4 Open economy macroeconomics

Introduction

So far in this part of the book, we have been dealing with international linkages between macroeconomic variables. As such, we have looked at these relationships in isolation – the so-called partial equilibrium approach. In this chapter, we step back from the fine detail to consider the overall context within which open economy relationships operate to determine the exchange rate.

The general equilibrium model here will look familiar to most readers, since it is simply an extension of the closed economy aggregate supply and demand framework found in most modern undergraduate textbooks. Nonetheless, an understanding of this apparatus is essential to what follows in Chapters 5, 6 and 7, when it will be used to present three important models of exchange rate determination.

We start in Section 4.1 with an introduction to the open economy version of the IS–LM model of aggregate demand. It will be seen that, as far as the IS curve is concerned, allowing for the openness of the economy involves introducing an additional shift variable, the real exchange rate. The LM curve needs no modification at all to accommodate a floating exchange rate. With a fixed exchange rate, however, allowance needs to be made for the endogeneity of the money stock, a task undertaken in Sections 4.2.2 and 4.2.3.

The remainder of the chapter is concerned with the aggregate supply curve, in its classical and Keynesian forms. It is assumed here that openness makes no difference to the supply side of the economy model,[1] and so the analysis is identical to that found in macroeconomics texts.

4.1 IS–LM model of aggregate demand

4.1.1 IS curve

In an open economy, the national income identity can be written as follows:

$$y \equiv C + I + G + B \tag{4.1}$$

where y is (real) national income, C and I are expenditure on consumption and investment, respectively, G is net government purchases of goods and services, and B is the surplus on the current account of the balance of payments:

$$B \equiv exports - imports \tag{4.2}$$

that is, the excess of the country's exports over its imports.

If we subtract C, I and B from both sides of Equation 4.1, we have:

$$Savings - I - B \equiv G \tag{4.3}$$

where we have made use of the fact that savings are, by definition, income less consumption.

Now Equation 4.3 is simply an accounting relationship; as it stands, it can tell us nothing about what determines the level of aggregate demand. In order to reach any conclusions about the determination of aggregate demand, we must make some assumptions about what factors actually impinge on the components of aggregate demand.

Start with saving, which, for simplicity's sake, we can identify with the unspent income of the household sector.[2] What kinds of element are likely to have the most significant impact on the household's savings decision? As far as the individual household is concerned, it seems obvious that the higher its income, other things being equal, the more it will save.[3] Aggregating households, then, total private sector saving is likely to depend positively on the (planned) level of economic activity.

However, that is not all. At any given level of income, it seems probable that a rise in the level of interest rates in the economy will stimulate savings. This is likely because when interest rates are high, the incentive to save (or refrain from consuming) is great and vice versa when interest rates are low. In fact, the interest rate can be viewed as the premium set by society on present relative to future consumption. At a zero interest rate, one might expect saving to drop to zero.

Investment spending by the corporate sector[4] will depend on a comparison of the cost of funds for the purchase of equipment and so on with the profits to be expected from the investment. Now although, in practice, this calculation is bound to be hedged in with uncertainties surrounding the prospective cashflows to investment projects, there is one statement that can be made with reasonable confidence: other things being equal, the higher the interest rate in the economy, the greater the cost of capital and, hence, the less likely it is that any given prospective investment will appear profitable to the decisionmaker. In general, therefore, aggregate investment will vary inversely with the interest rate.

Notice that we have concluded that savings will increase as interest rates rise, while investment falls, so that the difference (savings less investment) will be *related positively* to interest rates.

Government spending will be taken as exogenously given – determined outside the model by factors (political, social, technological, and so on) beyond the scope of a mere economist. That is not to say we assume it never changes. On the contrary, we shall be very much concerned with analysing the effect of a change in fiscal policy. We are only saying that we shall not be concerned with questions of *why* or *how* fiscal expenditure changes.

Finally, but most important of all for present purposes, we come to the current account of the balance of payments. What are likely to be the main influences here? Among the many factors that affect the demand by UK residents for other countries' products and, conversely, the demand by foreigners for goods from Britain, there is almost certain to be one overriding consideration: the competitiveness of domestic relative to foreign output.

In a way, the issue is one that has already been covered in Chapter 2, where the PPP hypothesis was discussed, compared with the facts and appeared to fail. It was concluded there that, whether because of deficiencies in the way that price indices are calculated, because different countries produce different goods or because of the failure of the law of one price, PPP did not apply other than in the very long run. At the same time, it was also argued that it would be grossly implausible to go to the opposite extreme of supposing trade volumes to be completely unaffected by relative prices.

Consider a compromise to take account of the obvious fact that, invariably, the composition of one country's exports differs from that of another, so that country A's exports are only an imperfect substitute for those of country B. For example, the UK exports relatively little in the way of agricultural products, whereas a significant proportion of US exports consist of wheat, rice, soya beans, fruit, and so on, and thus cannot be said to compete with UK output.

Moreover, as was pointed out in Chapter 2, even where similar types of good are concerned, product differentiation means that direct, head-on competition in international trade is quite rare: a Cadillac is by no means a perfect substitute for a Jaguar, and neither is Bourbon for Scotch. Even where international trade in services is concerned, *Dallas* is not the same soap opera as *Coronation Street* or Disneyworld the same kind of attraction as Tower Bridge.

The conclusion reached in Chapter 2 was that indirect international competition between imperfect substitutes would mean that PPP would not necessarily be obtained as an equality at all times, but that, instead, *there would be an equilibrium price ratio fixed by the market at any moment*. In other words, there is an equilibrium relationship between the price of Bourbon and the price of Scotch, which is not necessarily one of equality. It might be, for example, that one bottle of Scotch equals one and a half bottles of Bourbon. If the price of Scotch were to rise to double that of Bourbon, then consumers in both countries would switch in increasing numbers to Bourbon.

What this implies for price levels in general is that it is relative competitiveness, that is, the *real exchange rate*:

$$Q \equiv \frac{SP^*}{P}$$

that determines the state of a country's current account. Recall that the numerator of Q is the price of foreign (that is, US) produced goods, measured in pounds. The greater Q is, therefore, the more competitive is domestic output. It follows that, *at higher levels of Q*, the *current account surplus is likely to be greater (or the deficit smaller) than at low levels*.

There is one other macroeconomic factor likely to influence the current account balance. Just as consumption of domestically produced goods and services rises with national income, the same is bound to be true of expenditure on imports, at a rate determined by the country's *marginal propensity to import*. The higher the income, the smaller is likely to be the surplus (or greater the deficit) on external trade, *other things being equal* – in other words, at any given real exchange rate. Notice that symmetry requires *our* (UK) exports to be greater when *their* (US) national income is higher, so that our external balance depends on the level of economic activity in the USA – which is one of the major channels through which booms and slumps spread from one country to its trading partners. However, since US national income is very much exogenous to a model of the UK economy, we ignore this relationship here.

We now look back at Equation 4.3 and incorporate our conclusions about how the components of aggregate demand are determined. To keep the analysis simple, suppose the relationship takes the following simple form:

$$by + zr - hQ = G_0 \tag{4.4}$$

where b, z and h are behavioural parameters, the coefficients of the unknowns in the equation and are all positive. The first term on the left-hand side summarizes the dependence of savings and imports on the level of economic activity, the second incorporates the positive relationship between interest rates and the private sector's saving net of its investment, and the third represents the current account as a function of the real exchange rate. On the right-hand side is the exogenous policy variable, G, fixed initially at the level G_0.

Now Equation 4.4 represents the equilibrium condition in the goods market. In other words, it shows the relationship that must hold between the three variables, y, r and Q, for there to be no excess demand for goods and services.

We shall proceed by taking a given value of Q, say Q_0, allowing us to rewrite the equation as follows:

$$by + zr = G_0 + hQ_0 \tag{4.5}$$

which reduces it to a relationship between y and r, for the given values of G and Q on the right-hand side.

Obviously, there are an infinite number of possible combinations of y and r that satisfy Equation 4.5. Consider plotting them on a graph, with r on the vertical axis and y on the horizontal.

Start by picking any arbitrary value for the interest rate, say r_0 (Figure 4.1), and ask yourself the question: at that level of the interest rate, what would have to be the level of national income if we were to have equilibrium in the product market? The answer is, of course, given by the value of y that solves Equation 4.5, when r takes the value r_0. If that value is y_0, then the combination (r_0, y_0) at the point A in Figure 4.1 is the first point we have found on the curve that we set out to plot. In other words, one possible solution of Equation 4.5 is given by:

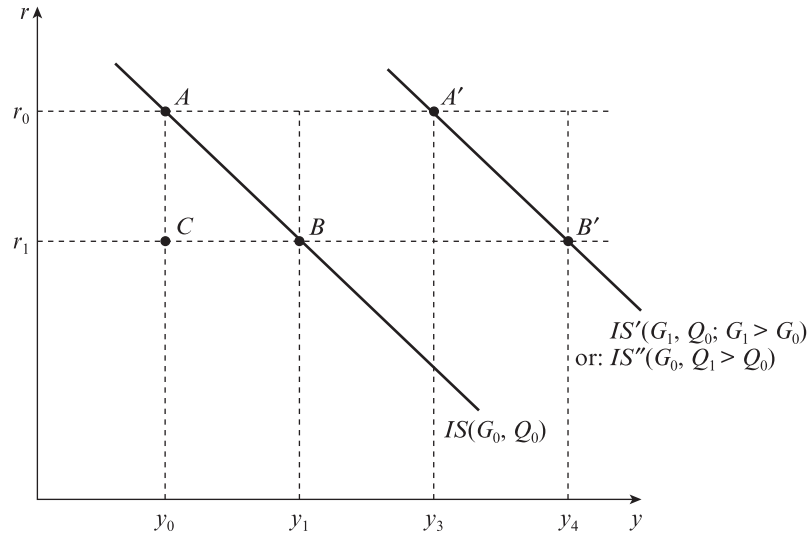

Figure 4.1 The IS curve

$$by_0 + zr_0 = G_0 + hQ_0 \qquad \textbf{(4.5a)}$$

To generate more points on the curve, we simply repeat the process, starting from a different interest rate, say r_1, lower than r_0. Now, at r_1, the term zr_1 will be smaller (remember that z is positive, as are all the parameters) so, at the previous value of y (that is, y_0), the left-hand side of Equation 4.5 will be smaller than before and hence smaller than the right-hand side. In terms of the economics, at the lower interest rate, the volume of saving will be smaller and the volume of investment spending greater than at A. Therefore, since A was an equilibrium, with saving net of investment just sufficient to finance the given deficits in the public and external trade sectors of the economy (as in Equation 4.3), net saving must be inadequate at the point C. In other words, there must be an excess demand for goods and services. The crucial point is that the *additional savings will forthcoming only be if the level of national income is greater than y_0, say, y_1*. B will be the next point on our locus of solutions to Equation 4.5 if it happens to be the case that by_1 is just great enough to offset the impact of the lower value of zr (that is, zr_1), so as to leave the left-hand side unchanged. In that case, y_1 will be the level of economic activity that stimulates a flow of savings sufficient to offset the otherwise reduced level of net saving associated with the lower interest rate, so that Equation 4.5 is satisfied at point B by:

$$by_1 + zr_1 = G_0 + hQ_0 \qquad \textbf{(4.5b)}$$

We see from this argument that the curve we are plotting, which is universally known as the IS curve, will be downward-sloping, always associating lower levels of the interest rate with higher levels of y.

Note, however, that the exercise we have just undertaken involved seeking solutions of Equation 4.5, for a *given value* of the right-hand side of the equation $G_0 + hQ_0$. For this reason, we have taken care to label the IS curve with the values of G and Q to which it relates. Obviously, an increase in net government spending, G, would make

this term larger. So also would a rise in the real exchange rate, Q (that is, a real devaluation), since, as we have already seen, h will be positive, reflecting the fact that the UK current account will have a larger surplus (or smaller deficit) the more competitive are British prices relative to those of the USA. Conversely, a lower value of Q will reduce the right-hand side of Equation 4.5, by making UK output less competitive.

Now consider the effect of an increase in the right-hand side of Equation 4.5 on the solution values for r and y. For example, go back to the interest rate r_0 and repeat the question: what would the value of y have to be were we to have equilibrium in the product market at this interest rate, now that the right-hand side of Equation 4.5 has increased? With a larger right-hand side, the left-hand side of the equation will need to be greater than before. In other words, where previously we saw that y_0 was the answer to our question, making the left-hand side into $by_0 + zr_0$, we must now have a greater value of y, say y_3. Similarly, when the interest rate is at the lower level, r_1, equilibrium now requires y_4 instead of y_1.

The logic of these conclusions is easy to follow. The right-hand side of Equation 4.5 is the sum of the public sector deficit and the foreign sector surplus. This represents the total finance required out of net saving by the domestic private sector. An increase in that requirement can be satisfied only by a change in y and/or r that serves to increase net savings in the economy. The change could be a rise in y, stimulating greater saving by households, or a rise in r, which would have the same effect but which would also cause a fall in investment by the corporate sector or a rise in both.

We conclude that the IS curve will shift to the right whenever there is any change that increases the right-hand side of Equation 4.5, whether a rise in net government spending, G, or an improvement in UK competitiveness (increase in Q).

Before proceeding, we summarize our conclusions about the open economy IS curve.

- The (open economy) IS curve is a downward-sloping line joining all combinations of the interest rate and the level of income, such that the flow of net savings is sufficient to cover the total financing requirements of the public and foreign sectors.
- It is drawn for given values of net government spending, G, and the real exchange rate, Q. Any increase in either G or Q or both will shift the IS curve outwards.

4.1.2 Money market

We now turn to a consideration of the conditions necessary for equilibrium in the money market. Before we go ahead, however, there is an important question to be settled: what do economists mean by the term 'money'?

The problem arises from the fact that 'money' is another of those words such as 'demand', 'supply', 'scarce' and a number of others used both by economists to mean something very precise and by non-economists to refer vaguely to something rather imprecise. Readers who are unfamiliar with the distinction will find it difficult to understand what is meant by concepts such as the 'demand for money', because in laymen's terms money is often a mere synonym for wealth or even, on occasion, for income.

From now on, we shall restrict ourselves to using the word in the way that economists do:

> **Money** refers to the asset or assets that are commonly used as a means of payment.

In other words, money is the name given to the assets used to finance transactions such as to pay for goods and services, to discharge debts and to make loans.

There are a number of things to notice about this definition. In the first place, nothing has been said about which assets actually *are* used as a means of payment, for the simple reason that this will vary from country to country and from one period to another. In medieval Europe, for example, gold and silver were the only widely accepted means of payment. Nowadays, at least in industrialized countries, transactions are conducted using either coins or banknotes or, most frequently, by cheque or electronic funds transferred directly between bank accounts.[5] In the latter case, it is the actual bank deposits that serve as the means of payment.

Notice also that this definition does not exclude the possibility that money assets may have other important properties, in addition to their usefulness as means of payment. Typically, money will also serve, to some extent at least, as a store of value, particularly in situations where short-term considerations are paramount. However, this feature of money is, of course, one it shares with all other assets (this is precisely what we mean by the word 'asset'), and so it can hardly be used to distinguish money from non-money assets. In fact, money is a small part of total wealth, both for typical individuals and for the economy as a whole.

In the context of wealth in general, it is not only the case that money has, to some extent at least, the same property as other assets – that is, its usefulness as a store of value. The opposite is also true: other assets may to a greater or lesser degree share the essential characteristic of monetary assets, functioning as a means of payment. Obviously, some assets (for example, real estate, capital equipment, and so on) are virtually useless for conducting transactions. On the other hand, many short-term financial assets other than bank current account deposits are so easily realized, so 'liquid', as to be strong candidates for inclusion in the definition of money – for example, deposits in building societies (savings and loan associations in the USA) and time deposits in banks.

Not surprisingly, a number of different operational definitions of money can be and have been used in practice: the narrowest, known as $M0$, or the monetary base, is currency in circulation, $M1$, which is simply $M0$ plus demand deposits, and, broader still, $M3$, which includes time deposits, and $M4$ more recently in the UK, which adds deposits in building societies as well as some wholesale deposit instruments.[6] How far one should stretch the definition of money to embrace near-money assets is essentially an empirical question and not one that is directly relevant to the subject of exchange rate determination. None the less, in so far as the reader may, on occasion, find it a help in understanding what follows, it is worth specifying more or less arbitrarily a particular measure of the money stock. That being the case, it is suggested that all statements in this book about the demand for or supply of money be interpreted as relating to the total of currency in circulation plus demand deposits ($M1$ in the official statistics) unless otherwise specified.

Demand for money

Perhaps no subject in macroeconomics has received as much attention from researchers as the demand for money, and so what follows can only be a greatly simplified overview of this vast literature. (See the reading guide for further references.)

Start with the following fundamental (and deceptively simple) question: why do people hold money? Why not hold all one's wealth in the form of other, non-money assets? Remember that currency earns no interest at all and neither, until very recently, did cheque accounts.[7]

The answer obviously relates to the special qualities of money already mentioned. Other assets can be used to execute payments, but nowhere near as easily as money. No other asset can offer as attractive a collection of characteristics as money: near-universal acceptability, portability, storability. Money as an asset is easily realized (that is, converted into other assets) and its value, although variable, is usually a lot easier to assess than it is for other assets, with the result that the transaction costs of using money are lower than for other assets. In general, to use an expression that summarizes all these features, we say that money is the most *liquid* of all assets.

It follows that, in holding a proportion of their wealth in the form of money, people are able to enjoy the advantage of liquidity. The greater their money balances, the more of this intangible, immeasurable, but very real benefit they enjoy. The less they hold, the more frequently they have to realize other assets such as stocks or bonds, life insurance policies, real estate, and so forth, in order to pay for transactions.

We can state, then, with complete confidence that *liquidity is always an inherently desirable property*. The statement is virtually a truism. Why then do people ever choose to hold illiquid assets? Why not hold all wealth in its most liquid form – money?

The obvious answer is that non-money assets offer a counterattraction to the liquidity of money. In exchange for the sacrifice of liquidity, non-money assets offer as compensation a return, which appears in a number of different guises: interest (on savings deposits, for example), yield (on bonds), dividend and, possibly, capital gain (on shares), rent (on property), and so on. Sometimes the reward for illiquidity is completely intangible – like the psychological benefit in the peace of mind given by risk insurance or the satisfaction of owning a prestige make of motor car. In other cases, it may be an 'own' return, such as the benefit of having a roof over one's head, which is part of the return on house ownership, or the cooling services yielded by an air conditioner. In the case of physical assets, the returns are often part pecuniary, part non-pecuniary. A painting hanging in one's home, for example, may yield an intangible return every time one sees it. It may also yield a capital gain when sold. Similarly, a house may yield pecuniary and/or non-pecuniary benefits.

Note that since there is a rental market for houses and for many consumer durables, it is possible to put a reasonably accurate value on the services they yield simply by seeing how much people are willing to pay in order to rent them.

In general, then, illiquidity is rewarded by the return on an asset. There is another way of putting the same point: liquidity involves sacrificing the return that could have been earned by holding a less liquid asset. In particular, holding money means sacrificing the return that could be earned on non-money assets. In the jargon of monetary economics:

The **opportunity cost of holding money**[8] is the return that could have been earned by holding a asset less liquid than money.

Notice that, in principle at least, there are as many possible measures of the opportunity cost as there are non-money assets. In particular, in an open economy context, the alternatives may include foreign securities or currency, as we shall see in Chapters 8 and 9.

We can now see the broad outline of how the demand for money mechanism works. On the one hand, economic agents need a stock of money balances in order to transact efficiently. The more transactions they want to conduct, the more money they would like to hold, other things being equal. On the other hand, holding one's wealth in the form of money involves a sacrifice, in the form of a forgone return. The choice as to how much money to hold involves a trade-off between the benefit in terms of transactions convenience and the opportunity cost.

So we have a theory stating that, broadly speaking, the demand for money will be greater the larger the volume of transactions and will be smaller the higher the return on non-money assets. All that remains at this stage is to make the theory operational by specifying, if possible, observable macroeconomic variables to act as indicators of the volume of transactions and of the opportunity cost.

Take the volume of transactions first. A transaction, for present purposes, is any activity where money normally changes hands – either in exchange for goods and services or in exchange for repayment at a later date.

Consider the relationship between transactions and the level of economic activity. Macroeconomics uses as its index of economic activity the aggregate known as national income, which includes only those transactions involving the generation of value added. National income thus excludes, among other things, purchases of raw materials, loans, gifts, government transfers, gambling, and so on, because they involve no net output in the domestic economy. However, all these activities are likely, at some stage, to involve the transfer of money balances and they are all therefore potentially relevant to the demand for money. Neither is that the only problem.

Even where transactions are properly associated with national income, factors such as money transfer technology, established payment practices, the structure of the economy, and so on are likely to influence the relationship between the number of transactions in the economy and the use of money balances. For example, other things being equal, the more infrequently employees are paid, the higher their *average* money balances will need to be.[9] Similarly, any change in the acceptability of near-money assets or in the efficiency with which money balances can be used (because of increased availability of credit cards, for example) would also be likely to affect the demand, for any given volume of transactions.

It follows that we can neither take national income as an infallible indicator of the volume of transactions nor rely on a completely fixed relationship between transactions and the demand for money.

However, suppose that the structure of the economy is fairly stable over some period. In particular, suppose that the kinds of institutional and technological factor that determine the way money is used in the economy are fairly stable and that the structure of the various industries is such that the volume of transactions bears some

stable relationship to national income. It may then be reasonable to assume that there would be a stable relationship between the level of economic activity and the volume of transactions.

If we ignore the opportunity cost argument for the moment, what we have arrived at is a relationship between the demand for money and national income, which could be summarized as:

$$M^d = kY \quad k > 0 \tag{4.6}$$

where M^d is the demand for money and Y is national income, both measured in *nominal* terms,[10] and k is a positive parameter.[11] The reason why M^d and Y are both defined in nominal terms should be obvious. Other things being equal, one would expect, say, a 10% increase in the real volume of transactions to have the same effect on the demand for money as a 10% increase in the price level at which the transactions are conducted. In fact, if we define:

$$Y \equiv Py$$

which just says nominal income, Y, is by definition the product of real income, y, multiplied the price level, P, at which it is traded, then we can rewrite Equation 4.6 as:

$$M^d \equiv kPy \tag{4.7}$$

This formulation is known as the Cambridge quantity equation. The quantity theory, of which it is one version, was the orthodox approach to what we now call the demand for money until well into the twentieth century. It is characterized by a concentration on the role played by the volume of transactions, to the exclusion of any other variables, in particular the return on non-money assets.

The corollary of ignoring the return on other financial assets is that we implicitly assume that economic agents choose between money and goods in general, and not between money and near-money assets. In this sense, Equation 4.7 is oversimplified. None the less, there will be occasions in the next chapter (Sections 5.1 and 5.2) when it will be convenient to make use of this formulation, because of its simplicity and because, very often, it is sufficient as it stands to generate important insights.

Before moving on, notice that by dividing both sides of Equation 4.7 by P, it can be rewritten as follows:

$$\frac{M^d}{P} = ky \tag{4.8}$$

which is, for some purposes, a more useful way of looking at the quantity equation.

The left-hand side is the demand for real money balances, in other words the quantity of purchasing power that the agents in the economy wish to hold in the form of money. The right-hand side is the constant k multiplied by the real income generated in the economy. Now if, as was often assumed by the classical economists, the level of economic activity can be regarded as fluctuating more or less randomly in a fairly narrow region around its long-run equilibrium level, then it follows that the right-hand side of Equation 4.8 must be reasonably stable and hence the demand for real balances must equally be stable.

To see why this is so important a conclusion, consider the effect of an increase in the money supply in this context.

Equilibrium in the money market involves a situation where the demand for money is equal to the supply. So, in the aftermath of an increase in the stock of money by, say, 10%, equilibrium can occur only when the demand has risen by the same amount.

Now look back at Equation 4.8. With the right-hand side broadly constant, the demand for real balances must, as we have seen, be more or less fixed – once the dust has settled, at least. And the demand for *real* balances will be constant only if the 10% rise in the demand for *nominal* money, M^d, is offset by an increase of equal proportions in the price level, P, thereby keeping the ratio M^d/P constant. In other words, each increase in the money supply generates an equiproportionate rise in the price index.

Not only that, but the converse is also true: in a quantity theory world, no increase in the general level of prices can occur without an accommodating rise in the money stock. Hence Milton Friedman's famous assertion, albeit on the basis of a far more sophisticated version of the quantity equation, that 'inflation is always and everywhere a monetary phenomenon'. With the real demand fixed, the real money stock must be pegged, which, in turn, means the numerator and denominator must move in parallel.

The power of these conclusions all stems from one critical simplification, which is the assumption, already mentioned, that agents choose between holding money and goods, rather than between money and bonds (or long-term deposits, savings accounts, and so on). It follows that money balances can be reduced only by spending, creating a direct transmission mechanism from excess money supply to additional demand for goods, which, with output fixed, must drive up the price level. Conversely, excess demand for money can be satisfied only by reducing spending on goods and not by selling other financial assets. Hence, excess demand for money is directly associated with excess supply of goods, as agents in aggregate attempt to replenish their balances. The result must be a fall in goods prices, on average at least.

It turns out that these conclusions all have to be modified as soon as we take account of the opportunity cost of holding money.

In order to decide how to measure the opportunity cost, one has, in principle, to decide first what are the relevant assets competing with money for a share in economic agents' portfolios. Again, at this theoretical level, any or all other assets are potential candidates such as savings deposits, bonds, stocks and shares, real estate or even consumer goods held as inventories (assuming they are not perishable). Much research and not a little controversy has centred on this issue, but it is for the most part not directly relevant to exchange rate determination,[12] and so we shall not open up this particular can of worms.

Instead, make the following assumption: suppose that whenever the return on one non-money asset goes up by 1%, all the other rates of return do the same, other things being equal. This is not as unrealistic an assumption as it might at first appear. Broadly speaking, the difference between the yield on, for example, government long-term securities and the same kind of paper issued by a private sector company is determined by factors unrelated to macroeconomics, and so there is no reason to expect it to change simply because interest rates in the economy rise. In general, if the relative liquidity of the various assets is unchanged, one would expect the returns they offer to stay the same. In fact, changes in the returns on different assets are so

closely correlated that it is very difficult in practice to identify a separate impact on the demand for money for more than two assets.

The advantage of making this assumption is that, if it holds good, it makes little difference which rate of return we choose in order to measure the opportunity cost of holding money. The simplest way to proceed, then, is to take an easily observable interest rate (the yield on treasury bills, for example) and refer to it from now on as *the* interest rate. If we do that, we can modify Equation 4.8 as follows:

$$\frac{M^d}{P} = ky - lr \quad k, l > 0 \tag{4.9}$$

The equation now expresses our contention that the demand for real balances will increase with the volume of transactions, but decrease with the opportunity cost, as measured by the interest rate, r.

Notice that our simple quantity theory conclusions about the impact of money supply changes no longer apply, unless interest rates can be assumed to remain constant, which is unlikely.

Moreover, if over some period the price level can be regarded as constant, an increase in the nominal money supply must amount to a rise in the value of real balances in the economy and this, in turn, must cause an increase on the right-hand side of Equation 4.9. Again, if changes in real income are ruled out this must imply a fall in r, so as to reduce the damping effect of the opportunity cost on the demand for real balances.

We can summarize the implications of this more sophisticated demand for money equation as follows. It states that the impact of an increase (or decrease) of x% in the supply of money will be either to cause the price level to rise (fall) by x% if the interest rate is unchanged *or* to push the interest rate down (up) if the price level is constant *or some combination of the two* – that is, a price change of less than x% in addition to an interest rate change.

Government budget constraint and money supply

In order to understand how the supply of money is determined, it will be helpful to start by making a detour to consider the mechanics of government budget finance.

Suppose that in some year the government decides to spend more than it receives in tax revenue. In other words, suppose the government wants to run a budget deficit. How can it finance spending in excess of its tax revenues?

Essentially, the answer to this question is that, just like you or I, a government can live beyond its means (that is, spend more than its income) only by reducing its net assets, in other words by borrowing from others, thereby increasing its liabilities, or by running down its accumulated assets ('borrowing from itself', so to speak). Since governments rarely have much in the way of accumulated assets with which to finance spending,[13] we shall assume any budget deficit is financed by borrowing of one form or another. Let us summarize this fact in the following identity:

$$G - T \equiv \textit{the budget deficit} \equiv \textit{total government borrowing}$$

where G is government expenditure on goods and services during the year and T is government tax revenue during the year.

Now there are many different forms in which a government can borrow, as indeed there are for an individual, depending on whether the borrowing is long- or short-term, secured or unsecured, indexed or unindexed, in negotiable or non-negotiable instruments, and so on. With one vital exception, we shall not be concerned with the particular form that government borrowing takes.

The single exception is that governments have one borrowing option open to them that is not available to the ordinary individual: they can issue, via the agency of the central bank, a kind of security that the public is willing to accept as money. The fact that some of the state's liabilities are universally acceptable in order to settle debts between parties outside the government sector gives the authorities another degree of freedom, an additional avenue for financing overspending that is not open to any other agency.[14]

In recognition of this fact, we can rewrite the identity, breaking down the right-hand side into components that reflect government borrowing in money and non-money terms:

$$G - T \equiv \Delta MB + \Delta B^s$$

where MB is the quantity of currency in existence, B^s is the quantity of nonmonetary government debt in existence ('bonds')[15] and Δ is, by convention, an operator denoting the change in a variable over any period.

This identity is usually given a special name:

> The **government budget constraint** is the identity that expresses the fact that all government spending over any given period must be financed by taxation, by issuing currency or by issuing non-money securities (typically, long-term debt, called from now on 'bonds').

Although it is an identity, in other words a truism, the government budget constraint is important because it summarizes the necessary relationship between fiscal policy, affecting net expenditure on the left-hand side of the equation, and monetary policy, determining the way in which spending is financed on the right-hand side.

Armed with an understanding of the broad outline of the government's funding problem, we can now proceed to a consideration of the money supply mechanism.

Supply of money in an open economy

The first thing to notice about the supply of money is that it is not really a supply at all – at least, not in the sense we use the term in elementary microeconomics. What we mean by the supply of money is nothing more than the quantity of money in existence in the economy at a particular point in time. It is not very helpful for present purposes to think of money actually being *supplied* to the market. Rather, think of the supply process as being simply a mechanism whereby the stock of money (currency plus demand deposits) gets determined.

It turns out that, in order to understand the process, we need to take a look at the structure of the banking system. The fine detail of the institutional framework is, for most industrial countries, bewilderingly complicated and characterized by all sorts of peculiar features, some attributable to legal or regulatory constraints, others

Central bank

Assets		Liabilities	
Gold and foreign currency reserves	FX	Currency issued ('monetary base')	MB
Lending to government	LG		

Commercial banks

Assets		Liabilities	
Currency *plus* deposits with central bank	MB^b	Deposits by public	D
Loans advanced to personal and corporate sectors	L		

Consolidated banking sector

Assets		Liabilities	
Gold and foreign currency reserves	FX	Currency in circulation: $MB - MB^b =$	MB^p
Domestic credit: $L + LG =$	DC	Deposits of public	D
Money supply: $FX + DC =$	M^s	Money supply: $MB^P + D =$	M^s

Figure 4.2 Balance sheet of the banking system

to custom and practice. To make matters worse, the whole subject is often shrouded in a fog of esoteric jargon.

We shall avoid most of the complications by dealing with the banking system in broad outline only, avoiding much of the fine detail or relegating it to footnotes that the reader can safely ignore without losing the thread of the argument.

One notable simplification made here and throughout the book, with the exception only of Chapter 9, is to ignore foreign holdings of domestic currency. As far as money supply control is concerned, it is hard to see why foreign holdings of sterling balances should create a problem. Certainly, any difficulties caused are likely to pale into insignificance beside those resulting from demand instability – or even beside the problems the authorities have created for themselves in the past.

Look at Figure 4.2, which lays out in schematic form the balance sheets of two kinds of institution.

The first balance sheet is for a central bank such as the Bank of England, the Federal Reserve Bank of the USA, the Bank of Japan or the People's Bank of China.[16]

Now, although it is the linchpin of a country's financial system, most people's only perception of their central bank is as the issuer of the banknotes they use. Certainly,

serving as the bank of issue is an important function,[17] but it is by no means the only one performed by the central bank. In the first place, the central bank has the job of holding the nation's reserves of gold and foreign currency[18] – hence the first entry on the asset side in the table, labelled *FX*. Second, and most important of all, the central bank differs from a commercial bank in having one large customer, the government, to whom it is forced to provide the main banking services of facilitating transactions and, inevitably, providing credit.[19]

At this point, we are back to the question of how the government funds its deficit spending, but this time we concentrate on the mechanics of the process. Suppose a government decides to spend £100m, without raising any additional taxation. Suppose, furthermore, that it intends to do so without borrowing from the non-bank public – in other words, it is determined to avail itself of the option to increase the money supply in order to pay for its spending. The first step in the process involves the government approaching the central bank for a loan, which is granted more or less automatically. In return, we can think of the treasury being compelled to sign IOUs, promising to repay the central bank at some uncertain future date. These IOUs (government securities) are then locked away in the vaults of the central bank.

This transaction is represented on the asset side of the accounts by the item 'Lending to government', consisting of the accumulation of government securities held by the central bank. The form the loans take is the printing of currency, which can then be used by the government for its additional expenditure. Notice that currency printed counts as a *liability* of the central bank, although neither the Bank of England nor the Federal Reserve Bank is actually obliged to redeem banknotes in any meaningful sense at all.[20]

Notice, also, that it is the supply of *currency* that is increased when the government borrows in this way from the central bank, and not the money supply in any but the narrowest sense. The notation ΔMB in the government budget constraint refers to the change in the quantity of currency issued or, as it is sometimes called, base money.

However, this is not the only way the monetary base can be changed. Just as you or I sometimes borrow in order to spend (or consume), we may also, on occasion, take a loan so as to buy another asset. In effect, we voluntarily take on a liability (the loan) in order to purchase an asset, such as a house, perhaps, some shares or a partnership in a business. In the same way, the government sometimes borrows in order to spend, as we have seen already, but may also at other times decide simply to raise a loan from the central bank in order to buy assets. Typically, the government may use the loan to buy back its own IOUs (mainly gilt-edged securities or bonds) from their holders among the non-bank public.

In the jargon of central banking, this type of activity is known as an *open market purchase* to distinguish it from transactions conducted behind closed doors – that is, between the government and the Bank of England.

As far as the impact on the monetary base is concerned, the effect will be to cause an expansion just as if it had used the loan to finance increased spending. Conversely, the government could decide to reverse the process, by issuing debt unnecessarily so as to raise cash direct from the non-bank public (an open market sale). If it subsequently spends the proceeds, the effect on the monetary base is nil. If, by way of contrast, it uses the proceeds so as to repay its debts to the central bank (reducing *LG*), it effectively retires money, reducing *MB*.

Now consider the commercial banking sector.[21] Start this time on the liabilities side of the balance sheet. There is only one major item here, deposits by the public with the banks, D. Clearly, your deposit with the bank in the high street is an asset to you. However, it must equally be a *liability* of the bank, since it represents an obligation on the part of the bank to repay a fixed amount on demand.[22] Thus, that part of the money stock that consists of demand deposits is in fact a liability of the commercial banking system.

What do banks do with the funds deposited with them by the public? If they are to be able to pay any interest on deposits, or indeed cover their costs,[23] they must put the funds deposited with them to work. The way they achieve this is, of course, by lending, mainly to corporate borrowers, although often also to private individuals. At any rate, the important thing to note is that these loans constitute *assets* of the banking system, since they are obligations by the non-bank sector of the economy to repay debts to the banking sector.

Now it would be very helpful, from the bank's point of view, if it could put out on loan every last penny it received on deposit. Unfortunately, it would also be grossly irresponsible, since it would be left with nothing in the till with which to repay depositors who wanted cash. As far as the public is concerned, it is on the understanding that they can liquidate their deposits for cash on demand that they agree to deposit in the first place. An inability to meet a demand for cash would almost certainly precipitate a crisis of confidence and a run on the bank. Prudence requires, therefore, that a significant proportion of the funds deposited with the banks be held as a cash reserve in order to cover demands for withdrawals.[24]

Hence, the first item on the asset side of the commercial banks' balance sheet, MB^b, which represents precisely these precautionary reserves. Currency in the commercial banks' vaults is kept at a manageable level by regular trips to deposit cash at the central bank – hence the armoured vehicles cluttering up the roads outside our high-street banks. By convention, no balance sheet distinction is made between the currency actually in a commercial bank's till at any moment and the amount it has available on deposit with its bank, the central bank. The total cash reserve is the commercial bank's share of the monetary base, MB^b in the figure.

The bottom third of Figure 4.2 contains a consolidated balance sheet for the banking sector as a whole, commercial banks plus central bank. To see how that is derived, we need to write the balance sheets of the central bank and the commercial banks respectively as equations. First the central bank:

$$FX + LG \equiv MB$$

Then the commercial banks:

$$MB^b + L \equiv D$$

Now, if we add the two balance sheet equations, we get:

$$FX + MB^b + (LG + L) \equiv MB + D$$

The term in parentheses represents the total of the consolidated banking sector's lending, both to the government and to the private, non-bank sector. It is known as *domestic credit*, DC for short. If we rewrite the equation making use of this definition and also taking MB^b away from both sides, we get:

$$FX + DC \equiv (MB - MB^b) + D$$

On the right-hand side, in parentheses, we have the total monetary base, MB, net of that part that is in the coffers of the banks themselves, MB^b. Obviously, currency not in the banks must be in circulation with the non-bank public. Call this part of the monetary base MB^p. Thus, we can rewrite our equation as:

$$FX + DC \equiv MB^p + D$$

The right-hand side now consists of the total of currency in circulation plus deposits in the banks – which is precisely how we defined the money stock at the beginning of Section 4.1.2. We can therefore conclude that:

$$FX + DC \equiv M^s \quad \textbf{(4.10)}$$

In words: the money supply is identically equal to the sum of the domestic credit generated by the banking system plus the value of the country's reserves of gold and foreign currency held at the central bank.

There are a number of points worth noting about this very important equation. Notice that it is an *identity*, in other words a relationship that is true by definition. It cannot be disputed or disproved. What, if anything, does it tell us? If we regard the left-hand side as showing the ultimate process by which the money stock is generated or backed, it tells us that every unit of domestic money must originate either in lending by the banking system (domestic credit) or in the reserves of foreign money. In terms of an expansion of the money stock, it says that each additional unit of domestic currency must have been generated by an expansion of domestic credit or, alternatively, by an increase in the reserves of gold and foreign exchange.

Now these two kinds of asset are qualitatively different. For present purposes, the essential difference is in the way they are created. As far as domestic credit is concerned, it is all directly or indirectly amenable to control by the home country's monetary authorities. To see that this is the case, look back at its two main components: central bank lending to the government and commercial bank advances to the non-bank sector. The former is self-evidently within the control of the auth orities – it is determined by the government's budget constraint in conjunction with any central bank open market operations that may be undertaken.

As far as commercial bank lending is concerned, the situation is not quite as clear cut.[25] Recall the discussion of how commercial banks were forced to keep reserves broadly in proportion to their liabilities. This amounts to saying that their lending is constrained by the available reserves of cash. We shall assume that, either by virtue of its control of the quantity of base money printed, or via its ability to fix, legally or otherwise, the relationship between commercial banks' lending and their till money (that is, the ratio MB^b/L), the authorities can determine the total domestic credit in the economy, DC. In other words, we shall assume that domestic credit is an exogenous, policy variable.

By contrast, we cannot possibly regard the foreign exchange reserves as a policy variable. For one thing, the Bank of England may not print US dollars or yen to put in the reserves – that is the prerogative of the Federal Reserve and the Bank of Japan respectively.

To understand this important distinction, recall what was said about the function of the reserves in Chapter 1. They represent a buffer stock of international money available, when necessary, to purchase the domestic currency so as to support the exchange rate. Their size is determined by the accumulation of gold and foreign

currency bought by the central bank on previous occasions to support other countries' currencies relative to the domestic.

Changes in the reserves therefore come about as a result of imbalances between the endogenous demand and supply in the currency markets. If there is an excess supply of the domestic currency (that is, excess demand for the foreign currency), a fall in the domestic country's exchange rate can be prevented only by using some of the reserves to buy the domestic currency, thereby satisfying the demand for foreign money. Conversely, the reserves will increase when there is an excess supply of foreign currency that must be bought with domestic money in order to prevent an appreciation.

Two conclusions follow. First, and most obviously, there is no point in a country's hoarding reserves if it does not intend to peg its exchange rate. At the very least, if it is content to see its exchange rate determined in the freely operating currency markets, it need never use its reserves. Second, since the balance of payments accounts represent, at least in principle, an analysis of the flow supply and demand in the currency markets, a deficit in those accounts will need to be offset by reductions in the reserves and vice versa for a surplus. Thus, whatever factors determine the balance of payments also determine the change in the reserves.

To see the importance of these conclusions, go back to our open economy money supply identity, Equation 4.1, and rewrite it in terms of discrete changes:

$$\Delta FX + \Delta DC \equiv \Delta M^s \tag{4.11}$$

The change in a country's money supply is identically equal to the sum of the change in the volume of domestic credit issued by its banking system *plus* the change in its stock of international reserves.

Now, if the country's authorities either have no reserves or choose not to use them, in other words if they are happy to see the value of their currency determined in world markets, then the first term in Equation 4.11 is always zero and changes in the domestic money stock originate exclusively in domestic credit expansion or contraction. Putting the point in terms of the balance of payments account, under these circumstances the balance for official financing is kept at zero.

It follows from all this that we can rephrase the definitions given in Chapter 1:

> A (**pure** or **cleanly**) **floating exchange rate regime** is one such that the balance of payments for official financing is identically zero, because the monetary authority either holds no foreign currency reserves or never uses them to intervene in currency markets.

This is not all, however. Since, in this case, any changes in the money supply are due entirely to domestic credit expansion or contraction, we can add the following to our definition of a floating exchange rate:

> Under a floating exchange rate, the domestic money stock changes only as a result of changes in the **lending behaviour of the domestic banking system**. Since the volume of lending can be controlled by the authorities, so can the money stock. In this sense, the money supply is a policy instrument at the disposal of the home country's monetary authority.

By contrast:

> A **fixed exchange rate regime** (including a '**managed**' or '**dirty**' **float**) is one such that the balance of payments for official financing is *not* identically zero, the surplus or deficit being covered by the domestic monetary authority's use of the foreign currency reserves to intervene in currency markets.

The implications of fixed exchange rates for the money stock have already been mentioned in this section.

Under any form of fixed exchange rate, the domestic money stock may change either as a result of a change in the volume of bank lending or as a result of a change in the foreign currency reserves (that is, a balance of payments deficit or surplus for official financing) – *or both*. Since the change in the reserves cannot be directly controlled by the domestic monetary authority, neither can the change in the domestic money stock.

It follows that, under a fixed exchange rate regime, the money supply cannot be regarded as a policy variable but instead will be an endogenous variable determined by whatever factors influence the balance of payments.

In a nutshell, then, under a floating exchange rate regime, the balance of payments surplus and the consequent increase in the reserves are fixed at zero, making the money supply exogenous and leaving the exchange rate to be determined endogenously by market forces. Under fixed exchange rates, the balance of payments, the change in the reserves and in the money supply are all endogenous. Only the change in the exchange rate is exogenously fixed at zero.

It follows that, in an open economy context, the term 'monetary policy' has two different meanings. Under floating exchange rates, 'monetary policy' means exactly the same thing that it does in textbook models of a closed economy – that is, the management of the money supply. By contrast, under fixed exchange rates, 'monetary policy' means control of the banking system's lending – in other words, what we have called *domestic credit*.

LM curve

Equilibrium obtains in the money market when demand is equal to supply. Combinations of income and the interest rate consistent with equilibrium are plotted along the LM curve.[26] Looking back at Equation 4.9, this means when:

$$\frac{M^s}{P} = \frac{M^d}{P} = ky - lr \quad k, l > 0 \tag{4.12}$$

where we bear in mind what was said in the last section about the determination of the money supply and its relationship to the exchange rate regime.

Now, if the money stock is fixed initially at the level M^s_0, Equation 4.12 is a single equation in three unknowns: r, y and P. To proceed, take a given level of prices, say P_0. If we do that, we are left with the equation:

$$ky - lr = \frac{M^s_0}{P_0} \quad k, l > 0 \tag{4.13}$$

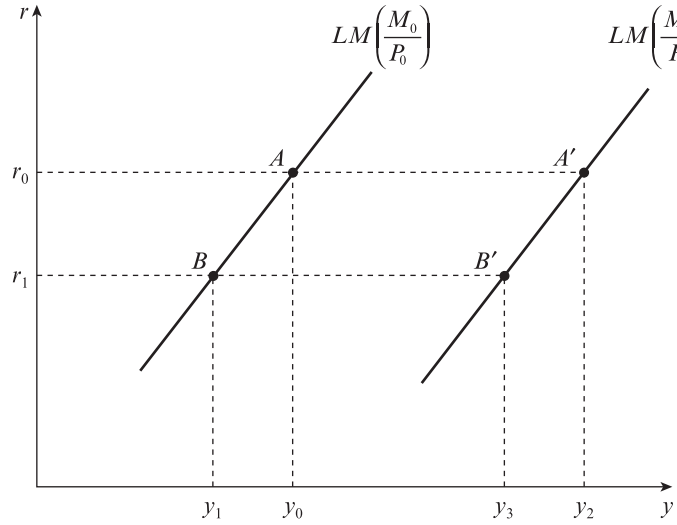

Figure 4.3 The LM curve

as the condition necessary for the money market to be in equilibrium – in other words, for the demand for money to be equal to the supply.

Notice that Equation 4.13 now has the same general form as Equation 4.5, the equation of the IS curve. The right-hand side is given, as before, and we can again generate the equilibrium conditions by picking an interest rate at random and asking ourselves the question: if the interest rate is r_0, what must be the level of y if the money market is to clear? The answer, y_0,[27] is the first point we have located on our LM curve, because it is the first combination of r and y we have found that generates a demand for money just equal to the given supply (point A in Figure 4.3).

Now consider a lower interest rate, r_1. The implication of a lower interest rate is a smaller penalty for holding non-interest-bearing assets and hence less of an incentive to economize on money holdings. It follows that, at r_1, the demand for money would be greater than the supply at the previous level of activity, y_0. Equilibrium therefore requires the lower interest rate to be offset by a lower level of activity, so as to damp down the demand for money for transactions purposes and thereby compensate for the reduced incentive to make money balances 'work harder'. It follows that money market clearing requires a lower interest rate (like r_1) to be associated with a lower level of real activity (like y_1).

We conclude that the LM curve drawn for a given level of real money balances must be upward-sloping.

As before, consider a change in the right-hand side of Equation 4.12. Clearly, an increase in the supply of real money balances could come about as the result of a *rise* in the nominal money stock or a *fall* in the price level at which it is valued. In fact, as we saw, the impact of, say, a 10% rise in the numerator, M^s, would be identical to that of a 10% fall in the price index, since it would have an identical effect on the ratio M^s/P. Anything that leaves the ratio unchanged, for example a 25% increase in both the nominal money stock and the price level, will have no effect on the LM curve.

From whatever source, an increase in the real money stock will mean that previous equilibrium combinations of r and y are now associated with an excess supply of real balances. At any given interest rate, say r_0, there will have to be a compensating change in order to increase the demand. Clearly, this must mean a higher level of activity than at the previous equilibrium. y_0 will no longer be associated with a high enough transactions demand. An income level like y_2 will be required. Putting the same argument somewhat differently, at any level of activity (and associated transactions demand), the larger real money stock makes possible a lower interest rate, because it reduces the required equilibrium opportunity cost of money holding.

We can summarize our conclusions about the LM curve as follows:

- The (open or closed economy) LM curve is an upward-sloping line joining all combinations of the interest rate and the level of income, such that the demand for real money balances is exactly equal to the supply.
- It is drawn for given values of the nominal money stock, M^s, and the price level, P. Any increase in the ratio M^s/P, whether associated with a rise in the nominal money stock, M^s, or a fall in P, will cause an outward shift in the LM curve and vice versa in the case of a fall in M^s or rise in P.

4.1.3 Aggregate demand

Figure 4.4(a) combines the IS and LM curves so as to generate a unique solution to the two simultaneous Equations 4.5 and 4.13. In other words, the value of y at A, y_0, is the answer to the question: what would be the equilibrium level of income if government fiscal spending amounted to G_0, the money supply were set at the level M_0^s, the price level were P_0 and the real exchange rate were Q_0?

However, the term 'equilibrium income' has to be interpreted with care in this context. It means nothing more than 'the level of income consistent with equality between planned spending and income, on the one hand, and also consistent with equality between the demand for money and the supply, on the other'. While this may seem quite a mouthful, it is not yet sufficient for a complete equilibrium. What is missing is some consideration of how the output y_0 relates to the productive capacity of the economy – it might, for example, be way beyond maximum potential output levels. Conversely, it might be so low as to imply mass unemployment of labour and/or capital.

These considerations lie behind the determination of the aggregate supply schedule, which will be dealt with in Section 4.2. For the moment, in recognition that the value of y that corresponds to equilibrium in the IS–LM model may not be the output actually produced, we shall refer to it as aggregate demand, denoted y^d.

Now ask yourself the question: suppose that aggregate demand is greater than the economy's productive potential – what is likely to happen? The intuitively obvious answer is that the price level will rise – and, on this occasion, the intuition is correct. How will a rise in the price of goods and services affect aggregate demand? The answer is given by the aggregate demand schedule, which plots how aggregate demand changes as the general price level changes.

In Figure 4.4(b), aggregate demand is plotted against the price index, taking G, M^s and SP^* as given. Hence the equilibrium at A in the IS–LM diagram is associated with the point on the aggregate demand curve labelled a.

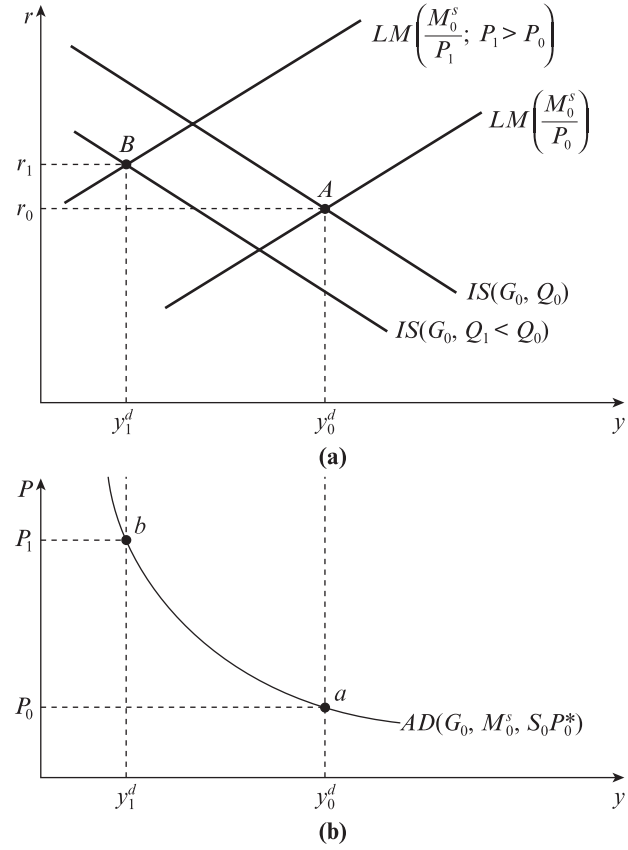

Figure 4.4 Derivation of the aggregate demand curve

The remainder of the points on the aggregate demand schedule are derived simply by varying the price index upwards or downwards from P_0. Take, for example, the higher price level, P_1. As far as the LM curve is concerned, we have seen that a higher price level reduces the value of the given nominal money stock, because it must be the case that:

$$\frac{M_0^s}{P_1} < \frac{M_0^s}{P_0}$$

if P_1 is greater than P_0. It follows that the LM curve for P_1 is further to the left (that is, higher up) than for P_0.

What about the IS curve? The only effect of a higher price level on the IS curve arises via the international relative price term, Q. *Other things being equal*, and, in particular, assuming that neither the nominal exchange rate nor the foreign price level changes, the higher domestic price level reduces the competitiveness of UK production, pushing the IS curve downwards and to the left.

Since both shifts are to the left, the outcome must be a lower level of aggregate demand, say y_1^d instead of y_0^d.[28] Plotting the outcome on to the bottom diagram we conclude that the aggregate demand curve will be downward-sloping, for given values of the money stock, fiscal expenditure, the nominal exchange rate and the foreign price level.

The aggregate demand curve AD_0 was drawn for given values of G, M^s and SP^*. Now consider the effect of changes in each of these variables in turn.

At any given price level, an increase in the nominal money supply will, as we have seen, shift the LM curve outwards, moving the equilibrium down the IS curve to the southeast, in the direction of higher aggregate demand and lower interest rates. It therefore follows that it will shift the aggregate demand curve outwards, accommodating higher demand at all price levels.

An increase in net government expenditure has effects that are similar, but not identical. As we have already established, a fiscal expansion shifts the IS curve outwards, other things being equal, so that the new equilibrium involves a higher level of both the interest rate and aggregate demand in the economy. The greater the rise in interest rates, the smaller the increase in aggregate demand. On the one hand, interest rates have to rise so as to persuade agents to economize on money balances so as to finance the new, higher volume of transactions. On the other hand, the more interest rates rise, the lower the level of private sector expenditure in the new equilibrium. To the extent that the net outcome is a smaller increase in demand than would have taken place had interest rates remained unchanged, private sector expenditure is said to have been *crowded out* by the additional government spending.

Most important, consider the effect of an increase in SP^*, the home currency price of foreign products. At any given domestic price level, P, an increase in SP^* translates directly into a rise in $Q = SP^*/P$ – in other words, an improvement in UK competitiveness. This is the case whether the change is brought about by a rise in S (sterling devaluation) or in P^* (foreign inflation). Hence, a rise either in S or in P^* leads, via the increase in Q, to a rightward shift in the IS curve and consequent rise in aggregate demand, other things being equal. The aggregate demand schedule is therefore moved to the right by the change.

To summarize our conclusions about the aggregate demand curve:

- The aggregate demand curve shows how the level of income that clears the goods and money markets (that is, the equilibrium level of income in the IS–LM curve context) varies as the (domestic) price level changes. It is downward-sloping because, as the domestic price level rises, other things being equal, the real money stock is reduced in value and also the competitiveness of UK output falls, at given levels of S and P^*.
- The aggregate demand curve is drawn for given levels of government spending, G, the nominal money supply, M^s, the nominal exchange rate, S, and the price of foreign output, P^*. An increase in any of these variables will shift the schedule outwards, since they will induce greater aggregate demand at any domestic price level.

4.2 Aggregate supply

The exposition given here of the underpinnings of the aggregate supply curve is brief in the extreme. A fuller explanation taking account of the possibility of short-run output variations can be found in almost any undergraduate macroeconomics textbook.

As a starting point in the derivation of the aggregate supply curve, assume that the quantity of equipment, plant and machinery, and so on (the economy's capital stock) is given at any moment in time. The implication is that, with one factor of production fixed, output depends only on the manpower employed.[29]

Now consider the situation in the labour market.

4.2.1 Flexible prices

A profit-maximizing firm will demand labour up to the point at which the value of its marginal product is just equal to the money wage;[30] that is, up to the point where the last man hour adds just as much to revenue as it does to costs. If marginal product diminishes as labour utilization increases and if employers are competitive in the labour market, then, subject to the usual provisos with respect to aggregation conditions, the demand for labour across the economy will be a decreasing function of the nominal wage at any given price level.

For example, at the price level P_0, the demand for labour will be given by the downward-sloping line labelled $n^d(P_0)$ in Figure 4.5(a).

Clearly, if the price level were 20% higher, firms would seek to employ the same amount of labour at a money wage 20% greater than previously because, other things being equal, the marginal man hour would result in a proportionately greater addition to revenue than previously. Hence, the demand curve would shift vertically upwards so as to associate with any given employment level a nominal wage 1.2 times higher.

By way of contrast, households are presumed to be intent on choosing a utility-maximizing combination of work and leisure. If they are able to do so *in freely competitive labour markets*, they will offer their services up to the point at which the monetary benefit (in terms of consumption) they derive from the last man hour supplied, as measured by the wage rate, is equal to their subjective assessment of its leisure value.[31] The marginal utility of leisure can safely be assumed to increase as the number of hours worked rises and leisure time is reduced. It follows that, at any price level, the quantity of man hours supplied will be an increasing function of the nominal wage, as can be seen from the line marked $n^s(P_0)$ in Figure 4.5(a).

Again, at a higher price level $P_1 = 1.2P_0$, any given money wage buys 20% less consumption. The household will therefore be in equilibrium supplying any given number of man hours at a money wage 20% higher than before.

At the initial price level, the combination of a downward-sloping demand for labour with respect to the nominal wage and an upward-sloping supply of labour guarantees a unique equilibrium level of employment, C, and a corresponding real wage, $I_0 = S_0/P_0$. Following the accepted jargon, we shall call these market-clearing values the equilibrium or *natural* real wage and rate of employment.

At the higher price level $P_1 = 1.2P_0$, as can be seen at point B in Figure 4.4(a), the equilibrium money wage is 20% higher, but *both the real wage and the level of employment are unchanged.*

If the labour market clears with employment equal to C, what will be the level of output? Given the level of the total capital stock, the quantity of equipment per worker will be given, at any employment level. Assuming a given state of technology,

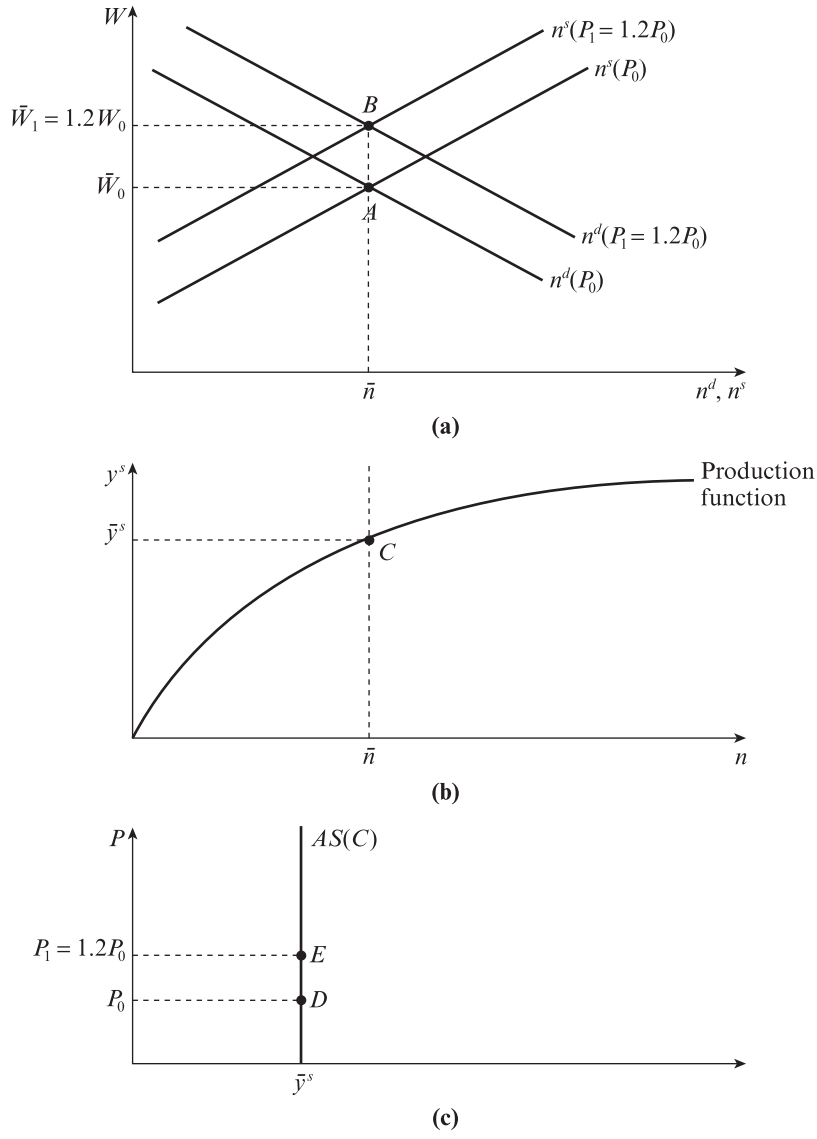

Figure 4.5 Derivation of the classic aggregate supply curve

it follows that the productive capacity of any quantity of labour will be determined. The functional relationship between the quantity of labour utilized and the output produced is normally called the short-run production function (Figure 4.4(b)). In particular, the employed labour force, C, will produce the output, J^s, which we can, in turn, regard as the equilibrium or natural level of real national product.

Now what does this analysis imply for the aggregate supply function? Recall that, symmetrically with the aggregate demand schedule, we are looking for the relationship between aggregate supply and the price level in the economy.

It is plain from what has already been said that the price index has no part to play in determining equilibrium output, since both the demand for labour and the supply

depend on the *real* wage. As long as the corporate sector's demand for labour and the household sector's supply are both unaffected by price level changes (as long as neither side suffers from money illusion), the aggregate supply curve will be a vertical line at the level J^s, so that the actual output produced in the economy will be independent of the price level (Figure 4.4(c)).

Notice the implication: if there is no connection between the price level and the aggregate supply, the latter cannot determine the former. It follows that the price level must depend on aggregate demand alone. In other words, any of the factors that have been shown to shift the aggregate demand curve upwards cause the price level to increase, and vice versa for a price level decrease.

One point that should be absolutely clear from this brief analysis is that the role of price and wage flexibility is absolutely critical. If money wages rose by more (or less) than 20% as the price level rose from P_0 to P_1, the outcome would be an increase (or decrease) in the real wage and consequent fall (rise) in employment and output. Everything hinges, therefore, on the smooth functioning of the labour market.

Now since a faith in the ability of markets to adjust rapidly to shocks is the hallmark of classical economics,[32] we conclude:

> The classical or flexible-price aggregate supply curve is vertical at the long-run capacity output level of the economy. This is because, as the price level fluctuates up or down, the money wage adjusts to keep the real wage constant. Hence employment and output never vary and the price level itself is determined by aggregate demand.

4.2.2 Fixed prices

Now consider what would happen if money wages were fixed.

This situation might arise in a number of ways. It could be regarded as the very short-run reaction to a sudden change in the price level – after all, money wages can hardly respond overnight. In fact, wages may be fixed over a longer horizon by the existence of employment contracts, explicit (in other words, written) or implicit (unwritten, but none the less binding). In many cases, the contracts might be between employers and employees represented by trade unions, in which case the mechanism for adjusting money wages is likely to be even more cumbersome.[33]

In any case, the implication is that the *effective* labour supply curve is horizontal at the current fixed money wage, $\bar{W}_0$. Assuming firms continue to react to price changes in the way described in Section 4.2.1, it follows that employment and output are no longer constant (Figure 4.6). A rise in the price level is associated with a fall in the real wage and a consequent increase in employment from $\bar{n}_0$ to $\bar{n}_1$ and output from $\bar{y}_0^s$ to $\bar{y}_1^s$. The result is the upward-sloping aggregate supply curve labelled AS(K) in Figure 4.6(c).

In the limit, if wages form the bulk of the corporate sector's costs or if non-labour costs are constant, competitive producers will never raise prices. Firms will react to higher demand for their output by simply taking on more labour at the fixed money

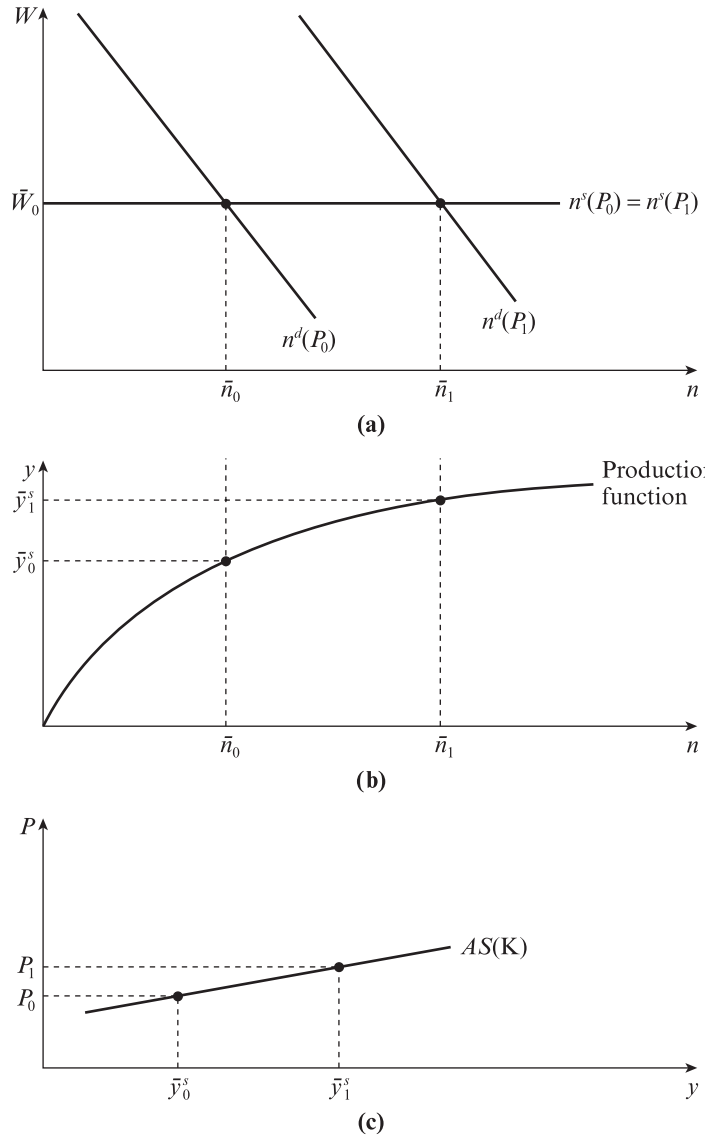

Figure 4.6 Derivation of the Keynesian aggregate supply curve

wage and scaling up production at constant unit costs. The outcome will be a completely flat aggregate supply curve.

Now nobody believes this scenario to be the complete truth about aggregate supply in the long run as; after all, it would imply that a policy of demand expansion (for example, by printing money) could lead to a permanent increase in employment and output. None the less, the account given here will serve as a conveniently simplified description of a commonly held view of aggregate supply in the short run, one we shall henceforth refer to as Keynesian or fixed price.

We conclude:

> The Keynesian or fixed-price aggregate supply curve is flat at the currently given price level, reflecting the fact that demand fluctuations are associated with equal and opposite variation in the real wage. Hence, with fixed prices, employment and output are demand-determined.

With the price level fixed in this scenario, we need concern ourselves only with aggregate demand and the position of the IS and LM curves – which, as we shall see in Chapter 6, is exactly the focus of the Mundell–Fleming model.

4.2.3 A compromise: sticky prices

If the flexible wage model seems an implausible description of the very short-run reaction of labour markets, while the fixed wage alternative is equally unrealistic as a description of long-run behaviour, consider a compromise between the two extremes.

Suppose the assumptions underlying the Keynesian model apply to the short run, while wages and prices are completely flexible in the long run, as in the classic model. In other words, suppose the aggregate supply curve is flat in the immediate impact phase, but gets steeper as time elapses, ultimately becoming vertical in the steady-state equilibrium.

According to this scenario, the economy passes through three phases in the aftermath of a disturbance. Take, for example, the case of an increase in demand from AD_0 to AD_1 (Figure 4.7).

The impact effect is felt purely on output and employment. Workers willingly supply more labour at the going wage, whether because they are unsure how long the increased demand will persist, or because they are bound by fixed wage contracts, or simply as a result of very short-run inertia. The economy therefore moves more or less immediately from point E to point A, with an increase in output from $\bar{y}_0^s$ to $\bar{y}_1^s$.

As time passes, however, contracts (implicit or explicit) are renegotiated and money wages are progressively bid up as employers seek to increase output. In turn, the corporate sector recoups the rise in labour costs by raising output prices, thereby giving added impetus to the wage inflation, as households seek to increase their pay to take account of the higher cost of living. The adjustment phase is therefore characterized by rising prices and falling output, as the aggregate supply curve tilts towards the vertical. At each stage of the adjustment process, we can think of the economy being in a temporary equilibrium, like B in the diagram, with output at y_2^s still above its initial long-run level and the price index somewhat higher, at P_2.

Eventually, the economy reaches its long-run steady state at C, where all real quantities are back where they started. The price level is at its new long-run level of $\bar{P}_s$, but since the money wage has increased pro rata, the real wage and hence employment are unchanged. Output is back at its natural level.

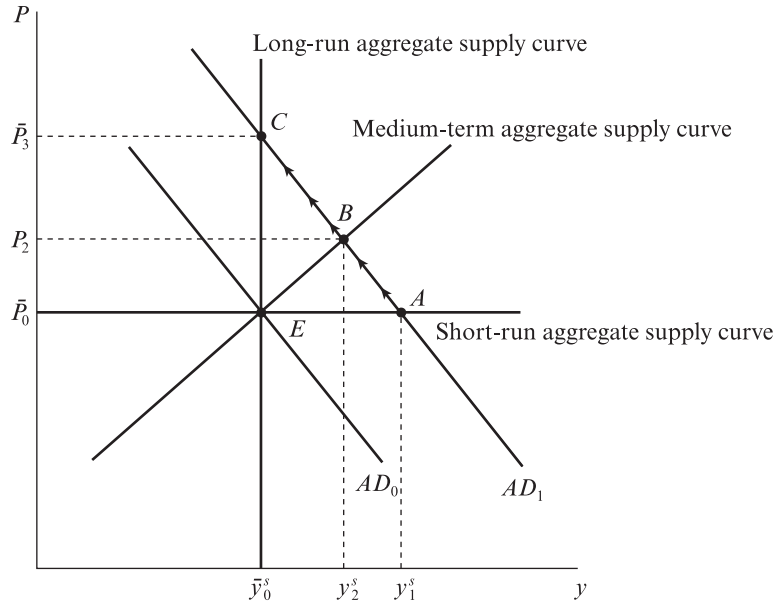

Figure 4.7 The sticky price aggregate supply curve

We can summarize this model as follows:

> The sticky price model of the supply side of the economy is one where the aggregate supply curve is instantaneously flat, becoming steeper as time passes and eventually vertical. It follows that the reaction to an increase in aggregate demand involves initially an expansion in output with no inflation. As time passes, the price level rises and output falls back towards its long-run level.

This compromise provides the background to the Dornbusch model, which will be the subject of Chapter 7.

4.3 Conclusions

For those readers who are familiar with the basic textbook model of the closed economy, this chapter will have contained only two new features. First, the introduction of the real exchange rate into the IS curve. The higher the real exchange rate, the more competitive domestic output on world markets and hence the greater the demand in the economy, other things being equal.

The second innovation is in the treatment of the determination of the money stock. Under a fixed exchange rate regime, the money supply is no longer a policy variable, because, as we saw, it contains a component that is the direct counterpart of the country's foreign currency reserves. Since the latter are clearly not within the direct control of the *domestic* monetary authority, it follows that the money stock can no

longer be regarded as simply a policy lever that can be operated at will by the governor of the central bank or the minister of finance.

With the end of this chapter, we can now proceed to the analysis of models of exchange rate determination, the subject matter of Part II of the book.

Summary

- The open economy IS curve links combinations of the interest rate and income consistent with equilibrium in the market for flows of goods and services. It is downward-sloping. Increases in the real exchange rate (real devaluation) or in government net expenditure cause it to shift to the right.
- *Money* means the asset or assets that are usually used as means of payment. For most purposes, it consists of currency in circulation *plus* demand deposits (cheque accounts).
- The *supply of money* is equal to the sum total of the domestic credit created by the central bank in its lending to the government and the commercial banks by their advances to the public *plus* the country's reserves of gold and foreign exchange, if any.
- Increases in the supply of money can arise from policy decisions to permit increased domestic credit or from growth in the reserves. The latter are the result of maintaining a fixed exchange rate under conditions of balance of payments surplus. It follows that, under a fixed exchange rate, changes in the money stock are not necessarily the result of policy decisions. Only under a freely floating rate is the money stock an exogenous variable, since, in that case, the whole of the money stock (or, at least, the whole of any change in it) consists of domestic credit.
- The *demand for money* is to be construed as the demand to hold a stock of purchasing power in the form of monetary assets. It is, therefore, a demand for real balances. As such, it will be greater, other things being equal, when national income is higher, since then the volume of transactions to be financed is larger and it will be smaller the higher are interest rates in the economy, because higher rates of return on non-money assets mean a higher *opportunity cost of holding money*.
- The LM curve links combinations of income and the interest rate consistent with equilibrium in the money market. It is upward-sloping. An increase in the real money stock, whether originating in one of the components of the nominal supply or in a fall in the price level, causes the LM curve to shift to the right.
- The aggregate demand curve shows how the IS–LM equilibrium value of national income ('aggregate demand') varies as the price level changes, holding everything else constant. Since, at higher price levels, a given stock of nominal balances represents a lower real money supply, the aggregate demand curve slopes downward.
- On the supply side of the economy, the demand for labour is a downward-sloping function of the money wage, at any given price level.
- If households are concerned only with the real value of any wage offer, then their labour supply will be homogeneous in prices and wages. If all prices adjust smoothly, it will follow that the level of employment and output are constant,

other things being equal, and in particular independent of the price index. Hence the (classic) aggregate supply curve is vertical.

- If households behave as though subject to money illusion – whether as a result of long-term contracts, inertia, misinformation or simply delays in reacting – real wages will not be invariant and hence both employment and output will fluctuate as the price level changes. Hence the (Keynesian) aggregate supply curve will be upward-sloping. In the limit, it could well be horizontal.
- It could be argued that the classic aggregate supply curve is a plausible picture of the long-term behaviour of the economy, while the Keynesian version relates to the short term. A compromise would therefore be to take the aggregate supply curve as flat in the immediate term, pivoting upwards as time passes in the aftermath of a demand disturbance, until ultimately the economy returns to its steady state along the vertical long-run supply curve.

Reading guide

The closed economy aggregate demand–supply framework is covered in practically all modern textbooks, though the extensions given here are harder to locate. Among the most widely used macroeconomics texts are Mankiw (2007) and Blanchard (2005), both of which cover these and many other topics.

The approach taken to the demand for money in this section is by no means the only one possible, nor even the most plausible in some respects. Interested readers, and for present purposes that includes all those with ambitions to progress beyond this book in the study of exchange rates, would be well advised to consult Laidler (1985) on the demand for money, and the more important references given therein.

There are a number of channels through which the external sector might impinge on the supply side of the economy – for example, via a cost of living index with a significant import component. However, the result of incorporating these linkages is usually to complicate matters very substantially. The adventurous reader with sufficient mathematical background might like to look at the different approaches taken by Daniel (1982), Bruno and Sachs (1982), Marston (1982) and Copeland (1983).

Web page: **www.pearsoned.co.uk/copeland**.

Notes

1 See the reading guide for references to research on the implications of openness for aggregate supply.
2 In reality, the corporate sector saves by not distributing profits. We shall ignore this fact, without running any risk of its affecting our conclusions.
3 Notice that it is not being claimed that, as a household's income rises, it will actually save a higher proportion of its income, although that may also be a reasonable conjecture. All that is being asserted here is that a household's savings will be greater in absolute terms, the higher its income.
4 For simplicity, we assume investment is an activity exclusive to firms, which means ignoring investment by other sectors of the economy.
5 It should be noted that, even in industrialized countries, there can arise situations where the normal means of payment are no longer universally accepted, either as a result of bank failures or, more frequently, as a result of a lack of confidence in the internal or external purchasing

power of the currency. In such circumstances, the spontaneous response may be to return to a commodity money (cigarettes, gold, silver, and so on), as in the German hyperinflation of 1923, or, in some cases, to adopt a foreign currency as an unofficial means of payment (for example, Israel in the early 1980s). The latter phenomenon is known in the literature as 'currency substitution', and its implications will be examined in Chapter 9.

6 For the sake of simplicity, the definitions given here ignore a number of the less important financial instruments. For a detailed description of how these aggregates are computed in the UK, where *M*4 and *M*5 money supply definitions are also given, see the *Bank of England Quarterly Bulletin*, May 1987.

7 Until, say, the early 1970s, it was almost universally true that demand deposits earned no interest. In fact, there was some debate in the profession as to whether one could regard this as an intrinsic feature of money or whether it was simply a result of the way the banking system had developed, the monopoly power of banks, low inflation rates, and so on. During the 1970s and 1980s, the increased competition in banking, combined with higher inflation rates, led to a situation where explicit interest payments were frequently made to holders of demand deposits. What is even more common, however, is implicit interest on cheque accounts. This involves paying interest by the back door: offering depositors 'free' banking by refunding any charges they would otherwise incur for normal banking operations, such as cheque clearance, use of automated teller machines, and so on. This amounts to paying interest in kind – the bank pays its customers in units of its own output – banking services. Economists refer to this type of arrangement as the payment of an 'own rate' of interest.

8 In general, opportunity cost is the term used by economists to refer to the cost of a resource, viewed as the loss incurred by not employing it in its best alternative use.

9 To see why, compare two employees, one paid monthly and the other paid the same annual salary once a year. Suppose they both spend their income at a smooth rate throughout the year. The monthly paid employee would have an average balance over the year equal to one-half the monthly wage, while the second employee's average money balance would be one-half the annual wage or six times the monthly wage.

10 The distinction between nominal and real values is very important. A real variable is one that measures a physical quantity or volume of goods or services. A nominal variable is one that is valued in money terms. So a nominal value rises when prices rise and falls when they fall, while a real variable changes only when the quantities it measures change. In practice, real variables are computed by taking the values we actually observe – which are always nominal magnitudes – and deflating them by an appropriate price index. For example, if Y is nominal national income and P is the relevant price index (called, in this case, the 'GNP deflator'), then real income, also known as income in volume terms, is given by $y = Y/P$.

11 Strictly speaking, k is the product of two factors of proportionality: one relating national income to the volume of transactions, and the other relating the demand for money to transactions. In fact, k is the reciprocal of what is known as the (income) velocity of circulation. The demand for money is presented here in linear form, for simplicity. In fact, researchers nowadays tend to favour a demand for money equation that is linear in logs, as we shall see in Chapter 7.

12 Broadly speaking, Keynesian monetary theory tends to emphasize substitution between money and other financial assets (typically bonds) while monetarists are more inclined to concentrate, at least in theoretical argument, on the margin between money and real assets, such as physical capital and consumer goods. In practical applications (that is, empirical work), these issues have proved hardly worth the controversy (see Laidler (1985)). The only aspect of this question relevant to exchange rates is, as has already been mentioned, the possibility that the appropriate opportunity cost involves foreign assets (see Chapters 8 and 9).

13 Although governments often have enormous physical assets, they are rarely willing to use them to finance budget deficits, except as part of privatization programmes undertaken, ostensibly at least, for different reasons altogether. Incidentally, readers in the UK will note that, for once, British jargon is particularly appropriate in referring to the budget deficit as the public sector borrowing requirement.

14 The power to print money is referred to as seigniorage.

15 From now on, we shall use the term 'bonds' as a shorthand to refer to all non-monetary financial instruments issued by the authorities.

16 What's in a name? Not a lot, in this instance – not the Bank of Scotland, the Royal Bank of Scotland or even the Bank of America is a central bank, whatever their names may suggest.

17 But not one that is always the exclusive preserve of the central bank. The Scottish commercial banks, for example, are permitted to issue their own banknotes. They do so, however, under licence to the Bank of England. For our purposes here, we shall ignore such complications and assume that the central bank is the sole issuer of currency.

18 A point to note with regard to jargon: in the context of the open economy, the reserves can be taken to refer to the reserves of gold and foreign currency. However, as we shall see shortly, in the context of domestic banking, we often need to refer to a very different type of reserve, the precautionary cash (that is, domestic currency) reserves kept by commercial banks in their vaults. Since it is only in the present section that any possibility for confusion could arise, the convention will be followed in this chapter and throughout the rest of the book that 'the reserves' when unspecified always means the stock of gold and foreign currency held by the central bank.

19 The word 'forced' probably overstates the case somewhat, at least for some countries. Suffice to say that the degree of independence of the central bank depends entirely on constitutional considerations and/or custom and practice in the country in question. The whole issue of the feasibility and desirability of central bank independence is an extremely vexed one and is related to a number of questions regarding monetary and fiscal policy that will crop up at various points in what follows.

For present purposes, we shall treat the central bank as simply an arm of government, so that transactions between it and the rest of the government are a purely formal, bookkeeping exercise.

20 Under the old Gold Standard, banknotes could in principle be exchanged for gold on demand, although by the time the system collapsed (in August 1971, in the case of the USA), the right to do so was restricted almost exclusively to foreign central banks. British banknotes still carry the meaningless legend: 'I Promise to pay the Bearer the Sum of . . .', signed, it should be noted, not by a minister or government official but by the cashier of the Bank of England.

21 A commercial bank is best thought of as the kind of financial institution with which we are all familiar: the retail or high-street bank, which takes deposits from the public at large and lends the proceeds to local businesses. For present purposes, we can safely ignore the existence of the wholesale banks, discount houses and other quasi-banks as unnecessarily complicating factors.

22 Only in the case of demand deposits (that is, current or cheque accounts). For time deposits ('deposit accounts') and the like, the balance may not have to be repaid on demand. However, in order to keep matters simple we shall restrict our attention to demand deposits, as is consistent with the definition of the money stock we chose at the beginning of Section 4.1.2.

23 The costs of running a bank are covered in two ways: through explicit charges and through the gap between the interest rate that the bank earns on its assets and the rate it must pay on its deposits (which is often zero in the case of cheque accounts).

24 Not only prudence. In some countries, and at times in the UK, the authorities have fixed the proportion of cash the banks are required to hold, not because they mistrusted the banks' judgement but as an alternative means of controlling the volume of bank lending. In other words, instead of controlling the quantity of cash in existence, they control the ratio of cash to loans.

Another complication we shall ignore is that banks are often allowed to include some types of highly liquid earning assets (for example, treasury bills) in their precautionary reserves. Another fact of banking life over which we draw a veil is that, in addition to making provision for the repayment of deposits to normal, solvent customers, banks also have to set aside a reserve in order to cover bad debts. Finally, we also choose to ignore the possibility that banks can (and do) issue shares and use the proceeds to expand their balance sheets.

25 In fact, in a closed economy context, control of domestic credit is all there is to money supply control. Nevertheless, there have been long, acrimonious debates about how far it is possible to control the money stock, even in a closed economy. The issues, which revolve around the minutiae of the institutional arrangements for control of the banking system, need not concern us here. It should suffice to say that it is certainly possible to imagine a banking system such that control of domestic credit would be possible to 100% accuracy. Whether it is desirable to have

such a system is another matter, as indeed is the question of whether domestic credit can be controlled under the regimes actually in force in, say, the UK or the USA today.

26 So called because it joins points at which the money stock is equal to the demand for money, originally denoted by the letter L, in deference to Keynes' name for it: liquidity preference.

27 Not, of course, the same y_0 as in Figure 4.1.

28 And, almost certainly, a higher interest rate, since the effect of the upward shift in the LM curve is unlikely to be offset by the downward shift in the IS curve.

29 Ignoring other textbook factors of production such as land, managerial expertise, and so on. Notice also that employment ought, in principle, to be measured in man hours of homogeneous labour.

30 Strictly speaking, the product wage. We assume that the distinction can be ignored, at least at the aggregate level. We also make all the usual assumptions necessary to guarantee a downward-sloping marginal physical product curve.

31 Viewed either as the marginal utility of leisure or (the negative of) the marginal disutility of work. We assume leisure is, in the jargon of microeconomics, a normal good.

32 Known nowadays as new classic economics.

33 It is not our concern here as to why or how money wages might be fixed. The original Keynesian view was that money wages were not so much constant as 'downward invariant' – that is, incapable of being pushed down, even in the aftermath of price deflation, but quite easily pushed upward. This asymmetrical money illusion was held to be, at least in part, a cause of the mass unemployment of the 1930s.