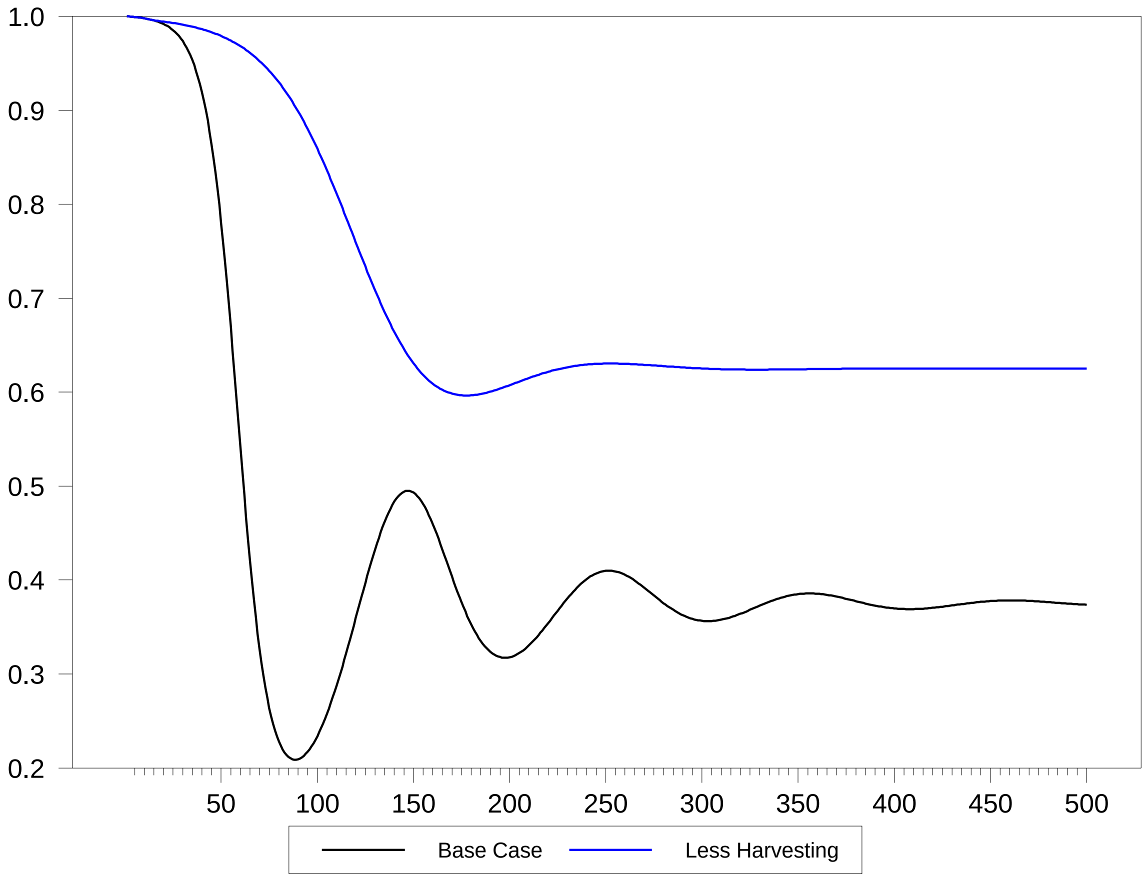
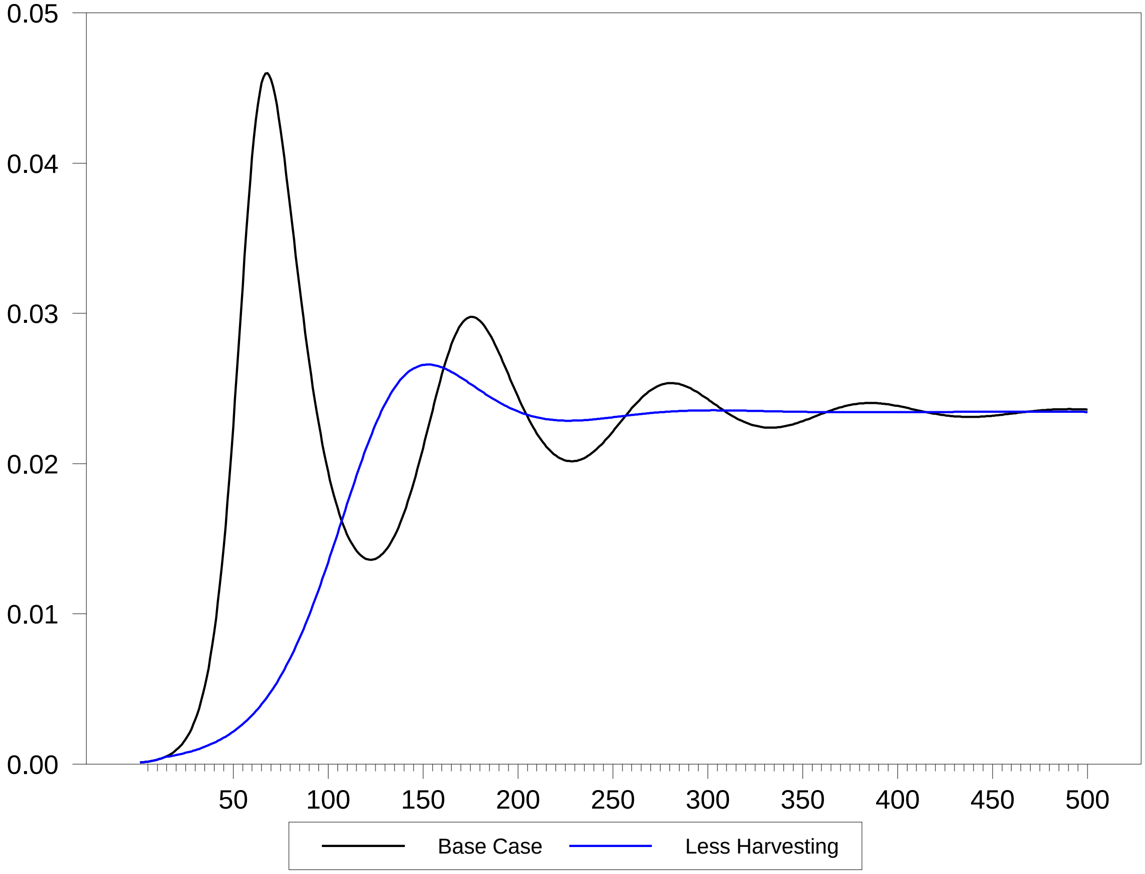


Figure 16.11: Population with Less Harvesting



Why Doesn't Someone Shout Stop?

The pattern demonstrated in the model—in which the economy far overshoots its long-run level before collapsing to an equilibrium with lower population and depleted resources—may seem to fit what happened at Easter Island. But it raises plenty of questions: Why did the residents of the island allow this to happen? Why didn't they establish better governance rules to prevent the deforestation that proved so devastating? And could this model possibly be a warning that today's global economy could represent an overshooting with a significant collapse awaiting us all?

In his book, *Collapse*, Jared Diamond discusses Easter Island and a number of other cases in which societies saw dramatic collapses, many triggered by long-term environmental damage. Diamond points to a number of potential explanations for why societies can let environmental damage occur up to the point where they trigger disasters.

- **The Tragedy of the Commons**: It may simply never be in anyone's interests at any point in time to prevent environmental degradation. A fisherman may acknowledge that excess fishing will eventually put him out of business but there may be little he can do to prevent others fishing and today he needs to earn an income. Some societies can put in place centralised political institutions to prevent environmental disasters and some cannot. At present, the society called The Earth is not known for its efficient centralised political decision making.

- **Failure to Anticipate**: Societies may not realise exactly how much damage they are doing to their environment or what its long-term consequences will be. Up until the point at which Easter Island's environment failed to support a growing population, there was probably a limited realisation among the population of the damage being done. Once the population began to shrink and the tribes turned against each other (there's some evidence of cannibalism during this period) the likelihood of a common negotiated solution to cut down less trees to preserve the environment was unlikely. Similarly today, the future effects of climate change are unpredictable and the costs (and even potential benefits) may be unevenly distributed.

- **Failure to Perceive, Until Too Late**: Diamond notes that environmental change often occurs at such a slow pace that people fail to notice it and plan to deal with it. The Easter Islanders of 1500 probably couldn't remember (and certainly had no written record of) their island being covered in palm trees. The islander who eventually cut down the last tree probably had little idea that these trees had once been the mainstay of the local economy. Similarly, global climate change has occurred at such a slow pace that, despite the mountain of scientific evidence that it is real, many simply choose to deny it.

Appendix: Programme For Easter Island Simulation

Figures 16.8 and 16.9 were produced using the programme below. The programme is written for the econometric package RATS but a programme of this sort could be written for lots of different types of software including Excel.

```
allocate 10000

set d = 0.075
set r = 0.075
set gamma = 2
set theta = 0.1

set s = 1
set n = 0.0001
set h = 0

do k = 2,10000
comp s(k) = s(k-1) +  r(k)*s(k-1)*(1-s(k-1) ) - h(k-1)
comp h(k) = gamma(k)*s(k-1)*n(k-1)
comp n(k) = (1-d(k))*n(k-1) +  theta(k)*h(k)
end do k

graph 1
# s 1 500

graph 1
# n  1 500
```