14 The Trade Cycle: Keynesian and Monetarist Interpretations

The long-run expansion of industrialised market economies has been accompanied by cyclical fluctuations in economic activity. This type of fluctuation is known as the business or trade cycle. The general feature of the cycle is that an expansion of economic activity is followed by a contraction, which is in turn succeeded by a further expansion. Explaining the occurrence of trade cycles has been a major preoccupation of macroeconomics for a long time.

Measuring cyclical fluctuations in economic activity is not particularly easy because a large number of variables, such as industrial production, total employment, fixed private investment, the price level and interest rates, are involved. Each one has its own specific cycle and does not move exactly in phase with the others. An aggregate measure of the trade cycle is obtained by recording a large number of specific cycles, comparing their respective phases and dating their turning-points. Those variables which then appear to cycle in a consistent way in relation to the others are then used to date the turning-points and phases of the general cycle in economic activity. This is termed the *reference cycle*. Figure 14.1 shows the UK reference cycle as dated by the Central Statistical Office. The top line gives the dates of troughs (T) and peaks (P) for the reference cycle.

In the nineteenth century and up to the Second World War trade cycles in the United Kingdom involved absolute increases and decreases in real GDP. The post-war cycles up to the mid-1970s involved no decline in absolute GDP during recessions, only a retardation in its rate of growth. This was in complete contrast to the inter-war cycles, which had been more severe than the nineteenth-century business cycles. In the post-war period cyclical fluctuations have been milder than the nineteenth-century ones. However, the two recessions since 1973 have been more severe and have involved absolute falls in GDP from one year to the next.

Another difference between pre-war and post-war trade cycles is that the price level fell during the pre-war recessions. In fact there was a downward trend in UK prices throughout the 1921–38 period. Since the Second World

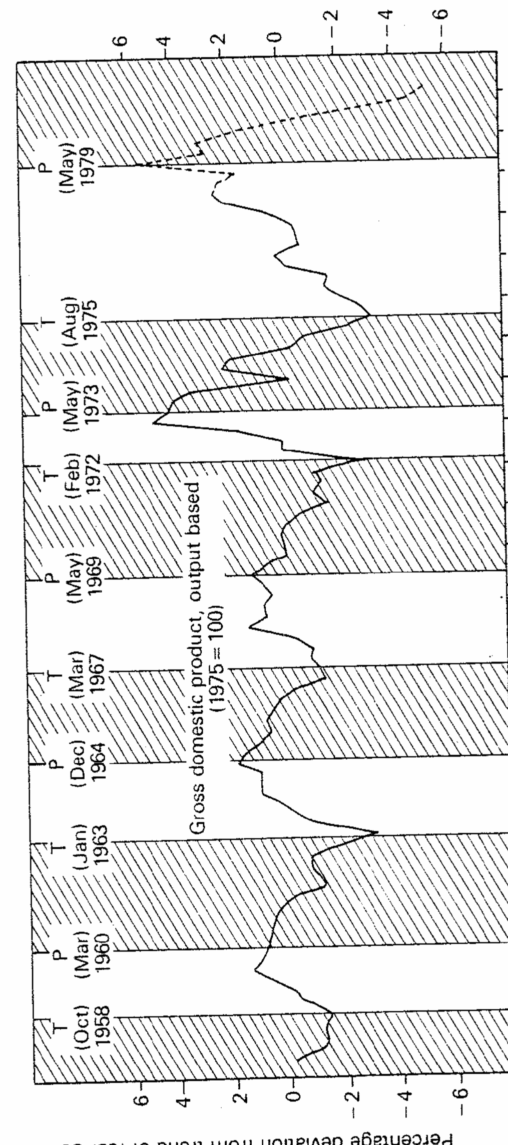

FIGURE 14.1 *The reference cycle and the cycle of GDP for the United Kingdom, 1957–81*

Sources: various issues of *Economic Trends*.

War prices have risen at all stages of the cycle, though there has been some tendency for inflation to slow down in the latter part of the recession phase (see Shapiro [1]).

A good deal of macroeconomic theory relates in one way or another to explaining trade-cycle behaviour. Theories of the determination of output, unemployment, investment, the price level, interest-rate levels and other relevant variables all play their part in explaining fluctuations in the general level of economic activity. In this chapter we narrow the focus in order to examine some theoretical approaches which are explicitly directed at explaining the general configuration of the cycle. The study of particular cycles would require more detailed empirical analysis.

Trade-cycle theorists have sought to explain the central feature of the cyclical behaviour of industrialised, market economies. This is the regular recurrence of expansion and contraction in the process of long-run economic growth. Thus it appears that expansion generates factors which bring about its own end and induce a period of contraction. Similarly, the process of contraction generates the conditions for recovery. These aspects of the cycle have encouraged the development of theories concentrating on factors endogenous to the cycle that will explain the cycle in terms of its internal dynamics.

There is another approach to trade-cycle theory which does not rely so strongly on internal factors. It analyses cyclical adjustment paths that are generated by the impact on the economic system of exogenous factors, such as population changes, the accumulation of new inventions (Schumpeter [2]), the opening up of new territories or changing patterns of international trade. These two approaches are not dissimilar and a very clear classification of endogenous and exogenous factors cannot be made, as this distinction depends on what behavioural relationships are included in the model of the economy with which one is working.

For Keynesian economists the existence of trade cycles is *prima facie* evidence of the failure of market co-ordination and so provides a rationale for active government intervention intended to stabilise the economy. Keynesian explanations of the trade cycle emphasise the part played by disturbances in the real variables, particularly in private-sector investment. In contrast, monetarists see changes in the supply of money which originate from actions by the monetary authorities as a primary causal factor.

14.1 Keynesian trade-cycle theory

Keynesian trade-cycle theory grew out of the 'General Theory', starting in the late 1930s. Samuelson [3] first used the multiplier relationship together with the key role played by unstable investment, which he expressed in terms of the accelerator theory of investment, to construct cumulative upwards and downwards movements in real output. This process works as follows. An expected increase in output which generates a demand for additional capital stock leads to an increase in investment. The increase in investment causes output to rise by an amount equal to the increase in investment *times* the income multiplier. The increase in income causes investment to rise further,

and so the multiplier accelerator process continues. The money supply is implicitly assumed to adjust to the quantity of output. There is also no mention of price changes, as in the Keynesian tradition these are assumed fixed. Prices do not adjust: all adjustment is by quantities and hence gives rise to sizeable fluctuations in real output and investment. The supply of labour is assumed, in the typical Keynesian manner, to be perfectly elastic. As much labour as firms want to employ at existing prices is always forthcoming. The use of the multiplier–accelerator relationship to derive cyclical fluctuations is now examined in more detail.

The first-order multiplier–accelerator interaction

The simplest specification of the accelerator function for net investment, I_t^n, is the following first-order difference equation (see Chapter 13 for its derivation):

$$I_t^n = v(y_t - y_{t-1}) \tag{14.1}$$

In this version the capital stock is fully adjusted to its desired level at the end of each period. When firms enter the current period their capital stock is not optimally related to current output, as it was adjusted to last period's output. Investment (or disinvestment) therefore takes place but not until the end of the period. Consumption is assumed to depend proportionately on the current level of income:

$$C_t = (1 - s)y_t \tag{14.2}$$

where s is the marginal propensity to save. The model is solved by assuming short-run equilibrium is achieved in each period whereby aggregate demand equals national output:

$$y_t = C_t + I_t + AE_t \tag{14.3}$$

where AE is autonomous expenditure. Therefore, for short-run equilibrium we must have

$$y_t = (1 - s)y_t + v(y_t - y_{t-1}) + AE_t. \tag{14.4}$$

Output and demand vary from one period to another because aggregate demand depends on last period's income as well as on current income. When last period's income differs from this period's income aggregate demand changes from period to period. Solving equation 14.4 for y we obtain

$$y_t = \left(1 + \frac{s}{v - s}\right)y_{t-1} - \frac{AE_t}{v - s} \tag{14.5}$$

Equation 14.5 is an example of a first-order linear difference equation as it is lagged just one period. Using equation 14.5 we can derive the time path of income. We start from a static equilibrium level of income which equals AE/s. (This is derived from equation 14.4 by setting $y_t = y_{t-1}$ and solving for the static equilibrium level of income.) Income then diverges from its static equilibrium level. In the initial period 0 income is y_0: thus the initial divergence is $y_0 - AE/s$. From equation 14.5 we therefore obtain that income in period 1 is

$$y_1 = \left(1 + \frac{s}{v - s}\right) y_0 - \frac{AE}{v - s}$$

$$y_2 = \left(1 + \frac{s}{v - s}\right) y_1 - \frac{AE}{v - s}$$

$$= \left(1 + \frac{s}{v - s}\right)\left[\left(1 + \frac{s}{v - s}\right) y_0 - \frac{AE}{v - s}\right] - \frac{AE}{v - s}$$

$$= \left(1 + \frac{s}{v - s}\right)^2 y_0 - \frac{AE}{v - s}\left[1 + \left(1 + \frac{s}{v - s}\right)\right] \tag{14.6}$$

Therefore, continuing on for y_3, y_4, etc., by substituting into equation 14.5 we obtain

$$y_t = \left(1 + \frac{s}{v - s}\right)^t y_0 - \frac{AE}{v - s}\left[1 + \left(1 + \frac{s}{v - s}\right) + \left(1 + \frac{s}{v - s}\right)^2 + \dots + \left(1 + \frac{s}{v - s}\right)^{t-1}\right]$$

$$= \left(1 + \frac{s}{v - s}\right)^t y_0 - \frac{AE}{s}\left[1 - \left(1 + \frac{s}{v - s}\right)^t\right]$$

$$= \left(1 + \frac{s}{v - s}\right)^t \left(y_0 - \frac{AE}{s}\right) + \frac{AE}{s} \tag{14.7}$$

$$= (1 + g)^t \left(y_0 - \frac{AE}{s}\right) + \frac{AE}{s} \tag{14.7a}$$

where $g = s/(v - s)$.

Now g will be positive provided that v is greater than s. This is what one would expect given that s is less than 1.0, and v, the capital–output ratio is normally greater than 1.0. Since s is a fraction, $g = s/(v - s)$ will also be less than 1.0. We can interpret g as the growth of national output. Any initial divergence of output from its original level, y_0 is continually magnified at a rate of $1 + g$ each period.

In general there are four basic types of adjustment path that a variable can follow. The adjustment path can converge towards equilibrium, in which case the model is stable, or diverge from equilibrium, in which event the model is

unstable or explosive. The adjustment path is further characterised as oscillating or monotonic. The latter path always moves in the same direction. The four kinds of adjustment path are shown in Figure 14.2 and illustrated using the general form of equation 14.7a in which $\alpha = 1 + g$:

$$y_t = \alpha^t \left(y_0 - \frac{AE}{s}\right) + \frac{AE}{s} \tag{14.7b}$$

The four types of adjustment path are:

1. *Stable and monotonic* when α lies between 0 and +1.0. α^t gets steadily smaller as t increases so that the divergence of income from equilibrium eventually becomes negligibly small.
2. *Stable and oscillating* when α lies between −1.0 and 0. As t tends to infinity α^t tends to zero but in doing so alternates between being positive when t is an even number and being negative when t is an odd number. The adjustment path oscillates around the equilibrium value of income, which is therefore overshot a number of times. As the oscillations gradually die away, the cyclical path is said to be damped.
3. *Explosive and monotonic* if α is greater than 1.0. As time progresses α^t gets larger. In the model discussed above the adjustment path is explosive because g lies between 0 and +1.0 so that $\alpha = 1 + g$ is greater than 1.0. Whether income is continually rising and falling is determined by whether the initiating movement was an increase or a decrease in output respectively.
4. *Explosive and oscillating.* If α is less than −1.0, α^t again tends to infinity as time passes. Since α is a negative number, α^t is alternatively positive and negative depending on whether t is an even or an odd number. Because α^t gets absolutely larger as time passes, equilibrium is overshot by ever-increasing amounts. Cycles of ever-increasing amplitude are said to be anti-damped.

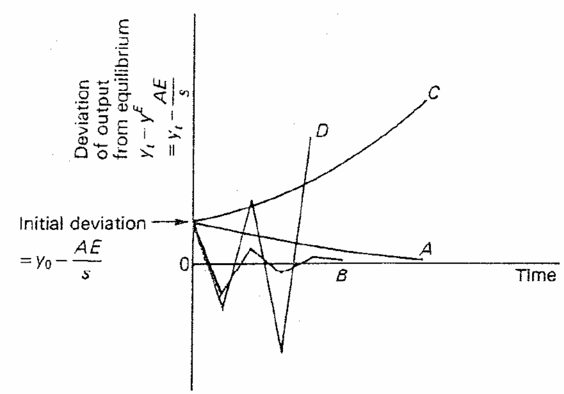

FIGURE 14.2 *The four types of adjustment path*

Since economically plausible values for s and v give a value of α in equation 14.7b which exceeds 1.0, a first-order multiplier–accelerator interaction cannot by itself generate cycles. It only produces continuous (that is, monotonic) upward or downward movements.

Second-order multiplier–accelerator interaction

Given the values of s and v that occur, we require income to be lagged by more than one time period if the multiplier–accelerator relationship is to generate cycles. A second-order multiplier–accelerator interaction contains both current income and income lagged by one and two periods. A number of models of this type, based on varying assumptions, can be set up. In the model outlined below consumption is assumed to be a function of income lagged one period. The accelerator equation assumes that, because of lags, net investment makes good last period's discrepancy between the desired and actual capital stock. This means that current investment is related to last period's change in output and not this period's. It is this factor that causes cyclical fluctuations:

$$C_t = (1 - s)y_{t-1} \qquad (14.2a)$$

$$I_t^n = v(y_{t-1} - y_{t-2}) \qquad (14.1a)$$

Substituting equations 14.2a and 14.1a into 14.3 we obtain

$$y_t = (1 - s + v)y_{t-1} - vy_{t-2} + AE \qquad (14.8)$$

This of course has the same static-equilibrium solution as equation 14.7, namely that income equals AE/s. Equation 14.8 is a second-order difference equation because it has two lags.

The type of adjustment path followed by output in this model depends on the numerical values of v and s. Figure 14.3 shows the various combinations of values for v and s which will give the four types of adjustment path. Cycles will occur if v and s lie within the areas B or D. For example, if $s = 0.3$, v must lie between 0.205 and 2.395 for the adjustment path to exhibit oscillations. For this particular model (though not for some others; see Samuelson's model [3] for a

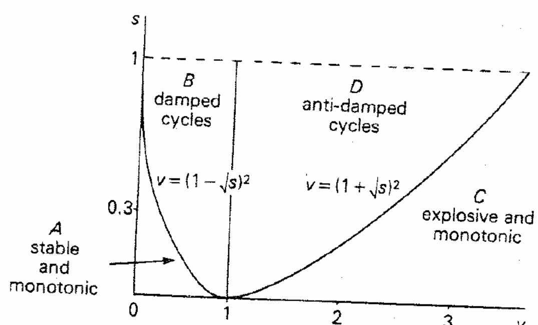

FIGURE 14.3 *The combinations of v and s which give the four types of adjustment path*

TABLE 14.1 *Adjustment path of national income following a one-period change in autonomous expenditure of 30*

Period	A: $s = 0.3, v = 0.2$ y_t	$y_{t-1} - y_{t-2}$	I_t^n	B: $s = 0.3, v = 0.8$ y_t	$y_{t-1} - y_{t-2}$	I_t^n	C: $s = 0.3, v = 3$ y_t	$y_{t-1} - y_{t-2}$	I_t^n	D: $s = 0.3, v = 1.5$ y_t	$y_{t-1} - y_{t-2}$	I_t^n
1	1,030	0	0	1,030	0	0	1,030	0	0	1,030	0	0
2	1,027	30	6	1,045	30	24.0	1,111	30	90	1,066	30	45
3	1,018.3	−3	−0.6	1,043.5	15	12.0	1,321	81	243	1,100.2	36	54
4	1,011.1	−8.7	−1.74	1,029.3	−1.5	−1.2	1,854	210	630	1,121.4	34.2	51.3
5	1,006.7	−6.8	−1.4	1,009.1	−14.3	−11.4	3,196	533	1,599	1,116.9	21.2	31
6	1,003.7	−4.4	−0.9	990.2	−20.2	−16.2	6,563	1,342	4,026	1,074.9	−4.6	−6.9
7	1,002	−2.9	−0.6	978.0	−18.9	−15.1				989.6	−41.9	−62.9
8	1,001.1	−0.9	−0.18	974.9	−12.5	−10.0				864.8	−85.4	−128.1
9	1,000.6	−0.5	−0.11	979.9	−3.1	−2.5				718.2	−124.8	−187.2
10	1,000.4	−0.2	−0.03	999.0	5.0	4.0				582.8	−146.6	−219.9
11				1,001.1	10.1	8.0				504.0	−135.4	−203.1
12				1,010.3	11.5	9.2				534.6	−78.8	−118.2
13				1,014.5	9.2	7.4				720.1	30.6	45.9
14				1,013.6	4.3	3.4				1082.3	185.5	278.3

different result) a capital–output ratio of less than 1.0 produces convergence, while a v greater than 1.0 causes divergence, irrespective of the size of the marginal propensity to save. In general a small capital–output ratio is required to produce anti-damped cycles.

Table 14.1 presents numerical examples which illustrate the four types of adjustment path by taking pairs of values for s and v from each of the areas A, B, C and D of Figure 14.3. The marginal propensity to save is taken to be 0.3 in all four cases and v is varied to give different adjustment paths. Autonomous expenditure is assumed to be equal to 300, which gives an equilibrium level of output of 1,000. The difference between 1,000 and income in period t gives the divergence from equilibrium.

The initial deviation from equilibrium is caused by an increase in period 1 of 30 in autonomous expenditure which returns to 300 in all subsequent periods. In the divergent monotonic case income moves steadily away from equilibrium. Cycles occur in cases B and D. After the initial shock, equilibrium is approached but is overshot. Income falls below equilibrium and returns, again overshooting. The cycles in case B diminish with time and there is convergence towards equilibrium, unlike case D.

The lagged capital-stock adjustment mechanism, linked with particular values of v and s is responsible for this cyclical behaviour. When income rises so does the desired capital stock, but the actual capital stock is not increased until the next period. Take case B. Income rises by 30 in period 1 but net investment does not rise (by 24) until the next period. This increase in investment demand stimulates an increase in income but, because autonomous expenditure falls back to 300, income rises less in period 2 than in period 1. Consequently net investment is lower in period 3 than in period 2. The fall in net investment causes income to descend from its peak level of 1,045. Since income has fallen, firms wish to hold less capital stock than actually exists. This is indicated by a desired disinvestment of −1.2. The fall in investment promotes a further fall in output which is carried below its equilibrium level. As income falls its rate of decline diminishes and disinvestment eventually reduces the actual capital stock below its desired level. Net investment becomes positive in period 10 and as a consequence output rises.

Modifications to the multiplier–accelerator interaction

There are several important features of actual trade-cycle experience which the multiplier–accelerator mechanism considered on its own fails to explain:

(1) While the accelerator theory could explain net fixed investment when it is positive it cannot explain *dis*investment. When the actual capital stock exceeds the desired stock, the rate of disinvestment is determined not just by the decline in output but also by the rate at which the capital stock depreciates. It is for this reason that the accelerator principle works better as an explanation of inventory investment than of fixed investment since the rate of both increases and decreases in stocks is directly linked to the behaviour of actual and anticipated sales. The accelerator principle therefore can quite plausibly be applied to the analysis of inventory cycles (as by Metzler [4]).

(2) Positive net investment, let alone gross investment, has not ceased entirely during depressions, as occurs in the multiplier–accelerator model outlined above. Even in the very severe depression of 1929–33, when US GNP fell by 30 per cent, real gross private investment fell by 90 and not by 100 per cent.

(3) The cycles generated by the multiplier–accelerator interaction are either damped or explosive. Only if v and s have particular values is a regular cycle of constant amplitude produced. To achieve this in the model developed here, v must equal 1.0 for all values of s. Such stringent requirements are unlikely to be fulfilled in practice, yet actual cycles have been fairly regular and have displayed no tendency to either die out or explode.

In order to provide a more satisfactory theory of the trade cycle the multiplier–accelerator relationship has been modified and supplemented in various ways. The relationship can be made non-linear by allowing s and v to vary over the cycle. The theory of permanent income rationalises pro-cyclical (that is, moving with output) movements in the MPS. The gap between the desired and actual capital stock can be expected to get smaller as the boom proceeds. Both factors reduce the rate at which aggregate demand grows and can thus rationalise an upper turning-point in the cycle.

An alternative but mutually compatible way of generating cycles is to limit the explosive path of the multiplier–accelerator interaction by imposing floors and ceilings to the level of real output. This is Hicks's [5] solution. The floor is set by autonomous investment and the ceiling is determined by limitations on the quantity of labour supplied and on the capacities of the capital-goods industries.

The role of money was not entirely rejected by Keynesian trade-cycle theorists. When included (as by Hicks [5]) it played a secondary role. In a boom the growth in the demand for money relative to its supply caused interest rates to rise and so helped to dampen investment. In the depression when the demand for money had fallen relative to income, low interest rates might help to stimulate investment once profit expectations had recovered somewhat.

14.2 The monetarist interpretation of trade cycles

The monetarist view, which stretches back to classical writings, holds that, while money has no long-run effect on real national output, in the short run monetary disturbances can exert powerful independent influences on real output. The relationship between money and income in the USA has been thoroughly documented by Friedman and Schwartz [6]. They single out four periods of economic stability, 1882–92, 1903–13, 1923–9 and 1948–60, during which real income grew at similar rates but the rate of change of the price level varied between minus 1 per cent and plus 2 per cent. The variation in the rate of change of the price level is explained by differences in the rate of growth of the money supply. They consider that the long-run rate of growth of real output and the money supply are largely independent.

Short-run monetary fluctuations are associated with similar fluctuations in real output. Friedman and Schwartz consider that the primary causal link goes from monetary changes to variations in real output. They believe this is particularly important in the explanation of severe contractions. Six of the contractions are classified as severe: 1873–9, 1892–4, 1907–8, 1921–2, 1929–33 and 1937–8. In the severe contractions the money supply decreased and four of them were accompanied by a banking crisis.

Since the money supply has generally continued to increase during less severe recessions. Friedman and Schwartz relate fluctuations in the rate of change of the money supply to variations in economic activity. One such exercise involves comparing the reference cycle with a step function in the rate of change of the money supply. Each step occurs when the rate of change of the money supply passes from a high to a low rate and vice versa. On average decreases in the rate of change of the money supply precede the reference peak by seven months, while increases in the rate of change of the money supply lead the reference-cycle trough by four months. The standard deviation of the lead in the money supply series is eight months at the peak and six months at the trough.

From this and other evidence Friedman and Schwartz conclude that changes in the rate of growth of the money supply cause changes in the same direction in real output which occur after quite a long and variable lag. It is this type of evidence which leads Friedman to eschew the use of discretionary monetary policy since it is likely in these conditions to be unsuccessful.

The arguments supporting the Friedman position have been the subject of controversy, particularly relating to whether changes in the money supply primarily depend on national income or occur independently. Another issue concerns the timing evidence. Friedman himself recognises that the timing of a relationship whereby variable X leads variable Y by no means justifies the conclusion that X causes Y. Tobin (see [8] and [9]) and others have devised models in which income is the causal factor but money leads income, or in which money is the prime-mover but income leads money. Although Friedman's theoretical underpinning of a causal and timing relationship that goes from money to income is not fully worked out, the alternative possibilities, such as an increase in output inducing a rise in the money supply several months before the increase in output actually occurs (Kalder [10]) seem less plausible.

In support of his contention that money supply changes occur independently of output and produce disturbances in output Friedman cites three occasions (January to June 1920, October 1931, and July 1936 to January 1937), when the Federal Reserve System[1] deliberately followed a restrictionary policy. This led to a sharp contraction in the money supply to be followed by a severe contraction in industrial output.

Cagan [11] provides further detailed evidence for the USA of the cyclical behaviour of the determinants of the money stock. He estimates that about one-half the variation of the money supply about its trend is accounted for by the currency–money supply ratio, one-quarter by the reserve ratio and one-fifth by 'high-powered' money. The latter variable has become more important in recent cycles. Since 'high-powered' money is subject to government control it can vary independently of income, unless the government chooses to allow it

[1] The central bank of the USA.

to vary with income. This leaves us with the currency and reserve ratios as variables which could depend on income and thus cause money to dance to the tune played by income.

As the demand for money grows more (less) rapidly in relation to its supply, the rate of interest is predicted to rise (fall). If a rise (fall) in the interest rate decreases (increases) the currency or reserve ratio, the money supply will expand (contract), given a constant 'high-powered' money base (see Chapter 9). In this event the direction of causality is from income to the money supply so that the latter is endogenous.

Cagan considers the evidence to support the interest-rate mechanism outlined above to be weak but finds other ways in which the state of business activity influences the currency and reserve ratios. As the expansion approaches the peak the currency ratio tends to rise. (The reasons for this are not clear.) The rate of growth of the money supply falls and checks the growth of output. Once the contraction sets in both ratios rise. Banks become more cautious about lending and wish to strengthen their liquid assets position. The public, experiencing similar uncertainties, raise their desired currency ratio. Both these factors lower the money supply (or alternatively its rate of growth) and strengthen the forces of contraction. This process may become self-generating, as the attempt by banks and the public to make their asset portfolios more liquid drains banks of cash reserves. Banks then need to liquidate more assets and the public, losing confidence in the banks, draw out more cash. A full-scale bank panic can be triggered off which exacerbates the contraction. Monetary factors can therefore contribute to cyclical disturbances in economic activity, both by occurring independently and by being related to changes in national income.

There are two aspects to the Keynesian argument that money is of little importance in influencing the level of economic activity in the short run. One, which we have just discussed, is that the money supply should be determined endogenously by national income. The other is that national income should be unresponsive to changes in the money supply. This necessitates velocity changing so that money supply changes are rendered ineffective. An increase in the money demand, accompanied by a reduction in the rate of growth of the money supply such as might occur at the peak, causes higher interest rates, which drive up velocity. This means people are financing each pound's worth of annual income with less money. If money is to have little influence, the changing demand for money with respect to output that occurs over the trade cycle must be accompanied by large pro-cyclical variations in velocity. Also, the change in interest rates which accompanies the change in velocity should have little effect on expenditure.

The observation that velocity rises in the upswing and falls in the downswing therefore weakens the case for the powerful influence of money. Friedman [12] distinguishes between measured velocity, the usual calculation of velocity, which is current income divided by the stock of money, and desired velocity, which is permanent income divided by the money stock. Since permanent income changes less over the cycle than measured income, desired velocity will vary less than it appears to do when measured inappropriately.

14.3 Conclusion

The Keynesian emphasis on the causal role of real variables, particularly investment, in generating trade cycles places the responsibility for cyclical fluctuations firmly with the private sector. This contrasts with the monetarist view that monetary disturbances are the major factor in causing cyclical fluctuations. The private sector is inherently stable but is subjected to monetary shocks brought about by the authorities. For instance, the 1929–33 depression is primarily attributed to the inappropriate policy of monetary contraction adopted by the Federal Reserve System following the Wall Street crash.

References

[1] E. Shapiro, 'Cyclical Fluctuations in Prices and Output in the UK 1921–71', *Economic Journal*, 86 (December 1976).
[2] J. Schumpeter, *Business Cycles* (New York: McGraw-Hill, 1939).
[3] P. Samuelson, 'Interactions between the Multiplier Analysis and the Principle of Acceleration', *Review of Economics and Statistics*, 21 (1939).
[4] L. A. Metzler, 'The Nature and Stability of Inventory Cycles', *Review of Economics and Statistics* (1941); reprinted in Gordon and Klein (eds), *Readings in Business Cycles* (London: Allen & Unwin, 1966).
[5] J. R. Hicks, *A Contribution to the Theory of the Trade Cycle* (Oxford University Press, 1950).
[6] M. Friedman and A. J. Schwartz, *A Monetary History of the United States 1867–1960* (Princeton University Press, 1963).
[7] M. Friedman and A. J., Schwartz, 'Money and Business Cycles', *Review of Economics and Statistics*, 45, 1, 2 Supplement (February 1963).
[8] J. Tobin, 'The Monetary Interpretation of History', *American Economic Review*, 55 (June 1965).
[9] J. Tobin, 'Money and Income: Post Hoc Ergo Propter Hoc', and M. Friedman, 'Comment on Tobin', *Quarterly Journal of Economics*, 84 (May 1970).
[10] N. Kaldor, 'The New Monetarism', *Lloyds Bank Review* (July 1970).
[11] P. Cagan, *Determinants and Effects of Changes in the Stock of Money, 1875–1960* (New York: National Bureau of Economic Research, 1965).
[12] M. Friedman, 'The Demand for Money: Some Theoretical and Empirical Results', *Journal of Political Economy*, 66 (1958).

15 Economic Growth

The short-run general macro model developed in section I was based on the following simplifying assumptions:

(1) The time period involved in the analysis was so short that the net investment taking place had no significant effect on the size of the economy's capital stock. This allowed us to concentrate on the effect which that net investment has on the size of aggregate demand while neglecting its effect in increasing the economy's productive capacity.

(2) A constant population. This allowed us to concentrate on the factors underlying the demand and supply of labour from a constant population. However, during the last century and a half the population of almost every country has been increasing at varying geometric rates. This has resulted in an increase in the labour force over time and has been another factor in increasing the productive capacity of the economy.

(3) An unchanging level of technical knowledge. This assumption also needs to be relaxed as increasing technical knowledge has resulted in a continual decrease in the real resources required per unit of output as well as in the introduction of new products. Technical change has been a major historical factor in increasing the economy's productive capacity. It has also stimulated the rate of growth of demand through its impact on the rate of investment.

It is now necessary to extend the earlier analysis in order to examine the effect that capital accumulation, population growth and technical progress have on the equilibrium solution for the economy. The reason for concentrating on equilibrium solutions are as follows:

(a) Equilibrium dynamic solutions are relatively easy to set up and solve, while disequilibrium dynamic systems are much more difficult to deal with as their behaviour depends on the error-adjusting mechanisms and lag structures specified.

(b) Neoclassical economists believe that the economy is inherently stable and that it tends to return to some full-employment growth path if it is displaced from this path. On the other hand, some Keynesians and neo-Keynesians believe that a capitalist economy is inherently unstable, with such stringent sufficient conditions for equilibrium full-employment growth that it is unlikely to be achieved. In order to throw some light on this difference of opinion it is necessary to examine the conditions necessary for the existence of full-employment equilibrium growth and the stability of this equilibrium growth path.

15.1 The impact of capital accumulation and population growth

We now consider what impact capital accumulation and population growth have on the economy. For the time being technical knowledge is assumed fixed. It must therefore be kept in mind that the conclusions of this section relate to an economy in which technical change is absent and some of them would be modified by the presence of technical progress. This is examined later in the chapter.

Capital accumulation and population growth mean that the quantity of available inputs in the aggregate production function increases over time. The concept of an aggregate production function was introduced in Chapter 5 when discussing the supply side of the economy. It is written as

$$y = f(L, K, T) \tag{15.1}$$

where y is the real output per period of time, L is the flow of labour services per period of time, K is the capital stock which yields a proportionate flow of capital services per period of time, and T is the state of technical knowledge which we assume fixed. The law of variable proportions is expected to hold in this economy, so that in the absence of technical change an increase in only one type of input, the other remaining fixed, will lead to a less-than proportionate change in output and further increases in this input will lead to diminishing increases in output.

For simplicity we assume that the economy as a whole experiences constant returns to scale, so that a given proportionate increase in both capital and labour inputs per period of time will lead to the same proportionate increase in output per period. This implies that in the economy the size of the market for each industry's products is much larger than the capacity output of the plant with the lowest unit costs in the industry. Constant returns to scale for the economy as a whole implies that the aggregate production function is linearly homogeneous in capital and labour inputs. (See the mathematical appendix to this chapter, *note 1*, for the implications of linear homogeneity in the production function and for the proofs relating to the following discussion.) This means that in the absence of technical change the production function can be written as

$$\frac{y}{L} = F\left(\frac{K}{L}\right) \quad \text{or} \quad y = LF\left(\frac{K}{L}\right) \tag{15.2}$$

Thus the average product of labour in the economy, y/L, is an increasing function of the capital–labour ratio, given that factor proportions are variable.

However, the operation of the law of variable proportions allows us to deduce that y/L will increase at a diminishing rate as the capital–labour ratio

increases. Therefore, a proportionate increase in both labour and capital inputs will keep average output per head constant, while a greater increase in labour inputs to capital inputs will decrease average product per head. This is illustrated in Figure 15.1.

The law of variable proportions implies that the marginal product of a factor decreases as relatively more of that factor is applied to the production of output. This leads to the deduction (see appendix, *note 1*) that the marginal product of capital decreases with increases in the capital–labour ratio, while the marginal product of labour decreases when the capital–labour ratio decreases. Therefore, in a competitive economy where profit-maximising entrepreneurs hire factors up to the point where their price is equal to their marginal value product, the equilibrium real wage and real interest rate will change with changes in the capital–labour ratio.

If the capital stock is growing faster than the labour force, real interest rates will fall and real wages will rise, while in the case where the labour force is growing faster than the capital stock, real interest rates will tend to rise and real wage rates fall. In both cases the increase in inputs into the production process results in a growth in output. However, output only grows at the same rate as any of the inputs if both inputs are growing at the same rate. In this case relative factor prices will remain constant as the capital–labour ratio stays unchanged.

The analysis in earlier chapters concentrated on the impact of net investment on aggregate demand and neglected its effect on the economy's productive capacity. In growth models explicit account is taken of the productive impact of positive net investment which equals the increase in the capital stock. Therefore

$$I = dK/dt \tag{15.3}$$

A constant labour force in the presence of net investment means that the rate of growth in output capacity will be below the rate of growth of the capital stock because of decreasing marginal returns. If incentives for entrepreneurs to

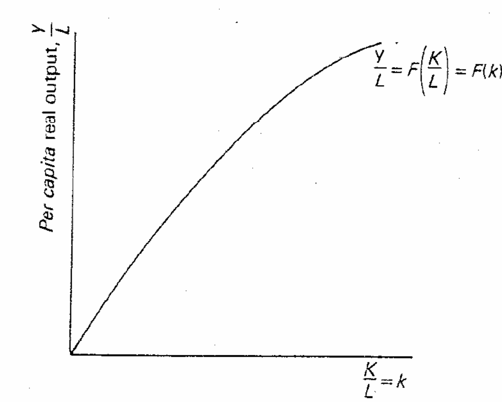

FIGURE 15.1 *Average output per unit of labour input as a function of the capital–labour ratio*